Johannes Bitter, BSc

Characterization of the *Komagataella phaffii CTA1* promoter

MASTER THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree program: Biotechnology

submitted to

Graz University of Technology

Supervisor

Ao.Univ.-Prof. Mag.rer.nat. Dr.rer.nat. Anton Glieder

Institute of Molecular Biotechnology

Faculty of Technical Chemistry, Chemical and Process Engineering, Biotechnology

Graz, April 2019

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the used sources. The text uploaded to TUGRAZ online is identical to the master thesis.

Graz, April 2019 

(signature)

EIDESSTATTLICHE ERKLÄRUNG

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

Graz, im April 2019 

(Unterschrift)

DANKSAGUNG

An erster Stelle möchte ich mich bei meinem Betreuer Ao.Univ.Prof. Mag.rer.nat. Dr.rer.nat. Anton Glieder für die Begleitung während meiner Masterarbeit und das Verschaffen der Möglichkeit diese am Institut für Molekulare Biotechnologie durchzuführen bedanken. Weiters möchte ich mich an meine Kollegen Antonia Volpini de Maestri und Katharina Ebner richten, mit denen ich das Projekt teilen durfte und wir somit in einem Boot saßen. Zahlreiche kreative Inputs, fruchtbare Diskussionen und gegenseitiges Hochhalten der Moral waren unumgänglich für ein erfolgreiches Bestreiten der Arbeit.

Weiters gebührt der gesamten Glieder-Group, den Mitarbeitern des Instituts, sowie bisy e.U. Dank. Mit euch gestaltete sich der Laboralltag so leichtfällig und unterhaltsam wie man es sich nur wünschen kann. Im Besonderen bedanke ich mich bei Anna Hatzl, Christian Schmid, Bianca Huber und Amneris Schlager-Weidinger für die stetige Unterstützung und ein offenes Ohr bei Fragen.

Ich möchte mich bei all meinen Studienkollegen und Freunden, die mich in den letzten sechs Jahren begleitet haben und Teil dieses großartigen Lebensabschnittes waren, bedanken. Ihr habt auch außerhalb des Labors die Zeit zu einer schönen gemacht.

Zuletzt kann ich mich bei meiner Familie gar nicht genug bedanken. Dies gilt nicht nur für die notwendige finanzielle Unterstützung durch meine Eltern Julius und Ulrike, sondern auch für die Geduld, persönliche Annahme und das Schaffen von Möglichkeiten in den vielen Jahren. Vielen Dank an meine „kleine“ Schwester Veronika, die mich mit ihrer „Hands-on“ Mentalität und ihrem Durchhaltevermögen immer wieder dazu inspirieren konnte, meine Ziele zu erreichen.

ABSTRACT

The methylotrophic yeast *Komagataella phaffii* (*Pichia pastoris)* is already established in industrial biotechnology as a well renowned microorganism for recombinant protein production. Recently, the attention had shifted more and more towards applying the yeast in metabolic engineering and synthetic biology. These applications require a set of differently regulated promoters, diverse in sequence to avoid *in vivo* homologous recombination. The methanol utilization pathway (MUT) offers more than capable tools with a set of highly diverse promoters, which are repressed on current carbon sources (e.g. glucose or glycerol) and exhaust their full potential when methanol is added. Though, the application of methanol poses a risk due to the substance's hazardous nature. Some promoters of the *P. pastoris* MUT pathway reach promising expression levels by simple carbon source depletion connected to promoter de-repression. Thereof, the *CTA1* (*CAT1)* promoter is most outstanding, competing with the prominent *AOX1* promoter on several levels and even exceeding it in terms of extracellular secretion for specific targets such as lipase production. Thus, it was the goal of this thesis to learn more about the *CTA1* promoter, testing it in different implementations and evaluating it for methanol-free applications. Human growth hormone (hGH) and *Candida albicans* lipase B (CalB) were used as models for different target proteins and successfully secreted into the media employing *CTA1* promoter-based expression cassettes.

KURZFASSUNG

Der methylotrophe Hefeorganismus *Komagataella phaffii* (*Pichia pastoris)* ist in der Biotechnologieindustrie mittlerweile als herausragender Mikroorganismus für rekombinante Proteinproduktion etabliert. Letztlich wird der Organismus immer mehr in den Bereichen „Metabolic Engineering" und „Synthetic Biology" implementiert. Diese neuen Anwendungsgebiete erfordern einer Palette an unterschiedlich regulierten Promotern, die sich in ihrer Sequenz unterscheiden, um *in vivo* homologe Rekombination zu vermeiden. Der „Methanol Utilization Pathway" (MUT) bietet das nötige Werkzeug mit einer Fülle an vielfältigen Promotern, die unter gängigen Kohlenstoffquellen (z.B. Glucose oder Glycerin) reprimiert und durch Methanol stark induziert werden. Die Anwendung von Methanol stellt jedoch durch seine gefährliche Natur ein gewisses Risiko dar. Daher gilt besondere Aufmerksamkeit den Promotern des *P. pastoris* „MUT Pathways", die unter einfacher Erschöpfung der Kohlenstoffquelle und der einhergehenden Derepression vielversprechende Expressionslevel erreichen, allen voraus der *CTA1* (*CAT1)* Promoter. Dieser erzielt Resultate vergleichbar mit dem prominenten *AOX1* Promoter und übertrifft ihn sogar hinsichtlich Anwendungen wie etwa der extrazellulären Lipase Produktion. Darum galt es in dieser Masterarbeit den *CTA1* Promoter zu charakterisieren, ihn in unterschiedlichen Bedingungen zu testen und für Methanol-freie Anwendungen zu evaluieren. Humanes Wachstumshormon (hGH) und *Candida albicans* Lipase B (CalB) wurden als Modellenzyme verwendet und mit *CTA1* Promoter basierten Expressionskassetten erfolgreich ins Zellmedium sekretiert.

Table of Contents

1. INTRODUCTION

Over the last 40 years, the methylotrophic yeast *Komagataella phaffii* (previous name: *Pichia pastoris*) has gradually developed into one of the most prominent microbial organisms for heterologous gene expression. It has been used to clone and express over 1,000 genes encoding for recombinant proteins (1) and is by now even more favored for single heterologous protein production than the well-established and genetically better characterized yeast *Saccharomyces cerevisiae* (2). The growing field of *P. pastoris* applications however requires an expanded set of molecular tools, including strong promoters for the efficient expression of valuable enzymes.

1.1 *Pichia pastoris:* a promising host for recombinant protein production

1.1.1 *Pichia pastoris* and the *Komagataella* genus

At the beginning of the 21st century two yeast species were identified as closely related to *P. pastoris* and they were all grouped under the *Komagataella* genus. Thus, *P. pastoris* was renamed into *Komagataella pastoris*, alongside *Komagataella phaffii* and *Komagataella pseudopastoris* (3). Later on, it was discovered that early *P. pastoris* strains actually contained two different strains, deposited at the US Agricultural Research Service (ARS) strain collection under the numbers NRRL-Y11431 (*K. pastoris*) and NRRL-Y11430 (*K. phaffii*). So, *K. phaffii*, which was originally thought to be an independent yeast species, was actually the *P. pastoris* strain preferentially used for biotechnological and cloning processes (4). In the following thesis I will continue to use the nomenclature "*Pichia pastoris*" for reasons of simplification.

1.1.2 Growth to high cell densities

P. pastoris was first applied in the 1970s by Phillips Petroleum for the production of single cell protein (SCP). SCP was manufactured as an animal feed additive, which was produced in a continuous fermentation by exploiting two major advantages of *P. pastoris*, i.e. the cultivation to high cell densities (> 130 g dry cell weight per liter fermentation broth) (5) leading to a high amount of product and the ability of *P. pastoris* to utilize cheap methanol as its sole carbon and energy source (6). The former is due to the reason that *P. pastoris* is an obligate aerobe. When a microorganism is grown to high cell densities, the metabolic demand is high, and oxygen can be a limiting factor. In this case facultatively aerobic organisms, such as *Saccharomyces cerevisiae*, alter their metabolism to fermentation and start producing ethanol as

a side product. This can swiftly rise to toxic levels, inhibiting further cell growth and protein production or even leading to cell death. On the other hand, *P. pastoris* remains at respiratory growth even under oxygen limitation, allowing high cell culture concentrations (7,8). Since according to Cereghino and Cregg (9) the level of secreted protein in the medium is approximately proportional to the density of the culture, this is of tremendous advantage for high product titers. Also, the fact that *P. pastoris* belongs to the group of Crabtree negative yeasts, which are known for their low energy maintenance demand, is favorable for a high cell density growth (10).

1.1.3 The Methanol Utilization Pathway

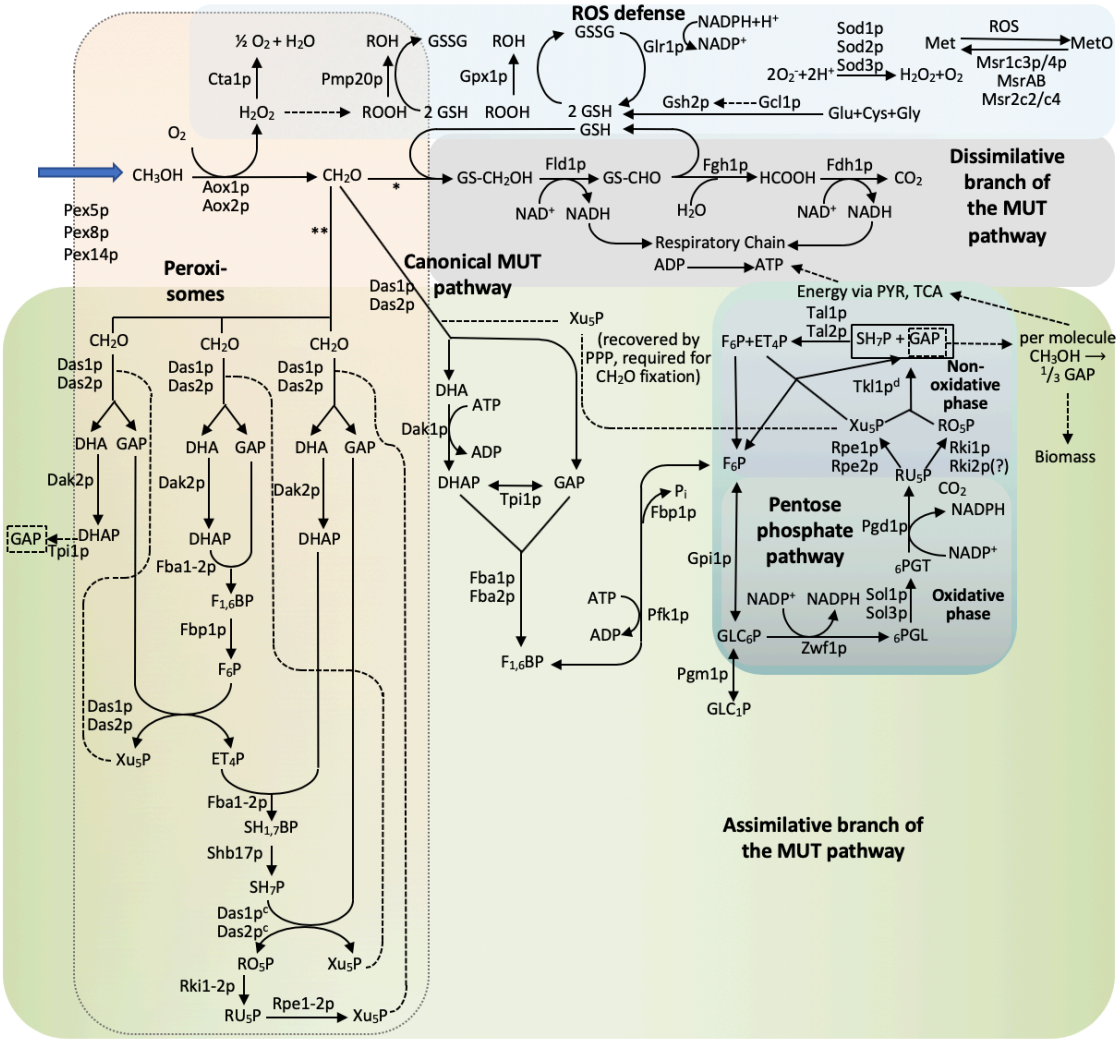


Figure 1.1 Methanol utilization (MUT) pathway of *Pichia pastoris* divided into canonical and noncanonical parts (Figure based on drawings from Vogl *et al.* (11) was complemented with an alternative Xu_5P forming pathway located in the peroxisome based on the publication from Rußmayer *et al.* (2015) (12)). The canonical MUT pathway annotation (13) and physiological interactions (14) were drawn according to different studies. The canonical MUT pathway is constituted of the compartmentalized peroxisomes, the assimilative branch – containing the pentose phosphate pathway (PPP) – and the dissimilative branch. The PPP is divided into the oxidative phase regenerating NADPH and the non-oxidative phase. An alternative Xu_5P forming pathway entirely located in the peroxisome is displayed. The reactive oxygen species (ROS) defense was annotated according to reference (15). Proteins associated with peroxisome biogenesis Pex5p, Pex8p and Pex14p are also shown. The alternative steps of the dissimilative MUT pathway are overlooked in this figure. Abbreviation of main enzymes and metabolites: AOX, alcohol oxidase; CTA, catalase; FLD, formaldehyde dehydrogenase; FGH, formyl glutathione hydrolase; FDH, formate dehydrogenase; DAS, dihydroxyacetone synthase; TPI, triosephosphate isomerase; DAK, dihydroxyacetone kinase; FBA, fructose-1,6-biphosphate aldolase; FBP, fructose-1,6-biphosphatase; SHB17, seduheptulose-1,7-bisphosphatase; RKI, ribose-5-phosphate ketol-isomerase; RPE, ribulose-5-phosphate 3-epimerase; DHA, dihydroxyacetone; DHAP, dihydroxyacetone phosphate; $F_{1,6}BP$, fructose-1,6-bisphosphate; P_i, inorganic phosphate; F_6P, fructose-6-phosphate; ET_4P, erythrose 4-phosphate; GLC_6P, glucose-6-phosphate; GLC_1P, glucose-1-phosphate; $_6PGL$, 6-phosphogluconolactone; $_6PGT$, 6-phosphogluconate; RU_5P, ribulose-5-phosphate; Xu_5P,

xylulose-5-phosphate; RO_5P, ribose-5-phosphate; GAP, glyceraldehyde-3-phosphate; SH_7P, sedoheptulose-7-phosphate; TCA, tricarboxylic acid cycle; PYR, pyruvate; GSH, glutathione; GSSG, oxidized glutathione self-dimer; MetO, methionine sulfoxide; $SH_{1,7}BP$, seduheptulose-1,7-bisphosphate. *, The reaction of CH_2O and GSH is occurring nonenzymatically; **, In the alternative Xu_5P forming pathway three methanol molecules are required for the formation of one GAP molecule.

The latter advantage of *P. pastoris*, the capability of metabolizing methanol, is enabled by the methanol utilization (MUT) pathway (**Figure 1.1**). This is a cascade of enzymes found in all methylotrophic yeasts, partly governed by very strong and tightly regulated promoters, which offer a set of tools for the production of recombinant proteins (16,17). The first step is catalyzed by alcohol oxidase (AOX), where the C1-alcohol methanol is oxidized to formaldehyde and hydrogen peroxide under oxygen consumption (18). Since the side product, hydrogen peroxide (H_2O_2), is toxic and produced in equimolar amounts to formaldehyde, catalase (CAT or CTA) is required to break down the compound into oxygen and water, as part of the stress response against reactive oxygen species (ROS) (15,19). Therefore, methanol oxidases and catalase are sequestered in close vicinity to each other and the first two reaction steps are compartmentalized in peroxisomes (18,20). Those cell organelles can make up to 80% of the cytoplasm when cell metabolism is changed to methanol utilization to counteract to cell stress (21). In order to confine H_2O_2 to the organelles, peroxisomal catalase enzymes are primarily gathered at the peripheral part of the peroxisomes, forming a barrier like cluster to the cytoplasm (22,23). In contrast to baker's yeast *S. cerevisiae*, there seems to be only one catalase gene in the *K. phaffii* genome, which shows similarity to the peroxisomal catalase CTA1 of *S. cerevisiae* (11,24). After formaldehyde production, the pathway then separates into an assimilative and dissimilative branch. The former is to convert C1 to C3 sources, which are further processed in the cytosol for biomass formation and the latter produces carbon dioxide, NADPH as energy form and detoxifies residual formaldehyde. In the assimilative branch of the MUT pathway xylulose 5-phosphate (Xu5P), which is originally thought to be solely generated by the pentose phosphate pathway, is condensed with formaldehyde to the C3-compounds dihydroxyacetone (DHA) and glyceraldehyde 3-phosphate (GAP). The reaction is catalyzed by dihydroxyacetone synthase (DAS), a special transketolase (25–28). Recent studies suggested an alternative Xu5P forming and methanol assimilation pathway located in the peroxisomes (12). This makes for a more efficient metabolism, since the sugar phosphate is not required to be transferred into the peroxisome for formaldehyde fixation. The novel cycle uses an isoform set of the enzymes fructuose-1,6-bisphosphate aldolase (13), transaldolase, ribose-5-phosphate ketol-isomerase

and ribulose-5-phospate 3-epimerase, which is distinctively designated to peroxisomes with a PTS1 peroxisomal targeting sequence (29). These enzymes are labeled Fba1-2, Tal1-2, Rki1-2 and Rpe1-2, respectively. In the pathway suggested by Rußmayer (12) DAS catalyzes the interconversion of fructose-6-phosphate (F6P) and GAP to Xu5P and erythrose-4-phosphate (ET4P). In a consequent reaction Fba1-2 or Tal1-2 initiates the condensation of ET4P and dihydroxyacetone phosphate (DHAP) to sedoheptulose-1,7-bisphosphate (SH1,7BP). This is then dephosphorylated and eventually converted into two additional Xu5P molecules by enzymes Rki1-2 and Rpe1-2. Therefore, a total of three Xu5P molecules are at disposal for fixation of three formaldehyde molecules resulting in a net reaction of one GAP molecule per three methanol molecules. In the dissimilative branch formaldehyde undergoes two subsequent reactions, each catalyzed by dehydrogenases, to be oxidized into CO_2. Both reactions take place in the cytosol and produce NADH, which is utilized for energy production when the methylotrophic yeast is cultured on methanol (30,31).

1.1.4 Inducible Promoters

AOX is produced by the expression of two genes, *AOX1* and *AOX2*. Although both protein sequences are more than 90% identical in their protein-coding areas, the former is transcriptionally regulated by a far stronger promoter (P_{AOX1}) (32). In *P. pastoris*, P_{AOX1} is tightly repressed and no AOX enzymes are produced when the methylotrophic yeast is cultured on glucose, glycerol or ethanol, putatively because the cell is in no need of metabolizing methanol and avoids exposing itself to unnecessary stress (33). The promoter is slightly de-repressed upon carbon source depletion, so the cell can prepare for a switch to methanol metabolism (34–38). This includes foremost the production of minor amounts of AOX to avoid accumulation of and intoxication with methanol, which would lead to AOX inhibition and cell death (9). However, P_{AOX1}, as well as many other MUT pathway promoters, is only fully induced upon methanol addition (17). Compared to promoter repression, where mRNA levels of the *AOX* gene are undetectable, they make up about 5% of total mRNA and AOX protein makes up 30% of the total soluble proteins in induced state (39). This overexpression is associated with the poor affinity of AOX for oxygen, which is required for the conversion of methanol to formaldehyde. Thus, to counterbalance, the cell is programmed to produce extensive amounts of the protein to avoid hazardous methanol accumulation (40). This biphasic behavior of P_{AOX1}, i.e. the strong induction on methanol and the tight repression on any other carbon sources, is ideal for a two-phase cultivation. Growth and production phase are separated, offering the expression system an

effective channeling of the metabolic flux. In the first phase biomass is accumulated without posing additional stress on the cells due to recombinant protein production. Subsequently, proteins are synthesized, allowing even the production of proteins toxic to the host organism (41,42).

1.1.5 Applications of *P. pastoris* for recombinant protein production

Based on the properties mentioned above, P_{AOX1} is well-suited for the expression of recombinant genes and is hence to date the most applied promoter and was also the first being used with the *P. pastoris* recombinant expression system (41,43). Since production of SCP turned uneconomical because of increasing gas prices during the oil crisis in 1973, Phillips Petroleum Company consolidated with Salk Institute Biotechnology/Industrial Associates, Inc. (SIBIA, La Jolla, CA, USA). They developed an expression system, which combined the initial SCP fermentation process, strains and vectors designed by SIBIA, and the tightly regulated P_{AOX1}. This led to remarkably high titers of recombinant proteins (43,44) and the first large-scale implementation in industry in the 1990s. One of those was the recombinant production of the plant enzyme hydroxynitrile lyase from *Hevea brasiliens* (Hb HNL) with titers of > 20 g target protein per liter fermentation broth. This recombinant enzyme was used to catalyze the stereoselective synthesis of enantiopure m-phenoxybenzaldehyde cyanohydrin, which was further used for the production of synthetic pyrethroids (45). The accessibility of the expression system as a simple kit form, commercialized by Invitrogen (by Thermo Fisher Scientific, Inc., Waltham, MA, USA), allowed the scientific community to research the organism and develop new fields of application (41). Hence, it was not long until further commercial usage was made of *P. pastoris*. Major breakthroughs took place in 2006, as the Food and Drug Administration (FDA) classified phospholipase C, an animal feed additive protein, as GRAS (generally recognized as safe). This step was distinctive for the benign nature of *P. pastoris* recombinant products, which are free of endotoxins, as well as oncogenic and viral DNA (46,47). Further, Kalbitor®, a kallikrein inhibitor, was FDA approved as first recombinant biopharmaceutical product in 2010, acting as launching pad for many more to come (48). Other prestigious products are Phytase (Phytex, Sheridan, IN, USA), which is an animal feed additive that processes plant derived phytate to release a natural phosphate source for the cattle. Trypsin (Roche Applied Science, Germany) is applied as protease in the study of proteomics. Nitrate reductase (The Nitrate Elimination Co., Lake Linden, MI, USA), helps with water analyses and waste water treatment and Collagen (Fibrogen, San Francisco, CA, US) is utilized in medical research. An extensive list of

heterologous proteins produced by *P. pastoris* is featured at the web page *www.pichia.com* (15.03.2019), as well as in the report of Cregg and Cereghino (1). Besides developing new *P. pastoris* recombinant products, research has also been focusing on optimizing strains to facilitate the implementation of expression systems into the host. This includes protease-deficient strains to ensure better product stability, glyco-engineered strains for improved human-like structures and auxotrophic strains for easier selection (49–52).

1.1.6 Combining prokaryotic and eukaryotic advantages

P. pastoris is broadly applied and a subject of research, because it perfectly combines both positive prokaryotic and positive eukaryotic features. That is, the ease in cultivation and genetical modification on the one side and the ability to perform typically eukaryotic posttranslational modification and extracellular secretion on the other.

As opposed to mammalian or insect cells, *P. pastoris* can be cultivated to high cell densities using only a small set of media components and trivial culture conditions. This includes biotin, salts, certain trace elements, water, and a primary carbon source, such as glycerol, glucose or methanol. Additionally, the fermentation can be conducted at a rather acidic pH, so contamination with other species is rather improbable, facilitating sterilization. These conditions allow a relatively cheap production process. Also, since the medium constitution lacks components that might degrade to toxins or pyrogens, heterologous production of human biopharmaceuticals is possible and hazards are minimized (9,49).

When trying to secrete heterologous proteins derived from a eukaryotic genome, limitations are met using prokaryotes as an expression host. That is, because some of these proteins require correct folding to obtain an active state, modifications such as glycosylation or phosphorylation, or extracellular secretion when the product is toxic (53). Conventional microorganisms such as *Escherichia coli* do not offer the necessary tools for that kind of processing. The reducing milieu of the cytoplasm and insufficient concentration of chaperons exacerbate disulfide bond formation and the absent glycosylation machinery alters the function of the protein (54–57). This can lead to the formation of inclusion bodies or inactive enzymes (45), reducing process efficiency, and thus limiting the variety of producible proteins. *P. pastoris* offers a feasible solution to this problem, since it possesses a secretory pathway (47). Therein, a secretion signal guides the translated pre-mature protein through the endoplasmic reticulum (ER) and the Golgi apparatus. In the lumen of the ER, glycosylation is initiated and final disulfide bonds are formed and rearranged, leading to the active final protein structures

(53,58). In the Golgi the signal sequence is cleaved off and the readily protein is packed in vesicles for extracellular transport (59).

The extracellular secretion is the second most prominent posttranslational modification to proteins, succeeding glycosylation, and is concurrently a significant advantage of the *P. pastoris* expression system. The recombinant protein is released directly into the culture medium and the amount of secreted native proteins in *P. pastoris* is relatively small. Thus, the downstream processing is facilitated enormously, reducing time and cost. Further, the extracellular secretion allows the production of compounds toxic to the organism, since no intracellular accumulation takes place (47,60). For a proper guidance of the nascent protein into the culture medium a signal sequence is required. In the *P. pastoris* expression system, the *S. cerevisiae* derived alpha mating factor (α-MF) is mostly preferred over the native *P. pastoris* acid phosphatase (PHO 1) signal peptide, since the former has been reported to reach stronger secretion levels (9,53). The α-MF is constituted of a pre- and a pro-sequence, with a length of 19 and 60 amino acids, respectively (53). The pre-region guides the translated pro-protein into the ER lumen and is cleaved off by signal peptidase (61), while the pro-region is responsible for directing the residual protein into the Golgi apparatus and is subsequently processed by the endo-protease *Kex2p* (62). Before the mature protein is directed onto the cell surface, the separation of those regions is essential to avoid signal peptide residues on the N-terminus, which might decrease product yield or activity (63). An alternative to the expected posttranslational ER import is the use of the signal peptide of the OST1 protein, which was described to be responsible for a co-translational ER targeting (64).

1.2 New applications require new tools

After 40 years of successful establishment for its use in industry and academia, the boundaries of possible applications for *P. pastoris* have been pushed. Besides conventional expression of one recombinant gene, the attention has been shifting towards utilizing the yeast for metabolic engineering and implementing biosynthetic pathways (65). Namely, introducing a set of genes into the organism to comprise a novel pathway that can convert simple substrates into valuable products (66,67). In many areas, such as the synthesis of fatty alcohols, this approach could substitute the production via chemical synthesis, which is strongly dependent on petroleum resources and hence damaging the environment (68). Therefore, the microbial production of refined products needs to be optimized to become a competitive solution for the industry and make petroleum-based processes redundant. In order to enhance the meta-

bolic flux and thus increase product yield, it is essential to engineer and control the cascade of enzymes (69). In eukaryotes this is done mainly on the transcriptional level by using promoters of different strength (42), since it is not desired to have each gene in the pathway expressed at the same level. For example, to remove bottlenecks, an enzyme with poor affinity for the substrate should be governed by a stronger promoter and vice versa. One approach to receive promoters with a wide range of regulatory strength is creating a mutant library. By changing small parts of the wild type sequence, a vast set of variants can be created. Although they are highly similar in sequence and frequently also in respect to their mode of regulation, their strength can fluctuate from 10- to 1000-fold of the parental sequence (34,42). However, the use of a mutant promoter library for governing the expression of a set of genes can pose certain problems. The highly similar promoters can misalign *in vitro* in assembly cloning and also cause production strain instabilities, since parts of the vector or whole expression cassettes can be lost due to homologous recombination *in vivo* (70–72). As opposed to the mutant libraries, the promoters of the MUT pathway from *P. pastoris* are extremely diverse in their sequence and thus more suitable for overlap-directed DNA assembly (11). The identification of those MUT pathway genes and the isolation of their governing promoters was enabled by a number of full genome annotations in the last few years (13,73,74). Further, Vogl *et al.* conducted a comprehensive characterization of a total of 45 promising promoters derived from the *P. pastoris* MUT pathway – a set of promoters even vaster than *S. cerevisiae* was provided (11). They displayed diversity within their sequence, a large palette of different expression levels, and a great number was tightly regulated on methanol. One of the most outstanding promoters, was the one governing the expression of the peroxisomal catalase gene (*CAT1* or *CTA1*). However, although it is relatively strong, promoters exceeding the strength of the P_{AOX1} are desired, preferably even without relying on methanol as inductor.

1.2.1 *CTA1* Promoter

In older publications, the gene encoding for peroxisomal *P. pastoris* catalase was referred to as *CAT1*. However, in recent papers the term *CTA1* became rather established due to homologies to the peroxisomal *S. cerevisiae* catalase gene *CTA1* (13,24). To cohere with modern standards, I will continue with the latter nomenclature and abbreviate the peroxisomal *P. pastoris* catalase promoter with P_{CTA1}. Noteworthy, in the plasmid maps the catalase promoter sequences are still depicted as PCAT1.

The *CTA1* promoter (11) regulates the production of catalase (CTA), an enzyme that is part of the ROS defense and converts toxic hydrogen peroxide into water and oxygen (15,19). As H_2O_2 is produced as a consequence of methanol oxidation, P_{CTA1} is therefore highly upregulated on that substance. According to Vogl *et al.*, the expression levels are comparable to the tightly regulated P_{AOX1}, even exceeding it for example in case of recombinant production of *Candida antarctica* lipase B (CalB). Moreover, when using P_{CTA1}, expression levels elevate faster upon methanol addition (11). Catalase already needs to be present in the peroxisome before AOX converts methanol into formaldehyde, since also POX1 produces H_2O_2 as a side product of ß-oxidation (75) in the initial phase of starvation after depletion of easily convertible carbon sources such as glycerol or glucose. In that manner the cell is prepared to catabolize H_2O_2 immediately after its production when such stress is increased due to methanol induction and avoid hydrogen peroxide accumulation, which might harm the cell (76,77). The fast response of P_{CTA1} upon methanol induction could be linked to the increased concentration of TATA box motifs. Those are often enriched in stress-regulating genes, e.g. in *S. cerevisiae*, because they are highly and tightly regulated on the transcriptional level in presence of their coactivator SAGA (Spt-Ada-Gcn5-acetyltransferase). Overall, there is an abundance of TATA box motifs in MUT pathway core promoters to be found (11). This could be an adaption of the organism to cope with the stress of methanol metabolization or carbon source starvation by having tightly regulated genes – which again highlights their great qualification for heterologous protein production (11,78). Other than that, little is known about the regulatory mechanisms of P_{CTA1}, especially how the promoter is de-repressed upon carbon source depletion and induced after methanol or oleic acid addition. There can only be assumptions based on knowledge of similar promoters. For example, the promoter of the *S. cerevisiae* homologue is controlled by transcription factors (Adr1p, Oaf1p) that are sensitive to carbon source levels. Therefore, analyzing the catalase promoter of *P. pastoris* for possibly similar transcription factor binding sites within the P_{CTA1} sequence of *P. pastoris* and a possible dependency on carbon source-responsive transcription factors and stimuli might be interesting.

Besides the potential to become an alternative to the well-established P_{AOX1}, the *CTA1* promoter offers a possible methanol-free solution, due to its relatively strong de-repressed state. This is especially of interest for companies that conduct large-scale fermentations with *P. pastoris*, since they favor a methanol-independent process, due to the substance's flammable and health-damaging nature (41). So far, most applications of the *P. pastoris* expression system without methanol involved the use of a constitutive promoter (11),

although there is an increasing number of publications about alternatives (79–81). The prominent *glyceraldehyde-3-phosphate* promoter (P_{GAP}) is reported to be the strongest and most used amongst a meanwhile broad set of *P. pastoris* constitutive promoters, but nonetheless cannot match up to P_{AOX1} expression levels. More recent reports described the *GCW14* promoter as a more efficient alternative (82). An additional downside of constitutive promoters is that growth and production phase cannot be uncoupled (83). Thus, a strong and tightly regulatable promoter that reaches high expression levels without using methanol is required. P_{CTA1} is a suitable candidate, since it showed strong induction on oleic acid as an alternative substance. Moreover, among all MUT pathway promoters tested by Vogl *et al.*, P_{CTA1} was the strongest in de-repressed state after carbon source depletion, omitting the use of methanol (11). This is why the *CTA1* promoter was set as focus and starting point for engineering within this thesis and previous coherent projects.

1.2.2 Promoter engineering

The simplest and an efficient way to control metabolic flux and protein yield in yeast is regulation at the transcriptional level. Promoter engineering is thereby the appropriate tool to establish a vast range of expression levels for fine-tuning pathways or upregulate the expression of a single gene for heterologous protein production (42).

Eukaryotic promoters are quite complex and are usually characterized by two definite regions, the core promoter and the upstream regulatory sequence (URS), as displayed in **Figure 1.2**. The former is the essential part of the promoter which allows for transcription initiation by binding RNA polymerase II (RNAPII) and general transcription factors (TFs). Those rearrange into the pre-initiation complex (PIC) that unwinds the promoter DNA, proceeds downstream and screens for a suitable transcription starting site (TSS) where the elongation commences (84–88).

In a large number of genes from the MUT pathway, a well-organized an efficient recruitment of the basal transcriptional apparatus is associated with the TATA box motif (89). It is situated at the upstream end of the core promoter and has a consensus sequence of TATA(A/T)A(A/T)(A/G). First, transcriptional activators localize TATA binding proteins (TBP), such as SAGA, which then form a complex with the TATA box (78). This complex is able to delegate RNAPII and general transcription factors, required for assembling the PIC (90). Subsequently, the PIC moves downstream to find a TSS. In *P. pastoris* this is mostly between 25-30 bp downstream of the TATA box (91). In general, only genes regulating stress

response and therefore in need for a tight regulation are privileged to a TATA box in combination with the SAGA co-activator. Thus, only about 19% of regulatory elements from the yeast genome are equipped with a TATA box and 10% are SAGA dependent (88,92).

The latter part of the promoter, the URS, is responsible for transcriptional frequency and the promoter strength which is imparted onto the core promoter. It is referred to as a cis-regulatory module (CRM) that bares transcription factor binding sites (TFBS). Those localize trans-acting regulatory elements, i.e. TFs, that can either enhance transcript production or reduce it. If the specific DNA sequence increases the transcript it is regarded as an upstream activation sequence (UAS), or as an upstream repressive site if it diminishes it (88).

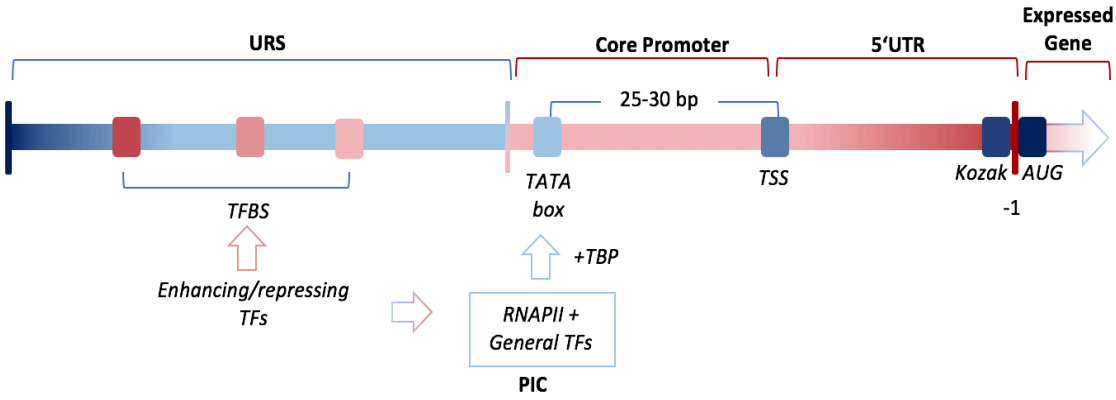


Figure 1.2 The eukaryotic promoter is separated into an upstream regulatory sequence (URS) and the core promoter. The URS contains transcription factor binding sites (TFBS) that localize transcription enhancing or repressing trans-regulatory elements. The most conserved motif of the core promoter, the TATA box, plays a central role for recruiting the basal transcription machinery, namely, the RNA polymerase II (RNAPII) and general transcription factors (TFs). Those bind to the core promoter with the help of the TATA binding protein (TBP) and form the pre-initiation complex (PIC). The PIC then moves 25-30 bp downstream to the transcription starting site (TSS). The end of the 5' untranslated region (5'UTR) is marked by the start codon AUG. Just upstream of it, is the Kozak sequence, which plays a central role in translation. The promoter is not drawn proportionally.

Enhancing TFs can have numerous functions, such as aiding in the assembly and recruitment of the basal transcriptional machinery or stabilizing RNAPII and the PIC (42,93,94). As the transcript of a specific gene is not continuously required, most TFs are activated upon certain conditions. For instance, the carbon source-responsive binding factors Adr1p and Oaf1p upregulate the genes responsible for peroxisome proliferation and β-oxidation in *S. cerevisiae* only under oleic acid induction (95).

For promoter engineering several approaches have been applied in the past. The most trivial one is to produce a randomized promoter library via error-prone PCR, namely, randomly

mutating small parts of the promoter sequence in disregard of damaging a TFBS or core promoter motif. Although in most of the cases those mutations reduce the strength of the promoter, the site of the disruption might reveal a putative binding site or motif. Additionally, weaker promoters with a change in expression activity might be used for fine-tuning metabolic pathways (96,97). A more sophisticated approach is the identification and modification of TFBS. Those can be utilized to alter the promoter strength by either adding additional binding sites or deleting existing ones. For instance, if the goal is to enhance promoter performance, upstream activation elements are added as tandem repeats - increasing overall promoter strength with each additional element - and repressive binding sites are deleted (34,42). With this strategy whole promoter libraries can be established. For instance, Hartner *et al.* (34) created a set of P_{AOX1} variants with activities ranging from 6% to above 160% compared to the wild type. The engineered promoters were successfully tested with industrially relevant enzymes, leading to enhanced product titers. This proves that direct modification of TFBS is a suitable tool for creating a set of promoters with diversified sequences that facilitate the introduction of novel pathways into host organisms. For inducible promoters originating from the MUT pathway, especially three kinds of TFBS are essential (34) *i)* those that are repressing promoter activity when the host is cultivated on glucose, so the gene expression is tightly regulated, and growth and production phase are separated (e.g. Adr1p) *ii)* those that impart upregulation on the promoter upon stress exposure, i.e. methanol induction or carbon-source depletion (e.g. Msn2p/Msn4p) *iii)* general TFs (e.g. abaA, Rap1p, QA-1F). However, for the catalase promoter of *P. pastoris* so far, no specific transcription factor binding sites are known. In addition, it was uncertain, whether this promoter might be enhanced by engineering, or if its sequence can be diversified by exchanging non-essential bases of the promoter region. Also, the length of the promoter region carrying potential binding sites for regulatory factors was not known, but a 500 bp DNA element upstream of the start codon of the catalase gene showed similar promoter strength compared to longer tested fragments.

When it comes to modifying the core promoter of the P_{AOX1}, the structure seems to be rather resilient to small mutations with a few exceptions. Obviously, the TATA box motif is sensitive to changes, resulting in a decreased promoter activity for most of the cases. Moreover, alterations in the transition region between the URS and the core promoter, the downstream part of the transcription start site, and the upstream sequence in close vicinity to the start codon decrease the promoter strength (98). The latter might be due to the Kozak sequence, which is important for translational activity and can be found directly upstream of the start

codon (99). Since the sequences in between are apparently not crucial for basal promoter activity, they could be substituted by beneficial elements, such as mRNA stabilizing introns. It is yet to be tested, whether these findings also apply on other MUT pathway promoters than the P_{AOX1}.

For the sake of completeness, also *de novo* synthesized promoters have to be mentioned as a possibility for promoter engineering (100). In case of *P. pastoris* and *S. cerevisiae*, fully synthetic core promoters had been designed and their efficiency exceeded even natural promoter in some cases (101)

1.3 Objectives and previous work

Over the past 40 years *P. pastoris* has been established as one of the best host organisms for high titer recombinant protein production. This swift advance was mostly owed to the large palette of advantages of *P. pastoris*, such as the capability of high cell density growth, easy maintenance, the post-translational modification apparatus and the tight and strong P_{AOX1} expression system, just to name a few. Recently, this methylotrophic yeast received increasing recognition for being applied in metabolic engineering and synthetic biology. However, new applications require new tools. For efficient fine-tuning of metabolic pathways, a vast set of promoters with different strengths is necessary. In addition to that, it is essential that they differ in sequence to avoid unintended *in vivo* or *in vitro* recombination. The MUT pathway of *P. pastoris* is a valuable source of a large number of such promoters. Especially P_{CTA1}, governing the expression of catalase for the detoxification of H_2O_2, was standing out in previous studies (11). While reaching comparable expression levels to P_{AOX1}, it was the best promoter in de-repressed state. After further promoter optimization, this feature might open up a possibility to produce large amounts of protein only by carbon source starvation, making methanol induction redundant. In order to rationally engineer a promoter, one needs to be acquainted with its structure. Since little to none is known about the regulatory elements of P_{CTA1}, an efficient approach to optimize it is challenging. Therefore, this master thesis and preceding ones covering this topic and invention, made it their endeavor to get a better understanding of P_{CTA1}. This included to examine and determine best cultivation conditions and to find out if it is eventually possible to optimize the promoter by sequence modifications in order to produce outstanding heterologous protein levels.

So far, the challenging topic has been addressed by a series of studies, including this one. All had in common that a truncated 500 bp long version of the wildtype promoter was used. In

a study by Vogl *et al.* (11), problems were encountered when using the native promoter or an elongated version of 1000 or even 692 bps. Transformation backgrounds were observed, that showed no promoter activity when being re-cultivated. The 500 bp version was proven to be fully active just like longer fragments but uniform, stable colonies were obtained after cell transformation. Therefore, this version was used within these studies and in the remaining parts of this thesis P_{CTA1} is actually referring to the truncated version $P_{CTA1\text{-}500}$.

1.3.1 Previous work

The first chapter of this comprehensive project was introduced by Mohamed Hussein and concentrated on creating a P_{CTA1} promoter library by applying different mutation strategies. The first approach was to introduce sequence changes in different areas of the promoter and to analyze whether or not changes in activity levels were observed. By that, first crucial and neutral promoter regions could be distinguished. Two approaches for creating the library were used: a more the rational approach and a systemic scanning approach. For the former, the bioinformatic tool MatInspector (103) within the GenomatixSuite using Genomatix Software GmbH servers was used as by Hartner *et al.* (34), previously. It determined putative TFBS in the wildtype promoter and compared those with the homologous peroxisomal *S. cerevisiae* CTA1 promoter (**Figure 1.3**). By deleting one predicted TFBS per variant, 54 mutant promoters were established. The mutations were labeled 1 to 54. For some constructs, double mutations were accidently inserted, e.g. MM13b. The latter approach replaced every 10 bp in the promoter with a neutral sequence that had no TFBS present, according to MatInspector (**Figure 1.4**). In a 500 bp long promoter this amounted to 50 different areas being substituted. To be sure, two different synthetic random sequences, which were supposed to be neutral, were introduced, A (ATCCTTTTAG) or alternatively B (GATAACCGTG), producing a scanning library of 100 variants. Those were named 1A to

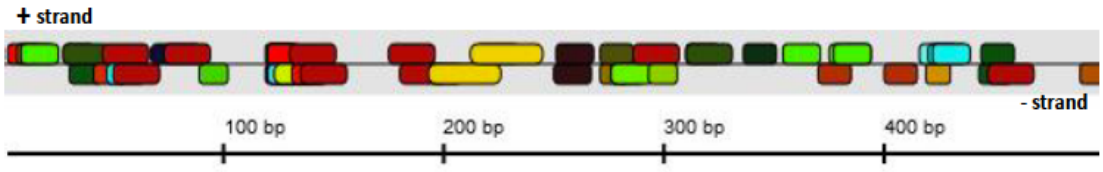


Figure 1.3 The truncated 500 bp P_{CTA1} with 54 putative transcription factor binding sites (TFBS) (Figure obtained from Hussein (102)) Predicted by MatInspector using the Genomatix Software GmbH servers and GenomatixSuite. Each box shows a TFBS and mutual colors represent coherent families. The graphic predicts binding on the upper or lower strand.

50A and 1B to 50B, whereas for instance "1A" would be bp 1-10 of P_{CTA1} substituted with sequence A. To ensure comparability between the total of 154 promoter variants certain conditions had to be met. Therefore, site-directed homologous integration was performed to avoid locus effects, as well as multicopy integration. The position effect is the phenomenon that depending on where in the host genome the recombinant gene is integrated to, tight or leaky and strong or weak transcription is achieved. This for example is a result of the host chromatin surrounding the locus, making it more or less available for the transcription machinery leading to different expression levels (104). Hence, site-directed integration into the GUT1 locus by using 5' and 3' homologous regions was performed. To even further favor homologous recombination the *P. pastoris* strain CBS7435 Δ*KU70* was used in this and the following study. During my work, the *P. pastoris* strain BSY11dKU70, which was acquired from bisy, e.U. (Graz, Austria), was used as the host organism for recombinant protein production. The *KU70* gene, which encodes for a protein involved in non-homologous end joining (NHEJ), was knocked-out, facilitating integration of the vector into a defined locus (105). The second requirement, avoiding multicopy integrations, is even more important, because a different number of integrated genes might lead to differently strong activities (106). This results in poor comparability between the promoter variants. Thus, double cross-over events were targeted, where the linear expression cassette replaced the glycerol kinase gene *GUT1* instead of being inserted at this site. This measure easily leads to single copy transformants (41). In addition, after successful transformation, the strain loses

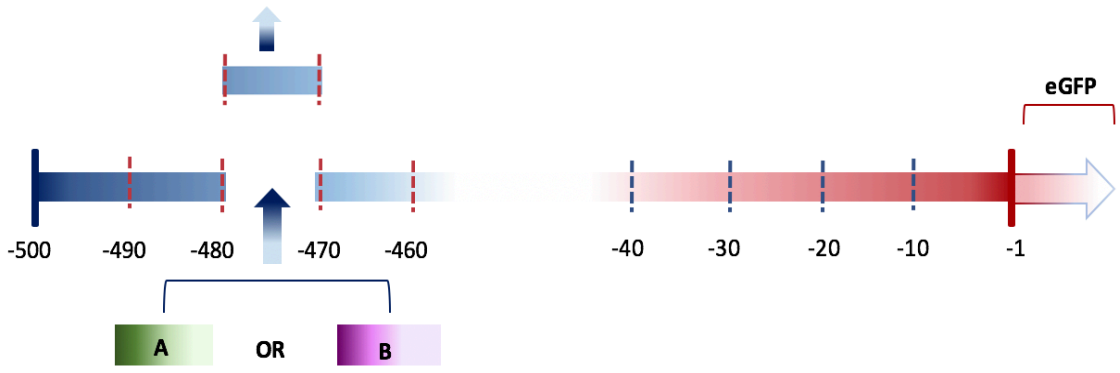


Figure 1.4 In **the walking library** 10 bp fragments were substituted with either sequence A (ATCCTTTTAG) or sequence B (GATAACCGTG). Both were supposedly neutral (no transcription factor binding sites present according to MatInspector) and equally long (10 bp). In the truncated 500 bp P_{CTA1} promoter a total of 50 fragments were substituted (1-10 bp; 11-20 bp; etc.) with either sequence A or B, amounting to a total of 100 promoter variants.

the *GUT1* gene and its ability to grow on glycerol as the sole carbon source. For selection of transformed *P. pastoris* colonies the integrated vector carried a Zeocin™ resistance gene and cells were grown on agar plates supplemented with antibiotics. Furthermore, transformed *P. pastoris* strains were cultivated on media solely containing glycerol as a carbon source for growth counter selection. Afterwards, the transformants were picked, cultured and the promoter variants screened for their strength using enhanced green fluorescence protein (eGFP) as a reporter gene. The P_{AOX1} promoter acted as benchmark to indicate the strength of the mutant promoters and the activity levels were measured in de-repressed, as well as in induced state. When compared to the wildtype P_{CTA1}, activity increased significantly for some mutants. For instance, variants from the first scanning library reached about 200% of the wildtype fluorescence levels in de-repressed state and 160% after methanol induction.

After Mohammed Hussein, the project was undertaken by Katharina Ebner, who focused on further pushing promoter strength by selecting the best variants and combining their mutations for an improved effect. Therefore, the acquired data by Hussein was thoroughly analyzed in order to avoid mistakes due to sample identifiers or sequence results, which were mixed up, and to identify mutations that allegedly were responsible for promoter activity elevation in de-repressed or in induced state. Ten mutations were picked, among them four were prominent for increased promoter strength upon carbon source depletion, three were associated with better activity after methanol induction and three mutations displayed better promoter performance in both states. Those mutation sequences were isolated and recombined to double and triple combinations. Additionally, promoter sequences with duplications of certain regions presumed to be beneficial for promoter activity were produced. A total of 54 new variants was cloned and they were integrated into the *AOX1* locus via double-crossover events, instead of the *GUT1* locus. This led to the deletion of the *AOX1* gene and a phenotype change of the host from fast methanol utilization (Mut^+) to slow methanol utilization (Mut^S) (41). This alteration was made to allow for future experiments with glycerol as comparative carbon source to glucose. The wildtype *P. pastoris* Mut^+ strain tends to pose certain problems, especially in large-scale cultivations. For one, it requires a high feeding rate of methanol, which is hazardous to work with due to its high heat of combustion (-727 kJ C-mol^{-1}). Therefore, expensive explosion-proof storage facilities are needed, adding unnecessary costs to the process. Furthermore, it has a high specific oxygen consumption rate, leading to limitations (107,108) and heat production by fast methanol metabolization can result in cooling problems during the cultivation. These issues cause operational difficulties that eventually reflect in decreased strain productivity.

Additionally, screening in small scale methanol containing media shows a low reproducibility for clones which have a fast methanol metabolism. The mutated Mut^S strain can ease those problems. Due to the deletion of the *AOX1* gene, *P. pastoris* strains carrying the Mut^S phenotype are dependent on the transcriptionally weaker *AOX2* gene, which metabolizes methanol substantially slower (109). Less methanol is required, and the related difficulties decreased, which eventually leads to higher expression levels of the recombinant gene compared to the wild-type strain (107).

The transformed Mut^S *P. pastoris* cells were cultivated and screened for in the same manner as previously done. The results showed that by accumulating favorable mutations, the promoter strength was even further elevated in several cases. To be precise, in de-repressed state activity levels of up to 300% and in induced state activity levels of up to 190% were achieved when being compared to the parental wildtype P_{CTA1} sequence.

1.3.2 Objectives

The promising results from Katharina Ebner set the base for my research and objectives, which were *i)* test the best P_{CTA1} variants with extracellularly secreted proteins *ii)* further optimize P_{CTA1} variants by additional mutation combinations *iii)* test P_{CTA1} variants on different carbon sources *iv)* perform up-scaling in shake flasks and characterize promoter variants via an online monitoring system. One of the main advantages of *P. pastoris* is its ability to extracellularly secret heterologous proteins. Combined with the low amount of native secreted proteins, downstream processing is enormously facilitated, which makes it only natural that a great number of industrial products expressed in *P. pastoris* are secreted. Therefore, I tested, whether the high promoter activity of the established variants can be maintained when extracellular proteins are being expressed instead of intracellularly produced eGFP reporter protein. Regarding further promoter optimization, the aim of this study was to combine up to five mutations in one promoter variant to answer the following question: Is the enhancing effect of upregulating mutations even more accumulative or will the integrity of the promoter be de-stabilized at some point? Further, all promoter variants were tested in different conditions to optimally find a cultivation protocol that makes methanol induction redundant. Finally, suitable promoter variants were selected for up-scaling in shake-flask cultivations. The growth was followed by an online monitoring system from the PreSens Precision Sensing GmbH in order to learn more about strain and promoter behavior. This was done as a preliminary experiment for future up-scaling in bioreactor fermentations.

2. MATERIALS AND METHODS

All chemicals, reagents, instruments, devices, enzymes, primers, software and webtools used in this master thesis are listed in the Appendix section.

2.1 Media and Buffers

Table 2.1 The amounts in this table correspond to the preparation of one liter, i.e. the components in the media were filled up to one liter by adding ddH_2O. The different carbon source concentrations and the corresponding amounts in the buffered minimal media (BM_) are separated with a "/"; e.g. BMD1% contains 11 g of α-D(+)-glucose monohydrate and BMD0.5% contains 5.5 g. Zeocin™ was added to the cultivation media in concentrations of 50 µg/mL for *P. pastoris* and 25 µg/mL for *E. coli*.

	Component	Amount
BMD1/0.5% (Buffered minimal Dextrose medium)	α-D(+)-glucose monohydrate	11/5.5 g
	10x YNB	100 mL
	1M Potassium phosphate buffer, pH 6	200 mL
	500x Biotin	2 mL
BMG2.5/1/0.5% (Buffered minimal Glycerol medium)	50% glycerol	50/20/10 mL
	10x YNB	100 mL
	1M Potassium phosphate buffer, pH 6	200 mL
	500x Biotin	2 mL
BMM5/2/1.25/1% (Buffered minimal Methanol medium)	Methanol (≥ 99.8%)	50/20/12.5/10 mL
	10x YNB	100 mL
	1M Potassium phosphate buffer, pH 6	200 mL
	500x Biotin	2 mL
YPD2% (Yeast Extract Peptone Dextrose Medium)	Bacto Yeast Extrat	10g
	BactoPeptone	20g
	10xDextrose	100 mL
	Agar (only for plates)	20 g

10xDextrose	α-D(+)-glucose monohydrate	220g
10xYNB (Yeast Nitrogen Base)	Yeast nitrogen base	134 g
	filter sterilized	
500xBiotin	Biotin	200 mg
	filter sterilized, stored at 4°C	
LB (Luria-Bertani Medium)	Yeast Extract	10 g
	Tryptone	5 g
	NaCl	10 g
	Agar (only for plates)	15 g
	adjust to pH 7 with NaOH	
50xTAE Buffer	Tris	242 g
	EDTA	14.6 g
	Acetic acid (conc.)	57.1 mL
SOC Medium	α-D(+)-glucose monohydrate	3.46 g
	Tryptone	20 g
	Yeast extract	5 g
	NaCl	0.58 g
	MgCl2	2 g
	KCl	0.16 g
	MgSO4	2.46 g
BEDS	Bicin (10mM)	1.632 g
	Ethylenglycol	30 mL
	DMSO	50 mL
	Sorbitol (1M)	182 g
	filter sterilized, set pH to 8.3 with 2M NaOH	

10xMOPS/SDS	MOPS (50mM)	209.2 g
	Tris (50mM)	121.2 g
	SDS (0.1%)	20 g
	EDTA (3.0g)	6 g
	pH 7.7	
10xTBS (Tris Buffered Saline)	Tris (0.5M)	60.5 g
	NaCl (1.5M)	87.6 g
	adjust pH with 1M HCl	
TBS/TMP	10xTBS	100 mL
	Milk powder	50 g
TBST	10xTBS	100 mL
	TritonX-100	1 mL
TBST/TMP	10xTBS	100 mL
	Milk powder	50 g
	TritonX-100	1 mL
Primary antibody solution	10xTBS	100 mL
	Milk powder	5 g
	TritonX-100	1 mL
	GH Antibody (E-7): sc-374266	2 mL
Secondary antibody solution	10xTBS	100 mL
	Milk powder	5 g
	TritonX-100	1 mL
	m-IgGK BP-HRP: sc-516102	0.2 mL

2.2 Kits and Protocols

2.2.1 Kits

Table 2.2 Kits used in this master thesis.

Name	Company
Wizard® SV Gel and PCR Clean-Up System	Promega Corporation, Fitchburg, WI, USA
Wizard® *Plus* SV Minipreps DNA Purification System	Promega Corporation, Fitchburg, WI, USA
Q5 High-Fidelity Polymerase Kit	New England Biolabs® Inc., Ipswhich, MA, USA
GoTaq® kit	Promega Corporation, Fitchburg, WI, USA
FastDigest™ kit	Thermo Fisher Scientific Inc., Waltham, MA, USA
NuPAGE 4-12% Bis-Tris Gel system	Invitrogen by Thermo Fisher Scientific, Inc., Waltham, MA, USA

2.2.2 Plasmid DNA isolation from *E. coli*

E. coli plasmid isolation was performed using the Wizard® *Plus* SV Minipreps DNA Purification System (Promega Corporation, Fitchburg, WI, USA). The cells were plated on selective LB-agar plates. After overnight incubation at 37°C the grown cell layer was directly scraped off the agar plates with a sterile round wooden toothpick. The biomass was processed accordingly to quick protocol (110) with minor alterations to improve DNA yield, i.e. after the washing steps an additional centrifuge interval of one minute at top speed and room temperature was performed and a drying period of five to ten minutes at 37°C was introduced to remove residual ethanol from the spin column. The DNA pellet was eluted with 30-50 µL ultrapure ddH_2O (Fresenius Kabi, Graz, Austria). The DNA concentration as well as the purity of the DNA was determined using a UV/VIS spectrophotometer (Nanodrop™ 2000, Thermo Fisher Scientific Inc., Waltham, MA, USA). The DNA quality was described by the 260/280 and 260/230 ratio. The plasmid DNA was either directly used or stored at -20°C.

2.2.3 PCR procedures

2.2.3.1 Standard PCR

During this thesis standard polymerase chain reactions (PCR) were used to produce DNA amplicons, which were further applied in several plasmid cloning steps. For the reaction, the Q5 High-Fidelity Polymerase Kit from New England Biolabs® Inc. (Ipswhich, MA, USA) was used. Template DNA was diluted to 10 ng/µL using ultrapure nuclease-free ddH_2O (Fresenius Kabi). The extension time varied between 30 seconds and two minutes depending on the produced amplicon length (it was calculated with approximately one minute per one kilo base pair (kb)). For the detailed master template and the complete PCR temperature profile see **Table 2.3** and **Table 2.4**, respectively.

Table 2.3 Master template for standard PCR reactions using Q5 High-Fidelity Polymerase Kit from New England Biolabs®, Inc. (Ipswhich, MA, USA)

PCR Components	Volumes
Buffer Q5 (5x)	10
GC Enhancer Q5 (5x)	10
dNTP mix (2 mM)	5
Phusion Polymerase Q5	0.5
ultrapure ddH2O	14
Primer fwd (5 pmol/mL)	4
Primer rev (5 pmol/mL)	4
Template DNA (10 ng/µL)	2.5
total	50

Table 2.4 Cycle parameters for standard PCR reactions

Cycle Step	Temperature	Time [sec]	N° of Cycles
Initial denaturation	98°C	60	1
Denaturation	98°C	10	
Annealing	58°C	20	30
Extension	72°C	30 per 1 kb	
Final extension	72°C	540	1
Holding	4°C	∞	∞

2.2.3.2 Colony PCR (cPCR)

cPCR was conducted with the *GoTaq*® kit (Promega Corporation, Fitchburg, WI, USA) to screen for the correct assembly of insert(s) and backbone in *E. coli* transformants. The master template used is described in **Table 2.5** and the cycling profile in **Table 2.6**. To obtain template DNA, cell material from colony forming units (CFU) was added via pipette tips directly from the transformation plate into PCR reaction tubes. Thus, the initial denaturation step was extended to six minutes to disintegrate the cell walls and make template DNA available for the polymerase. The primers were designed to amplify distinctive parts of the plasmid. So, when separated by agarose gel electrophoresis, missing or multi-copy inserts could be determined by amplicon bp length. The gel electrophoresis was conducted as described, however gel loading dye was obsolete, since *GoTaq*® Flexi Buffer Green was already equipped with fluorescent. When a band of the correct size was visible the *E. coli* transformant was cultivated on selective LB-agar plates for plasmid DNA isolation. Subsequently, the plasmid DNA was prepared and sent for sequencing according to the standards of the respective sequencing company (MycroSynth AG, Balgach, Switzerland; GATC Biotech AG, Konstanz, Germany).

Table 2.5 Master template for colony PCR reactions using the *GoTaq*® Kit from Promega Corporation (Fitchburg, WI, USA)

PCR Components	Volumes
GoTaq Flexi buffer green (5x)	5
GoTaq Flexi Polymerase (5 U/µL)	0.5
$MgCl_2$ (50 mM)	0.65
dNTPs (2 mM)	2.5
ultrapure ddH_2O	9.85
Primer fw (10 µM)	2
Primer rv (10 µM)	2
Template DNA (via pipette tips)	0.5
total	25

Table 2.6 Cycle parameters for colony PCR reactions

Cycle Step	Temperature	Time [sec]	Number of Cycles
Initial denaturation	98°C	300	1
Denaturation	98°C	15	
Annealing	58°C	20	30
Extension	72°C	60 per 1 kb	
Final extension	72°C	300	1
Holding	4°C	∞	∞

2.2.4 Purification of DNA fragments

DNA fragments produced in amplicon PCR and plasmids processed by enzyme restriction were purified and retrieved in two essential steps: separation by agarose gel electrophoresis and purification with the Wizard® SV Gel and PCR Clean-Up System (Promega Corporation, Fitchburg, WI, USA). Fifty µL of the DNA fragment were mixed with ten µL gel loading dye 6x (New England Biolabs® Inc., Ipswhich, MA, USA) and were loaded in an 1% agarose gel containing ethidium bromide as fluorescent. As DNA ladder, 2.5 µg GeneRuler™ DNA

Ladder Mix or GeneRuler™ 1 kb DNA Ladder (Thermo Fisher Scientific Inc., Waltham, MA, USA) were used. The gel was run at 110 volt (V) for approximately 50 min or up to 90 min, when the DNA fragments were longer and needed more time to separate. Appropriate bands were cut out under UV-light and the DNA was purified with the Wizard® SV Gel and PCR Clean-Up System following the Quick protocol (111).

2.2.5 Transformation protocols

2.2.5.1 *Escherichia coli*

Electrocompetent *E. coli* TOP10F' cells were prepared according to Seidman *et. al* (112) and 40 µL aliquots were stored at -80 °C for further use. For one transformation two to four µL of the assembly cloning mix or isolated plasmid DNA were mixed with 40 µL competent cells, previously thawed on ice. The reaction mixture was transferred to a chilled two-mm electroporation cuvette. The electroporation was conducted with the Gene Pulser (Bio-Rad Laboratories, Inc., Hercules, CA, United States) using 2.5 kV, 25 µF and 100 Ω as settings. Afterwards, the cells were immediately regenerated in one mL SOC medium at 37°C and 600 rpm for 60 min. Following the regeneration step, the cells were centrifuged, resuspended into a smaller volume, plated out on selective LB-agar plates and incubated overnight at 37°C.

2.2.5.2 *Pichia pastoris*

Prior to transformation, electrocompetent *P. pastoris* cells were freshly prepared according to Lin-Cereghino *et al.* (113). Forty µL of the competent cells and one µg of linearized plasmid were mixed in a pre-chilled 0.2 cm cuvette. After an incubation period of five minutes on ice, the transformation was conducted with the Bio-Rad Gene Pulser at 1.5 kV, 25 µF and 200 Ω. The cells were immediately resuspended in 500 µL of a 1:1 mixture of YPD2% glucose and 1M sorbitol and regenerated for two hours at 30°C and 700-800 rpm. Subsequently, the cells were centrifuged and resuspended in a smaller volume, before being plated on selective YPD2% glucose agar plates. The plates were incubated for three days at 30°C.

2.2.6 Assembly cloning

The preparation of an assembly master mix (**Table 2.7**) and the reaction procedure were followed according to Gibson *et al.* (71), however slight alterations were applied; i.e. for the assembly reaction 100 ng of linearized vector DNA and *x* ng insert DNA were mixed in a molar ratio of 1:3 instead of equimolar for improved results. The total DNA load should not

surpass 250 ng. The DNA mix was filled up to 5 µL with ultrapure ddH_2O and added to 15 µL of the prepared assembly master mix for a final reaction volume of 20 µL. After an incubation period of one hour at 50°C three µL of the assembly reaction mix were used for *E. coli* transformation.

Table 2.7 Assembly master mix using reagents from New England Biolabs®, Inc. (Ipswhich, MA, USA) and Thermo Fisher Scientific, Inc. (Waltham, MA, USA)

Reaction Components	Volume
ISO reaction buffer (5x)	320 µL
T5 exonuclease (10 U/µL)	0.64 µL
Phusion® High-Fidelity DNA Polymerase (2 U/µL)	20 µL
Taq DNA ligase (40 U/µL)	160 µL
ddH_2O	699.36 µL
total	1200 µL

2.2.7 Restriction endonuclease reactions

2.2.7.1 Plasmid DNA linearization

Linearization steps of plasmid DNA were performed either to prepare plasmids for *P. pastoris* transformations or vector backbones for assembly cloning. For all the reactions reagents from the FastDigest™ kit (Thermo Fisher Scientific Inc., Waltham, MA, USA) were used and according protocol was followed with slight alterations. In general, one µL of FastDigest™ enzyme per one µg of DNA was applied. The manufacturer's proposed digestion times of the enzymes at 37°C were doubled to ensure complete plasmid digestion and the inactivation times were followed as stated. The linearization reaction for *P. pastoris* transformations was conducted at half the volume. For a detailed description of the reaction composition see **Table 2.8**.

Table 2.8 Typical reagent composition of restriction endonuclease reactions using Thermo Fisher Scientific Inc. (Waltham, MA, USA) reagents

Reaction Components	Volumes
FastDigest Green Buffer (10x)	5 µL
DNA (5 µg end concentration)	x µL
FastDigest Enzyme	5 µL
ultrapure ddH_2O	40-x µL
total	50 µL

2.2.7.2 Restriction endonuclease control cut

To screen for correct DNA fragment assembly in the cloning process, isolated plasmid DNA was digested using the FastDigest™ kit (Thermo Fisher Scientific Inc., Waltham, MA, USA). In general, one µL of FastDigest™ enzyme per one µg of DNA was applied. The reagent composition depicted in **Table 2.9** was used and apart from extending the incubation phase at 37°C to 60 minutes the manufacturer protocol was followed. By comparing obtained fragment size with expected values missing or multicopy inserts could be determined. When a band of the correct size was acquired the corresponding *E. coli* transformant was cultivated on selective LB-agar plates for plasmid DNA isolation. When isolated, the plasmid DNA was prepared and sent for sequencing according to the standards of the respective sequencing company (MycroSynth AG, Balgach, Switzerland; GATC Biotech AG, Konstanz, Germany).

Table 2.9 Typical reagent composition of restriction endonuclease control cut using Thermo Fisher Scientific Inc. (Waltham, MA, USA) reagents

Reaction Components	Volumes
FastDigest Green Buffer (10x)	2 µL
DNA (500 ng end concentration)	x µL
FastDigest Enzyme I	0.5 µL
FastDigest Enzyme II	0.5 µL
ultrapure ddH_2O	17-x µL
total	20 µL

2.2.8 Cultivation methods of *P. pastoris*

2.2.8.1 Cultivation of *P. pastoris* in Deep-well-plates

Deep-well-plate (DWP) cultivation was used for *P. pastoris* strains expressing eGFP, hGH or CalB under the control of P_{CTA1} variants. Depending on expressed gene or promoter variant minor cultivation condition differences were used and described onwards. As a general protocol following was used: each well of a 96-DWP was filled with 300 µL of BMD or BMG and inoculated with a *P. pastoris* variant. The variants were transferred either from selective YPD agar plates using sterile pipette tips or sterile wooden toothpicks or from pre-culture DWP using multichannel pipettes. The plates were incubated at 28°C and 320 rpm in a HT Infors Multitron Shaker. Upon carbon source depletion after estimated 60 hours the BMD cultures were induced with 250 µL BMM1% and the BMG cultures were administered a 250 µL BMG0.5% pulse per well. After 72 hours of incubation time a reinduction with 50 µL BMM5% and a second pulse with 50 µL BMG2.5% were applied to each well of the respective DWP. This was repeated for a final induction/pulse at 84 hours incubation time. The cultivation was finished after 98 hours.

After 60 hours of cultivation and prior to induction, cell material of all clones was transferred to BMM2% containing agar plates via a singed 96-well stamp. In a growth counter selection, the transformed *P. pastoris* cells were tested, whether the integration of the expression cassette took place at the correct locus (*AOX1*). If a clone was able to grow in described conditions, it resembled a Mut^+ phenotype and was discarded.

For strains expressing eGFP under the control of P_{CTA1} variants 1 to 54 quadruplicates were cultivated in BMD1% or BMG1% (1% glucose or glycerol concentration in the media). For all other constructs 14 transformants per variant were cultivated in media with 0.5% concentration of the prime carbon source.

2.2.8.2 Pre-culture DWP

For *P. pastoris* strains expressing eGFP or CalB pre-culture DWPs were used. Therefore, one DWP per media (BMD or BMG) was inoculated with colonies from transformation plates and incubated as described. After estimated 60 hours carbon source depletion and cell growth saturation were reached. At this point, 10 µL per well of the pre-culture were transferred to a new main culture DWP for inoculation using a multichannel pipette. The starting volume and type of the cultivation media in the main culture was kept identical to the pre-

culture; e.g. BMD0.5% cultivated *P. pastoris* cells form the pre-culture were transferred to a well filled with 300 µL BMD0.5%. The onward cultivation was continued as described above.

2.2.8.3 Sampling of DWP

For *P. pastoris* strains expressing eGFP samples were taken after 24, 48, 60 and 98 hours. Ten µL of each well were transferred into a Nunc™ MicroWell™ 96-Well Optical-Bottom Plate with Polymer Base and used immediately for the determination of eGFP fluorescence levels.

For *P. pastoris* strains expressing hGH or CalB the supernatant had to be harvested since the recombinant protein was extracellularly expressed. Therefore, the DWP was centrifuged at 4000 rpm and 4°C for five minutes and the supernatant was transferred to a 96-well microtiter plate using a multichannel pipette. The sample was directly used for the distinctive assays or stored at -20°C.

2.2.8.4 Cultivation of *P. pastoris* in shake flasks using an online monitoring system

The *P. pastoris* strains and controls selected for shake flask cultivation were first streaked out on selective agar plates and then consecutively cultivated in pre-pre-cultures, pre-cultures and main cultures. The pre-pre-culture was inoculated with colonies from the agar plate, the pre-culture with the pre-pre-culture and the main culture with the pre-culture.

Selective YPD agar plates were used and incubated at 30°C for 48 hours with the respective strain plated out selectively. For pre-pre-cultures, 5 mL YPD medium in 50 mL Greiner centrifuge tubes were inoculated with a single large colony from the agar plate. The cultivation media should be pre-warmed before an inoculation step. The pre-pre-cultures were incubated at 28°C and 100 rpm for 24 hours with a slightly opened cap to provide sufficient aeration. For the pre-cultures and main cultures 250 mL sterile baffled shake flasks covered with cotton sheets and filled with either 50 mL BMD0.5% or BMG0.5% were used. They were incubated at identical conditions, such as 28°C and 150 rpm. The OD_{600} of the inoculate was measured previously to inoculation to determine the appropriate volume. The pre-cultures were inoculated to an OD_{600} of 0.01 and incubated for 21 hours prior to inoculating the main cultures to an OD_{600} of 0.2. The main culture cultivations were carried out with BMD0.5% and BMG0.5% in biological duplicates per variant and control.

Due to limitation of instruments, only one WT cultivation per media was followed with an online monitoring system from the PreSens Precision Sensing GmbH (Regensburg, Deutschland) measuring relative biomass, oxygen uptake rate (OUR) and O_2 (%). This served as growth reference to all the other variants cultivated on the same media and facilitated

more accurate sampling and induction. Regarding the latter, in case of BMD0.5% cultivations induction was performed either by adding BMM1.25% or 100% methanol to reach an end concentration of 0.125% methanol. The induction time points were chosen when according to the online monitoring system the previously administered carbon source was depleted. For the BMG0.5% cultivations four glycerol feed disks (PS Biotech GmbH, Aachen, Germany) per 50 mL medium in the shake flask were added. This kept the cells de-repressed as well as nourished them just enough by constantly supplying small amounts of glycerol.

Sampling was performed under sterile conditions. The sample volume was always related to an OD_{600} of 10 for comparable results; e.g. if the measured OD_{600} was 20, 0.5 mL cultivation broth were removed. The sample was kept in 2 mL Eppendorf tubes or in 15 mL Greiner centrifuge tubes, if the sampling volume exceeded 1.5 mL. The samples were kept on ice until centrifugation at maximum rpm and 4°C for five minutes. Supernatant and cell pellet were separated and frozen at -80°C until RNA isolation and CalB activity determination was conducted.

2.2.8.5 Controls in cultivation methods

For DWP and shake flask cultivations several controls were required. *P. pastoris* strains bearing plasmids identical to the tested variants but with different promoter elements were used as controls. Instead they incorporated no promoter as negative, a WT P_{CTA1} promoter as positive or the strong de-repressed formate dehydrogenase promoter (P_{DF}) from *Hansenula polymorpha* as benchmark control. In DWP cultivations the controls were added in triplets and in shake flask cultivations in duplicates. Furthermore, sterile samples were carried along in each cultivation to check for contamination. The controls were treated identical to the tested subjects for comparability.

2.3 Assays and analytical methods

2.3.1 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

For SDS-PAGE the ready-to-use NuPAGE™ 4-12% Bis-Tris Gels and Mini Gel Tank system (Invitrogen by Thermo Fisher Scientific, Inc., Waltham, MA, USA) were used. The samples were mixed as described in **Table 2.10** and following the instructions of the manufacturer, they were denatured, 1x MOPS SDS running buffer was prepared, the gels were installed and 12 µL of the mixed samples were loaded. PageRuler™ Prestained Protein Ladder 10 to 180 kDa from Thermo Scientific, Inc. (Waltham, MA, USA) was used as reference.

After running the electrophoresis with 200 V and 48 W for 50 min, the gel holder was broken up and the gel rinsed with ultrapure ddH_2O. Subsequently, the gel was washed three times for ten minutes each with ultrapure ddH_2O before staining with PageBlue™ Protein Staining Solution (Thermo Scientific, Inc., Waltham, MA, USA) for 60 minutes. The gel was rinsed two times and destained for five minutes with ultrapure ddH_2O. The washing, staining and destaining steps were performed under modest shaking of the gel. The protein bands were visualized with the G:BOX HR16 imaging system and GeneSyn software (Syngene International Ltd., Bangalore, India). The distinctive protein bands were quantified with Image Lab™ Version 6.0.1 by Bio-Rad Laboratories, Inc. (Hercules, CA, United States).

Table 2.10 Reagent composition of sample preparation for SDS-PAGE loading using the NuPage® kit (Thermo Fisher Scientific, Inc., Waltham, MA, USA)

Components	Volumes
Sample	7.8 µL
NuPage® LDS Sample Buffer (4x)	3 µL
NuPage® Reducing Agent (10x)	1.2 µL
Deionized Water	7.8 µL
Total	12 µL

For CalB protein band quantification the supernatant harvested by centrifugation could be directly used, hGH had to be precipitated first. Therefore, 300 µL supernatant, 480 µL methanol and 160 µL chloroform were thoroughly mixed in a 2 mL Eppendorf tube using a vortex mixer (Vortex Genie 2, Scientific Industries Inc., Bohemia, NY, United States) and centrifuged at maximum rpm for five minutes at room temperature. The supernatant was carefully removed and discarded while keeping the interphase holding the protein intact. After adding 300 µL of methanol another mixing and centrifugation step at maximum rpm and 4°C for 30 minutes was applied. The supernatant was discarded, and the pellet dried in a heating block (Eppendorf Thermomixer Comfort, Eppendorf AG, Hamburg, Germany) at 60°C for ten minutes. The dried pellet was ten resuspended in 7.8 µL deionized water, ready for preparation with the NuPage® reagents for gel loading.

2.3.2 Fluorescence measurement of eGFP in *P. pastoris* strains

After vigorous mixing of the cultures by short vortexing of the deep well plates, ten µL of *P. pastoris* DWP cultivations were transferred to a Nunc™ MicroWell™ 96-Well Optical-Bottom Plate with Polymer Base (Thermo Scientific, Inc., Waltham, MA, USA) using a multichannel pipette and then diluted with 190 µL ddH_2O (1:20 dilutions). Fluorescence and optical density were determined with the SynergyMx Plate Reader (BioTek Instruments Inc., Winooski, Vermont, USA). The fluorescence for eGFP was measured at 488 nm excitation and 507 nm emission as well as the optical density at 600 nm for normalization of fluorescence levels to biomass. The obtained fluorescence data of the variants were subtracted by the mean blank fluorescence. For OD_{600} normalized fluorescence data the resulting values were divided by the respective OD_{600} values. For comparison to the P_{CTA1} wildtype, the OD_{600} normalized fluorescence data of the P_{CTA1} variants were divided by the mean wildtype value.

2.3.3 CalB activity measurement in *P. pastoris*

The CalB activity was determined based on the colorimetric assay described by Zhang *et al.* (114). Twenty µL of the supernatant containing CalB enzyme were transferred to Greiner 96-well Multiwell plates followed by the addition of 180 µL reaction solution to each well. The reaction solution was constituted of 1.8 µL 480 mM p-nitrophenol butyrate *(p*-NPB) dissolved in 100% dimethyl sulfoxide (DMSO) and 178.2 µL 300 mM Tris buffer, pH 7. The pH of the buffer was adjusted with HCl. With the addition of the reaction solution *p*-nitrophenolate was formed resulting in a color change, which was followed for 5 minutes at 405 nm with the SynergyMx Plate Reader (BioTek Instruments Inc., Winooski, Vermont, USA) at room temperature. In case of an overly fast enzyme reaction visible in a flattening absorbance curve, the culture supernatant containing CalB was diluted with 300 mM Tris buffer, pH 7. The difference between first and last measured absorbance was determined, and the mean values of variants, controls and blank were calculated. Subsequently, resulting values were subtracted with the mean value of the blank. For normalization to the WT, the obtained data was divided by the WT absorbance change, for the calculation of absolute enzyme activity the Lambert-Beer equation was applied.

2.3.4 hGH activity measurement in *P. pastoris*

For the hGH activity determination an enzyme-linked immunosorbent assay (ELISA) was performed in Greiner 96-well ELISA Microlon® microplates. As primary antibody an anti-

hGH mouse monoclonal IgG kappa light chain antibody was used. The secondary antibody was a mouse IgG kappa binding protein suitable for binding mouse IgG kappa light chains and was conjugated with horse radish peroxidase (HRP). Both antibodies were obtained from Santa Cruz Biotechnology, Inc. (Dallas, Texas, USA). Primary and secondary antibodies were diluted to a concentration of 400 ng/mL and 80 ng/mL, respectively, with a TBS solution containing 0.5% (w/v) milk powder and 1 mL/L TritonX-100. All incubation steps were conducted at room temperature and modest shaking. After each step the wells were emptied by inversion and gently tapping the plate on microfiber towels to remove residual liquid.

First, 150 µL of the *P. pastoris* culture supernatant containing hGH enzyme were transferred into each well. They were incubated for three hours in order to bind the protein to specific linkers located on the surface of the well. Afterwards the wells were washed for 30 seconds with 200 µL TBS solution to remove unbound proteins. Subsequently, all empty linkers were blocked by incubating each well with 200 µL TBS/TMP solution. A 30 seconds washing step with TBS solution was repeated. Then, 150 µL of the primary antibody solution were applied and incubated. After one hour of incubation the wells were successively washed with 200 µL TBST/TMP and 200 µL TBS/TMP for five minutes each. The secondary antibody solution was added with 150 µL per well. After 45 minutes of incubation three consecutive washing steps of five minutes each with 200 µL TBST/TMP, then TBS/TMP and finally TBS were conducted. After successful antibody binding staining and detection was commenced by adding 100 µL 1-Step™ Ultra 3,3',5,5'-tetramethylbenzidin (TMB) ELISA solution (Thermo Fisher Scientific Inc., Waltham, MA, USA) to each well and incubate it for 15 to 30 minutes. The reaction was stopped by 50 µL 2M sulfuric acid and absorbance was measured at 450 nm and room temperature with the SynergyMx Plate Reader (BioTek Instruments Inc., Winooski, Vermont, USA). The mean values of variants and controls was calculated and subtracted with the mean blank. For WT normalization the obtained data were divided by the mean WT.

2.3.5 Total protein concentration measurement

To determine the total protein concentration in the *P. pastoris* culture supernatants the Bradford method (115) was performed with reagents from Bio-Rad Laboratories, Inc. (Hercules, CA, United States). The reaction was performed in a Greiner 96-well Multiwell plate in duplicates in wells next to each other. Therefore, in each well ten µL of supernatant were mixed with 40 µL of 5x dye reagent concentrate and 150 µL ddH_2O and incubated for 15 to 60 min at room temperature under light protection. The end-point absorbance was measured at 595

nm using a high-performance microplate spectrophotometer (Eon™, BioTek Instruments, Inc., Winooski, VT, USA). In order to relate absorbance to protein concentration, a calibration curve was established by measuring a geometric serial dilution of bovine serum albumin with known concentrations (2 mg/L, 1 mg/L, 0.5 mg/L, 0.25 mg/L, 0.125 mg/L, 0.0625 mg/L).

2.3.6 Sequencing

To verify all cloning products, plasmid DNA was isolated as described previously and sent for Sanger sequencing to the companies MycroSynth AG (Balgach, Switzerland), GATC Biotech AG (Konstanz, Germany), or LGC Genomics (Berlin, Germany). To evaluate the sequencing results, they were aligned with *in silico* designed plasmid maps using SnapGene® and CLC Main Workbench. The plasmid samples were pre-mixed with primers to meet the companies' requirements.

2.4 Strains and plasmids

2.4.1 Construction of hGH and CalB plasmids

The design of plasmids expressing hGH or CalB under the control of P_{CTA1} variants and control plasmids was based on the vector pPp_AOX1-pUC-ZeoR-NotI-TT_AOX1 (pKEB#1) and is depicted in **Figure 2.1**. The vector backbone contained an up- and downstream *AOX1* homologous region for double-crossover insertion at the *AOX1* locus, which created a Mut^S *P. pastoris* strain. Further, a pUC ori element for autonomous replication in *E. coli*, a Zeocin™ resistance cassette for selection and a *Not*I restriction site for linearization was present. CalB with the α-MF (P17666/P17667) and hGH with a shortened α-MF (Dα) (P17668/17669) were amplified in order to introduce an *Eco*RI restriction site and an overhang to *ILV5* on the 5' and an overhang to the AOX1_TT on the 3' end. CalB was amplified from the pPpT4-pCAT1- α_CalB plasmid (IMBT collection) and hGH was ordered from IDT Integrated DNA Technologies, Inc., Coralville, Iowa, United States. The vector backbone pKEB#1 was linearized with *Not*I and ligated with either EcoRI- α_CalB or EcoRI-Dα_hGH via assembly cloning, resulting in plasmids pPp_AOX1-pUC-ZeoR-EcoRI-α_CalB-TT_AOX1 and pPp_AOX1-pUC-ZeoR-EcoRI-Dα_hGH-TT_AOX1, respectively. Three µL of the assembly mix were used for an *E. coli* TOP10'F transformation. The transformed cells were plated and cultivated on selective agar plates and after incubation four

transformants were used for cPCR and their plasmids were isolated and sent for sequencing. Plasmids from transformants with positive cPCR and sequencing results were linearized with *Eco*RI restriction enzyme to prepare them for the assembly with promoter variants. For this step, selected variants from Katharina Ebner (P_{CAT1_38B+53}, $P_{CAT1\_1A+14B+26B}$, $P_{CAT1\_38B+41B+48B}$, $P_{CAT1\_14B+26B+48B}$ and $P_{CAT1\_MM13b+38B}$) and the control promoters $P_{CTA1\text{-}WT}$ and P_{DF} were amplified one time attaching an overhang to the CalB associated α signal sequence (P17947-P17951) and one time attaching an overhang to the hGH associated shortened α signal sequence (P17947, P17952-P17954). Further, with selected primers a *Spe*I restriction site was introduced. For a detailed list of produced amplicons and corresponding primers see **Table 2.11**. The inserts with an overhang to the CalB signal sequence were ligated with the linearized pPp_AOX1-pUC-ZeoR-EcoRI-α_CalB-TT_AOX1 vector in assembly reactions. Inserts with an overhang to the hGH signal sequence were ligated with the pPp_AOX1-pUC-ZeoR-EcoRI-Dα_hGH-TT_AOX1 vector. This procedure resulted in plasmids pPp_AOX1-pUC-ZeoR-SpeI-pCTA1_variant-α_CalB-TT_AOX1 and pPp_AOX1-pUC-ZeoR-SpeI-pCTA1_variant-Dα_hGH-TT_AOX1 as well as plasmids with control promoters P_{DF}, $P_{CTA1\text{-}WT}$ and the negative control without a promoter. The obtained assembly mix was processed as already described above and after positive sequencing results the vector was linearized with *Smi*I and used for *P. pastoris* transformations.

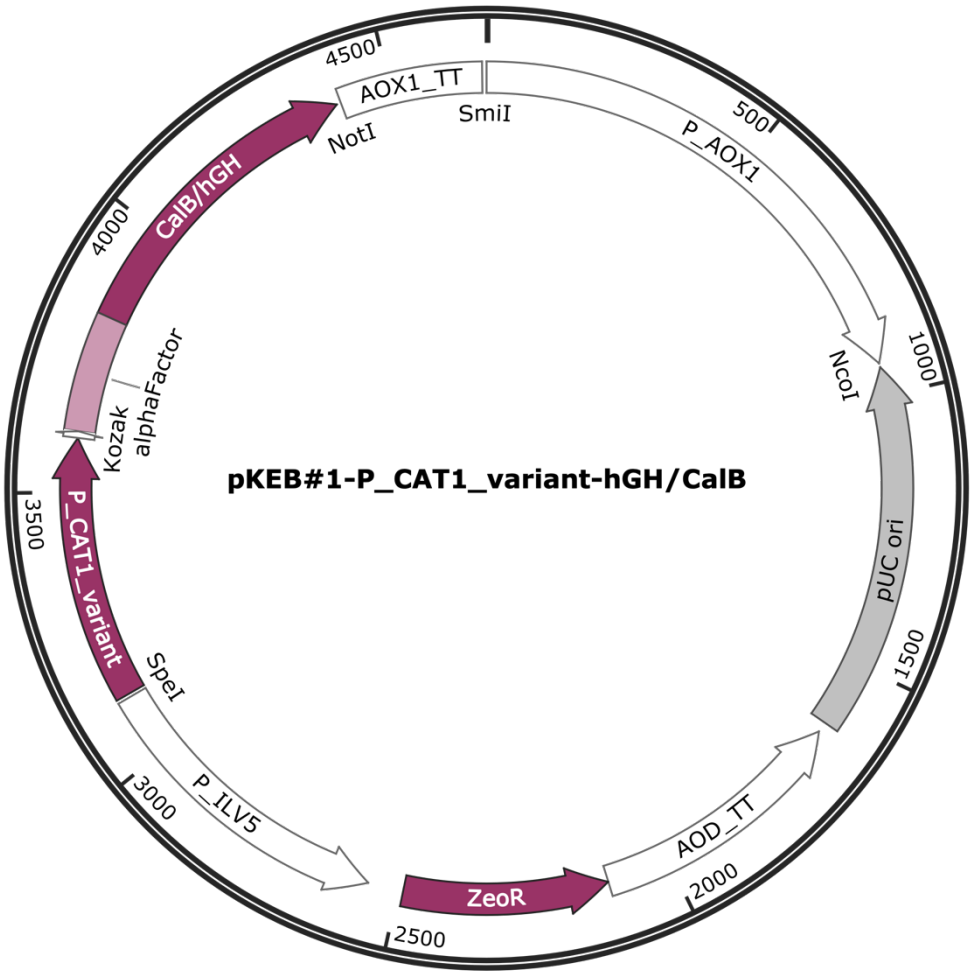


Figure 2.1 Plasmid construct based on the pKEB#1 vector expressing CalB or hGH. It contains two *AOX*1 homologous regions for *P. pastoris* genome integration, pUC ori for replication in *E. coli* and a Zeocin™ resistance cassette governed by P_{ILV5} for selection. The backbone was assembled with a P_{CTA1} promoter variant, CalB or hGH as gene of interest and the corresponding α-factor.

Table 2.11 Primers used to amplify the GOIs with an *Eco*RI restriction site and the signal sequence as well as primers used to amplify promoter variants with corresponding overhangs are listed with name and internal number.

Amplicon	**Internal Primer Nr°**	**Primer Name**
Gene of interest	**#**	**Primer NAme**
EcoRI-α_CalB	P17666	alpha+CalB-ILV5-Gib-fw
	P17667	alpha+CalB-AOX1_TT-Gib-rev
EcoRI-Dα_hGH	P17668	Dalpha+hGH-ILV5-gib-fw
	P17669	Dalpha+hGH-AOX1_TT-rev
Promoter variants	**#**	**Primer Name**
Fwd promoter for all variants	P17947	ILV5-SpeI-fw
SpeI-PCAT1_38B+53-Dα	P17950	Dalpha-PCAT500-53-rv
SpeI-PCAT1_1A+14B+26B-Dα	P17948	Dalpha-PCAT500-WT-rv
SpeI-PCAT1_38B+41B+48B-Dα	P17949	Dalpha-PCAT500-48B-rv
SpeI-PCAT1_14B+26B+48B-Dα	P17949	Dalpha-PCAT500-48B-rv
SpeI-PCAT1_MM13b+38B-Dα	P17948	Dalpha-PCAT500-WT-rv
SpeI-PCAT1_WT-Dα	P17948	Dalpha-PCAT500-WT-rv
SpeI-PDF-Dα	P17951	Dalpha-PDF-rv
SpeI-PCAT1_38B+53-α	P17954	Alpha-EcoRI-PCAT500-53-rv
SpeI-PCAT1_1A+14B+26B-α	P17952	Alpha-EcoRI-PCAT500-WT-rv
SpeI-PCAT1_38B+41B+48B-α	P17953	Alpha-EcoRI-PCAT500-48B-rv
SpeI-PCAT1_14B+26B+48B-α	P17953	Alpha-EcoRI-PCAT500-48B-rv
SpeI-PCAT1_MM13b+38B-α	P17952	Alpha-EcoRI-PCAT500-WT-rv
SpeI-PCAT1_WT-α	P17952	Alpha-EcoRI-PCAT500-WT-rv
SpeI-PDF-α	P17955	Alpha-EcoRI-PDF-rv

2.4.2 Construction of new P_{CTA1} promoter variants

The design of plasmids expressing eGFP with new promoter variants containing four or more mutations was based on the pKEB#1 vector with eGFP and an *Eco*RI restriction site already inserted (pPp_AOX1-pUC-ZeoR-EcoRI-eGFP-TT_AOX1). The elements of the vector backbone were described above. In order to create new promoter variants, those from Katharina Ebner containing the best mutations were used as templates in PCR reactions. Depending on where in the template the mutation was located, part A or part B of the promoter was amplified (**Figure 2.2)** and then fitting parts from different templates were recombined via assembly cloning. Five promoter variants were used as templates to create three distinctive part A's and three distinctive part B's to be re-ligated into five new promoter variants. In a PCR reaction, promoter part A (P17947/P18016) was provided with a *Spe*I restriction site, an overhang to the vector backbone (5' end) and an overhang to promoter part B (3' end). Promoter part B (P17673/P18017) was amplified to have overhangs to part A on the 5' and the vector backbone on the 3' end. Due to those overhangs part A and B could be recombined and assembled with the vector backbone in a single cloning step. The resulting plasmid is shown in **Figure 2.3** and for an assembly description of vector and fragments and associated mutations see **Table 2.12**. The procedure resulted in plasmids pPp_AOX1-pUC-ZeoR-SpeI-pCAT1_newvariant-eGFP-TT_AOX1 as well as plasmids with control promoters P_{DF}, WT and the negative control without a promoter. The obtained assembly mix was processed as already described above and after positive sequencing results used for *P. pastoris* transformations.

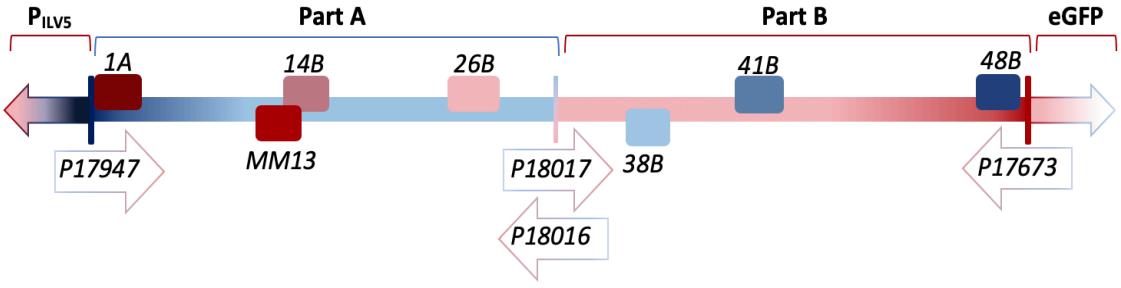


Figure 2.2 The P_{CTA1} promoter with the localization of the best mutations found by Katharina Ebner. Depending on where in the template the required mutation was situated the corresponding promoter pair was used to amplify either part A or part B. Overhangs were established at the 5' end to the P_{ILV5} and at the 3' end to part B (part A) and on the 5' end to part A and 3' end to eGFP (part B).

Table 2.12 The variants used as templates in PCR reactions are listed with their internal number, the corresponding isolated mutations and the produced fragment. Further, the pairing of the assembled fragments resulting in new plasmids is depicted with the plasmid name and P_{CTA1} promoter variant number.

P_{CTA1} promoter variant Nr°	**Isolated Mutations**	**Fragment**
19	1A+14B+26B	A1
46	MM13b	A2
11	26B	A3
23	38B+41B+48B	B1
46	38B	B2
34	48B	B3
#	**New Plasmids**	**Assembled Fragments**
55	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+41B+48B+MM13b-eGFP-TT_AOX1	A2+B1
56	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+41B+48B+26B-eGFP-TT_AOX1	A3+B1
57	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+14B+26B+38B-eGFP-TT_AOX1	A1+B2
58	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+14B+26B+48B-eGFP-TT_AOX1	A1+B3
59	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+14B+26B+38B+41B+48B-eGFP-TT_AOX1	A1+B1

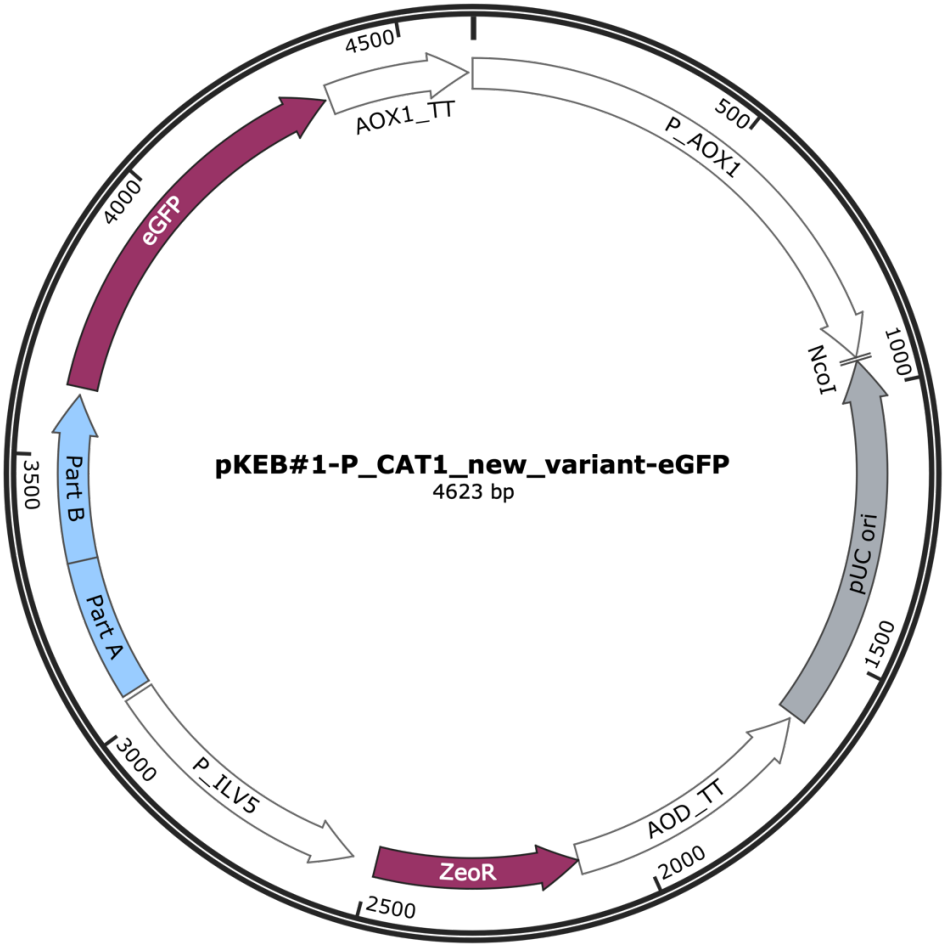


Figure 2.3 Plasmid construct based on the pKEB#1 vector expressing eGFP. It contains two *AOX*1 homologous regions for *P. pastoris* genome integration, pUC ori for replication in *E. coli* and a Zeocin™ resistance cassette governed by P_{ILV5} for selection. Assembled with a **new P_{CTA1} promoter variant** controlling the expression of **eGFP**.

3. RESULTS AND DISCUSSION

3.1 Rescreening of 54 P_{CTA1} promoter variants

In the previous master thesis addressing this topic Katharina Ebner established 54 P_{CTA1} promoter variants (unpublished) by combining the best mutations from Mohamed Hussein's work (102) into double and triple mutations. The variants are listed in **Table 3.1** with their internal numbering and corresponding mutations in the P_{CTA1} promoter sequence. They had been cloned into pKEB#1 based vectors and introduced into *P. pastoris* strains via homologous recombination at the *AOX1* locus. For selection of successfully transformed clones Zeocin™ resistance was encoded in the expression cassette and the cells were grown on selective YPD agar plates. Additionally, growth counter selection on gar plates containing BMM2% was conducted. Enhanced green fluorescent protein (eGFP) was used as a reporter gene to determine expression of the variants. The transformed *P. pastoris* clones were cultivated in DWP on BMD1% and induced with BMM1% and BMM5% over several days. Fluorescence levels were determined after 24, 48, 60 and 98 hours via a plate reader at 488 nm excitation and 507 nm emission. The respective values were normalized to OD_{600} and compared to the WT P_{CTA1}. The time points were chosen to have results from each day (24 and 48 hours), in de-repressed state and before induction (60 hours), as well as final values after induction (98 hours).

To verify the consistency of the fluorescence levels of the strains a re-screening using identical conditions was conducted, as well as adding an additional cultivation with BMG1% media. Based on the re-screening results in coordination with the screening of Katharina Ebner the best variants were picked for follow up experiments. In those experiments BMG was compared to BMD as alternative carbon source, especially for de-repressed cultivation strategies. Therefore, it seemed consequential to introduce a BMG cultivation round in the rescreening. For ensuring comparability to the previous results buffered minimal media containing 1% and a different inoculation method as described in the "Material and Methods" section was used, i.e. the BMD1% main culture DWP was inoculated directly from selective YPD agar transformation plates without using a pre-culture. After one-hour of incubation 10 µL per well from the BMD1% were transferred to the BMG1% DWP for inoculation using a multi-channel pipette. The remaining cultivation was conducted as described. Due to different metabolic loads of the cells, uneven biomasses between the wells

occurred (**Supplementary data S1**). To account for this, the measured fluorescence level of each well was normalized to the respective OD_{600} value, estimating the different biomass concentrations.

Table 3.1 Internal numbering of promoter variants with corresponding mutations in the P_{CTA1} promoter established by Katharina Ebner. The abbreviations "ind.336-360" and "derep.271-294" described the insertion of P_{CTA1} wildtype sections associated with induction (336-360 bp) or de-repression (271-294 bp), respectively. For variant 1 and 2 the sections were inserted upstream of the original part, doubling it. In variant 28 and 29 they substituted a promoter segment not associated with significant functions (99-122 bp). Numbers with capital letters "A" or "B" described the substitution of 10 bp of the native sequence with one of two neutral sequences A or B containing no putative TFBS, e.g. mutation 1A had the first 10 bp of P_{CTA1} exchanged with neutral sequence "A". Stand-alone numbers depicted the deletion of putative TFBSs and MM13b were multiple mutations around putative TFBS "13" caused by a faulty gBlock. A comprehensive overview for all the putative TFBSs predicted by MatInspector, is provided in the master thesis of Mohamed Hussein (102), unpublished.

#	P_{CTA1} mutation	#	P_{CTA1} mutation	#	P_{CTA1} mutation
1	double-ind.336-360	*19*	1A+14B+26B	*37*	1+41B
2	double-derep.271-294	*20*	14B+26B+53	*38*	1+53
3	1A+4	*21*	1A+26B+53	*39*	1+48B
4	1A+MM13b	*22*	1A+14B+53	*40*	4+26B
5	1A+14B	*23*	38B+41B+48B	*41*	4+38B
6	4+MM13b	*24*	MM13b+1+4	*42*	4+41B
7	4+14	*25*	14B+26B+41B	*43*	4+53
8	MM13b+14B	*26*	14B+26B+48B	*44*	4+48B
9	26B+38B	*27*	14B+26B+38B	*45*	MM13b+26B
10	26B+41B	*28*	flexi_ind.336-360	*46*	MM13b+38B
11	26B+53B	*29*	flexi_derep.271-294	*47*	MM13b+41B
12	26B+48B	*30*	1A+26B	*48*	MM13b+53
13	38B+41B	*31*	1A+38B	*49*	MM13b+48B
14	38B+53	*32*	1A+41B	*50*	14B+26B
15	38B+48B	*33*	1A+53	*51*	14B+38B
16	41B+53	*34*	1A+48B	*52*	14B+41B
17	41B+48B	*35*	1+26B	*53*	14B+53
18	48B+53	*36*	1+38B	*54*	14B+48B

Another matter of concern that had to be addressed was oxygen limitation. When cell cultures are provided with sufficient or too high nutrient concentrations, the most common limiting factor is the oxygen transfer capacity in the medium (116). Especially in DWP and microtiter plates the oxygen uptake rate into the medium is rather low due to the small surface of the boundary layer between liquid and gaseous phase (117). Thus, cell growth and protein production are merely a function of the mass transfer rate of oxygen. Additionally, centered wells in the DWP have worse oxygen supply, making for uneven growth conditions when oxygen acts as a limiting source. These aspects can complicate comparing the expression of promoter variants, because the protein production would not be defined by the differences in promoter strength, but the amount of available oxygen. Or in other words, an improved promoter could not exhaust its full potential, if the oxygen transferred to the cell was not sufficient. Hence, the determined protein production levels would not correspond to the actual promoter strength. In order to avoid such conditions several precautions were undertaken. The DWP cultivations were incubated at 28 instead of 30°C to mildly slow down cell growth. The DWP were placed on a shaker at high frequency to ensure maximum mixing and oxygen exchange at the boundary layer and vigorous diffusion in the cell suspension. Nonetheless, in the results discussed in sections 3.1.1 to 3.1.4, it seemed as if oxygen supply was still the limiting factor. Thus, when buffered minimal media were used in subsequent cultivations, they contained a maximum of 0.5% carbon source to slow down cell metabolism and hence the oxygen demand.

3.1.1 Reproducibility of the screening results

Two data-/screening-sets were compared, the re-screening results and the initial screening performed by Katharina Ebner. For both screenings, samples were taken after 24, 48, 60 and 98 hours and at each time point the same 54 promoter variants and three controls were measured. The promoter variants were measured in quadruplicates and controls in triplicates, and their mean values were established.

As a general benchmark for reproducibility, variants from same sampling time points were compared between re-screening and screening data and the standard deviation was calculated. If the standard deviation was under 20%, it was regarded as good.

Since the two datasets were measured on different days by different people using not exactly the same instruments for the procedure (e.g. pipettes), just to name a few deviating factors, statistical bias could not be entirely prevented. The issue was that for each sampling time

point one screening provided overall higher values over the other. So even though when variants behaved similar within re-screening and screening (the variants would be roughly in the same order when organized by performance), the datasets seemed inconsistent to one another, since one of the two was superior in terms of absolute numbers. To account for this, certain corrections were undertaken – separately per sampling time point. The average value of the variants was calculated for each screening set. Subsequently, the difference between the averages was subtracted from the screening set with the higher values. This should provide a more even base for data comparison by adjusting the dataset with higher values to the one with lower values.

For a third statistical tool a moving average trendline (**Figure 3.1**) was applied to counter another notable trend, especially visible in the 60- and 98-hour sample of the re-screening. It seemed, as if clusters of variants in the re-screening performed collectively superior or inferior compared to the initial screening, while within these clusters the relation between the variants stayed roughly identical. This bias could be referable to the fact, that the variants were divided into three separate DWP due to well number limitations. Thus, although all DWP were treated as identical as possible, they might have experienced slight, but notable differences, such as position in the shaker, order of inoculation and sampling, minor differences in the DWP itself or human error. Said trends in the clusters of variants correlated with their allocation to the DWPs, e.g. in the 98-hour sample, variants 1 to 21 showed superior results and they were all cultured in the same DWP. To level out this indifferencies same method as above (approximation by mean values) could not be applied, since compared groups contained different variants. Thus, moving average trendlines for both screenings were used, which illustrated that the correlation of the variants in both datasets was coherent by following the same progression. Only with a cluster of variants underlying a trend, the trendline from the re-screening would proceed above or below the one from the initial screening data, but in the same fashion.

Another factor taken into consideration was that the values compared between the two screenings were not absolute numbers but established by normalization to the relative fluorescence units of the parental P_{CTAI} wildtype (WT) positive control. Hence, if the WT values were distorted on one DWP, the bias could even be more amplified. Thus, for the normalization the average of the WT positive controls of all DWPs was used. This provided more data points and the possibility to exclude poor ones.

The re-screening results compared to the initial screening performed by Katharina Ebner differed depending on the sampling time points. The earlier the samples were taken, the more deviation between same variants was observed. The first sample was obtained after 24 hours and in this study, as well as in the results of Katharina Ebner (unpublished) the differences within the quadruplicates for each variant, and between the two screenings were too high to be useful. Thus, they were not evaluated and will not be discussed. In the following 48-hour sample slightly more correlation was found (**Figure 3.1A**). Between variants 1 and 21 only three exceeded a standard deviation of 20%, while from the residual 33 variants 17 did so. From those, the distribution of screening sets with the higher value seemed random. The average of the re-screening values was slightly higher, than that of the screening with 1.640 as opposed to 1.612 (RFU/OD_{600}/WT). The 60-hour sample showed some similar patterns as the previous one (**Figure 3.1B**). Between variants 1 to 21 a good reproducibility with only two standard deviations above 20% was observed and while in the residual variants 18 more were counted, the mean standard deviation was lower than in the previous 48-hour sample (14.84% versus 16.23%). This was supported by a more coherent behavior of the moving average trend line of re-screening and screening. Again, the re-screening data set had a higher average than the screening with substantial 1.867 to 1.578. Among variants with high standard deviation, no pattern was observed between variants 22 to 42, whereas screening values of Katharina Ebner were tendentially higher regarding variants 43 to 54. It is worth mentioning that the differences in the last section were consistent, meaning that the two moving average trend lines ran parallelly in the same pattern with the re-screening trend line situated lower. Further, there were eight variants with standard deviations above 20% in the 60-hour sample, which were inconspicuous in the 48-hour sample, suggesting inaccuracies in either cultivation or screening. For instance, optical density measurements were sometimes found to be challenging with *P. pastoris*, since the organism tended to sediment relatively fast and aggregate in the late phases of cultivation.

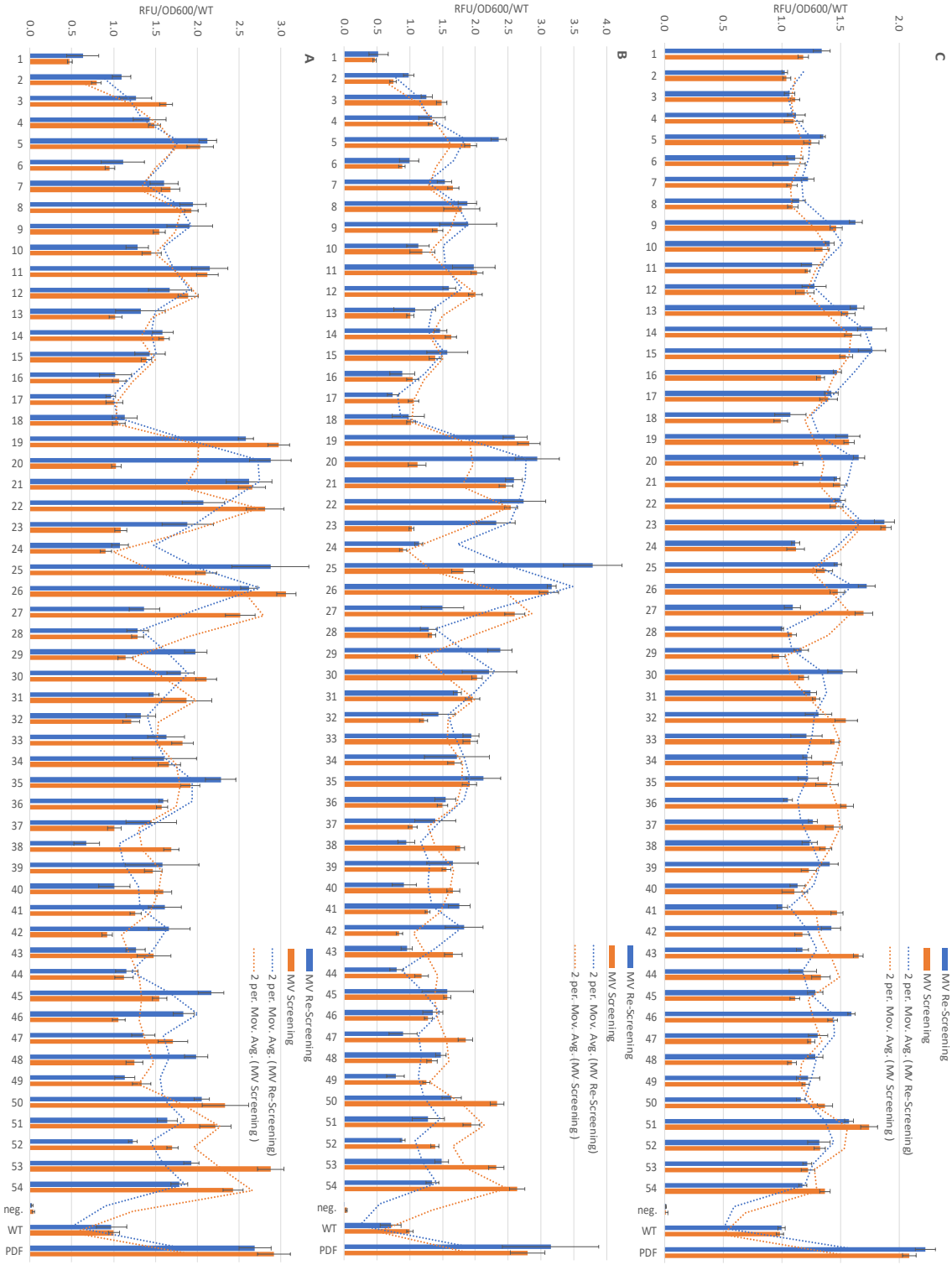


Figure 3.1 Comparison between re-screening and screening results of eGFP fluorescence levels expressed by P_{CTA1} variants in *P. pastoris*. The relative fluorescence units (RFU) were normalized to the respective OD600 value and the RFU of the P_{CTA1} wildtype (WT). The values were blotted against the 54 promoter variants, a variant with no promoter as negative, the P_{CTA1} WT promoter as positive and the P_{DF} promoter as benchmark control. The blue bars represent the mean values (MV) of the re-screening and the orange of the initial screening results. The dotted line shows the moving average in the respective color. Diagram **A**, **B** and **C** show results after 48, 60 and 98 hours, respectively. The cells were cultivated at 28°C in DWP containing BMD1%. After 60 hours they were induced to a methanol concentration of 0.5%, which was approx. maintained for another 38 hours of the cultivation.

The 98-hour sample displayed best coherence between re-screening and screening data sets (**Figure 3.1C**). Only five variants had an unsatisfactory standard deviation and the moving average trend lines were almost consistently in synchronization or at least running in same patterns in a parallel fashion. The standard deviation averaged out at 8.20% with only one variant exceeding 30% and 75% of the variants remaining under 10%. The averages of the data sets seemed nearly identical with merely 0.01% in difference, yet upon closer examination two trends were observed. In the re-screening the cluster of variants 1 to 21 was tendentially superior and 27 to 42 tendentially inferior to the initial screening results. This effect was already discussed and accredited to the allocation of promoter variant strains to three separate DWP, where variant 1 to 21, 22 to 42 and 43 to 54 were grouped together plus the controls.

The trend of more scattering and lower correlation between re-screening and screening in earlier samples (24, 48h) and more coherent results in later samples (60 and 98h) might be referable to unequal starting biomass and different metabolic loads per strain. The inoculation of the BMD1% main culture was performed by transferring transformed colonies from selective YPD agar plates into DWPs via sterile toothpicks. The amount of picked biomass could not be chosen precisely with this method, hence each well obtained different initial biomass concentrations. In wells with higher biomass concentration the initially provided carbon source depleted faster, resulting in an earlier promoter de-repression and recombinant protein production than in wells with less biomass. Additionally, the clones might have been picked in different growth phases, leading to delayed replication and protein production for some. While with the normalization of the RFU to the respective OD_{600} values the uneven biomass distribution could be compensated to some extent, the cell metabolism was not considered. As the cultivation went on, the cell cultures in each well reached growth saturation and the recombinant protein production was induced by methanol addition. This led to expression levels not governed by the amount of biomass, but the strength of the different promoter variants, making for more consistent results. Also, the higher overall values upon induction made small differences less significant. To provide additional statistical stability in earlier samples, the cultivation strategy in onward experiments was reconfigured. Instead of transferring colonies from the transformation plate into the main culture, a pre-culture was interposed. It was used to grow the cells for 60 hours until saturation, so the biomass in the wells was roughly equal. Then, a defined amount (mostly 10 µL) of each well was transferred with a pipette to the corresponding well in the main culture ensuring approximately same biomass concentrations and same growth phase state.

Another notable trend was that samples taken after 60 or 98 hours exhibited sections where either the re-screening or screening values were tendentially higher (e.g. samples taken after 98 hours showed higher values from the re-screening data for variants 1 to 21). This can be explained by the fact that the 54 promoter variants and three controls were divided onto three different DWP for cultivation due to limitation of well numbers on each plate. The segmentation corresponded with the trends and was as follows: variants 1 to 21, 22 to 42 and 43 to 54 plus controls on each plate. Although all DWP were treated as identical as possible, small influences, such as position on the shaker, order of inoculation and sampling, minor differences in the DWP itself or human error could have influenced the different plates, producing a statistical bias for each one. Additionally, the values compared between re-screening and screening were not absolute numbers but established by normalization to the RFU of the WT positive control. If the WT values were distorted on one plate, the bias could even be more amplified. Thus, for the normalization the average of the WT values of all plates was used. This provided more data points and the possibility to exclude poor ones. For the sake of completion **Supplementary data S2** shows eGFP values normalized to the WT of the respective DWP. A consensus was found between re-screening and screening regarding certain mutations associated with improved expression properties either upon de-repression or induction. This was crucial for the choice of promoter variants for further experiments. The prominent mutations will be discussed in the next section.

3.1.2 Re-screening results of P_{CTA1} variants cultivated on BMD1%

Although discrepancies with the screening results of Katharina Ebner occurred as described in the previous part, the key features stayed consistent. Most importantly mutations associated with strong expression either due to induction or de-repression were identical between both screens. The most prominent mutations linked to good expression values upon induction were 38B and 41B. Their partaking in almost all strong inducible promoter variants was best visible, when the fluorescence increase from the 60- to the 98-hour sample was observed (**Figure 3.2**). The difference corresponded to the eGFP produced after methanol induction, if we assumed that the previously produced protein was fully stable. The diagram showed that all variants reaching above 20,000 RFU/OD_{600} except one (variant 1) contained either 38B or 41B and when both combined, they made up the strongest inducible promoter variant (variant 13). Also noteworthy as mutation associated with mediocre to good inducible promoter variants was 48B. Mutations prominent for powerful de-repressed promoters were 1A, 14B and 26B. Although the combination of those three mutations did not lead to the

best de-repressed promoter variant (variant 25), they showed the most consistency. Eleven out of thirteen promoter variants (variants 5, 9, 11, 19-22, 25, 26, 30, 35) reaching a 2.5-fold expression of eGFP after 60 hours (de-repressed state) compared to the WT were affiliated with either one of the three mutations (**Figure 3.3**). Mentionable as mutation linked to mediocre to good de-repressed promoter variants was the deletion "53".

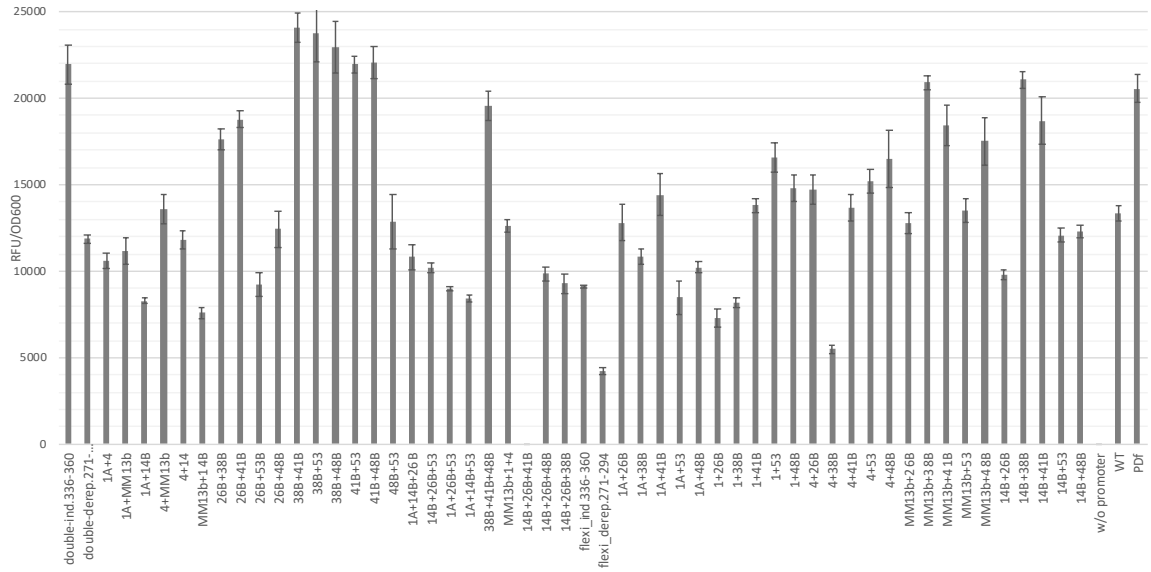


Figure 3.2 eGFP fluorescence levels expressed by P_{CTA1} promoter variants in *P. pastoris* between the 60- and the 98-hour sample upon methanol induction. The RFU were normalized to the respective OD_{600} values and blotted against variants and controls. The cells were cultivated at 28°C in DWP containing BMD1%. After 60 hours they were induced to a methanol concentration of 0.5%, which was approx. maintained for another 38 hours of the cultivation.

Since the P_{CTA1} promoter had only been putatively annotated by Hussein (102), the attribution of the mutations to specific TFBSs came with a challenge. However, a detailed discussion can be found in the master thesis of Katharina Ebner (unpublished). By mutating the promoter surprisingly so far unknown existing inhibitory regulatory sites were destroyed or new activating regulatory sites randomly established. It was striking that most prominent mutations were mainly substitutions with the randomized neutral sequence "B" (GATAACCGTG). This might be referable to the first four nucleotides "GATA", which build the core sequence of the GATA-TFBS (118). The GATA-sequence attracts transcription factors with zinc finger motifs based on nitrogen availability. In nitrogen rich media repressors such as *DAL80* inhibit the expression of the downstream gene, while in nitrogen poor media activators including *GAT1*, *GAT2*, or *GLN3* amplify gene expression (119). Hence, by introducing sequence "B" into the P_{CTA1} promoter, additional GATA binding motifs were established, which could lead to improved expression levels under the right conditions.

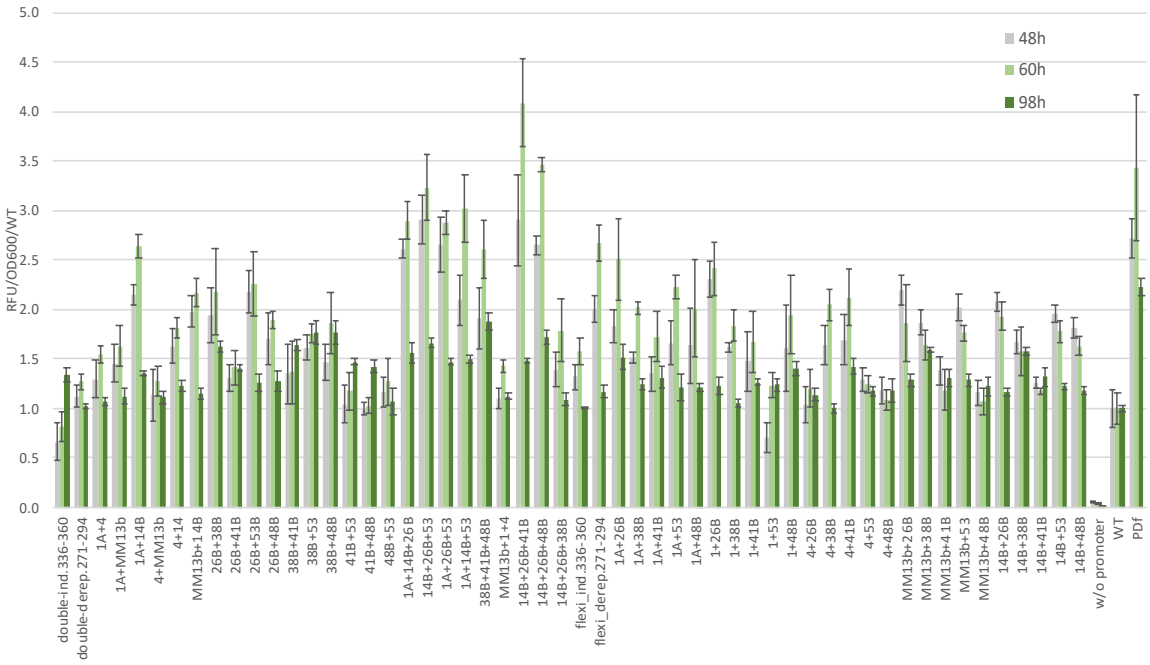


Figure 3.3 Relative eGFP fluorescence expressed employing P_{CTA1} promoter variants in BMD cultivations normalized to OD_{600} and the WT positive control. The bars in different colors represent results at respective sampling time points 48, 60 and 98 hours. The x-axis shows the controls and promoter variants with introduced mutations and the y-axis their eGFP fluorescence compared to the WT P_{CTA1} promoter (e.g. a value of two means double the amount of eGFP expressed). The cells were cultivated at 28°C in DWP containing BMD1%. After 60 hours they were induced to a methanol concentration of 0.5%, which was approx. maintained for another 38 hours of the cultivation.

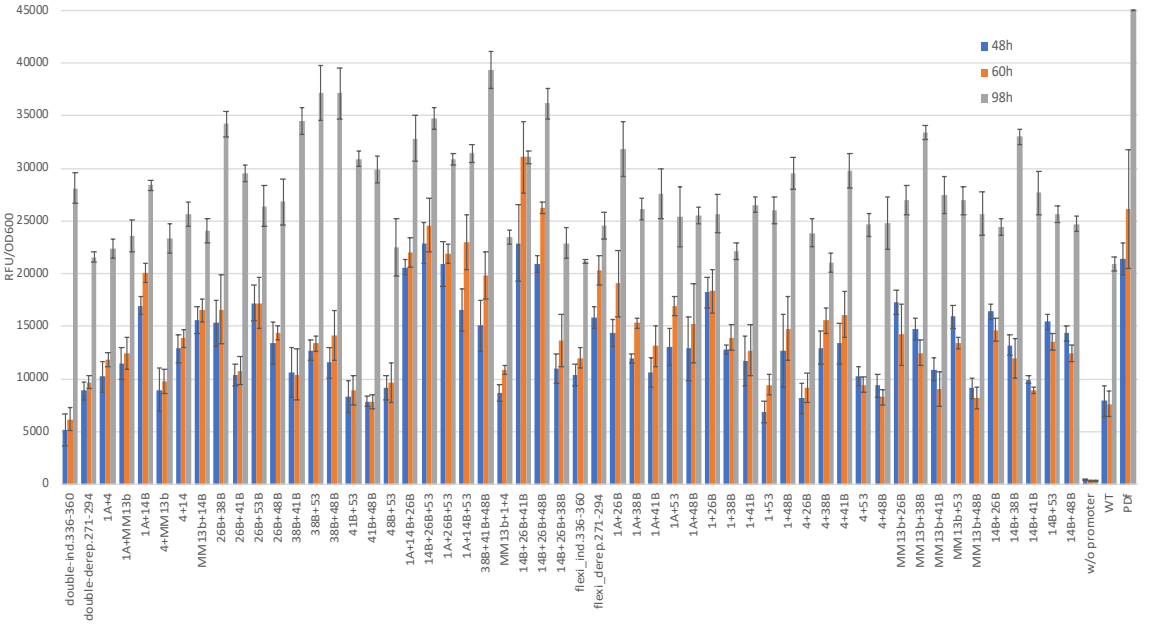


Figure 3.4 Relative eGFP fluorescence expressed employing P_{CTA1} promoter variants in BMD cultivations normalized to OD_{600}. The bars in different colors represent results at respective sampling time points 48, 60 and 98 hours. The x-axis shows the controls and promoter variants with introduced mutations and the y-axis their eGFP fluorescence normalized to OD_{600}. The cells were cultivated at 28°C in DWP containing BMD1%. After 60 hours they were induced to a methanol concentration of 0.5%, which was approx. maintained for another 38 hours of the cultivation.

These conditions were potentially unintentionally provided by using buffered minimal media with limited amounts of nitrogen. However, according to Weis *et al.* (120) salts including nitrogen sources were not limiting in DWPs under the performed conditions and earlier unpublished experiments with the P_{CTA1} promoter in our group showed no clear nitrogen starvation effects. Whenen blasting activating transcription factors of the GATA family against the *P. pastoris* genome several positive hits were obtained. This included the genes *GAT1*, *GAT2* and *GLN3* encoding for proteins (Accession Numbers: *F2QYN5*, *F2QZQ9* and *A0A1B2J6U0*, respectively), each protein containing a zinc finger domain, which binds specifically to DNA consensus sequence (A/T) GATA (A/G) as promoter element.

Overall the mutations lead to significantly higher fluorescence levels than the parental WT P_{CTA1} strain. Among the 54 promoter variants received by Katharina Ebner, there were only three with a lower eGFP fluorescence than the wildtype after 48 hours, one after 60 hours and none after 98 hours of incubation. One may think that combining mutations led to overall improved strains. However, looking at the eGFP reporter production some variants performed poorly after methanol induction. Regarding protein produced between the 60- and 98-hour sample, 29 promoter variants were weaker compared to the wildtype in terms of RFU/OD_{600} (**Figure 3.4**). Of all variants 53.7% produced more than half of the protein before induction and 40.7% of the variants even before de-repression (48-hour sample). While this ensured an overall high fluorescence level, since promoters started with gene expression early on, some mutations seemed to impair tightness of repression and the induction strength of the promoter. This was especially true for variants bearing triple mutations (variant 19-27). They delivered impressive results in the early stages of the cultivation, such as the five strongest promoter variants after 48 hours of cultivation (variant 19-21, 25, 26) and the four best de-repressed promoters (variant 22, 23, 25, 26). But the performance flattened out after 60 hours of cultivation, where eGFP production in all triple mutation variants but one was lower than in the wildtype control. The one better than the wildtype contained mutations 38B and 41B supporting previous statements. Variant 25 entirely lacked protein production in the induction phase, if we assumed that the previously produced protein was stable (Verkhusha *et al.* (121) report an eGFP fluorescence loss of 7% over 100 hours at 37°C).

3.1.3 Screening results of P_{CTA1} variants cultivated on BMG1%

To test an alternative expression strategy *P. pastoris* strains bearing P_{CTA1} promoter variants where cultivated on glycerol and recombinant protein production was preferentially induced

by promoter de-repression. Therefore, the strains were cultivated in DWPs containing BMG1% as a primary carbon source. After 60 hours this carbon source was presumably depleted, bringing the promoter in a de-repressed state. This state was retained for another 38 hours (total cultivation time of 98 hours) by maintaining a glycerol concentration of 0.25% with the regular administration of glycerol pulses.

Overall, the fluorescence levels of eGFP under P_{CTA1} promoter variants was lower with this expression strategy than when grown on BMD1% and induced with methanol. To be more precise, after 98 hours cultivation time the average produced eGFP levels were only 42.15% compared to glucose cultivations and methanol induction, when looked at RFU/OD_{600}. Mentionable was that actually most of the promoter variants (72%) in glycerol cultivations and de-repressed conditions lost fluorescence activity from the 60- to 98-hour sample, as visible in **Figure 3.7**, although the transformed cells continued to grow during that period of time, according to OD_{600} values. Additionally, the variants which showed positive development between last two time points (60h and 98h) only increased by 6.74%, as opposed to BMD cultures, where the eGFP production was boosted by 113.28% upon methanol induction. With samples taken after 48 or 60 hours, the consequence of which expression strategy was used was not as striking. The former time point (48h) showed an average eGFP formation of 74.21% compared to BMD cultures and methanol induction and the latter (60h) 88.86%. In-between those two time points cultures based on glycerol were superior to the ones based on glucose with an average eGFP fluorescence increase of 3013 RFU/OD_{600} versus 1229. This might be explained by the different growth rates and time point of carbon source depletion depending on the choice of carbon source. According to absorption values obtained by a plate reader at 600 nm, after 24 hours *P. pastoris* cell concentrations were three times as high when grown on glucose than compared to glycerol (**Figure 3.5**). However, the growth of glucose-based cells seemed to be already declining at this point of time, while the growth of glycerol-based cells slowly stopped increasing approximately between the 48- and 60-hour sample. This suggested that most of the recombinant protein production in de-repressed state occurred between the 24- and 48-hour sample regarding cells cultivated on glucose as primary carbon source and the 48- and 60-hour period regarding cells cultivated on glycerol.

Taking into consideration that hardly any recombinant protein was produced after the 60-hour sample in glycerol cultivations, and expression in glucose cultivations coupled with

methanol induction worked seemingly better in the later period of recombinant protein production (60 to 98h) explained the differences in total fluorescence levels.

Prominent mutations associated with good eGFP production upon certain conditions were similar between strains grown on either glycerol or glucose. As expected, mutations 38B and 41B, which were involved in strong inducible promoters, where unremarkable in BMG cultivations, since no methanol induction phase was applied. Mutations leading to strong de-repressed promoters were mostly identical between both cultivation strategies. At the 48-hour sample in BMG cultivations 22 promoter variants exceeded 10,100 RFU/OD_{600} of which 20 contained either mutation 14B, 26B and/or 53. Regarding eGFP produced between the time points 48 and 60 hours of cultivation the mutations MM13b, 14B, 26B and 53 were most prominent. They were found in 18 of the 20 best variants during that cultivation period. Noteworthy is that strains containing triple mutations produced the best variants among the 48- and 60-hour samples, as well as the highest eGFP fluorescence increase between those two time points. In comparison to the WT, which is shown in **Figure 3.6**, the promoter variants were evidently superior at the 48-hour sample, when all but two promoter variants (38, 40) had higher expression levels, the highest of up to 3.51-fold (variant 21). On the other hand, merely 19 variants exceeded the WT in eGFP production between 48 and 60 hours of cultivation.

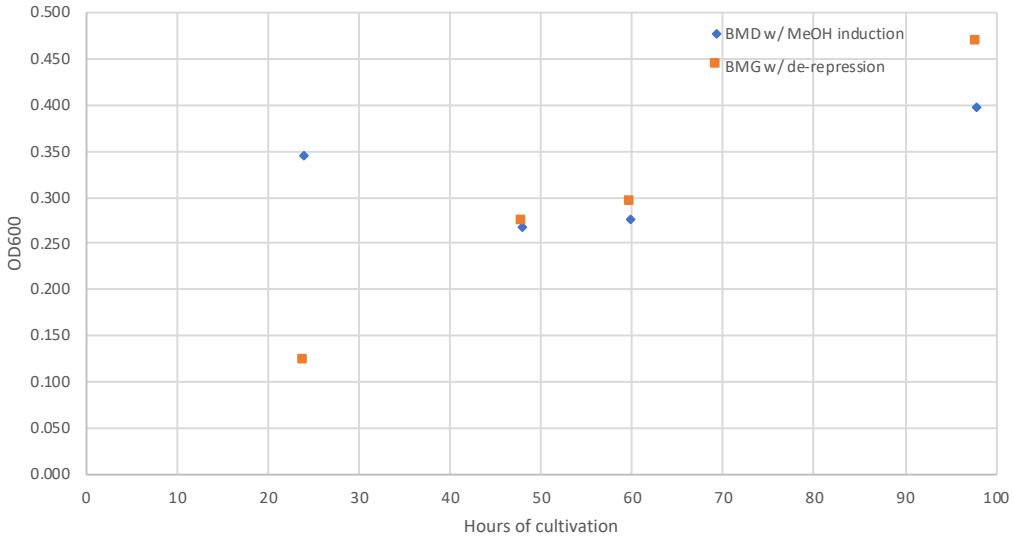


Figure 3.5 Cell density of *P. pastoris* cultivations in DWP. The figure confronts the progression of cell growth in glucose- and glycerol-based cultivations. Absorption values at 600 nm were determined by a plate reader after 24, 48, 60 and 98 hours. The blue symbols depict results from transformed cells cultivated in DWPs on BMD1% at 28°C. After 60h they were induced with methanol to a concentration of 0.5%, which was maintained for another 38h. The orange symbols show results from transformed cells that were cultivated at 28°C in DWPs containing BMG1%. After 60h of cultivation BMG pulses were applied to maintain a glycerol concentration of 0.25% for another 38h.

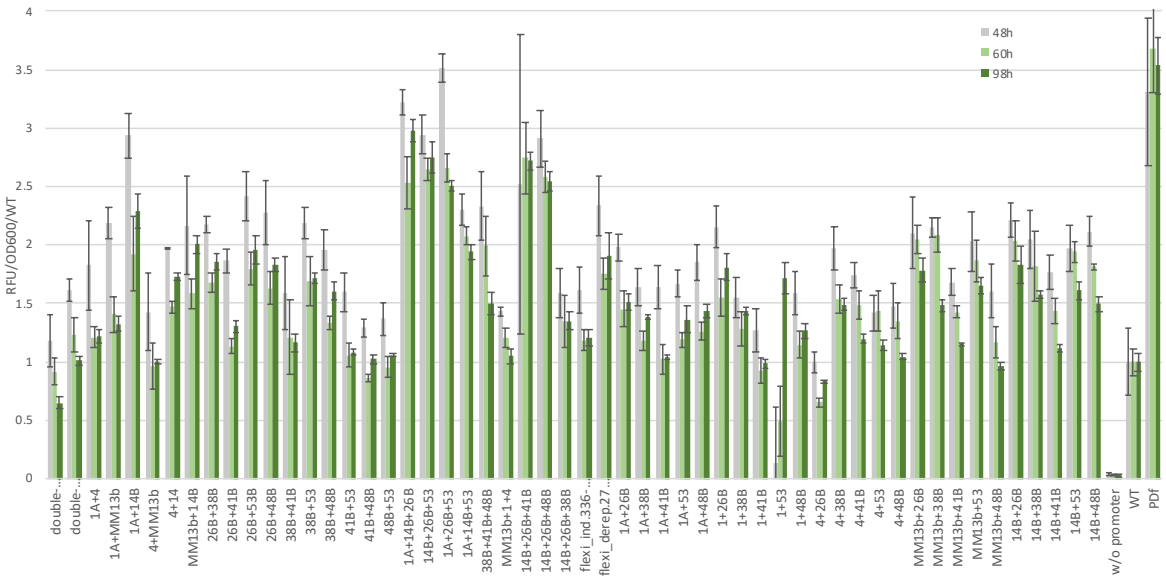


Figure 3.6 Relative eGFP fluorescence expressed under P_{CTA1} promoter variants in BMG cultivations normalized to OD_{600} and the WT positive control. The bars in different colors represent results at respective sampling time points 48, 60 and 98 hours. The x-axis shows the controls and promoter variants with introduced mutations and the y-axis their eGFP fluorescence compared to the WT P_{CTA1} promoter (e.g. a value of two means double the amount of eGFP expressed). Cells were cultivated at 28°C in DWPs containing BMG1%. After 60h of cultivation BMG pulses were applied to maintain a glycerol concentration of 0.25% for another 38h.

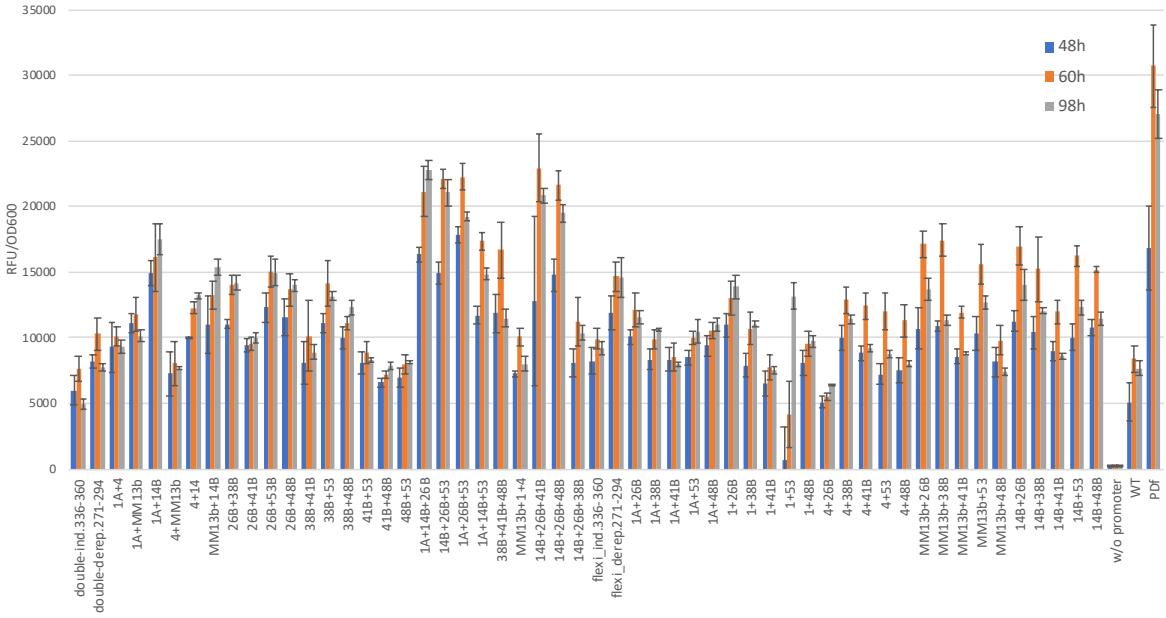


Figure 3.7 Relative eGFP fluorescence expressed under P_{CTA1} promoter variants in BMG cultivations normalized to OD_{600}. The bars in different colors represent results at respective sampling time points 48, 60 and 98 hours. The x-axis shows the controls and promoter variants with introduced mutations and the y-axis their eGFP fluorescence normalized to OD_{600}. Cells were cultivated at 28°C in DWPs containing BMG1%. After 60h of cultivation BMG pulses were applied to maintain a glycerol concentration of 0.25% for another 38h.

The results of eGFP expression with BMG grown cultures left some questions unanswered. It was hard to determine, whether eGFP production after 60 hours of cultivation truly stagnated or already produced protein was lost and further production was required to maintain relative fluorescence levels. The observation that most promoter variants showed weaker results at the 98- than the 60-hour sample supported the latter assumption. However, it was not possible to identify any source of protein degradation. On the one hand it did not seem that the cells degraded themselves to scavenge nutrients, since the OD_{600} values increased between time points (**Figure 3.5**). On the other hand, there are several reports confirming reliable stability of eGFP. Verkhusha*, et al.* (121) reported a relative fluorescence decrease of 7% over 100 hours at 37°C. Since we used milder cultivation conditions of 28°C the reporter protein should have been even more stable. eGFP was also proven to be highly resilient against exposition to detergents, high temperatures or proteases (122). Additionally, 22 promoter variants showed a fluorescence decrease of over 10% and up to 35%, exceeding the reported natural eGFP denaturation by far. However, also oxygen supply has an important influence on fluorescence intensity and eGFP maturation If still in doubt the use of a different reporter protein, such as *Discosoma DsRed*, a GFP like reporter protein, would be advisable, since it only shows around 2% relative fluorescence decrease over 100 hours (121). Also, faster maturing GFP variants or destabilized variants would be interesting in order to get more information about the promoter kinetics. mRNA levels would provide even more sophisticated information, but unfortunately this was not available as a high-throughput method, which would be needed in this context.

3.1.4 P_{CTA1} promoter variants for further experiments

In order to choose variants for the follow-up stages, the results from the re-screening and screening, as well as cultivation on BMG were taken into consideration. The elected promoter variants are listed in **Table 3.2**.

Table 3.2 P_{CTA1} promoter variants used in further experiments listed with internal numbering and introduced mutations.

#	Mutations
14	38B+53
19	1A+14B+26B
23	38B+41B+48B
26	14B+26B+48B
46	MM13b+38B

Variant 14 and 46 had a relatively tight regulation on BMD and were one of the strongest inducible promoters. The latter also showed outstanding production levels on BMG between 48 and 60 hours of cultivation. Variant 19 displayed high expression levels in the early stages on both media and the best overall production when BMG was used. Variant 23 had the highest production rate over 98 hours when cells were grown on BMD and variant 26 the best gradual improvements from one time point to another.

3.2 Extracellular expression of hGH and CalB using P_{CTA1} promoter variants

Extracellular secretion is one of the most prominent posttranslational modifications and advantages of *P. pastoris*. The recombinant protein is transported out of the cell and combined with low amounts of native protein, downstream processing is greatly facilitated. Thus, it was vital to test the improved P_{CTA1} variants with extracellular protein expression. The heterologous proteins used were human growth hormone (hGH) and *Candida antarctica* lipase B (CalB). First was a protein, which promotes growth of soft and skeletal tissue in humans, with a molecular weight of 22 kDa, containing 191 residues and two loops formed by disulfide bridges (123). Hence, it was a suitable subject to test the compatibility of the promoter variants with medium sized proteins requiring folding modifications. Additionally, hGH is a frequently expressed recombinant protein applied as clinical therapeutics, e.g. for growth-retarded children (124). The second protein was a lipase natively expressed by *Candida antarctica* with a molecular weight of 33 kDa and a length of 317 amino acids (125). Lipases in general are commonly applied in large-scale industrial processes due to their broad substrate scope and wide range of reactions (126). CalB in particular is favored, because of its

specificity in hydrolysis and synthesis (127,128). For our purposes, CalB was very practicable with its high rigidity. The enzyme is generally regarded as very stable and withstands a variety of conditions, such as pH ranges from 3.5-9.5, thermostability up to 50°C and tolerance to certain detergents (129). This provided the optimum requirement for our experiment, where it was crucial to have a stable enzyme, which would not be degraded during the 98 hours of cultivation

3.2.1 SDS-PAGE

Both enzymes were successfully expressed in DWP cultivation based on BMD0.5% and using methanol induction, which was made visible with an SDS-PAGE analysis (**Figure 3.8**), where the recombinant protein bands were displayed at the appropriate sizes (22 kDa for hGH and 33 kDa for CalB). The samples were taken from 98-hour cultivations on BMD0.5% with methanol induction. The intensity of the bands relative to each other was evaluated with the software Image Lab 6.0.1 (Bio-Rad Laboratories, Inc., Hercules, CA, United States) and expressed in adjusted band volume. On the SDS gel from hGH samples several other bands than those of the recombinant protein at 22 kDa were visible. The concentration of the heterologous product was rather low in the supernatant of the cell culture. Thus, it had to be purified, also focusing native proteins from *P. pastoris*, which made them visible on the gel. All variants showed stronger hGH bands compared to the WT, some even exceeded the P_{DF} benchmark. Unfortunately, these findings were not reflected by the hGH-ELISA, which was performed three times with separate cultures (**Supplementary data S3**). The reliability of the applied screening procedure was very low. Two times our benchmark promoter was equal to the WT and one time only surpassing it by 30%, which was highly uncharacteristic for the strong P_{DF} promoter. The strongest variant of all three screenings was merely 10% improved and the assays did not show any correlation between each other. Due to this apparent unreliability of the assay, hGH was excluded as promoter strength indicator from further experiments. Nonetheless, hGH expression was successfully improved and conducted with the P_{CTA1} promoter variants, as confirmed with SDS-PAGE. CalB bands were indicated in the SDS gel at 38 kDa, slightly higher than the actual molecular weight of 33 kDa. This could be explained by the frequent hyper-glycosylation in yeast, where a various number of mannosyl molecules extend the glycoprotein (130). Bands for all variants were visible and more prominent than the WT, according to Image Lab software. For CalB samples purification was not necessary because of high homologous protein concentrations. Thus, unpurified supernatant was used.

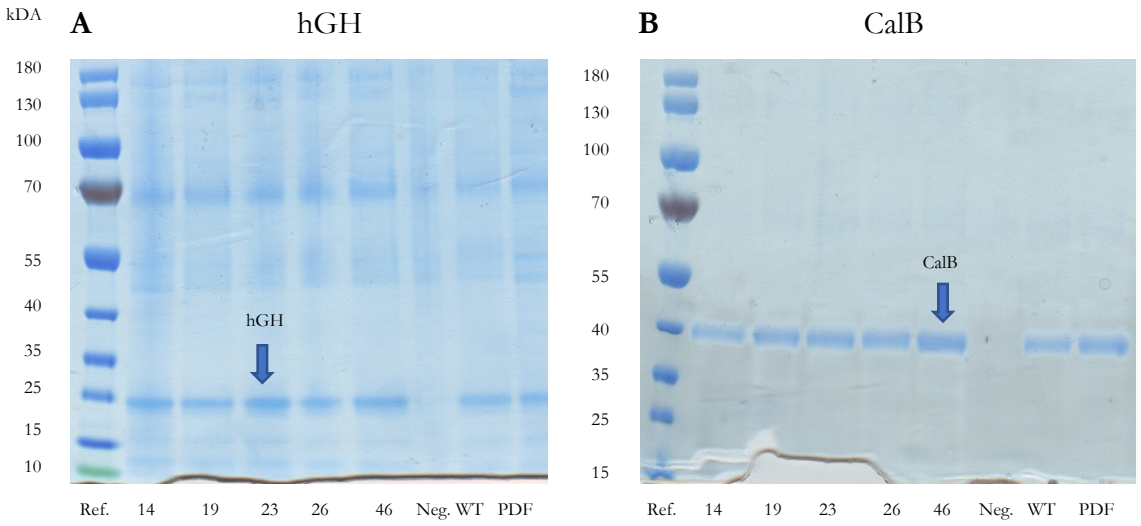


Figure 3.8 SDS-PAGE displaying recombinant protein bands expressed by P_{CTA1} promoter variants cultivated in DWPs for 98h on BMD0.5% with methanol induction. The columns from left to right show the reference for molecular weight, for which PageRuler™ Prestained Protein Ladder 10 to 180 kDa from Thermo Scientific, Inc. (Waltham, MA, USA) was used, five variants, negative control with no band at the appropriate recombinant protein size, positive control and benchmark promoter. hGH bands were situated at 22 kDa and other native *P. pastoris* bands were visible (**A**). This could be attributable to the need of purifying hGH samples due to its lower concentration in cell culture supernatant. Thus, other protein bands were also focused. CalB bands were shown at 38 kDa, surpassing the actual calculated theoretical weight (**B**). This might be explained by hyper glycosylation, often occurring in yeast. Untreated supernatant was used.

3.2.2 Deep-well-plate cultivations

In CalB cultivations only one measurement per DWP could be performed, since the supernatant containing the product had to be harvested via centrifugation. To overcome this problem and be able to determine enzyme activity at different time points four DWP cultivations per media were identically conducted. One set of four DWP with BMD0.5% and one with BMG0.5% were inoculated from two separate pre-cultures containing the respective media. The carbon source concentration was decrease from 1 to 0.5% in order to further minimize the oxygen limitation effect. The supernatants from the DWPs from each media were harvested after 48, 60 and 98 hours and the methanol induction and glycerol pulsing protocols were performed as previously discussed. The fourth DWP was used to observe the variant strains under carbon source starvation. Therefore, they were only supplied the initial amount of glucose or glycerol and harvested after 98 hours without any further media ad-

ministration. Determining enzyme activity at different time points was crucial to learn how much recombinant protein was produced by the promoter variant strains in the dependency to carbon source, carbon source concentration and time. The harvested supernatant of the different time points was added to the reaction solution containing p-nitrophenyl butyrate. The lipase in the sample catalyzed a hydrolyzation of the substrate into p-nitrophenolate and butyrate. The former created a yellow color change, which was proportional to the enzyme activity and measurable at an absorption maximum of 405 nm, if the reaction milieu was set to pH 7 to 7.5.

Equation 3.1 Beer-Lambert equation to determine enzyme activity with spectrophotometric measurement. ΔAbs=Absorption change, V=volume, df=dilution factor, ε=molar extinction coefficient, d=path of light

$$\frac{U[\mu mol/min]}{mL} = \frac{\Delta Abs}{\Delta t[\mathrm{min}]} * \frac{V_{total}[mL] * df}{V_{sample}[mL] * d[cm] * \varepsilon[mL * cm^{-1} * \mu mol^{-1}]}$$

The Beer-Lambert equation (**Equation 3.1**) was applied to calculate enzyme activity, using a light path *d* of 0.594 cm and a molar extinction coefficient *ε* for p-nitrophenyl of 18.5 $mM^{-1} \cdot cm^{-1}$.

Fourteen wells were inoculated per variant for an optimal statistical evaluation and the controls were carried along in triplets on each DWP. One clone of the WT positive control and one of variant 23 did not grow and was therefore excluded from the analysis. Other than that, the standard deviation between replicates on each DWP was constantly under 15%. This was attributable to the use of pre-cultures, which enabled identical initial biomass concentration in wells, making comparability between variants more feasible. In that manner, normalization to OD_{600} was not necessary. Furthermore, the curves of the measured absorption changes over time were inspected for linearity. In some cases, they leveled off towards the end indicating either product inhibition or rather too fast depletion of the substrate. For those events the samples were diluted with reaction buffer 1:2 to 1:4 to ensure a slower progression of the reaction. The dilution factor was taken into consideration in the Beer-Lambert equation.

At the first two time points (48- and 60-hour sample) of the BMD cultures, variant 19 was most prominent with a 2.1-fold enzyme activity compared to the WT, even exceeding the P_{DF} benchmark promoter at the 60-hour sample (**Figure 3.9a**). Variant 23 was the only one

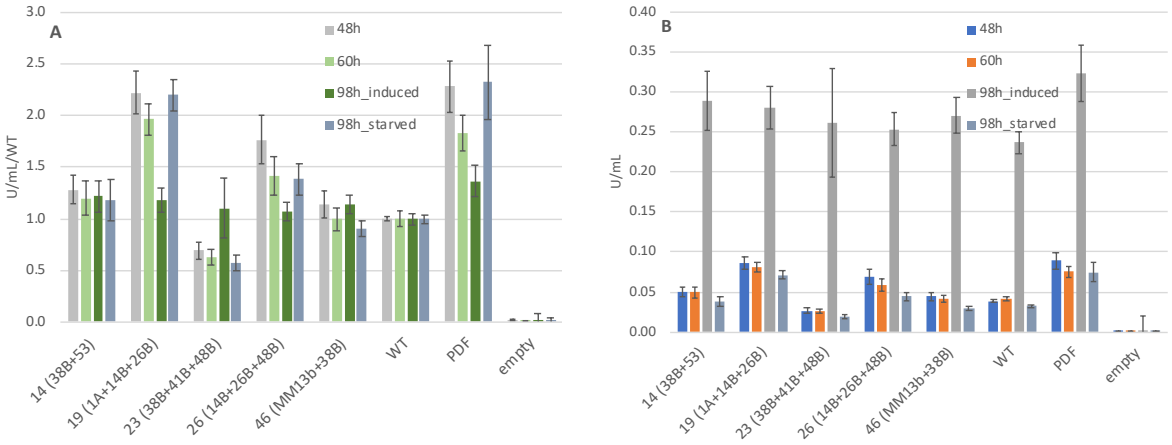
with a significant decrease (34%) and variant 46 was rather unchanged in activity. From timepoint one (48h) to two (60h) no increase was observed for any variant, in fact there was a slight drop of an average 6.7%. As expected, most enzyme was formed upon induction, visible with an average activity increase by approx. 500% after methanol addition. Variant 23 even exceeded 900%. The highest overall enzyme activity from the 98-hour cultivation with induction was exhibited by variant 14 with 0.288 U/mL (**Figure 3.9b**). While all variants were improved under induction and 98 hours of cultivation, the increased enzyme activity only ranged from 1.1 to 1.2-fold to the WT. The variants grown under starvation conditions displayed similar results as the 60-hour sample, so no expression boost was induced after carbon source depletion (**Figure 3.9b**).

The results obtained by cultivation of recombinant *P. pastoris* strains on glucose-based media and expression of CalB under the P_{CTA1} wildtype was comparable to previous findings by Vogl, *et al.* (11). Notable was that Vogl, *et al.* conducted the induction phase roughly twice as long (72 hours) reaching a total of 132 hours of cultivation and enzyme activity, which was three times higher (approx. 0.7 U/mL). So, for future experiments it would be considerable to prolong the induction phase and shorten the batch phase in order to highlight possible differences between the promoter variants in the later stage of recombinant protein production.

Variants cultivated on glycerol containing media behaved differently than glucose grown strains (**Figure 3.10**). They showed higher enzyme activity at earlier time points and under starvation, but inferior results after 98 hours and glycerol pulsing compare to glucose pre-grown and methanol induced strains. After 48 and 60 hours of cultivation the average enzyme activity was 42% and 80% higher, respectively. Also, from time point one to two CalB activity gained by 17% as opposed to BMD cultures, where some activity was lost. The 98-hour BMG samples in de-repressed state however displayed merely half of the average activity of methanol induced samples. This was attributable to the difference in promoter activation strategy. While induction with methanol lead to an enzyme activity boost of around 500%, de-repression with glycerol pulsing led only to 34%. The results under starvation were similar to the 60-hour sample with a range of change of under 4% for the variants, hence again no positive effect was observed. At every sampling time point only variant 19 was considerably increased (close to 2-fold), variant 14 and 26 were effectively unaltered and almost at all time points variant 23 and 26 were inferior to the WT (**Figure 3.9c**). In BMG cultivated variant strains none exceeded the P_{DF} benchmark, which reached 0.296 U/mL

after 98 hours of cultivation and glycerol pulsing as highest result (**Figure 3.9d**). Variant 19 was most comparable with 0.222 U/mL and thus, compatible to the P_{CTA1} wildtype on glucose-based media and under methanol induction. In that case the WT bearing *P. pastoris* strain reached 0.237 U/mL after 98 hours of cultivation, making variant 19 in combination with BMG and de-repression a methanol-free alternative with production levels only short of 6.7%. Furthermore, de-repressed variant 19 bearing strains on glycerol reached CalB activity levels that were even superior to P_{AOX1} or P_{GAP1} bearing strains on glucose-based cultivation (11).

Strains cultivated on BMD and induced with MeOH



Strains cultivated on BMG with glycerol pulsing inducing de-repression

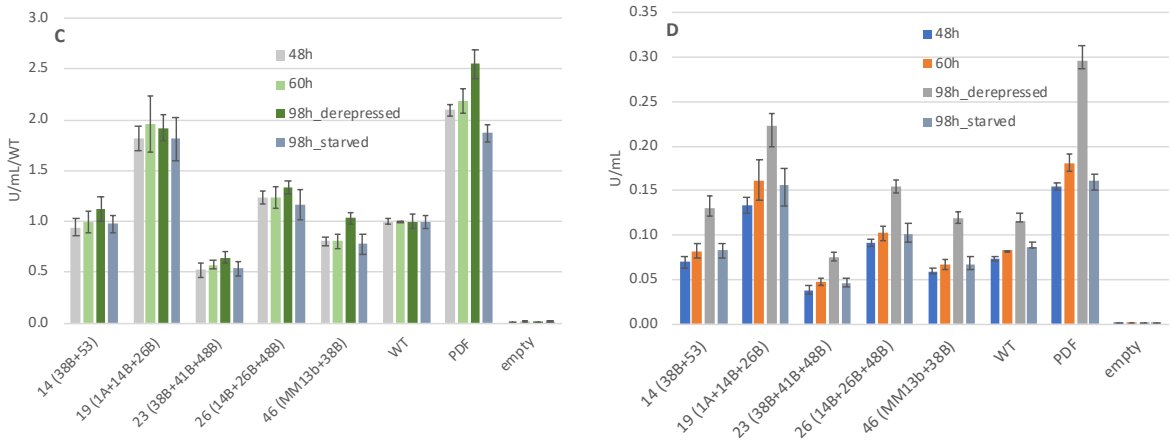


Figure 3.9 CalB produced by *P. pastoris* strains bearing expression cassettes with different P_{CTA1} promoter variants. They were cultivated in DWPs either on BMD0.5% with methanol induction or on BMG0.5% with glycerol pulsing. A group of columns resemble the same variant, positive control (P_{CTA1} wildtype), benchmark promoter (P_{DF}) or negative control (no promoter). The colors indicate the sampling time points 48, 60 and 98 hours or cultivation under starvation conditions. In **Figure A** and **C,** the volumetric units were normalized to the wildtype (y-axis) and blotted against the variants and controls (x-axis). In **Figure B** and **D,** the values on the y-axis show absolute results without normalization to the wildtype. **Figure A** and **B** depict CalB activity from promoter variant strains cultivated on BMD0.5% and induced with methanol, while **Figure C** and **D** are results from BMG0.5% based cultivations exposed to de-repressed conditions. Instead of labelling the variants with their internal numbering, the introduced mutations were used for better reference to the interpretation.

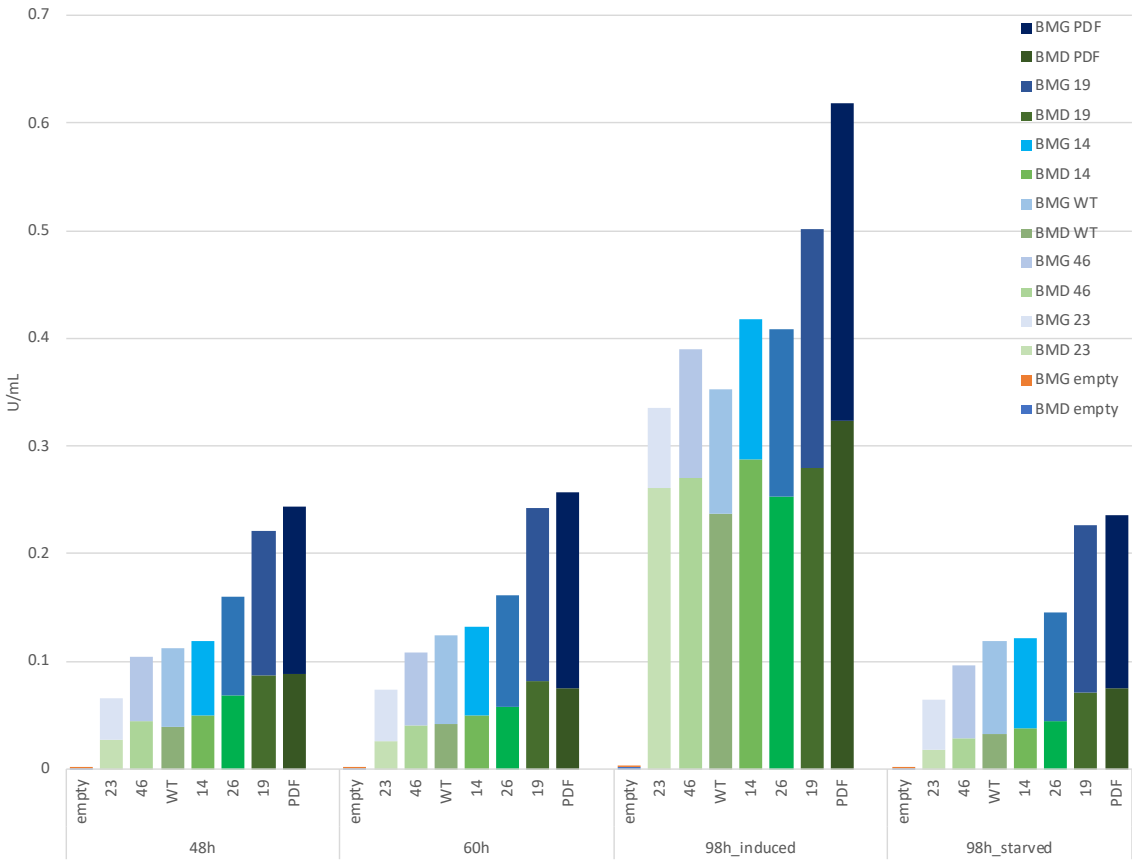


Figure 3.10 Comparison of CalB production by *P. pastoris* strains cultivated on different carbon sources. The upper blueish part of the bars sets CalB levels with glycerol as primary carbon source against glucose combined with methanol induction, indicated by the lower greenish part. The x-axis blots CalB enzyme activity per mL against the *P. pastoris* strains bearing P_{CTA1} promoter variants or controls on the y-axis. The bars are grouped by time points and ordered by increasing expression.

The findings in CalB expression showed some coherence with the eGFP screening results, but also deviations. Overall, cultivations on BMD induced with methanol resulted in higher recombinant protein production than when grown on BMG, which was determined with an enzyme activity assay. Furthermore, the characteristics at the different sampling time points stayed similar. BMD grown cultures received an enormous boost upon methanol induction and thus produced the greater fraction of the final amount of enzyme after 60 hours. Meanwhile BMG grown cultures did not rely on methanol induction and thus produced a great amount of the total recombinant protein in the earlier cultivation phases (48h and 60h) upon promoter de-repression. However, compared to eGFP, where barely any fluorescence increase could be determined after 60 hours under glycerol pulsing, CalB activity still rose by 34%. According to these results, it seemed as if CalB was more stable than eGFP, visible by

a smaller decrease in enzyme activity from one time point to another in de-repressed or starvation conditions. Also, lower initial carbon concentrations of 0.5% in CalB compared to 1% in eGFP expression produced less biomass after the batch phase, and thus a smaller metabolic load during the de-repression and induction phase (after 60h). This could have led to more nutrients being available for heterologous protein production. Noteworthy, the choice of carbon source did not affect the differences in CalB expression as significantly as with eGFP. While BMG was even superior to BMD as media of choice at the early sampling time points, it achieved 52% of the CalB activity compared to methanol induction strategies, while in terms of eGFP fluorescence it was only 42%. Additionally, CalB expression reacted more sensitive to methanol addition than eGFP with a boost of 506% in enzyme activity compared to 113% in fluorescence, respectively.

The mutations, which were associated with certain improved or impaired properties in the P_{CTA1} promoter variants in the eGFP screenings, were confirmed. The mutations 14B and 26B in the re-screening and 1A in the screening of Katharina were predominant for strong de-repressed promoters. This effect was resembled in CalB expression in both media especially by variant 19, which carried all three mutations and was continuously the best variant in samples after 48 and 60 hours of cultivation. Furthermore, if those mutations inclusively the deletion 53 were absent in a promoter variant, a tight regulation with low activity in the beginning was observed (e.g. variants 23 and 46), similar to the WT. The mutations 38B and 41B were affiliated with strong inducible promoters. Obviously, they led to negative effects with glycerol-based cultivations, where no induction phase took place. For instance, variant 26 containing both alterations scarcely reached half of the WT enzyme activity. On the other hand, upon methanol addition those mutations ensured strong effects. This was demonstrated by variant 14 and 23, the former reaching the highest overall enzyme activity among variants when glucose and methanol were used as carbon sources. Variant 23 was additionally lacking any de-repression associated mutations leading to a very tight and strongly inducible promoter variant – the enzyme activity was 9-fold elevated after methanol addition.

The most superior results were undoubtedly reached with induction strategies. Remarkable was that eGFP values at the 98-hour sample after methanol addition showed considerable differences while CalB activity seemed rather consistent. The percental standard deviation of eGFP fluorescence among variants, P_{DF} and WT was 22%, as opposed to only 10% for CalB activity. Further, the strong P_{DF} benchmark promoter was considered significantly superior to the unimproved WT positive control. This was reflected in the eGFP re-screening, where

P_{DF} ensured a 2.2-fold higher fluorescence than the WT when using glucose and methanol as carbon sources and 3.5-fold when using only glycerol. In the screening of Katharina there was also a 2.1-fold advantage. Therefore, it was rather uncharacteristic that the CalB activity of the P_{DF} benchmark control was merely 37% improved towards the WT positive control upon induction. Interestingly, the distinctive superiority of the P_{DF} promoter was observed at the earlier samples of CalB expression. The P_{DF} promoter delivered a considerable 2.3-fold and 1.8-fold CalB activity against the WT promoter at the 48- and 60-hour sample, respectively. Thus, the absent variety among promoter variants expressing CalB was specific for the induction phase. It appeared as if the promoters could not exploit their full potential, since a different bottleneck was the limiting factor and restraining the enzyme production rate to a collective pace, leading to roughly equalized enzyme activity. Since the recombinant protein was secreted extracellularly, it was reasonable to assume the bottleneck to be the secretion pathway and/or signal sequence processing. Among correct protein folding within the ER, these three are the most common restricting factors in recombinant protein production (47). That this limiting effect appeared only during the induction phase was made justifiable by the fact that this phase was accountable for 80% of the total CalB activity resulting in a tremendous metabolic burden and stress on the cell. When this stress exceeds the means of the cell in terms of folding, protein translocation and signal peptide cleavage, certain consequences are encountered. Noteworthy is that the poor performance of the P_{DF} promoter might also be linked to the growth rate. The P_{DF} promoter requires a certain growth rate to act optimally and in the later cultivation stages, cell growth was rather moderate. Generally, in DWPs growth rate is limited and hard to control, so to exploit the full potential of the P_{DF} promoter bioreactor experiments would be needed.

Stress regarding protein folding triggers the unfolded protein response (UPR). When too much protein is directed into the ER for folding, unfolded protein accumulates in the lumen and stress is signaled to the cell. General transcription levels are lowered and the expression of protein chaperons and foldases localized in the ER to assist with protein folding is enhanced. Additionally, partial breakdown of proteins in the lumen is induced by the ER-associated degradation (ERAD) mechanism (131). To increase the folding capacity of a cell a strain can be engineered to overexpress chaperons aiding in protein folding, such as BiP/Kar2p, DnaJ, PDI or PPIs (132).

Inaccurate cleavage of the signal peptide may cause that in subsequent steps the processed protein will not be recognizing and consequentially degraded (62). Our signal sequence of

choice was the *S. cerevisiae* α-mating factor (α-MF), because it has proven to be the most potent signal sequence applied with *P. pastoris*, even exceeding the native acid phosphatase (PHO 1) signal (1,53). The α-MF consists of two essential regions. First, the pre-region, which guides the nascent protein into the ER and is processed by a signal peptidase. Second, the larger pro-region that translocates the protein from the rough ER to the Golgi, where the last part of the signal sequence is cleaved off at the dibasic cleavage site KR (Arginine – Lysine) by the endo-protease Kex2p (62). As a final step the EA (Glutamate – Alanine) repeats as a remaining part of the linker between signal sequence and recombinant protein are trimmed by an aminopeptidase encoded by STE13 (63). If any of these steps are processed incompletely, e.g. due to excessive expression of the guided protein exceeding the means of the cell, the cleavage site may be altered and not recognized by the protease of the following step. This results in non-homogeneity of the N-terminus of the protein and often degradation (62). The signal peptidase processing can be optimized e.g. by enhancing the amino acid residues at the cleavage sites for a better binding of the peptidases. This was reported by Yang, *et al.* (133), who achieved increased secretory protein production by engineering the Kex2p site.

To conclude, it would be worth considering approaching strain engineering rather than promoter engineering, since it seemed as if the promoter optimization was saturated and limited by posttranslational protein modifications. Additionally, a longer cultivation and induction phase might reveal more substantial differences or even improvements between promoter variants at later time points. Furthermore, growth rate dependent analysis can show the full potential of the new promoters.

3.3 CalB expression in shake flask cultivations

In the industry the improved promoter variants will not be applied in DWP, but ideally used for large-scale recombinant protein production in bioreactors. For that, working volumes of up to hectoliters can be used. In this size range there are some factors, which have to be addressed special attendance, as opposed to in DWP, for an ideal execution of the process. Those factors are mostly of physical nature, including rheology, power input, geometry of the reactor tank and mixing, as well as heat-, mass- and oxygen transfer. Mixing, for example, is essential for an effective distribution of the the latter three and much more complex in larger volumes due to higher shear forces, a different behavior of fluidic currents and higher viscosity, because of a higher biomass concentration. In turn, the increased biomass

concentration demands more oxygen, which is supplied over a higher stirrer speed, producing additional mechanical heat. This heat then contributes to the already excessive heat load from the metabolic activity of the cells, which often limits reactor size (134). Just by naming a few of the challenges in scaling up, it is apparent that in real fermentation broths the optimum of each parameter is difficult to meet. Thus, it can occur that the conditions change from small- to large-scale cultivations, and the behavior of the engineered strain variants do so as well. Therefore, I conducted shake flask cultivations as an intermediate step before bioreactor fermentations, to obtain valuable knowledge of the P_{CTA1} promoter variant properties in larger volumes and partly controlled conditions. Additionally, samples obtained from the cultivations were processed by my colleague Antonia Volpini de Maestri for RNA isolation and subsequent real-time PCR. The transcriptome was used to determine the expression level of CalB on an RNA level. This should answer the question, whether a strong promoter variant achieved superior results because of upregulated transcription or by any other means.

The two *P. pastoris* strains carrying P_{CTA1} promoter variants 19 and 26 were chosen for shake flask cultivations. The WT positive control, the P_{DF} benchmark and the negative control bearing no promoter were carried along. The cultivations were conducted in two manners, one time with BMD0.5% and methanol induction and one time with BMG0.5% and glycerol feeding discs to establish a de-repressed state. For each strategy a sterile control was added. Due to room limitations in the shaker the cultivations were restricted to biological duplicates instead of triplicates. The duplicates were used to calculate the average of the OD_{600} values and CalB activity. Two culturing steps previous to the actual experiment were conducted, a pre-pre-culture and a pre-culture. The former was performed in YPD media to reach sufficient amount of biomass for further inoculations, which would not be possible on buffered minimal media. The latter was treated identically as the main culture to suffice as point zero for the real-time PCR. The cultivations were conducted as previously described and samples were drawn after 4, 17, 24, 42 and 65 hours.

Extensive information about the cultivation was drawn from the online monitoring device from PreSens Precision Sensing GmbH (Regensburg, Deutschland). A monitoring unit was comprised of a shake flask with an integrated sensor (SFS-PSt3) and a shake flask reader (SFR vario). The shake flask was equipped with an oxygen sensor at the bottom of the glass, which was mounted on the reader in a fashion so the sensor would be aligned with the reader optics. Additionally, the reader was able to determine biomass development via optical den-

sity (OD_{600}). Both signals were transmitted to a computer unit via Bluetooth, where the compatible PreSens software processed the information and used the slope of the online oxygen measurements to calculate the oxygen uptake rate (OUR). The retrieved information was used to adjust sampling as well as induction time points, i.e. when the OUR slope went down and the O_2 slope up, carbon source depletion was indicated and methanol or glycerol feeding discs added. To calibrate the oxygen sensor, the signal at the start of the cultivation was set to 100%, because the biomass and therefore the oxygen demand were the lowest. Consequentially, the initial OUR was zero. There were only two online monitoring devices available, so they were attached to one WT control per media as reference for all other variants grown on the same carbon source.

3.3.1 Cultivation Process

Unfortunately, the PreSens online monitoring system was not as precise as hoped for. To begin with, the sensors were overdue for calibration, which was only found out during usage. Additionally, problems were encountered, where no sensor signal was detected. This was affecting the biomass sensor in the early stages and the oxygen sensor in the last 20 hours of cultivation. The former could be explained by a very small initial biomass concentration producing a signal too small for the sensitivity of the probe. This was resembled in a stagnant (BMD) and declining (BMG) graph, which is displayed in **Figure 3.11A** and **3.11B**, respectively. Fortunately, the essential corner stones, such as the log-phase, cell reaction to induction, etc., were visible and could be used for analysis, interpretation and as indicator for a semi-controlled cultivation process. In the samples drawn after 4, 17, 24, 42 and 65 hours OD_{600} was determined to confirm correlation with the online monitoring system, which was found consistent.

In the BMD-based shake flask cultivation it was surprising that the initially provided carbon seemed to have been already depleted after 13 hours, whereas we expected it to last 20 hours, resulting merely in OD_{600} of approx. 8. This growth stop was indicated by a drop in OUR, hence an increase in O_2 (%) and the OD_{600} graph leveled off. For induction initiation we wanted to have a higher biomass concentration, so a 50% glucose pulse was added to which the cells responded immediately, with an increase in OUR. After a lag-phase of two hours they returned to the log-phase. The additional glucose was then depleted after 20 hours at an OD_{600} of approx. 12. Subsequently, the first methanol pulse was added to which the cells reacted rather slow. Metabolic activity indicated by an increase in OUR was only visible after three hours upon induction and biomass growth after five hours. This suggested that the

cells required a substantial adjustment period to the new carbon source, comparable to a prolonged initial lag-phase. In this phase the cells needed to adapt to their new surroundings and started sequestering enzymes to catabolize methanol (135), starting with AOX. The applied methanol provided resources for 21 hours and the biomass reached an OD_{600} of 16 by that time. Between methanol depletion and administration of the second pulse three hours passed in which the cells started degradation, decreasing biomass concentration. After 44 hours of cultivation a second methanol pulse was added to which the cells reacted instantaneously, since the cell metabolism was already set to methanol usage. The OUR increased immediately and the lag-phase lasted shorter than after the first induction. This supported the hypothesis that cells required an adjustment period to a carbon source exchange. After another 21 hours methanol seemed to have been depleted and the OD_{600} graph leveled off at 17.8. Unfortunately, the oxygen parameters were not reliable after 45 hours, because of a missing sensor signal. There was a significant OUR drop at 17 hours, which was probably created by sampling the culture broth. For that process the shake flasks had to be removed from the shaker and transferred to a laminar flow to provide a sterile environment for the sampling procedure. The online monitoring system was paused for that period of time. When the cultures were put back into the shaker and monitoring was continued, it required some minutes for the cells to resume growth and even longer for the oxygen parameters to re-stabilize, hence the disruption in the curve. Similar less severe interruptions in the progression of the graph were visible at the other sampling time points. As already discussed, initial problems with the optical biomass sensor occurred so the transition from lag- to log-phase was hard to detect. However, with a manual OD_{600} sample after four hours and a reliable monitoring value at the end of the log-phase, (after 13 hours) a growth rate of $\mu = 0.279$ during that period of time was determined using a rearranged **Equation 3.2**.

Equation 3.2. Exponential growth rate is a first order equation defined by the initial biomass concentration $\mathbf{X_0}$ (cell number, mass or OD_{600}), the specific growth rate $\boldsymbol{\mu}$ (t^{-1}) and the time of growth **t** (hours).

$$X = X_0 * e^{\mu * t}$$

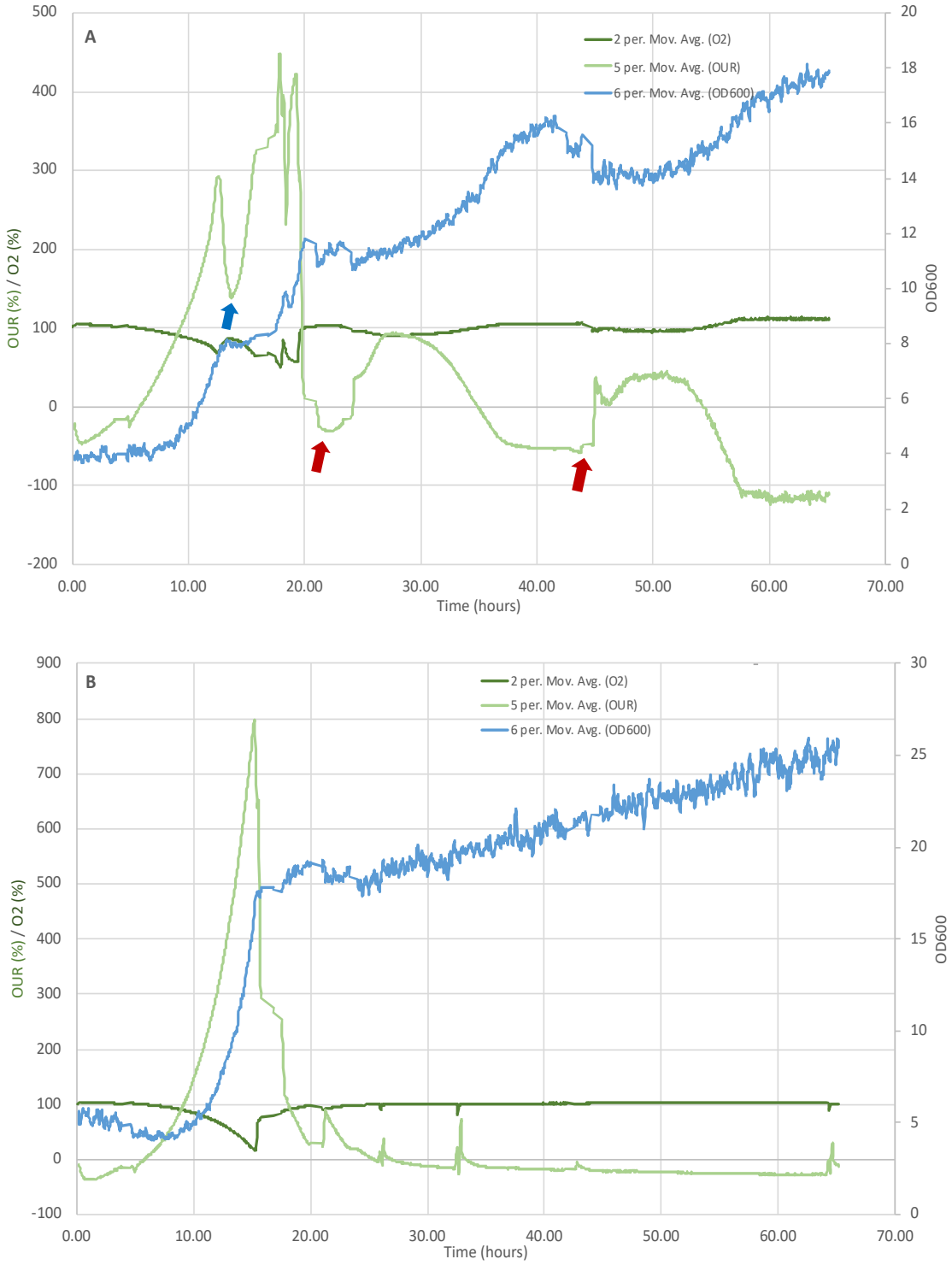


Figure 3.11 Shake flask cultivation of *P. pastoris* expressing CalB under P_{CTA1} WT control monitored via the PreSens system. For a clearer illustration moving average trend-lines were used, since the individual timepoints showed considerable deviations. The blue graph resembles OD_{600} of the cultivation broth, annotated on the secondary y-axis. The light green graph shows oxygen uptake rate (OUR) and the dark green O_2 saturation, both annotated at the primary y-axis in % and blotted against the time of the cultivation in hours on the x-axis. Results were obtained from cultivation on BMD0.5% and inoculation with methanol (**A**), as well as cultivation on BMG0.5% and glycerol pulsing to retain a de-repressed state (**B**). The blue arrow indicates a glucose pulse, while the red arrows indicate methanol pulses.

In the BMG-based cultivation it was striking that an amount of glycerol proportional to glucose in the BMD media lasted for 20 compared to only 13 hours and resulted in roughly twice as much biomass, i.e. OD_{600} of approx. 17. Although it was previously reported that glycerol would lead to a higher biomass yield than glucose, the vast difference could not be explained (136). It was hard to determine when the lag- to log-phase transition took place, since the optical biomass sensor of the PreSens system appeared to be sensitive to a minimum OD_{600} of four to five and starting concentration upon inoculation was 0.2. Nonetheless, the growth rate could be established utilizing a manual OD_{600} sample from after four hours and the monitoring value at the end of the log-phase. By using a rearranged **Equation 3.2**, $\mu = 0.279$ could be determined, which was identical to glucose-based cultivations during the same growth phase. Further on, the exponential curve of the OD_{600} graph with the transition to the stationary phase after around 16 hours was nicely displayed. In the stationary phase the cells continued with a decreased growth slope until glycerol depletion after roughly 20 hours. Consequently, glycerol feeding discs were added, which released nutrients based on a concentration gradient. The cells reacted rather fast to the administration, since no change in carbon source was applied. The OUR increased after roughly one hour and the biomass after a slightly longer delay. The feeding discs lasted throughout the experiment and caused a linear biomass increase with a slope of 0.172, reaching an OD_{600} of 26 after 65 hours of cultivation. Thus, approx. 71% of the total biomass were formed in the "batch-phase", while it was 67% for glucose-based cultivations. Disruptions in the OUR slope indicating sampling or administration of feeding discs were not as severe as in glucose-based cultivations, but still detectable. There were no problems with the oxygen sensor in the later stage of cultivation, but it seems as if it should have been set to higher sensitivity, since it reported almost constantly 100% saturation.

3.3.2 SDS-PAGE

SDS-PAGE of the shake flask cultivations was performed by my colleague Antonia Volpini de Maestri, since she carried on with the experiments after my finish. Only samples t4 (42 hours) and t5 (65 hours) produced a sufficient amount of protein to perform an SDS-PAGE analysis with. Due to the low amount of native proteins compared to recombinant ones, unpurified supernatant of the culture broth was used. The protein bands were analyzed using the software Image Lab 6.0.1 (Bio-Rad Laboratories, Inc., Hercules, CA, United States) to compare adjusted band volume relative to each other. CalB bands were visible slightly above expected values (38 instead of 33 kDa), probably due to hyper-glycosylation (**Figure 3.12**).

No other bands at all were visible. As expected, the negative control lacking a promoter showed no bands and the positive control bearing an unimproved P_{CTA1} WT promoter delivered the weakest signal, barely visible at time point t5. There was a consistent correlation spanning across time points, as well as carbon sources. Variant 19 displayed the strongest bands at least as prominent as the P_{DF} promoter and variant 26 was somewhere between the positive control and the benchmark. The SDS-PAGE results showed consistency towards the results from the DWP cultivations in terms of position and relative strength of the bands.

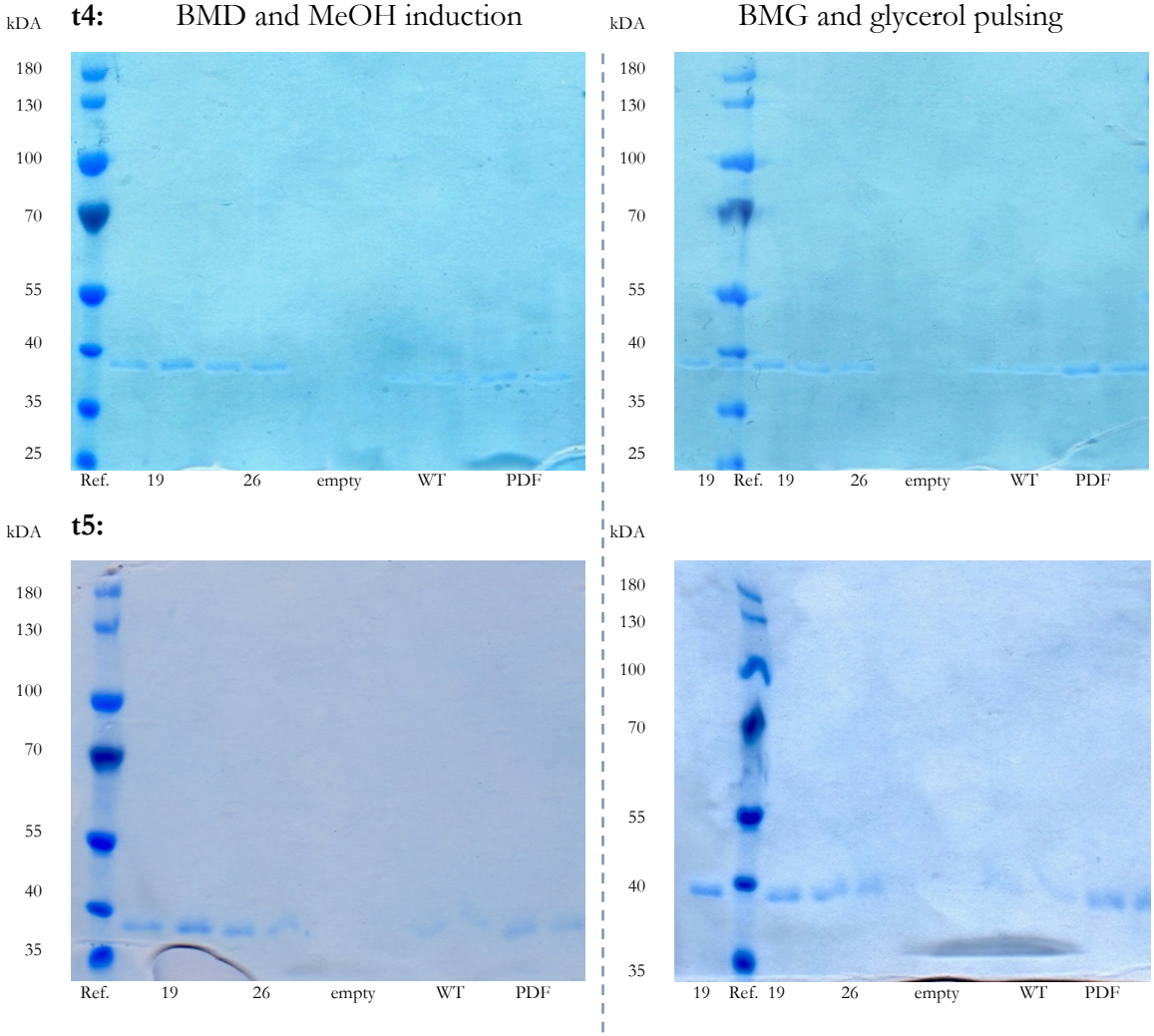


Figure 3.12 SDS-PAGE displaying CalB bands expressed under P_{CTA1} promoter variants by *P. pastoris* strains, which were cultivated in shake flasks. The columns display variants, controls and reference (PageRulerTM Prestained Protein Ladder 10 to 180 kDa from Thermo Scientific, Inc., Waltham, MA, USA). The latter indicates the respective protein sizes in kDa. The left gel pictures show bands from BMD cultivations plus methanol induction, whereas the right pictures show results from de-repressed BMG cultivations. T4 and t5 were samples taken after 42 and 65 hours, respectively. On each gel CalB bands were found at 38 kDa, slightly higher than expected due to N-glycosylation.

3.3.3 CalB activity

CalB activity assays of the shake flask cultivations were performed by my colleague Antonia Volpini de Maestri, since she carried on with the experiments after my leave. She kindly provided me with the results, so I was able to interpret and discuss them. In the assay technical quadruplicates per biological duplicate of the variants and controls were performed. The mean value and standard deviation were calculated and used for graphical display, as well as the mean value for enzyme activity calculation. In both media and time points t3 (24-hour sample), t4 (42-hour sample) and t5 (65-hours samples) the standard deviation never exceeded 17%. Henceforth in this section, only these time points were discussed, because t1 and t2, which were taken prior to induction/de-repression, did not show any CalB activity at all. This was a striking difference to DWPs, where a significant part of the CalB activity was already formed in the time period before methanol induction or further glycerol pulsing. However, one has to consider that this period was significantly longer in DWP (60h) than in shake flask cultivations (20h). Another notable difference was that at least regarding glycerol-based cultivations, the initially supplied carbon source lasted substantially longer in DWP (approx. 48h) than in shake flask cultivations (20h). That was mostly due to better mixing and a larger boundary layer between liquid and gaseous phase for oxygen exchange in the flasks. Thus, induction/de-repression was initiated already after 20 as opposed to 60 hours of growth and the "batch" phase turned out significantly shorter than in DWP cultivations.

Results from shake flask cultivations grown on glucose and induced with methanol are depicted in **Figure 3.13a** in absolute values (U/mL) and in **3.13b** normalized to the P_{CTA1} wildtype (U/mL/WT). While both mutated promoter variants were superior to the WT, variant 19 was more prominent, even exceeding the P_{DF} benchmark promoter. These findings resembled the SDS-PAGE and the DWP results. Variant 19 was especially strong at the 24-hour sample with a 4.5-fold activity compared to the WT, which was coherent with the mutations it carried (1A, 14B and 26B). In previous sections those mutations were already associated with promoters responding fast to de-repressed conditions. Also, at the 42- and 65-hour sample, as well as the percental increase between those, variant 19 was superior with a peak enzyme activity of 0.208 U/mL at the end of the cultivation.

In results from de-repressed BMG-based shake flask cultivations both mutated variants were improved towards the WT (**Figure 3.13c**). While there were only minor differences at the 24-hour sample, they extended especially with the more prominent variant 19 up to nearly three-fold at the end of the cultivation. These findings resemble the SDS-PAGE analysis and

previous association of mutations 1A, 14B and 26B with strong de-repressed promoters. In terms of absolute values, depicted in **Figure 3.13d**, variant 19 reached 0.205 U/mL slightly outperforming the P_{DF} promoter for the first time in de-repressed conditions. When the average enzyme activity of all five *P. pastoris* strains were compared between de-repressed or induced expression, the former was slightly superior with 0.148 to 0.145 U/mL. Of course, the higher average biomass concentration of 74 % at the end of the cultivation has to be considered.

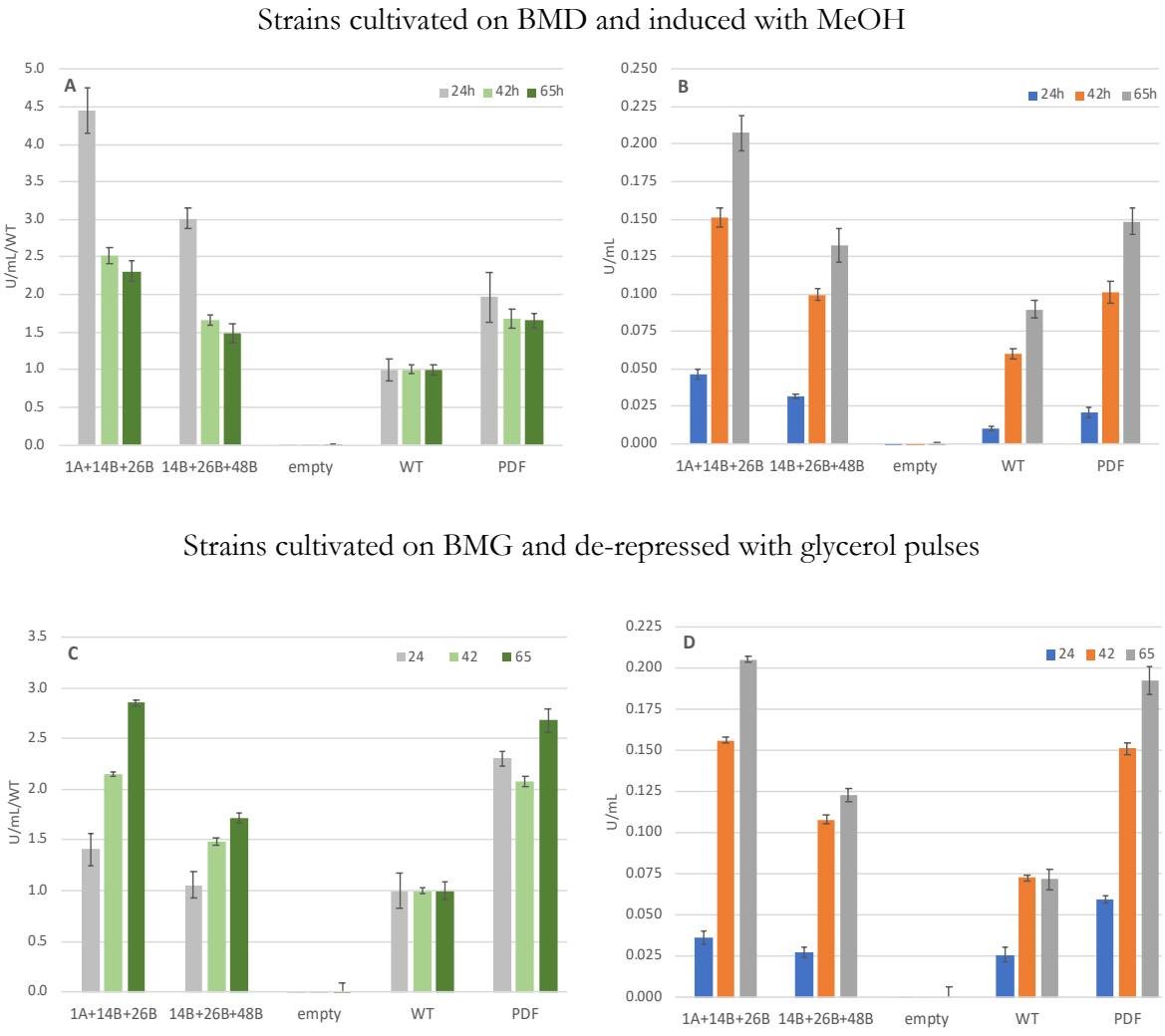


Figure 3.13 CalB expressed in shake flasks by *P. pastoris* strains under the control of P_{CTA1} promoter variants. A group of columns resemble the same variant, positive control (P_{CTA1} wildtype), benchmark promoter (P_{DF}) or negative control (no promoter). The colors indicate the sampling time points of 24, 42 and 65 hours. In **Figure A** and **C**, the volumetric units were normalized to the wildtype (x-axis) and blotted against the variants and controls (y-axis): In **Figure B** and **D**, the values on the y-axis show absolute results without normalization to the wildtype. **Figure A** and **B** depict CalB activity from promoter variant strains cultivated on BMD0.5% and induced with methanol, while **Figure C** and **D** are results from BMG0.5% based cultivations exposed to de-repressed conditions.

Regardless of the media no biomass and enzyme activity correlation within a variant was observed. Upon induction or de-repression already 68-71 % of the biomass were formed, so afterwards the slope of biomass increase continued rather flat and linear in overall average. Meanwhile, effectively no enzyme activity was detected before methanol addition or glycerol depletion, most was formed from 24 to 42 hours and the enzyme production seemed to decline towards the end. This resulted in a curve of slightly logarithmic shape not in coherence with the linear biomass progression. Also, when variants were compared to each other, the ones with higher enzyme activity did not necessarily show higher OD_{600} values. In fact, the OD_{600} values between variants at the respective time points were very consistent with a standard deviation of under 7 % (**Supplementary data S4**). Often recombinant protein production poses a significant metabolic burden on the cell, reducing biomass formation (137). However, the negative control strain, which expressed no heterologous gene, did not surpass the other strains in cell growth, showing that the production of CalB did not weaken proliferation.

Noteworthy, from the 24- to the 42-hour sample an 81 % higher enzyme activity increase was observed than from the 42- to the 65-hour sample in BMD-based cultivations. This was against the assumption that higher cell concentration in the end and an adjustment period to methanol addition in the beginning would lead to better results in the later cultivation phase. It also did not cohere with DWP results or results published by Vogl*, et al.* (11), where CalB production seemed to proliferate especially towards the end of the cultivation. The phenomenon was even more severe in BMG-based cultivations with an average 224 % higher enzyme increase in the first half of the de-repressed phase compared to the second. A possible explanation could be that the nutrients supplied by the amount of buffered minimal media did not suffice for the whole course of cultivation. To the starting media only a relatively small amount of additional BMM1.25% was added to the glucose-based shake flask cultivations after 20 hours of growth. Other than that, only pure methanol or glycerol for glycerol-based cultivations was used for the pulses. While with these pulses the carbon source was replenished, other essential nutrients contained in the media, such as biotin, nitrogen or ammonium sulfate were neglected but expected to be non-limiting. Same applied to glycerol pulsing, where feeding discs only releasing the carbon source were used. This problem was not encountered in the DWP cultivations, due to less biomass and the fact that all methanol/glycerol pulses were performed with BMM/BMG containing stated nutrients. As a further example, also in fed-batch fermentations the carbon source feed is usually mixed with biotin and trace salts to compensate for the depletion of initial amounts. The greater

difference in enzyme activity increase between the two time periods (24 to 42h and 42h to 65h) in BMG-based shake flask cultivations than glucose-based might be affiliated with a higher biomass concentration and thus a faster depletion of the nutrients contained in the broth.

The volumetric enzyme activity obtained by shake flask cultivations was inferior to DWPs under induction with an average of 0.145 versus 0.271 U/mL and only slightly superior in de-repressed conditions with 0.148 against 0.140 U/mL. This was surprising, since shake flask cultivations produced a higher biomass concentration, usually leading to more recombinant product (9). However, limitations and problems in the upscaling process were discussed earlier and possibly did already apply to the step up from DWP to shake flask cultivations, especially in the later phase, where the biomass had already increased. On the upside, the ratio of expressed enzyme per cell was so low that the metabolic burden did not impose the putative secretory bottleneck on recombinant protein production, as observed in DWP cultivations. Thus, considerable differences and improvements of the P_{CTA1} promoter variants towards the WT and even P_{DF} benchmark promoter could be observed.

3.4 P_{CTA1} promoter variants containing at least four mutations

Results from the first round of mutation by Mohammed Hussein (102), where only one mutation per promoter variant had been introduced, led to some slightly improved variants. In induced conditions those reached a maximum of 1.2-fold, except for one amounting to a 1.6-fold signal as compared to the wildtype P_{CTA1} promoter. Under de-repression the strongest ones ranged from 1.5- to 1.8-fold. When the mutations from the prominent promoter variants were combined to two or three mutations per variant, further notable enhancements were observed. Almost all variants showed consistently better results towards the wildtype and the performance of the strongest ones was 1.7- to 1.9-fold increased after 98 hours of cultivation and induction. At sampling time points 48 and 60 hours under de-repression the most outstanding promoter variants ranged from a 2.5 to 3-fold enhancement towards the WT. It was striking that the combination of three mutations associated with either good de-repression or induction properties led to the best variants under the respective circumstances. This raised the question, whether this was an accumulative effect and if so, how far could it be exhausted until the integrity of the promoter would be impaired.

In order to answer this question five new P_{CTA1} promoter variants combining four to six mutations per variant were established. PCR-based techniques were used to amplify the

DNA-sections containing best mutations and the produced fragments were assembled. The variants and the mutations they carried are listed in **Table 2.12** in the Material and Methods section. The new promoter variants were cloned into a pKEB#1 based vector expressing eGFP, which was introduced in *P. pastoris* by means of homologous recombination. Two cultivation protocols were performed in 96-DWP for the transformed cells. On one plate the cells were grown on BMD0.5% and induced with methanol after 60 hours of cultivation. On the other plate BMG0.5% was used and de-repressed conditions were kept. To avoid unequal growth between the wells and thus provide comparability a pre-culture for each protocol with the respective identical medium was used for inoculation of the main culture. Samples were drawn after 18, 24, 48, 60 and 98 hours. The levels of eGFP were measured in a microtiter plate with a plate reader at 488 nm excitation and 507 nm emission in relative fluorescence units (RFU).

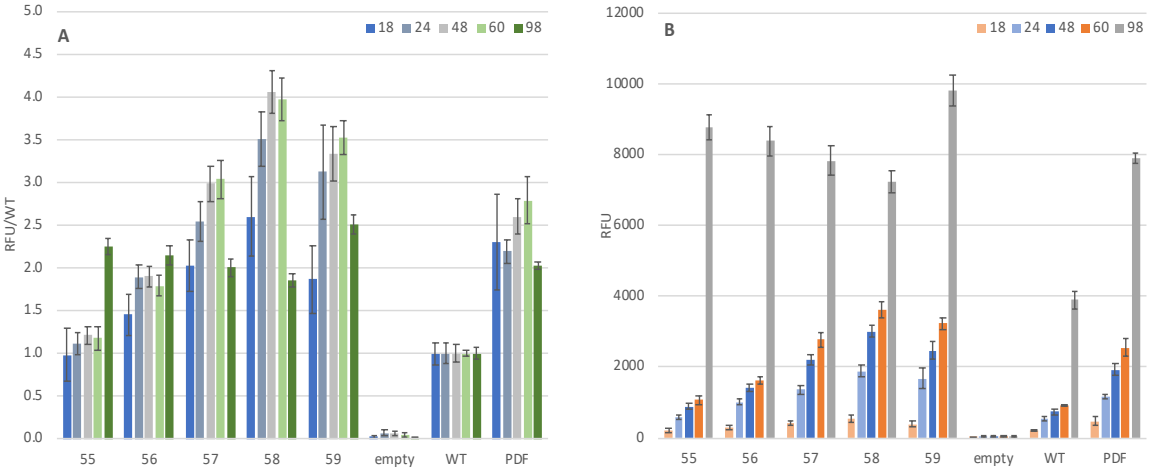
3.4.1 Screening results of P_{CTA1} promoter variants with at least four mutations

The screening results from BMD-based cultivations are depicted in **Figure 3.14a** in RFU normalized to the WT and in **Figure 3.14b** only in RFU without normalization. Overall, the addition of mutations led to further improvements of the promoter variants and not to deterioration. After 98 hours of cultivation and induction all variants were significantly improved towards the WT and three even exceeded the P_{DF} benchmark promoter (variant 55, 56, 59) with more than double of WT activity. The most prominent variant 59 reached a 2.5-fold fluorescence signal, outperforming the best variant from previous screenings with a 1.9-fold signal. As earlier observed, most recombinant protein was produced after 60 hours of cultivation with an average 308% increase upon induction. The most outstanding promoter variant in this phase was variant 55, which behaved nearly identical to the WT in the early stages of cultivation, while producing 2.6-fold more eGFP upon methanol addition. Under de-repressed conditions and before induction, no further beneficial accumulative effects were experienced, but also no impairments. The three promoter variants 55, 56 and 59 with highest production boost upon induction carried mutations 38B, 41B and 48B, which were previously associated with strong inducible promoters. The fourth mutation (MM13b) in variant 55 was a rather neutral mutation, so there was effectively no change in behavior towards the WT prior to induction. In variant 56 the fourth mutation (26B) was found in strong de-repressed promoters and led exactly to that effect – a slight increase upon carbon source depletion. The promoter variants 57 and 58 were equipped with de-repression affiliated mutations 1A, 14B and 26B. Thus, prior to methanol addition the variants were improved and

at least as productive as the P_{DF} benchmark promoter. But they provided overall inferior results, because in an induction-based expression protocol, de-repression was not as efficient. The fourth mutation in variant 57 was 38B, which was related to very strong induction properties. In variant 58 it was mutation 48B, only linked to mediocre improvement. This difference was resembled in the results. Both variants delivered elevated fluorescence levels towards the WT, but in variant 57 this difference was more significant. Variant 59 was the only one carrying six mutations. It combined 1A+14B+26B and 38B+41B+48B, the mutation triplets connected with strongest de-repressed or induced promoters, respectively. This resulted in the best promoter variant with outstanding production levels in the de-repression as well as the induction phase. The obtained data showed that at least up to six mutations per promoter variant would boost the recombinant protein production and not interfere with the integrity of the promoter structure. Further, mutations with either de-repression or induction properties, did not deteriorate each other, when being combined in one promoter variant. In fact, their beneficial effects were stacked.

The eGFP activity from new P_{CTA1} promoter variant strains expressed on glycerol under de-repression is displayed in **Figure 3.14c** in RFU normalized to the P_{CTA1} wildtype promoter and in **Figure 3.14d** in RFU without normalization. Overall, the stacking of mutations led to superior variants with merely one showing no notable improvement towards the WT (variant 55). The most prominent variant (58) came even close to the P_{DF} benchmark promoter after 98 hours of cultivation and de-repression with a difference of only 4% and a 3.3-fold enhancement. The eGFP fluorescence was measured at 5251 RFU, 34% more than from the WT, when cultured on BMD and induced. This suggested a potential for methanol-free cultivation procedures resulting in high production levels without using constitutive promoters, such as the *GAP* promoter. Variant 59 came close to 3-fold and variant 57 exceeded a 2.5-fold increase. As opposed to previous screenings, where the best promoter variants did not exceed the 3-fold threshold, the addition of mutations meant an improvement. Compared to BMD-based cultivations the measured values were practically identical prior to induction with an average 2512 RFU for strains grown on glycerol and 2256 RFU when grown on glucose. It was only until methanol administration that the induction strategy prevailed significantly with 7687 to 3639 RFU.

Strains cultivated on BMD and induced with Meoh



Strains cultivated on BMG and de-repressed by glycerol pulsing

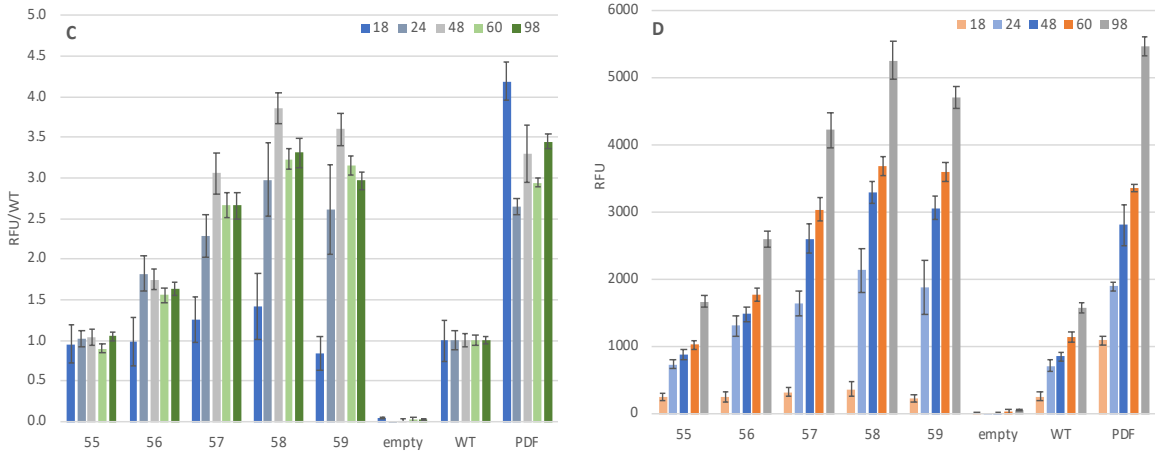


Figure 3.14 eGFP expressed under the control of P_{CTA1} promoter variants by transformed *P. pastoris* strains cultivated in 96-DWP. A group of columns resemble the same variant, positive control (P_{CTA1} wildtype), benchmark promoter (P_{DF}) or negative control (no promoter). The colors indicate the sampling time points 18, 24, 48, 60 and 98 hours. In **Figure A** and **C**, the relative fluorescence units (RFU) were normalized to the wildtype (y-axis) and blotted against variants and controls (x-axis). **Figure B** and **D** show values without normalization. **Figure A** and **B** depict results from promoter variant strains cultivated on BMD0.5% and induced with methanol, while **Figure C** and **D** are results from BMG0.5% based cultivations exposed to de-repressed conditions

The previous observations in terms of mutations and their effect were found consistent in BMG-based cultivations of *P. pastoris* strains. Mutations 38B, 41B and 48B showed little effect, since they were associated with strong inducible promoters and induction was not performed. On the other hand, 1A, 14B and 26B led to the best promoter variants due to de-

repressed conditions. The more prominent mutations were carried by a variant the greater the positive effect. For example, variant 56 was equipped with only one de-repressive mutation (26B), while variant 58 carried all three, making the latter a considerably more potent variant.

In **Figure 3.15** the increase of average eGFP fluorescence from the variants between sampling time points is depicted. The least productive phase was the first 18 hours of cultivation, showing that the mutations did not lead to leaky promoter variants when grown on glycerol. Also, the measured RFU of the variants were comparable to the WT at that time point. In the very short period from the 18- to the 24-hour sample a majority of eGFP was formed in only six hours (34%). Apparently enough carbon source had been consumed to de-repress the promoters and initiate rapid heterologous gene expression. The assumption was supported, when the OD_{600} values indicating biomass content were looked at and no increase was observed after 18 hours of cultivation (**Supplementary data S5**). It was only until the 98-hour sample when glycerol pulses were previously supplied that biomass growth could be measured. After the prompt eGFP increase at the 24-hour sample a continuous and linear production was measured with no decline until the end of cultivation. Thus, it would be advising to prolong the cultivation in future experiments to determine how long recombinant protein production in the established expression system was viable. Especially production at the late growth phase is commercially interesting to improve final product titers.

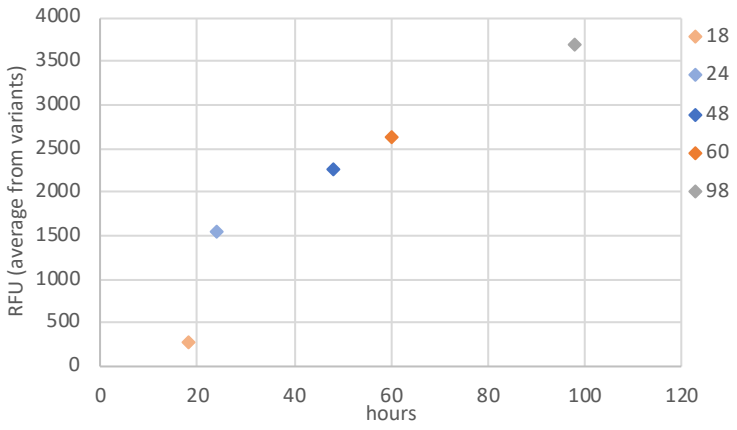


Figure 3.15. Average eGFP RFU increase between sampling time points 18, 24, 48, 60 and 98 hours. The recombinant protein was expressed under the control of P_{CTA1} promoter variants by *P. pastoris* cultivated in BMG0.5% under de-repressed conditions. The average RFU from the variants (y-axis) was blotted against cultivation time in hours (x-axis). The squares indicate sampling time and fluorescence signal.

4. CONCLUSION AND FUTURE OUTLOOK

Within this thesis the reproducibility of recombinant proteins expressed by already established P_{CTA1} promoter variants were reevaluated and confirmed. A special focus was put on evaluation of possible advantages for the expression of secreted proteins, where high transcript levels are not always providing advantages. Glycerol was tested as an alternative carbon source to avoid the necessity to use hazardous ethanol for induction. New and essential aspects for up-scaling and further improved promoter variants were found and described. The most essential findings were:

i.) P_{CTA1} promoter variants showing strong intracellular eGFP expression were found to provide also advantages for extracellular secretion.

ii.) Using an inventive methanol-free expression strategy, P_{CTA1} promoter variants exceeded the methanol-induced parental P_{CTA1} strain in terms of extra- and intracellular recombinant protein production.

iii.) The accumulation of beneficial mutations within one P_{CTA1} promoter variant showed significant additive and enhancing effects without interfering with the promoter integrity or tightness.

Although small discrepancies to initial screenings (observed in previous unpublished work in our group) occurred, the key findings about positive and regulatory regions of the P_{CTA1} promoter were highly reproducible and novel combinatorial mutations with beneficial properties were developed. Mutations affiliated with improved promoter properties either upon induction with methanol or sole de-repression were identified and consistent throughout the experiments. Most of those mutations contained the core sequence of the putative GATA transcription factor binding site. This motif was accidently introduced into the affected promoter variants by substituting every 10 bp in the parental P_{CTA1} WT promoter with a random neutral 10 bp DNA sequence carrying the GATA core sequence. According to publications about other hosts (118,119), these binding motifs were assigned to delegate transcription factors based on nitrogen availability, which aid in gene expression, thus amplifying protein production. However, previous studies in our lab did not reveal any significant differences in expression when varying concentrations of ammonia sulfate were applied (11). The positive effects of the GATA core sequence were already observed by Hussein (102) and Ebner (unpublished).

Surprisingly, the P_{CTA1} promoter variants showing strong intracellular eGFP expression were found to provide also advantages for extracellular secretion. Both test enzymes hGH and CalB were successfully produced and translocated to the culture supernatant. Limitations were met at high production titers. Upon methanol induction, a huge increase in recombinant protein secretion might have imposed a limiting metabolic burden on the cells. This caused that all variants were restricted at the same product maximum. The limitation was most likely affiliated with bottlenecks in the secretory pathway or folding. Limitations for CalB secretion at high transcript levels were observed previously. However, the stepwise induction by de-repression, followed by induction with methanol seemed to shift the upper threshold for secretion to higher titers compared to constitutive or *AOX1* promoter-based expression. High metabolic stress can exceed the means of the cell in terms of protein folding, membrane translocation and signal peptide cleavage (47). Thus, it might be advised to consider strain engineering as additional optimization resource. Possible approaches are co-/overexpression of protein folding chaperons as described for example by Guan *et al.* (138) or optimization of signal peptide cleavage sites for a better binding of the peptidases (133,139)

In this study recombinant protein production on glucose-based cultivations induced with methanol addition was compared to glycerol-based cultivations in de-repressed conditions. Further promising expression strategies to enhance observed effects and high productivity, but unfortunately could not be tested within this thesis, are simple pre-cultures with glucose followed by a limited glucose feed, simple starvation, growth on no repressing carbon sources such as sorbitol or induction by fatty acids including oleic acid. Yet under the tested conditions, the best P_{CTA1} promoter variant under de-repression was superior to the wildtype under induction in terms of CalB and eGFP expression. This confirmed the efficiency of this inventive methanol-free expression strategy with high product titers.

First up-scaling in shake flasks had proven to be challenging. However, valuable information could be obtained from the cultivation process. The carbon source shift form glucose to methanol put the cells in a prolonged lag-phase, in order to adapt to the new surrounding and sequester enzymes for catabolizing methanol. Further, the availability of nutrients aside from carbon, such as nitrogen, biotin or ammonium sulfate, was important for efficient protein production. Possible correlations between biomass and recombinant protein production, as well as a negative effect on cell proliferation due to heterologous gene expression, need to be evaluated by controlled bioreactor experiments.

Here, up to six mutations within a single P_{CTA1} promoter variant were combined, enhancing it significantly as opposed to single mutant variants. Interestingly, for the best promoter variants, the accumulation of mutations did not interfere with the integrity nor impair the tightness of the promoter. Also, mutations with either de-repression or induction properties, did not deteriorate each other, when being combined and showed high additivity. In fact, the beneficial effects were stacked upon each other and the more mutations enhancing one of the properties were applied, the stronger the promoter got.

Therefore, accumulation of beneficial mutations could even be further exhausted to extend the capability of the promoter. This can be evaluated for example by applying DNA shuffling methods combined with screening. In current cultivation protocols the induction strategy always led to better results than de-repressed conditions. However, the stronger the promoter variants will become, the more attractive the production titers under de-repression will get and saturation can be reached even by methanol free approaches. Further, longer cultivation periods would be interesting for future studies and evaluations. In several screenings the production level of recombinant protein did not decline towards the end. Thus, a prolonged induction/de-repression phase could determine for how long the established expression system was viable and productive. And the best variant for late phase expression still needs to be identified. This is an essential consideration, since longer production phases abbreviate dead times, which should be kept as short as possible in industrial applications.

As a follow-up study to this thesis the project was carried on by Antonia Volpini de Maestri, who isolated the transcriptome from shake flask cultivations and determined CalB mRNA concentration by real-time PCR experiments. This should answer the question, whether a P_{CTA1} promoter variant was also improved on a transcriptional level and if that improvement correlated with the measured enzyme activity in the culture supernatant.

5. BIBLIOGRAPHY

1. Cregg JM, Cereghino JL, Shi J, Higgins DR. Recombinant Protein Expression in Pichia pastoris. Mol Biotechnol [Internet]. 2000;16(1):23–52.
2. Schmidt FR. Recombinant expression systems in the pharmaceutical industry. Appl Microbiol Biotechnol. 2004;65(4):363–72.
3. Kurtzman CP. Description of Komagataella phaffii sp. nov. and the transfer of Pichia pseudopastoris to the methylotrophic yeast genus Komagataella. Int J Syst Evol Microbiol. 2005;55(2):973–6.
4. Kurtzman CP. Biotechnological strains of Komagataella (Pichia) pastoris are Komagataella phaffii as determined from multigene sequence analysis. J Ind Microbiol Biotechnol. 2009;36(11):1435–8.
5. Wegner GH. Emerging applications of the methylotrophic yeasts. FEMS Microbiol Rev. 1990 Dec;7(3–4):279–83.
6. Ogata K, Nishikawa H, Ohsugi M. A yeast capable of utilizing methanol. Agric Biol Chem. 1969;33(10):1519–20.
7. Trentmann O, Khatri NK, Hoffmann F. Reduced oxygen supply increases process stability and product yield with recombinant Pichia pastoris. Biotechnol Prog. 2004;20(6):1766–75.
8. Jahic M, Veide A, Charoenrat T, Teeri T, Enfors SO. Process technology for production and recovery of heterologous proteins with Pichia pastoris. Biotechnol Prog. 2006;22(6):1465–73.
9. Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast Pichia pastoris. FEMS Microbiol Rev. 2000;24(1):45–66.
10. Jahic M, Rotticci-Mulder J, Martinelle M, Hult K, Enfors S-O. Modeling of growth and energy metabolism of Pichia pastoris producing a fusion protein. Bioprocess Biosyst Eng [Internet]. 2002 Mar;24(6):385–93.
11. Vogl T, Sturmberger L, Kickenweiz T, Wasmayer R, Schmid C, Hatzl AM, et al. A Toolbox of Diverse Promoters Related to Methanol Utilization: Functionally Verified Parts for Heterologous Pathway Expression in Pichia pastoris. ACS Synth Biol. 2016;5(2):172–86.
12. Rußmayer H, Buchetics M, Gruber C, Valli M, Grillitsch K, Modarres G, et al. Systems-level organization of yeast methylotrophic lifestyle. BMC Biol [Internet]. 2015;13(1).
13. Küberl A, Schneider J, Thallinger GG, Anderl I, Wibberg D, Hajek T, et al. High-quality genome sequence of Pichia pastoris CBS7435. J Biotechnol. 2011;154(4):312–20.
14. Krainer FW, Dietzsch C, Hajek T, Herwig C, Spadiut O, Glieder A. Recombinant protein expression in Pichia pastoris strains with an engineered methanol utilization pathway. Microb Cell Fact [Internet]. 2012;11(1):22.
15. Yano T, Takigami E, Yurimoto H, Sakai Y. Yap1-regulated glutathione redox system curtails accumulation of formaldehyde and reactive oxygen species in methanol metabolism of Pichia pastoris. Eukaryot Cell. 2009;8(4):540–9.

16. Ellis SB, Brust PF, Koutz PJ, Waters AF, Harpold MM, Gingeras TR. Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast Pichia pastoris. Mol Cell Biol [Internet]. 1985;5(5):1111–21.

17. Hartner FS, Glieder A. Regulation of methanol utilisation pathway genes in yeasts. Microb Cell Fact. 2006;5:1–21.

18. van der Klei IJ, Yurimoto H, Sakai Y, Veenhuis M. The significance of peroxisomes in methanol metabolism in methylotrophic yeast. Biochim Biophys Acta - Mol Cell Res. 2006;1763(12):1453–62.

19. Yano T, Yurimoto H, Sakai Y. Activation of the Oxidative Stress Regulator PpYap1 through Conserved Cysteine Residues during Methanol Metabolism in the Yeast Pichia pastoris. Biosci Biotechnol Biochem [Internet]. 2009;73(6):1404–11.

20. Prielhofer R, Cartwright SP, Graf AB, Valli M, Bill RM, Mattanovich D, et al. Pichia pastoris regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. BMC Genomics [Internet]. 2015;16(1):1–17.

21. Gleeson MA, Sudbery PE. The methylotrophic yeasts. Yeast. 1988;4(1):1–15.

22. Keizer I, Roggenkamp R, Harder W, Veenhuis M. Location of catalase in crystalline peroxisomes of methanol-grown Hansenula polymorpha. FEMS Microbiol Lett. 1992 May;72(1):7–11.

23. Verduyn C, Giuseppin ML, Scheffers WA, van Dijken JP. Hydrogen peroxide metabolism in yeasts. Appl Environ Microbiol. 1988 Aug;54(8):2086–90.

24. Ghaemmaghami S, Huh W-K, Bower K, Howson RW, Belle A, Dephoure N, et al. Global analysis of protein expression in yeast. Nature. 2003 Oct;425(6959):737–41.

25. Roggenkamp R, Janowicz Z, Stanikowski B, Hollenberg CP. Biosynthesis and regulation of the peroxisomal methanol oxidase from the methylotrophic yeast Hansenula polymorpha. Mol Gen Genet. 1984;194(3):489–93.

26. Janowicz ZA, Eckart MR, Drewke C, Roggenkamp RO, Hollenberg CP, Maat J, et al. Cloning and characterization of the DAS gene encoding the major methanol assimilatory enzyme from the methylotrophic yeast Hansenula polymorpha. Nucleic Acids Res. 1985 May;13(9):3043–62.

27. Stewart MQ, Esposito RD, Gowani J, Goodman JM. Alcohol oxidase and dihydroxyacetone synthase, the abundant peroxisomal proteins of methylotrophic yeasts, assemble in different cellular compartments. J Cell Sci. 2001 Aug;114(Pt 15):2863–8.

28. Sakai Y, Yurimoto H, Matsuo H, Kato N. Regulation of peroxisomal proteins and organelle proliferation by multiple carbon sources in the methylotrophic yeast, Candida boidinii. Yeast. 1998 Sep;14(13):1175–87.

29. Neuberger G, Maurer-Stroh S, Eisenhaber B, Hartig A, Eisenhaber F. Prediction of peroxisomal targeting signal 1 containing proteins from amino acid sequence. J Mol Biol. 2003 May;328(3):581–92.

30. Sakai Y, Murdanoto AP, Konishi T, Iwamatsu A, Kato N. Regulation of the formate dehydrogenase gene, FDH1, in the methylotrophic yeast Candida boidinii and growth characteristics of an FDH1-disrupted strain on methanol, methylamine, and choline. J Bacteriol. 1997 Jul;179(14):4480–5.

31. Gellissen G. Hansenula Polymorpha: Biology and Applications. 1st editio. Germany: Weinheim: Wiley-VCH; 2002. 347 p.

32. Cregg JM, Madden KR, Barringer KJ, Thill GP, Stillman C a. Functional characterization of the two alcohol oxidase genes from the yeast Pichia pastoris. Mol Cell Biol [Internet]. 1989;9(3):1316–23.

33. Inan M, Meagher MM. Non-Repressing Carbon Sources for Alcohol Oxidase (AOXI) Promoter of Pichia pastoris Non-Repressing Carbon Sources for Alcohol Oxidase (AOXI) Promoter of Pichia pastoris. Pap Biochem Eng. 2001;92(6):585–9.

34. Hartner FS, Ruth C, Langenegger D, Johnson SN, Hyka P, Lin-Cereghino GP, et al. Promoter library designed for fine-tuned gene expression in Pichia pastoris. Nucleic Acids Res. 2008;36(12).

35. Kranthi BV, Balasubramanian N, Rangarajan PN. Isolation of a single-stranded DNA-binding protein from the methylotrophic yeast, Pichia pastoris and its identification as zeta crystallin. Nucleic Acids Res. 2006;34(14):4060–8.

36. Parua PK, Ryan PM, Trang K, Young ET. Pichia pastoris 14-3-3 regulates transcriptional activity of the methanol inducible transcription factor Mxr1 by direct interaction. Mol Microbiol. 2012;85(2):282–98.

37. Staley CA, Huang A, Nattestad M, Oshiro KT, Ray LE, Mulye T, et al. Analysis of the 5' untranslated region (5'UTR) of the alcohol oxidase 1 (AOX1) gene in recombinant protein expression in Pichia pastoris. Gene [Internet]. 2012;496(2):118–27.

38. Xuan Y, Zhou X, Zhang W, Zhang X, Song Z, Zhang Y. An upstream activation sequence controls the expression of AOX1 gene in Pichia pastoris. FEMS Yeast Res. 2009;9(8):1271–82.

39. Vogl T, Glieder A. Regulation of Pichia pastoris promoters and its consequences for protein production. N Biotechnol [Internet]. 2013;30(4):385–404.

40. Couderc R, Baratti J. chain . In this case the energy of methanol oxidation is available to the cell . 3 , 4) The higher yields observed with bacteria as compared to yeasts were attributed to the first step of methanol oxidation . 5 , 6) Since the yield on methanol of the ye. Agric Biol chemestry. 1980;44(10):2279–89.

41. Ahmad M, Hirz M, Pichler H, Schwab H. Protein expression in Pichia pastoris: Recent achievements and perspectives for heterologous protein production. Appl Microbiol Biotechnol. 2014;98(12):5301–17.

42. Blazeck J, Alper HS. Promoter engineering: Recent advances in controlling transcription at the most fundamental level. Biotechnol J. 2013;8(1):46–58.

43. Tschopp JF, Brust PF, Cregg JM, Stillman CA, Gingeras TR. Expression of the lacZ gene from two methanol-regulated promoters in Pichia pastoris. Nucleic Acids Res. 1987;15(9):3859–76.

44. Sreekrishna K, Potenz RH, Cruze JA, McCombie WR, Parker KA, Nelles L, et al. High level expression of heterologous proteins in methylotrophic yeast Pichia pastoris. J Basic Microbiol. 1988;28(4):265–78.

45. Hasslacher M, Schall M, Hayn M, Bona R, Rumbold K, Lückl J, et al. High-level intracellular expression of hydroxynitrile lyase from the tropical rubber tree Hevea brasiliensis in microbial hosts. Protein Expr Purif. 1997;11(1):61–71.

46. Ciofalo V, Barton N, Kreps J, Coats I, Shanahan D. Safety evaluation of a lipase enzyme preparation, expressed in Pichia pastoris, intended for use in the degumming of edible vegetable oil. Regul Toxicol Pharmacol. 2006;45(1):1–8.

47. Damasceno LM, Huang CJ, Batt CA. Protein secretion in Pichia pastoris and advances in protein production. Appl Microbiol Biotechnol. 2012;93(1):31–9.

48. Thompson CA. FDA approves kallikrein inhibitor to treat hereditary angioedema. Am J Heal Pharm. 2010;67(2):93.

49. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the Pichia pastoris expression system. Yeast. 2005;22(4):249–70.

50. Idiris A, Tohda H, Kumagai H, Takegawa K. Engineering of protein secretion in yeast: Strategies and impact on protein production. Appl Microbiol Biotechnol. 2010;86(2):403–17.

51. Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang Y, Rios S, et al. Humanization of yeast to produce complex terminally sialylated glycoproteins. Science. 2006 Sep;313(5792):1441–3.

52. Lin-cereghino GP, Jay A, Sunga J, Gleeson M, Lin GP, Lin J, et al. New selectable marker / auxotrophic host strain combinations for molecular genetic manipulation of Pichia ... New selectable marker / auxotrophic host strain combinations for molecular genetic manipulation of Pichia pastoris q. 2001;263(February):159–69.

53. Daly R, Hearn MTW. Expression of heterologous proteins in Pichia pastoris: A useful experimental tool in protein engineenring and production. J Mol Recognit. 2005;18(2):119–38.

54. Bardwell JC. Building bridges: disulphide bond formation in the cell. Mol Microbiol. 1994 Oct;14(2):199–205.

55. White CE, Kempi NM, Komives EA. Expression of highly disulfide-bonded proteins in Pichia pastoris. Structure. 1994 Nov;2(11):1003–5.

56. Cole PA. Chaperone-assisted protein expression. Structure. 1996 Mar;4(3):239–42.

57. Meldgaard M, Svendsen I. Different effects of N-glycosylation on the thermostability of highly homologous bacterial (1,3-1,4)-beta-glucanases secreted from yeast. Microbiology. 1994 Jan;140 (Pt 1:159–66.

58. Holst B, Bruun AW, Kielland-Brandt MC, Winther JR. Competition between folding and glycosylation in the endoplasmic reticulum. EMBO J [Internet]. 1996;15(14):3538–46.

59. Romanos MA, Scorer CA, Clare JJ. Foreign gene expression in yeast: a review. Yeast. 1992 Jun;8(6):423–88.

60. Cregg JM, Vedvick TS, Raschke WC. Recent advances in the expression of foreign genes in Pichia pastoris. Biotechnology (N Y). 1993;11(8):905–10.

61. Waters MG, Evans E a, Blobel G. Prepro-alpha-factor has a cleavable signal sequence. J Biol Chem. 1988;263(13):6209–14.

62. Julius D, Brake A, Blair L, Kunisawa R, Thorner J. Isolation of the putative structural gene for the lysine-arginine-cleaving endopeptidase required for processing of yeast prepro-α-factor. Cell. 1984;37(3):1075–89.

63. Brake AJ, Merryweather JP, Coit DG, Heberlein UA, Masiarz FR, Mullenbach GT, et

al. Alpha-factor-directed synthesis and secretion of mature foreign proteins in Saccharomyces cerevisiae. Proc Natl Acad Sci [Internet]. 1984;81(15):4642–6.

64. Fitzgerald I, Glick BS. Secretion of a foreign protein from budding yeasts is enhanced by cotranslational translocation and by suppression of vacuolar targeting. Microb Cell Fact. 2014;13(1):1–12.

65. Liu L, Redden H, Alper HS. Frontiers of yeast metabolic engineering: Diversifying beyond ethanol and Saccharomyces. Curr Opin Biotechnol [Internet]. 2013;24(6):1023–30.

66. Bailey JE. Toward a science of metabolic engineering. Science. 1991 Jun;252(5013):1668–75.

67. Stephanopoulos G. Synthetic biology and metabolic engineering. ACS Synth Biol. 2012;1(11):514–25.

68. Fitzherbert EB, Struebig MJ, Morel A, Danielsen F, Bruhl CA, Donald PF, et al. How will oil palm expansion affect biodiversity? Trends Ecol Evol. 2008 Oct;23(10):538–45.

69. Brophy JAN, Voigt CA. Principles of genetic circuit design. Nat Methods. 2014;11(5):508–20.

70. Montiel D, Kang H-S, Chang F-Y, Charlop-Powers Z, Brady SF. Yeast homologous recombination-based promoter engineering for the activation of silent natural product biosynthetic gene clusters. Proc Natl Acad Sci [Internet]. 2015;112(29):8953–8.

71. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods [Internet]. 2009;6(5):343–5.

72. Aw R, Polizzi KM. Can too many copies spoil the broth? Microb Cell Fact. 2013;12(1):1–9.

73. Mattanovich D, Graf A, Stadlmann J, Dragosits M, Redl A, Maurer M, et al. Genome, secretome and glucose transport highlight unique features of the protein production host Pichia pastoris. Microb Cell Fact. 2009;8:1–13.

74. De Schutter K, Lin YC, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, et al. Genome sequence of the recombinant protein production host Pichia pastoris. Nat Biotechnol. 2009;27(6):561–6.

75. Lefevre SD, van Roermund CW, Wanders RJA, Veenhuis M, Van der Klei IJ. The significance of peroxisome function in chronological aging of Saccharomyces cerevisiae. Aging Cell. 2013;12(5):784–93.

76. Eggeling L, Sahm H. Derepression and partial insensitivity to carbon catabolite repression of the methanol dissimilating enzymes inHansenula polymorpha. Eur J Appl Microbiol Biotechnol [Internet]. 1978 Sep;5(3):197–202.

77. Egli T, Haltmeier T, Fiechter A. Regulation of the synthesis of methanol oxidizing enzymes inKloeckera sp. 2201 andHansenula polymorpha, a comparison. Arch Microbiol [Internet]. 1982 Mar;131(2):174–5.

78. Basehoar AD, Zanton SJ, Pugh BF. Identification and distinct regulation of yeast TATA box-containing genes. Cell. 2004;116(5):699–709.

79. Prielhofer R, Maurer M, Klein J, Wenger J, Kiziak C, Gasser B, et al. Induction without methanol: novel regulated promoters enable high-level expression in Pichia

pastoris. Microb Cell Fact. 2013;12:5.

80. Wang J, Wang X, Shi L, Qi F, Zhang P, Zhang Y, et al. Methanol-Independent Protein Expression by AOX1 Promoter with trans-Acting Elements Engineering and Glucose-Glycerol-Shift Induction in Pichia pastoris. Sci Rep. 2017;7:41850.

81. Shen W, Xue Y, Liu Y, Kong C, Wang X, Huang M, et al. A novel methanol-free Pichia pastoris system for recombinant protein expression. Microb Cell Fact. 2016;15(1):178.

82. Liang S, Zou C, Lin Y, Zhang X, Ye Y. Identification and characterization of P GCW14 : a novel, strong constitutive promoter of Pichia pastoris. Biotechnol Lett. 2013 Nov;35(11):1865–71.

83. Waterham HR, Digan ME, Koutz PJ, Lair S V., Cregg JM. Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. Gene. 1997;186(1):37–44.

84. Liu X, Bushnell DA, Wang D, Calero G, Kornberg RD. Structure of an RNA Polymerase II–TFIIB Complex and the Transcription Initiation Mechanism. Science (80-) [Internet]. 2010 Jan 8;327(5962):206 LP – 209.

85. Smale ST, Kadonaga JT. The RNA polymerase II core promoter. Annu Rev Biochem. 2003;72:449–79.

86. Murakami KS, Darst SA. Bacterial RNA polymerases: the wholo story. Curr Opin Struct Biol. 2003 Feb;13(1):31–9.

87. Kostrewa D, Zeller ME, Armache K-J, Seizl M, Leike K, Thomm M, et al. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. Nature. 2009 Nov;462(7271):323–30.

88. Hahn S, Young ET. Transcriptional regulation in Saccharomyces cerevisiae: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. Genetics. 2011 Nov;189(3):705–36.

89. Benoist C, Chambon P. In vivo sequence requirements of the SV40 early promoter region. Nature [Internet]. 1981 Mar 26;290:304.

90. Orphanides G, Lagrange T, Reinberg D. The general transcription factors of RNA polymerase II. Genes Dev. 1996 Nov;10(21):2657–83.

91. Butler JEF, Kadonaga JT. The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes Dev [Internet]. 2002;16(20):2583–92.

92. Huisinga KL, Pugh BF. A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. Mol Cell. 2004 Feb;13(4):573–85.

93. Ptashne M, Gann A. Transcriptional activation by recruitment. Nature. 1997 Apr;386(6625):569–77.

94. Gaudreau L, Keaveney M, Nevado J, Zaman Z, Bryant GO, Struhl K, et al. Transcriptional activation by artificial recruitment in yeast is influenced by promoter architecture and downstream sequences. Vol. 96, Proceedings of the National Academy of Sciences of the United States of America. 1999. p. 2668–73.

95. Gurvitz A, Hiltunen JK, Erdmann R, Hamilton B, Hartig A, Ruis H, et al. Saccharomyces cerevisiae Adr1p Governs Fatty Acid β-Oxidation and Peroxisome

Proliferation by Regulating POX1 and PEX11. J Biol Chem. 2001;276(34):31825–30.

96. Nevoigt E, Kohnke J, Fischer CR, Alper H, Stahl U, Stephanopoulos G. Engineering of promoter replacement cassettes for fine-tuning of gene expression in Saccharomyces cerevisiae. Appl Environ Microbiol. 2006 Aug;72(8):5266–73.

97. Remans T, Grof CPL, Ebert PR, Schenk PM. Functional promoter analysis using an approach based on an in vitro evolution strategy. Biotechniques. 2005 Feb;38(2):209-210,212,214-216.

98. Portela RMC, Vogl T, Ebner K, Oliveira R, Glieder A. Pichia pastoris Alcohol Oxidase 1 (AOX1) Core Promoter Engineering by High Resolution Systematic Mutagenesis. Biotechnol J. 2018;13(3):1–8.

99. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 1987 Oct;15(20):8125–48.

100. Redden H, Alper HS. The development and characterization of synthetic minimal yeast promoters. Nat Commun [Internet]. 2015;6:1–9.

101. Portela RMC, Vogl T, Kniely C, Fischer JE, Oliveira R, Glieder A. Synthetic Core Promoters as Universal Parts for Fine-Tuning Expression in Different Yeast Species. ACS Synth Biol. 2017;6(3):471–84.

102. Hussein M. Komagataella phaffii promoter engineering. Graz University of Technology; 2017.

103. Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. 1995 Dec;23(23):4878–84.

104. Lewis EB. The phenomenon of position effect. Adv Genet. 1950;3:73–115.

105. Näätsaari L, Mistlberger B, Ruth C, Hajek T, Hartner FS, Glieder A. Deletion of the pichia pastoris ku70 homologue facilitates platform strain generation for gene expression and synthetic biology. PLoS One. 2012;7(6).

106. Rine J, Hansen W, Hardeman E, Davis RW. Targeted selection of recombinant clones through gene dosage effects. Proc Natl Acad Sci U S A. 1983 Nov;80(22):6750–4.

107. Cregg JM, Madden KR. Development of yeast transformation systems and construction of methanol-utilization-defective mutants of Pichia pastori by gene disruption. CRC Press; 1987.

108. Romanos M. Advances in the use of Pichia pastoris for high-level gene expression. Curr Opin Biotechnol [Internet]. 1995 Jan 1 [cited 2018 Jul 24];6(5):527–33.

109. Cregg JM, Tschopp JF, Stillman C, Siegel R, Akong M, Craig WS, et al. High–Level Expression and Efficient Assembly of Hepatitis B Surface Antigen in the Methylotrophic Yeast, Pichia Pastoris. Bio/Technology [Internet]. 1987 May 1;5:479.

110. Promega. Wizard Plus SV Minipreps DNA Purification System Protocol. Promega [Internet]. 2009;1–2.

111. Promega. Wizard ® SV Gel and PCR Clean-Up System. Promega [Internet]. 2010;1–12.

112. Seidman CE, Struhl K. Introduction of plasmid DNA into cells. Curr Protoc Neurosci. 2001;Appendix 1.

113. Lin-Cereghino J, Wong WW, Xiong S, Giang W, Luong LT, Vu J, et al. Condensed

protocol for competent cell preparation and transformation of the methylotrophic yeast Pichia pastoris. Biotechniques. 2005;38(1):44–8.

114. Zhang N, Suen W-C, Windsor W, Xiao L, Madison V, Zaks A. Improving tolerance of Candida antarctica lipase B towards irreversible thermal inactivation through directed evolution. Protein Eng Des Sel [Internet]. 2003;16(8):599–605.

115. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein using the principle of protein dye binding. Anal Biochem. 1976;72:248–54.

116. Place TL, Domann FE, Case AJ. Limitations of oxygen delivery to cells in culture : An underappreciated problem in basic and translational research. Free Radic Biol Med [Internet]. 2017;113(May):311–22.

117. Eibl D, Eibl R. Disposable Bioreactors II [Internet]. Springer Berlin Heidelberg; 2013. (Advances in Biochemical Engineering/Biotechnology).

118. Merika M, Orkin SH. DNA-Binding Specificity of GATA Family Transcription Factors. 1993;13(7):3999–4010.

119. Magasanik B, Kaiser CA. Nitrogen regulation in Saccharomyces cerevisiae. 2002;290:1–18.

120. Weis R, Luiten R, Skranc W, Schwab H, Wubbolts M, Glieder A. Reliable high-throughput screening with Pichia pastoris by limiting yeast cell death phenomena. FEMS Yeast Res. 2004;5(2):179–89.

121. Verkhusha V V, Kuznetsova IM, Stepanenko O V, Zaraisky AG, Shavlovsky MM, Turoverov KK, et al. Accelerated Publications High Stability of Discosoma DsRed As Compared to Aequorea EGFP †. 2003;7879–84.

122. Stepanenko O V, Verkhusha V V, Kazakov VI, Shavlovsky MM, Kuznetsova IM, Uversky VN, et al. Comparative Studies on the Structure and Stability of Fluorescent Proteins EGFP ,. 2004;14913–23.

123. Wilhelmi AE. Structure and function in growth hormone, in: Hormone Drugs, Proceedings of the FDA-USP Workshop on Drug and Reference Standards for Insulins, Somatropins and Thyroid-axis Drugs. USP Conv Inc. 1982;369–381.

124. Iglesias P, Díez JJ, Diez JJ. Clinical applications of recombinant human growth hormone in adults. Expert Opin Pharmacother [Internet]. 1999;1(1):97–107.

125. Høegh I, Patkar S, Halkier T, Hansen MT. Two lipases from Candida antarctica : cloning and expression in Aspergillus oryzae. Can J Bot [Internet]. 1995;73(S1):869–75.

126. Björkling F, Godtfredsen SE, Kirk O. The future impact of industrial lipases. Trends Biotechnol [Internet]. 1991 Jan 1 [cited 2019 Feb 15];9(1):360–3.

127. Kirk O, Björkling F, Godtfredsen SE, Larsen TO. Fatty acid specificity in lipase-catalyzed synthesis of glucoside esters. Biocatal Biotransformation. 1992;6(2):127–34.

128. Rogalska E, Cudrey C, Ferrato F, Verger R. Stereoselective hydrolysis of triglycerides by animal and microbial lipases. Chirality. 1993;5(1):24–30.

129. Anderson EM, Larsson KM, Kirk O. One biocatalyst - many applications: The use of Candida antarctica B-lipase in organic synthesis. Biocatal Biotransformation. 1998;16(3):181–204.

130. Lehle L. Protein glycosylation in yeast. Antonie Van Leeuwenhoek. 1992;61(2):133–4.

131. Mattanovich D, Gasser B, Hohenblum H, Sauer M. Stress in recombinant protein producing yeasts. J Biotechnol. 2004;113(1–3):121–35.

132. Rakestraw JA, Sazinsky SL, Piatesi A, Antipov E, Wittrup KD. Directed evolution of a secretory leader for the improved expression of heterologous proteins and full-length antibodies in Saccharomyces cerevisiae. Biotechnol Bioeng. 2009;103(6):1192–201.

133. Yang S, Kuang Y, Li H, Liu Y, Hui X, Li P, et al. Enhanced Production of Recombinant Secretory Proteins in Pichia pastoris by Optimizing Kex2 P1 ' site. 2013;8(9):1–11.

134. Charles M. Fermentation scale-up: Problems and possibilities. Trends Biotechnol. 1985;3(6):134–9.

135. Pepper I, Gerba C, Gentry T. Environmental Microbiology. 3rd Editio. Academic Press; 2014. 38–39 p.

136. Solà A, Maaheimo H, Ylönen K, Ferrer P, Szyperski T. Amino acid biosynthesis and metabolic flux profiling of Pichia pastoris. Eur J Biochem. 2004;271(12):2462–70.

137. Glick BR. Metabolic load and heterlougos gene expression. Biotechnol Adv. 1995;13(2):247–61.

138. Guan B, Chen F, Su S, Duan Z, Chen Y, Li H, et al. Effects of co-overexpression of secretion helper factors on the secretion of a HSA fusion protein (IL2-HSA) in pichia pastoris. Yeast. 2016 Nov;33(11):587–600.

139. Lin-Cereghino GP, Stark CM, Kim D, Chang J, Shaheen N, Poerwanto H, et al. The effect of alpha-mating factor secretion signal mutations on recombinant protein expression in Pichia pastoris. Gene. 2013 May;519(2):311–7.

6. LIST OF FIGURES

7. LIST OF TABLES

8. APPENDIX

List of materials

Following lists tabulate all materials, tools, strains and other things used within this master thesis.

Chemicals and Reagents

Reagent	Company
1-Step™ Ultra TMB ELISA solution	Thermo Fisher Scientific Inc., Waltham, MA, USA
Acetic acid (100%)	Carl Roth GmbH, Karlsruhe, Germany
Aqua bidest. "Fresenius"	Fresenius Kabi, Graz, Austria
Bacto™ Pepton	BD Biosciences, San Jose, CA, United States
Bacto™ Yeast Extract	BD Biosciences, San Jose, CA, United States
Bicine	Fluka Chemia AG, Basel, Switzerland
Bio-Rad Protein Assay Dye Reagent Concentrate	Bio-Rad Laboratories Inc., Hercules, CA, United States
Biotin	Fluka Chemia AG, Basel, Switzerland
bovine serum albumin	Sigma-Aldrich GmbH, Vienna, Austria
Chloroform	Carl Roth GmbH, Karlsruhe, Germany
dATP	New England Biolabs® Inc., Ipswhich, MA, USA
dCTP	New England Biolabs® Inc., Ipswhich, MA, USA
dGTP	New England Biolabs® Inc., Ipswhich, MA, USA
Difco™ Yeast nitrogen Base w/o Amino Acids	BD Biosciences, San Jose, CA, United States
disodium ethylenediaminetetraacetic acid (EDTA)	Carl Roth GmbH, Karlsruhe, Germany
Dithiothreiol (DTT)	Carl Roth GmbH, Karlsruhe, Germany
DMSO	Carl Roth GmbH, Karlsruhe, Germany
dTTP	New England Biolabs® Inc., Ipswhich, MA, USA

Ethidium bromide solution (1%)	Carl Roth GmbH, Karlsruhe, Germany
Ethylenglycol	Sigma-Aldrich GmbH, Vienna, Austria
Gel loading dye 6x	New England Biolabs® Inc., Ipswhich, MA, USA
GeneRuler TM 1 kb DNA Ladder	Thermo Fisher Scientific Inc., Waltham, MA, USA
GeneRuler™ DNA Ladder Mix	Thermo Fisher Scientific Inc., Waltham, MA, USA
GH Antibody (E-7): sc-374266	Santa Cruz Biotechnology Inc., Dallas, Texas, USA
Glycerol (≥ 98%)	Carl Roth GmbH, Karlsruhe, Germany
HCl (37%)	Merck KGaA, Darmstadt, Germany
ISO reaction buffer	New England Biolabs® Inc., Ipswhich, MA, USA
K2HPO4	Carl Roth GmbH, Karlsruhe, Germany
KCl	Merck KGaA, Darmstadt, Germany
KH2PO4	Carl Roth GmbH, Karlsruhe, Germany
KOH	Carl Roth GmbH, Karlsruhe, Germany
LB Agar	Carl Roth GmbH, Karlsruhe, Germany
m-IgGκ BP-HRP: sc-516102	Santa Cruz Biotechnology Inc., Dallas, Texas, USA
Methanol (≥ 99.8%)	Carl Roth GmbH, Karlsruhe, Germany
MgCl2	Carl Roth GmbH, Karlsruhe, Germany
MgSO4	Carl Roth GmbH, Karlsruhe, Germany
Milkpowder	Carl Roth GmbH, Karlsruhe, Germany
MOPS	Carl Roth GmbH, Karlsruhe, Germany
NaCl	Carl Roth GmbH, Karlsruhe, Germany
NaOH	Carl Roth GmbH, Karlsruhe, Germany
p-nitrophenolbutyrate	Sigma-Aldrich GmbH, Vienna, Austria
PageBlue™ Protein Staining Solution	Thermo Scientific, Inc. Waltham, MA, USA
PageRulerTM Prestained Protein Ladder	Thermo Scientific, Inc. Waltham, MA, USA
peqGOLD Univesal Agarose	PEQLAB by VWR International, Erlangen, Deutschland
peqGOLD Univesal Agarose	PEQLAB by VWR International, Erlangen, Deutschland

Phusion® High-Fidelity DNA Polymerase	New England Biolabs® Inc., Ipswhich, MA, USA
SDS	Sigma-Aldrich GmbH, Vienna, Austria
Sorbitol	Carl Roth GmbH, Karlsruhe, Germany
Sulfuric acid	Carl Roth GmbH, Karlsruhe, Germany
T5 exonuclease	New England Biolabs® Inc., Ipswhich, MA, USA
Taq DNA ligase	New England Biolabs® Inc., Ipswhich, MA, USA
Tris	Carl Roth GmbH, Karlsruhe, Germany
TritonX-100	Carl Roth GmbH, Karlsruhe, Germany
Tryptone	Carl Roth GmbH, Karlsruhe, Germany
Zecoin™	InvivoGen-Eubio, Vienna, Austria
D(+)-glucose monohydrate	Carl Roth GmbH, Karlsruhe, Germany

Instruments and Devices

The table lists all laboratory equipment used within this master thesis. Smaller items, such as pipette tips, 1.5 mL tubes or toothpicks are not mentioned.

Equipment	**Company**
Plate Reader	
Eon™	BioTek Instruments, Inc., Winooski, VT, USA
Synergy Mx Plate Reader	Biotek Instruments, Inc., Winooski, United States
Centrifuge	
Eppendorf Centrifuge 5810 R	Eppendorf AG, Hamburg, Germany
Eppendorf Centrifuge 5415 R	Eppendorf AG, Hamburg, Germany
Eppendorf Centrifuge 5415 D	Eppendorf AG, Hamburg, Germany
PCR Cycler	
GeneAmp® PCR system 2700	Applied Biosystems by Thermo Fisher Scientific, Inc., Foster City, California, United States
2720 Thermal Cycler	Applied Biosystems by Thermo Fisher Scientific, Inc., Foster City, California, United States

Pipettes

Denville XL 3000i™ 5-50 µL multichannel pipette	Denville Scientific, Inc., Charlotte, NC, United States
Denville XL 3000i™ 50-300 µL multichannel pipette	Denville Scientific, Inc., Charlotte, NC, United States
eppendorf Xplorer 15-300 µL automatic multichannel pipette	Eppendorf AG, Hamburg, Germany
eppendorf Xplorer 50-1200 µL automatic multichannel pipette	Eppendorf AG, Hamburg, Germany
peqlab peqPETTE 0.5-2 µL single channel pipette	PEQLAB by VWR International, Erlangen, Deutschland
peqlab peqPETTE 0.5-10 µL single channel pipette	PEQLAB by VWR International, Erlangen, Deutschland
Gilson Pipetman® 2-20 µL single channel pipette	Gilson, Inc., Mddleton, WI, United States
Gilson Pipetman® 20-200 µL single channel pipette	Gilson, Inc., Mddleton, WI, United States
Gilson Pipetman® 200-1000 µL single channel pipette	Gilson, Inc., Mddleton, WI, United States

Plates

Bel-Art 96-Well Deep Well Plate	Bel-Art Products, Wayne, NJ, United States
Greiner 96-well ELISA Microlon® microplates	Greiner Bio-One GmbH, Kremsmünster, Austria
Greiner 96-well Multiwell plate	Greiner Bio-One GmbH, Kremsmünster, Austria
Nunc™ MicroWell™ 96-Well Optical-Bottom Plate with Polymer Base	Thermo Scientific, Inc., Waltham, MA, USA

Gel electrophoresis

Sub-cell GT	Bio-Rad Laboratories Inc., Hercules, CA, United States
BioRad PowerPac Basic Power Supply	Bio-Rad Laboratories Inc., Hercules, CA, United States
GelDoc-It™ Imaging System	UVP®, Cambridge, UK
Digital printer UP-D897 6x	Sony, Vienna, Austria

Scales	
Bench Scale FKB	Kern & Sohn GmbH, Balingen, Germny
Sartorius Analytical BL120S	Sartorius Stedim Biotech GmbH, Göttingen, Germany
SDS-PAGE	
Mini Gel Tank system	Invitrogen by Thermo Fisher Scientific, Inc., Waltham, MA, USA
NuPAGE™ 4-12% Bis-Tris Gels	Invitrogen by Thermo Fisher Scientific, Inc., Waltham, MA, USA
G:BOX Chemi HR16 bioimaging system	Syngene, Frederick, MD, USA
Shaker & Incubator	
HT Infors Multitron shaker	Infors AG, Bottmingen, Switzerland
HT Infors Orbitron shaker	Infors AG, Bottmingen, Switzerland
HT Infors RS306 shaker	Infors AG, Bottmingen, Switzerland
Eppendorf Thermomixer comfort	Eppendorf AG, Hamburg, Germany
Heidolph Titramax 1000 Plate Shaker	Heidolph Instruments, Schwabach, Germany
Binder drying oven	Binder GmbH, Tuttlingen, Germany
Photometer and associated materials	
Eppendorf BioPhotometer plus	Eppendorf AG, Hamburg, Germany
Semi-micro cuvette 10x4x45 mm, Polystyrene	Sarstedt Aktiengesellschaft & Co., Nümbrecht, Germany
Nanodrop™ 2000	Thermo Fisher Scientific Inc., Waltham, MA, USA
Transformation	
BioRad Gene Pulser 1652076	Bio-Rad Laboratories Inc., Hercules, CA, United States
Capacitance Extender 1652087	Bio-Rad Laboratories Inc., Hercules, CA, United States
Pulse Controller P/N 1652098	Bio-Rad Laboratories Inc., Hercules, CA, United States
Electroporation cuvettes EP-102	Cell Projects Ltd, Kent, United Kingdom

Shake flask cultivation

250 mL baffled shake flask	Schott AG, Mainz, Germany
Shake flask with integrated sensor (SFS-PSt3)	PreSens Precision Sensing GmbH , Regensburg, Deutschland
Shake flask reader (SFR)	PreSens Precision Sensing GmbH , Regensburg, Deutschland
Glycerol feed disk	PS Biotech GmbH, Aachen, Germany

Other

Heidolph MR 2002 Mangentic Stirrer	Heidolph Instruments, Schwabach, Germany
Heidolph MR 3000 Magnetic Stirrer	Heidolph Instruments, Schwabach, Germany
Vortex Genie 2	Scientific Industries Inc., Bohemia, NY, United States
Fisherbrand™ accumet™ AB150 pH Benchtop Meters	Fisher Scientific by Thermo Fisher Scientific, Hampton, New Hampshire, United States
pH electrode	Sigma-Aldrich GmbH, Vienna, Austria
arium® basic ultrapure water system	Sartorius Stedim Biotech GmbH, Göttingen, Germany
UV-lamp	Laborgeräte Vetter, Wiesloch, Germany
0.22 µm Rotilabo® syringe filters, CME, sterile	Carl Roth GmbH, Karlsruhe, Germany

Enzymes

Enzyme	Type	Company
Q5® High-Fidelity DNA Polymerase	Polymerase	New England Biolabs® Inc., Ipswhich, MA, USA
GoTaq Flexi Polymerase	Polymerase	Promega Corporation, Fitchburg, WI, USA
T5 exonuclease	Exonucelase	New England Biolabs® Inc., Ipswhich, MA, USA
Phusion® High-Fidelity DNA Polymerase	Polymerase	Thermo Fisher Scientific Inc., Waltham, MA, USA
Taq DNA ligase	Ligase	Thermo Fisher Scientific Inc., Waltham, MA, USA

FastDigest SmiI	Restriction enzyme	Thermo Fisher Scientific Inc., Waltham, MA, USA
FastDigest NotI	Restriction enzyme	Thermo Fisher Scientific Inc., Waltham, MA, USA
FastDigest EcoRI	Restriction enzyme	Thermo Fisher Scientific Inc., Waltham, MA, USA
FastDigest XbaI	Restriction enzyme	Thermo Fisher Scientific Inc., Waltham, MA, USA

Software and webtools

Name	Company
GeneSyn software	Syngene International Ltd., Bangalore, India
Image LabTM Version 6.0.1	Bio-Rad Laboratories Inc., Hercules, CA, United States
SnapGene®	GSL Biotech LLC, Chicago, IL, USA
OligoAnalyzer	Integrated DNA Technologies Inc., Coralville, IA, USA
CLC Main Workbench 7	QIAGEN N.V., Spoorstraat, Netherlands
Expasy Translate Tool	http://web.expasy.org/translate/

Primer

Listed are all primers used within this master thesis. The primers were designed using SnapGene® and CLC Main Workbench 7 and were ordered from Integrated DNA Technologies Inc., Coralville, IA, USA. The delivered tubes containing lyophilized primers were diluted to 100 pmol/µL with Aqua bidest. "Fresenius" and frozen at -20°C.

Internal collection number	Primer Name	Sequence 5' --> 3'
CalB		
P17666	alpha+CalB-ILV5-Gib-fw	ACA AAA GAA ACA AGA CAT TAC TGA AGG ATC GAA TTC CGA AAC GAT GAG ATT CCC ATC TA
P17667	alpha+CalB-AOX1_TT-Gib-rev	AGG CAA ATG GCA TTC TGA CAT CCT CTT GAG GCG GCC GCT TAA GGG GTA ACG ATT CCT GAG CAA G
P17947	ILV5-SpeI-fw	CAA CAA AAG AAA CAA GAC ATT ACT GAA GGA TCA CTA G
P17954	Alpha-EcoRI-PCAT500-53-rv	GGA ATC TCA TCG TTT CGG AAT TCT TTA ATT GTA AGT CTT GAC TAG AGC AAG TGT TAA GTG
P17952	Alpha-EcoRI-PCAT500-WT-rv	GGG AAT CTC ATC GTT TCG GAA TTC TTT AAT TGT AAG TCT TGA CTA GAG CAA GTG TTA TGG
P17953	Alpha-EcoRI-PCAT500-48B-rv	GGG AAT CTC ATC GTT TCG GAA TTC TTT AAT TGT AAG TCT TGA CTC ACG GTT ATC TTA TGG
P17955	Alpha-EcoRI-PDF-rv	GAA AAT AGA TGG GAA TCT CAT CGT TTC GGA ATT CGA TTT GAT TGA TGA AGG CAG AGA GCG
hGH		
P17668	Dalpha+hGH-ILV5-gib-fw	ACA AAA GAA ACA AGA CAT TAC TGA AGG ATC GAA TTC ATG AGA TTC CCA TCT ATT TTC ACC GCT G
P17669	Dalpha+hGH-AOX1_TT-rev	AGG CAA ATG GC ATTC TGA CAT CCT CTT GAG GCG GCC GCC TAG AAG CCA CAG CTG CCC
P17947	ILV5-SpeI-fw	CAA CAA AAG AAA CAA GAC ATT ACT GAA GGA TCA CTA G

P17950	Dalpha-PCAT500-53-rv	CGG TGA AAA TAG ATG GGA ATC TCA TTT TAA TTG TAA GTC TTG ACT AGA GCA AGT GTT AAG
P17948	Dalpha-PCAT500-WT-rv	CGG TGA AAA TAG ATG GGA ATC TCA TTT TAA TTG TAA GTC TTG ACT AGA GCA AGT GTT ATG
P17949	Dalpha-PCAT500-48B-rv	CGG TGA AAA TAG ATG GGA ATC TCA TTT TAA TTG TAA GTC TTG ACT CAC GGT TAT CTT ATG
P17951	Dalpha-PDF-rv	CGG TGA AAA TAG ATG GGA ATC TCA TGA TTT GAT TGA TGA AGG CAG AGA GCG
New Variants		
P17673	P_CAT500-eGFP-Gib-rv	AGT TCT TCT CCT TTG CTA GCC ATC G
P18017	P_CAT500-partB-fw	GAT TAG TCA GGA AAG CTG AGC AAT TAA CTT CC
P17947	ILV5-SpeI-fw	CAA CAA AAG AAA CAA GAC ATT ACT GAA GGA TCA CTA G
P18016	P_CAT500-partA-rv	GGA AGT TAA TTG CTC AGC TTT CCT GAC TAA TC
Controls		
P13162	ZeoRTrev	CCG ATC TCG GTC ATA GCT GG
P13372	seqAOX1TT rev	CCG ATC TCG GTC ATA GCT GG
P13330	POXpILV5_R	ACT ACG GGT GGA ATG TTT GG
P15209	CalB_check_rv	CCC TCA AGG TCA GAG TAA CC
P15351	PEM72_rv	GGT TTA GTC CTC CTT ACA CCT TG
P14922	eGFP_BspEI_rev	GCA CTC TTG AAA AAG TCA TGC CGT T
P13948	seqRSFC-AOX1TTrev	CGA GAT AGG CTG ATC AGG AGC AAG
P17696	CalB2_r	AAC ATG AAT GGA GGT GGA GAG
P13330	POXpILV5_R	ACT ACG GGT GGA ATG TTT GG
P16259	Dalpha_seq_fw	TAC TCT GAC CTT GAG GGT GAT TTC
P17459	P17459	TGA TGG CAT TTG CTC CTG ACT A

P16143	AOXseq_rv	TCC CAA ACC CCT ACC ACA AG
P18018	seq-44-jo-fw	GTT TGG CCT CTG GAA ACA AC
P18019	seq-45-jo-rv	CAC CCT CTC CAC TGA CAG

Strains and Plasmids

All plasmids and strains used within this master thesis are listed. They were initially transformed with *E. coli* TOP10'F cells for replication, repairing DNA nicks and supercoiling. After verification of the integrity of the plasmids they were introduced into *P. pastoris* BSY11dKU70 strains for gene expression. The table shows host strains, plasmid name as well as promoter variant and antibiotic resistance encoded in the expression cassette.

P. pastoris **host strain**	***E. coli*** **Host Strain**	**Plasmid**	**Promoter Variant**	**Resistance**
	TOP10'F	pPpT4_pCAT1_α_CalB		Zeocin™
	TOP10'F	pPp_AOX1-pUC-ZeoR-NotI-TT_AOX1		Zeocin™
	TOP10'F	pPp_AOX1-pUC-ZeoR-EcoRI-eGFP-TT_AOX1		Zeocin™
	TOP10'F	pPp_AOX1-pUC-ZeoR-EcoRI-α-CalB-TT_AOX1		Zeocin™
	TOP10'F	pPp_AOX1-pUC-ZeoR-EcoRI-Dα-hGH-TT_AOX1		Zeocin™
Strains bearing expression cassettes encoding for selected P_{CTA1} promoter variants and CalB				
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+53-α-CalB-TT_AOX1	14	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+14B+26B-α-CalB-TT_AOX1	19	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+41B+48B-α-CalB-TT_AOX1	23	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_14B+26B+48B-α-CalB-TT_AOX1	26	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_MM13b+38B-α-CalB-TT_AOX1	46	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_WT-α-CalB-TT_AOX1	WT	Zeocin™

BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PDF-α-CalB-TT_AOX1	PDF	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-α-CalB-TT_AOX1	Negative	Zeocin™

Strains bearing expression cassettes encoding for selected P_{CTA1} promoter variants and hGH

BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+53-Dα-hGH-TT_AOX1	14	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+14B+26B-Dα-hGH-TT_AOX1	19	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+41B+48B-Dα-hGH-TT_AOX1	23	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_14B+26B+48B-Dα-hGH-TT_AOX1	26	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_MM13b+38B-Dα-hGH-TT_AOX1	46	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_WT-Dα-hGH-TT_AOX1	WT	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PDF-Dα-hGH-TT_AOX1	PDF	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-Dα-hGH-TT_AOX1	Negative	Zeocin™

Strains bearing expression cassettes encoding for new P_{CTA1} promoter variants and eGFP

BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+41B+48B+MM13b-eGFP-TT_AOX1	55	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+41B+48B+26B-eGFP-TT_AOX1	56	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+14B+26B+38B-eGFP-TT_AOX1	57	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+14B+26B+48B-eGFP-TT_AOX1	58	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+14B+26B+38B+41B+48B-eGFP-TT_AOX1	59	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_WT-eGFP-TT_AOX1	WT	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PDF-eGFP-TT_AOX1	PDF	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-eGFP-TT_AOX1	Negative	Zeocin™

Strains bearing expression cassettes encoding for P_{CTA1} promoter variants cloned by Katharina Ebner, and eGFP

BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_double-ind.336-360-eGFP-TT_AOX1	1	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_double-derep.271-294-eGFP-TT_AOX1	2	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+4-eGFP-TT_AOX1	3	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+MM13b-eGFP-TT_AOX1	4	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+14B-eGFP-TT_AOX1	5	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_4+MM13b-eGFP-TT_AOX1	6	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_4+14-eGFP-TT_AOX1	7	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_MM13b+14B-eGFP-TT_AOX1	8	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_26B+38B-eGFP-TT_AOX1	9	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_26B+41B-eGFP-TT_AOX1	10	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_26B+53B-eGFP-TT_AOX1	11	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_26B+48B-eGFP-TT_AOX1	12	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+41B-eGFP-TT_AOX1	13	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+53-eGFP-TT_AOX1	14	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+48B-eGFP-TT_AOX1	15	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_41B+53-eGFP-TT_AOX1	16	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_41B+48B-eGFP-TT_AOX1	17	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_48B+53-eGFP-TT_AOX1	18	Zeocin™

BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+14B+26B-eGFP-TT_AOX1	19	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_14B+26B+53-eGFP-TT_AOX1	20	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+26B+53-eGFP-TT_AOX1	21	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+14B+53-eGFP-TT_AOX1	22	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_38B+41B+48B-eGFP-TT_AOX1	23	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_MM13b+1+4-eGFP-TT_AOX1	24	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_14B+26B+41B-eGFP-TT_AOX1	25	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_14B+26B+48B-eGFP-TT_AOX1	26	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_14B+26B+38B-eGFP-TT_AOX1	27	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_flexi_ind.336-360-eGFP-TT_AOX1	28	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_flexi_derep.271-294-eGFP-TT_AOX1	29	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+26B-eGFP-TT_AOX1	30	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+38B-eGFP-TT_AOX1	31	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+41B-eGFP-TT_AOX1	32	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+53-eGFP-TT_AOX1	33	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1A+48B-eGFP-TT_AOX1	34	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1+26B-eGFP-TT_AOX1	35	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1+38B-eGFP-TT_AOX1	36	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1+41B-eGFP-TT_AOX1	37	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1+53-eGFP-TT_AOX1	38	Zeocin™

BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_1+48B-eGFP-TT_AOX1	39	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_4+26B-eGFP-TT_AOX1	40	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_4+38B-eGFP-TT_AOX1	41	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_4+41B-eGFP-TT_AOX1	42	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_4+53-eGFP-TT_AOX1	43	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_4+48B-eGFP-TT_AOX1	44	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_MM13b+26B-eGFP-TT_AOX1	45	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_MM13b+38B-eGFP-TT_AOX1	46	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_MM13b+41B-eGFP-TT_AOX1	47	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_MM13b+53-eGFP-TT_AOX1	48	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_MM13b+48B-eGFP-TT_AOX1	49	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_14B+26B-eGFP-TT_AOX1	50	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_14B+38B-eGFP-TT_AOX1	51	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_14B+41B-eGFP-TT_AOX1	52	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_14B+53-eGFP-TT_AOX1	53	Zeocin™
BSY11dKU70	TOP10'F	pPp_AOX1-pUC-ZeoR-SpeI-PCAT1_14B+48B-eGFP-TT_AOX1	54	Zeocin™

Promoter sequences

In following section all promoter sequences created within this thesis are listed in FASTA format.

>PCAT1_38B+41B+48B+MM13b
TAATCGAACTCCGAATGCGGTTCTCCTGTAACCTTAATTGTAGCATAGATCAC
TTAAATAAACTCATGGCCTGACATCTGTACACGTTTAGCAATCTTGAAGTCTT
TCTATTGTTCCGGTCGGTTACCTAATAAATTCGAATCGAGTTGCTAGTACCTG
TATCATATGAAGTAATCATCACATGCAAGTTCCATGATACCCTCTACTAATGG
AATTGAACAAAGTTTAAGCTTCTCGCACGAGACCGAATCCATACTATGCACCC
CTCAAAGTTGGGATTAGTCAGGAAAGCTGAGCAATTAACTTCCCTCGATTGG
CCTGGACTTTTCGCTTAGCCTGCCGCAATCGGTAAGGATAACCGTGCCAGCG
GGGTGATAGCCTCTGATAACCGTGAGGCCAAAATCATATATAAGCTGTAGAC
CCAGCACTTCAATTACTTGAAATTCACCATAAGATAACCGTGAGTCAAGACTT
ACAATTAAA

>PCAT1_38B+41B+48B+26B
TAATCGAACTCCGAATGCGGTTCTCCTGTAACCTTAATTGTAGCATAGATCAC
TTAAATAAACTCATGGCCTGACATCTGTACACGTTCTTATTGGTCTTTTAGCA
ATCTTGAAGTCTTTCTATTGTTCCGGTCGGCATTACCTAATAAATTCGAATCG
AGATTGCTAGTACCTGATATCATATGAAGTAATCATCACATGCAAGTTCCATG
ATACCCTCTACTAATGGAATTGAACAAAGTTTAAGCTTGATAACCGTGACCGA
ATCCATACTATGCACCCCTCAAAGTTGGGATTAGTCAGGAAAGCTGAGCAAT
TAACTTCCCTCGATTGGCCTGGACTTTTCGCTTAGCCTGCCGCAATCGGTAAG
GATAACCGTGCCAGCGGGGTGATAGCCTCTGATAACCGTGAGGCCAAAATC
ATATATAAGCTGTAGACCCAGCACTTCAATTACTTGAAATTCACCATAAGATA
ACCGTGAGTCAAGACTTACAATTAAA

>PCAT1_1A+14B+26B+38B
ATCCTTTTAGCCGAATGCGGTTCTCCTGTAACCTTAATTGTAGCATAGATCAC
TTAAATAAACTCATGGCCTGACATCTGTACACGTTCTTATTGGTCTTTTAGCA
ATCTTGAAGTCTTTCTATTGTTCCGATAACCGTGACCTAATAAATTCGAATCG
AGATTGCTAGTACCTGATATCATATGAAGTAATCATCACATGCAAGTTCCATG
ATACCCTCTACTAATGGAATTGAACAAAGTTTAAGCTTGATAACCGTGACCGA

ATCCATACTATGCACCCCTCAAAGTTGGGATTAGTCAGGAAAGCTGAGCAAT
TAACTTCCCTCGATTGGCCTGGACTTTTCGCTTAGCCTGCCGCAATCGGTAAG
GATAACCGTGCCAGCGGGGTGATAGCCTCTGTTGCTCATCAGGCCAAAATCA
TATATAAGCTGTAGACCCAGCACTTCAATTACTTGAAATTCACCATAACACTT
GCTCTAGTCAAGACTTACAATTAAA

>PCAT1_1A+14B+26B+48B
ATCCTTTTAGCCGAATGCGGTTCTCCTGTAACCTTAATTGTAGCATAGATCAC
TTAAATAAACTCATGGCCTGACATCTGTACACGTTCTTATTGGTCTTTTAGCA
ATCTTGAAGTCTTTCTATTGTTCCGATAACCGTGACCTAATAAATTCGAATCG
AGATTGCTAGTACCTGATATCATATGAAGTAATCATCACATGCAAGTTCCATG
ATACCCTCTACTAATGGAATTGAACAAAGTTTAAGCTTGATAACCGTGACCGA
ATCCATACTATGCACCCCTCAAAGTTGGGATTAGTCAGGAAAGCTGAGCAAT
TAACTTCCCTCGATTGGCCTGGACTTTTCGCTTAGCCTGCCGCAATCGGTAAG
TTTCATTATCCCAGCGGGGTGATAGCCTCTGTTGCTCATCAGGCCAAAATCAT
ATATAAGCTGTAGACCCAGCACTTCAATTACTTGAAATTCACCATAAGATAAC
CGTGAGTCAAGACTTACAATTAAA

>PCAT1_1A+14B+26B+38B+41B+48B
ATCCTTTTAGCCGAATGCGGTTCTCCTGTAACCTTAATTGTAGCATAGATCAC
TTAAATAAACTCATGGCCTGACATCTGTACACGTTCTTATTGGTCTTTTAGCA
ATCTTGAAGTCTTTCTATTGTTCCGATAACCGTGACCTAATAAATTCGAATCG
AGATTGCTAGTACCTGATATCATATGAAGTAATCATCACATGCAAGTTCCATG
ATACCCTCTACTAATGGAATTGAACAAAGTTTAAGCTTGATAACCGTGACCGA
ATCCATACTATGCACCCCTCAAAGTTGGGATTAGTCAGGAAAGCTGAGCAAT
TAACTTCCCTCGATTGGCCTGGACTTTTCGCTTAGCCTGCCGCAATCGGTAAG
GATAACCGTGCCAGCGGGGTGATAGCCTCTGATAACCGTGAGGCCAAAATC
ATATATAAGCTGTAGACCCAGCACTTCAATTACTTGAAATTCACCATAAGATA
ACCGTGAGTCAAGACTTACAATTAAA

Supplementary Data

S1: Cell density development of *P. pastoris* strains bearing P_{CTA1} promoter variants designed by Katharina Ebner (from the re-screening). Cells were cultured at 28°C in DWP containing either BMD1% or BMG1%. After 60h the BMD-cultures were induced with BMM to a methanol concentration of 0.5% and the BMG-cultures were pulsed to a glycerol concentration of 0.25%. With further induction pulses those concentrations were roughly held for another 38h. Cell density was measured at an absorption of 600nm with a plate reader (BioTek Instruments, Inc., Winooski, VT, USA) after 24, 48, 60 and 98h.

BMD w/ MeOH induction

	24h		48h		60h		98h	
P_{CTA1} variants	Abs 600nm	Stdev	Abs 600nm	Stdev	Abs 600nm	Stdev	Abs 600nm	Stdev
1	0.290	0.021	0.272	0.005	0.281	0.018	0.367	0.026
2	0.401	0.060	0.297	0.017	0.277	0.013	0.370	0.016
3	0.322	0.061	0.269	0.015	0.245	0.011	0.370	0.027
4	0.422	0.075	0.270	0.015	0.264	0.015	0.369	0.012
5	0.407	0.055	0.243	0.017	0.246	0.008	0.386	0.009
6	0.341	0.142	0.249	0.020	0.271	0.032	0.391	0.026
7	0.368	0.104	0.252	0.016	0.261	0.008	0.361	0.033
8	0.326	0.039	0.279	0.013	0.276	0.021	0.401	0.024
9	0.387	0.073	0.269	0.017	0.254	0.016	0.366	0.025
10	0.382	0.042	0.234	0.016	0.252	0.006	0.367	0.007
11	0.366	0.111	0.282	0.024	0.273	0.022	0.357	0.036
12	0.329	0.069	0.288	0.025	0.272	0.012	0.390	0.035
13	0.338	0.113	0.274	0.015	0.287	0.022	0.365	0.027
14	0.416	0.057	0.275	0.015	0.260	0.007	0.363	0.027
15	0.402	0.067	0.286	0.011	0.271	0.004	0.389	0.013
16	0.415	0.081	0.303	0.009	0.279	0.014	0.347	0.043
17	0.337	0.086	0.304	0.009	0.270	0.010	0.374	0.027
18	0.334	0.083	0.314	0.005	0.272	0.006	0.361	0.022
19	0.322	0.095	0.318	0.017	0.275	0.011	0.342	0.016
20	0.379	0.042	0.287	0.066	0.276	0.022	0.343	0.035
21	0.428	0.038	0.336	0.014	0.296	0.012	0.363	0.014
22	0.305	0.013	0.246	0.011	0.281	0.017	0.407	0.007
23	0.331	0.023	0.280	0.018	0.319	0.033	0.435	0.029
24	0.321	0.016	0.268	0.006	0.287	0.008	0.411	0.013
25	0.325	0.012	0.248	0.020	0.274	0.029	0.409	0.021
26	0.319	0.034	0.259	0.007	0.274	0.024	0.457	0.011
27	0.319	0.029	0.270	0.021	0.293	0.027	0.455	0.024
28	0.312	0.022	0.258	0.014	0.279	0.013	0.442	0.014

29	0.351	0.015	0.244	0.011	0.249	0.016	0.437	0.013
30	0.364	0.027	0.246	0.009	0.272	0.018	0.456	0.026
31	0.346	0.007	0.240	0.006	0.270	0.009	0.442	0.029
32	0.318	0.012	0.233	0.030	0.267	0.019	0.429	0.015
33	0.381	0.070	0.243	0.012	0.280	0.004	0.378	0.022
34	0.326	0.005	0.249	0.017	0.285	0.022	0.412	0.020
35	0.366	0.063	0.267	0.018	0.298	0.026	0.443	0.034
36	0.337	0.020	0.252	0.010	0.277	0.009	0.420	0.003
37	0.327	0.019	0.274	0.019	0.283	0.021	0.426	0.027
38	0.101	0.032	0.239	0.016	0.219	0.027	0.412	0.011
39	0.340	0.031	0.277	0.009	0.296	0.013	0.416	0.033
40	0.338	0.022	0.270	0.014	0.291	0.008	0.409	0.007
41	0.346	0.015	0.298	0.014	0.294	0.021	0.391	0.020
42	0.339	0.005	0.328	0.024	0.291	0.020	0.410	0.008
43	0.318	0.021	0.239	0.009	0.282	0.025	0.433	0.014
44	0.344	0.021	0.255	0.018	0.347	0.033	0.417	0.017
45	0.334	0.027	0.271	0.028	0.321	0.072	0.424	0.017
46	0.376	0.032	0.278	0.032	0.257	0.017	0.419	0.020
47	0.394	0.013	0.254	0.020	0.288	0.041	0.429	0.015
48	0.323	0.044	0.246	0.013	0.272	0.007	0.393	0.004
49	0.387	0.016	0.263	0.012	0.303	0.041	0.429	0.009
50	0.379	0.034	0.239	0.009	0.260	0.016	0.407	0.024
51	0.384	0.043	0.250	0.015	0.266	0.013	0.392	0.017
52	0.406	0.048	0.267	0.008	0.261	0.016	0.384	0.029
53	0.361	0.070	0.255	0.019	0.284	0.023	0.366	0.010
54	0.434	0.031	0.263	0.007	0.289	0.028	0.408	0.032
w/o	0.277	0.077	0.266	0.013	0.296	0.047	0.409	0.025
WT	0.293	0.088	0.262	0.024	0.274	0.023	0.409	0.032
PDF	0.176	0.051	0.233	0.094	0.240	0.081	0.332	0.119
average	0.346		0.267		0.277		0.398	

BMG w/ glycerol de-repression

	24h		48h		60h		98h	
P_{CTAI} variants	Abs 600nm	Stdev	Abs 600nm	Stdev	Abs 600nm	Stdev	Abs 600nm	Stdev
1	0.106	0.040	0.248	0.007	0.304	0.010	0.476	0.021
2	0.124	0.032	0.276	0.020	0.326	0.024	0.540	0.056
3	0.115	0.055	0.275	0.023	0.288	0.017	0.476	0.018
4	0.131	0.023	0.288	0.020	0.299	0.013	0.508	0.028
5	0.110	0.019	0.294	0.018	0.300	0.012	0.466	0.015
6	0.110	0.037	0.281	0.025	0.305	0.015	0.474	0.022
7	0.095	0.007	0.292	0.023	0.320	0.018	0.414	0.085

8	0.119	0.041	0.313	0.019	0.301	0.013	0.456	0.010
9	0.099	0.005	0.317	0.029	0.303	0.015	0.466	0.013
10	0.102	0.011	0.276	0.011	0.284	0.007	0.435	0.015
11	0.113	0.010	0.286	0.005	0.292	0.014	0.464	0.021
12	0.118	0.015	0.295	0.015	0.269	0.013	0.453	0.010
13	0.122	0.031	0.275	0.028	0.274	0.015	0.460	0.026
14	0.133	0.024	0.307	0.013	0.287	0.022	0.454	0.030
15	0.142	0.030	0.275	0.010	0.295	0.005	0.462	0.013
16	0.152	0.024	0.289	0.029	0.301	0.010	0.476	0.014
17	0.126	0.031	0.268	0.007	0.287	0.006	0.465	0.017
18	0.143	0.027	0.279	0.028	0.300	0.006	0.457	0.021
19	0.112	0.031	0.282	0.010	0.300	0.012	0.447	0.012
20	0.095	0.017	0.298	0.021	0.310	0.007	0.454	0.014
21	0.133	0.031	0.309	0.008	0.296	0.007	0.452	0.004
22	0.104	0.005	0.258	0.010	0.309	0.020	0.512	0.021
23	0.133	0.034	0.243	0.007	0.324	0.024	0.545	0.041
24	0.124	0.012	0.250	0.010	0.275	0.008	0.493	0.018
25	0.112	0.025	0.263	0.017	0.298	0.013	0.508	0.019
26	0.132	0.012	0.273	0.012	0.298	0.016	0.479	0.016
27	0.130	0.027	0.244	0.023	0.294	0.024	0.456	0.038
28	0.146	0.002	0.277	0.013	0.291	0.008	0.456	0.013
29	0.118	0.007	0.239	0.006	0.290	0.015	0.423	0.026
30	0.128	0.024	0.259	0.021	0.281	0.010	0.429	0.025
31	0.135	0.011	0.242	0.006	0.274	0.007	0.445	0.015
32	0.124	0.028	0.253	0.029	0.272	0.002	0.445	0.027
33	0.150	0.024	0.259	0.010	0.280	0.005	0.462	0.007
34	0.121	0.019	0.262	0.029	0.271	0.014	0.451	0.007
35	0.174	0.024	0.265	0.011	0.266	0.005	0.483	0.033
36	0.121	0.016	0.259	0.007	0.291	0.011	0.448	0.016
37	0.127	0.022	0.253	0.018	0.282	0.004	0.456	0.022
38	0.103	0.024	0.048	0.008	0.270	0.034	0.480	0.013
39	0.109	0.020	0.249	0.012	0.285	0.015	0.466	0.030
40	0.114	0.026	0.243	0.015	0.279	0.005	0.478	0.040
41	0.118	0.008	0.275	0.033	0.309	0.020	0.490	0.009
42	0.124	0.013	0.312	0.015	0.324	0.017	0.500	0.021
43	0.107	0.014	0.250	0.005	0.299	0.008	0.493	0.024
44	0.103	0.017	0.281	0.021	0.286	0.005	0.489	0.018
45	0.108	0.008	0.289	0.012	0.263	0.011	0.440	0.025
46	0.134	0.034	0.306	0.013	0.308	0.019	0.474	0.013
47	0.129	0.013	0.304	0.024	0.292	0.018	0.436	0.027
48	0.111	0.008	0.286	0.036	0.313	0.027	0.455	0.017
49	0.144	0.039	0.278	0.012	0.293	0.013	0.465	0.016
50	0.130	0.018	0.292	0.012	0.311	0.019	0.491	0.049

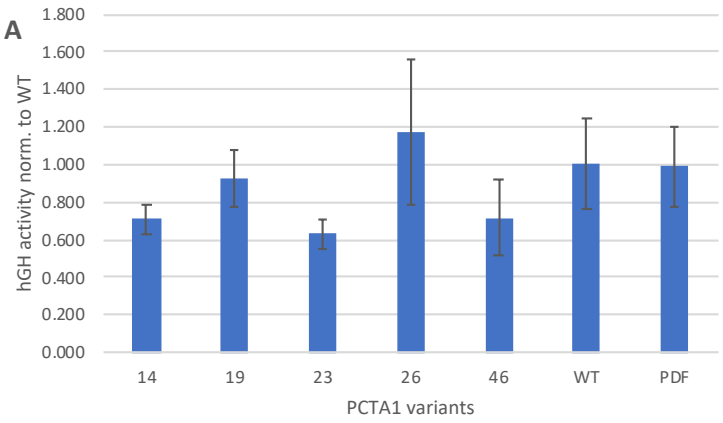
51	0.132	0.012	0.294	0.020	0.325	0.019	0.487	0.018
52	0.150	0.020	0.287	0.009	0.296	0.008	0.495	0.014
53	0.125	0.016	0.284	0.009	0.307	0.015	0.496	0.017
54	0.171	0.023	0.276	0.006	0.283	0.012	0.496	0.023
w/o	0.091	0.023	0.313	0.029	0.343	0.024	0.507	0.052
WT	0.104	0.019	0.323	0.030	0.328	0.014	0.473	0.012
PDF	0.094	0.008	0.274	0.102	0.277	0.102	0.399	0.161
average	0.122		0.273		0.295		0.469	

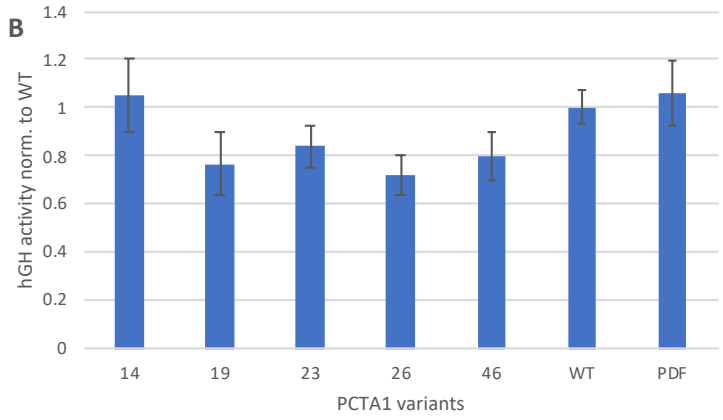
S2: eGFP fluorescence values expressed by *P. pastoris* strains employing P_{CTA1} promoter variants designed by Katharina Ebner (from the re-screening). The measured fluorescence was normalized to the respective OD_{600} value and the relative fluorescence units of the parental P_{CTA1} wildtype strain. Each variant was normalized to the WT cultured on the same DWP. Cells were cultured at 28°C in DWP containing either BMD1% or BMG1%. After 60h the BMD-cultures were induced with BMM to a methanol concentration of 0.5% and the BMG-cultures were pulsed to a glycerol concentration of 0.25%. With further induction pulses those concentrations were roughly held for another 38h. Samples were taken after 48, 60 and 98 hours.

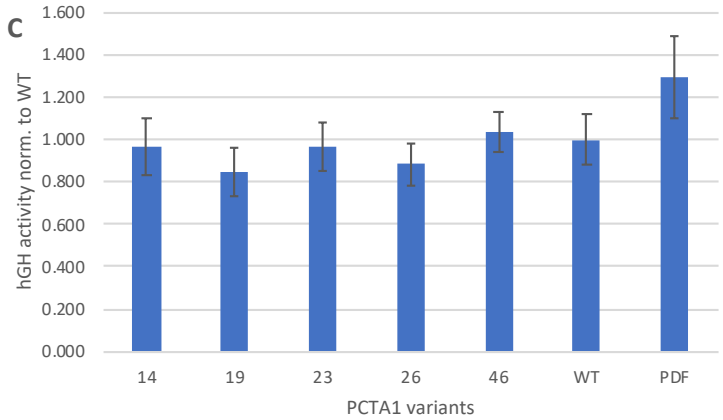
	BMD w/ MeOH induction			**BMG w/ glycerol pulsing**		
	48h	60h	98h	48h	60h	98h
P_{CTA1} variants	RFU/OD_{600}/WT			RFU/OD_{600}/WT		
1	0.77	0.72	1.36	1.67	0.97	0.59
2	1.31	1.13	1.04	2.29	1.31	0.92
3	1.51	1.37	1.08	2.59	1.29	1.12
4	1.70	1.45	1.14	3.10	1.49	1.21
5	2.51	2.34	1.37	4.15	2.05	2.09
6	1.33	1.13	1.13	2.02	1.02	0.92
7	1.90	1.61	1.24	2.79	1.56	1.58
8	2.31	1.92	1.16	3.07	1.69	1.84
9	2.26	1.93	1.65	3.08	1.78	1.70
10	1.53	1.25	1.43	2.64	1.21	1.20
11	2.54	2.00	1.28	3.42	1.91	1.79
12	1.98	1.68	1.30	3.23	1.73	1.67
13	1.57	1.21	1.67	2.25	1.29	1.06
14	1.88	1.56	1.79	3.10	1.79	1.57
15	1.71	1.65	1.80	2.77	1.42	1.47
16	1.22	1.04	1.49	2.26	1.12	0.99
17	1.16	0.91	1.45	1.83	0.91	0.94

18	1.36	1.12	1.09	1.94	1.02	0.97
19	3.05	2.57	1.59	4.55	2.69	2.72
20	3.39	2.86	1.68	4.16	2.81	2.52
21	3.09	2.55	1.49	4.97	2.83	2.30
22	1.58	1.63	1.48	1.97	2.13	2.00
23	1.44	1.41	1.86	1.99	2.05	1.55
24	0.83	0.77	1.11	1.22	1.24	1.08
25	2.19	2.21	1.47	2.15	2.82	2.80
26	2.00	1.87	1.71	2.49	2.66	2.62
27	1.05	0.97	1.08	1.36	1.38	1.39
28	0.99	0.85	1.00	1.38	1.22	1.24
29	1.51	1.44	1.16	2.00	1.80	1.96
30	1.38	1.35	1.50	1.69	1.49	1.56
31	1.14	1.09	1.23	1.40	1.21	1.43
32	1.02	0.93	1.30	1.41	1.05	1.07
33	1.25	1.20	1.20	1.43	1.22	1.41
34	1.23	1.09	1.20	1.58	1.30	1.48
35	1.74	1.30	1.21	1.84	1.59	1.86
36	1.22	0.99	1.04	1.32	1.32	1.48
37	1.11	0.90	1.25	1.09	0.95	1.02
38	0.66	0.67	1.23	0.12	0.50	1.77
39	1.21	1.05	1.40	1.36	1.18	1.31
40	0.78	0.65	1.13	0.85	0.67	0.86
41	1.24	1.11	0.99	1.68	1.58	1.53
42	1.28	1.15	1.41	1.49	1.52	1.23
43	1.35	1.43	1.17	1.99	1.44	1.22
44	1.23	1.25	1.18	2.07	1.36	1.12
45	2.28	2.14	1.28	2.95	2.05	1.91
46	1.94	1.89	1.59	3.01	2.09	1.59
47	1.44	1.36	1.31	2.36	1.43	1.23
48	2.10	2.02	1.28	2.85	1.87	1.77
49	1.21	1.24	1.22	2.25	1.17	1.03
50	2.16	2.21	1.16	3.11	2.04	1.96
51	1.74	1.80	1.57	2.87	1.83	1.68
52	1.31	1.35	1.32	2.48	1.44	1.20
53	2.04	2.04	1.22	2.77	1.95	1.72
54	1.89	1.87	1.18	2.97	1.82	1.60
NK	0.05	0.04	0.01	0.04	0.03	0.03
WT	1.00	1.00	1.00	1.00	1.00	1.00
PDF	2.72	3.44	2.23	3.31	3.67	3.54

S3: hGH expressed by *P. pastoris* strains employing P_{CTA1} promoter variants, which were cultivated in DWP containing BMD0.5%. The hGH activity was measured conducting an ELISA and the results were normalized to the parental P_{CTA1} wildtype strain. The strains were grown in DWP containing BMD0.5%. After 60h induction with BMM to a methanol concentration of 0.5% took place. With further induction pulses this concentration was held for another 38 hours. The cultivation and assay were performed three separate times (**A, B** and **C**) due to irregular results. No coherent results could be established.







S4: CalB activity and OD_{600} from P_{CTA1} strains cultured in shake flasks on either BMD0.5% or BMG0.5%. CalB activity was determined in a colorimetric assay using a plate reader (BioTek Instruments, Inc., Winooski, VT, USA) and the absorption change was converted into U/mL with the Lambert-Beer equation. Cells were cultured at 28°C in baffled shake flasks containing either BMD05% or BMG0.5%. After carbon source depletion at approx. 20h the BMD based cultures were induced with BMM1.25% to a final methanol concentration of 0.125% and the BMG-based cultures were pulse-fed with glycerol feeding discs. With further induction pulses these conditions were roughly held for a total of 65h cultivation time. Samples were drawn after 24, 42 and 65h under sterile conditions.

BMD w/ MeOH induction

	24h		42h		65h	
P_{CTA1} variants	U/mL	OD600	U/mL	OD600	U/mL	OD600
19	0.05	10.45	0.15	13.00	0.21	15.55
26	0.03	10.80	0.10	12.15	0.13	14.85
NK	0.00	11.30	0.00	13.40	0.00	15.85
WT	0.01	10.80	0.06	12.30	0.09	16.00
PDF	0.02	11.40	0.10	13.15	0.15	15.80

BMG w/ glycerol pulsing

	24h		42h		65h	
P_{CTA1} variants	U/mL	OD600	U/mL	OD600	U/mL	OD600
19	0.04	16.65	0.16	21.35	0.21	24.30
26	0.03	17.70	0.11	21.75	0.12	27.35
NK	0.00	17.80	0.00	22.35	0.00	28.70
WT	0.03	17.20	0.07	20.45	0.07	28.45
PDF	0.06	16.35	0.15	21.15	0.19	27.30

S5: Average cell density of *P. pastoris* strains bearing P_{CTA1} promoter variants with at least four mutations per variant. Cells were cultured at 28°C in DWP containing either BMD0.5% or BMG0.5%. After 60h the BMD-cultures were induced with BMM to a methanol concentration of 0.5% and the BMG-cultures were pulsed to a glycerol concentration of 0.25%. With further induction pulses those concentrations were roughly held for another 38h. Cell density was measured at an absorption of 600nm with a plate reader (BioTek Instruments, Inc., Winooski, VT, USA) after 18, 24, 48, 60 and 98h.

BMD w/ MeOH induction

	18h	24h	48h	60h	98h
	Average OD_{600} values				
MV	0.150	0.127	0.125	0.118	0.219
Std. Dev.	0.011	0.015	0.009	0.007	0.012
Std. Dev.%	7.056	11.782	7.363	6.224	5.443

BMG w/ glycerol pulsing

	18h	24h	48h	60h	98h
	Average OD_{600} values				
MV	0.195	0.171	0.176	0.163	0.266
Std. Dev.	0.015	0.014	0.013	0.009	0.016
Std. Dev.%	7.967	7.957	7.529	5.637	6.158