Expression and engineering of eukaryotic enzymes for polymer modification using *Pichia pastoris*

Dissertation

Dipl.-Ing. Andrea Mellitzer

Zur Erlangung des akademischen Grades einer Doktorin der
technischen Wissenschaften
an der Technischen Universität Graz
in Zusammenarbeit mit

Glieder, Anton, Ao.Univ.-Prof. Mag.rer.nat Dr.rer.nat
Institut für Molekulare Biotechnologie
Technische Universität Graz

2012

Deutsche Fassung:
Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008
Genehmigung des Senates am 1.12.2008

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 entnommene Stellen als solche kenntlich gemacht habe.

Graz, am …… ……..
(Unterschrift)

Englische Fassung:

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 marked all material which has been quoted either literally or by content from the used sources.

…… ……..
date (signature)

Danksagung

An dieser Stelle möchte ich meinem „Doktorvater“ Ao. Univ.-Prof. Mag. Dr.rer.nat. Anton Glieder danken, dass er mir die Möglichkeit gegeben hat an diesem interessanten Thema zu arbeiten. Des Weiteren möchte ich mich auch bei Karlheinz Flicker bedanken, der mir vor allem in der Anfangsphase des Projektes immer mit Rat und Tat zur Seite gestanden ist.

Ein großes Dankeschön geht auch an die ehemaligen bzw. aktuellen Mitarbeiter der Glieder Gruppe. Ohne eure freundschaftliche Unterstützung hätte die Arbeit bei weitem nicht so viel Spaß gemacht.

Danke auch meiner Familie und meinen Freunden, die mir immer beigestanden sind.

Abstract

Pichia pastoris is a simple but nevertheless important host for the expression of heterologous eukaryotic proteins since many posttranslational modifications such as glycosylation, disulfide bond formation, and proteolytic processing can be achieved. Additionally, the availability of strong and tightly regulated promoters makes this host very interestingly. However, the industrial application of *Pichia pastoris* is hampered in some cases since methanol is needed for protein production in the case of one of the most popular promoter (P(AOX1)). Additionally, strong promoters inducing high transcript levels can overload the cellular post-translational machinery resulting in misfolded, unprocessed or mislocalized proteins. Thus growth and productivity can be negatively affected. Therefore, alternatives are wanted. Recently, based on the P(AOX1) sequence a promoter library was generated to fine tune protein expression in *Pichia pastoris*.The main aim of the present study was to proof that protein production in *Pichia pastoris* can be still further improved by using these new promoter variants. Moreover, we wanted to demonstrate that *Pichia* serves as interesting alternative to more complex to handle fungal expression systems for the expression of lignocellulolytic enzymes. Therefore, four different lignocellulolytic enzymes were chosen as targets. A screening assay was adapted to micro-scale and its applicability as sensitive and reliable detection method for the engineering of lignocellulolytic enzymes was proven. Afterwards, the effects of synonymous codon substitution, promoter choice and gene dosage on protein production were further evaluated to get optimized conditions for protein production. We demonstrated *Pichia´*s ability to produce high levels of lignocellulolytic enzymes possessing high specific activities. Additionally, we showed that the new synthetic promoters positively influenced protein production in *Pichia pastoris*. *P. pastoris* strains were generated producing titers of 1.142 g/l *Tr*bMan, 1.2 g/l *Tl*XynA and 18 g/l *Tr*CBH2.

Zusammenfassung

Pichia pastoris ist ein einfacher aber dennoch wichtiger Wirt für die heterologe Expression von eukaryotischen Proteinen, weil viele posttranslationelle Veränderungen wie Glykosylierung, Disulfidbrücken Formierung und proteolytische Prozessierung erreicht werden können. Zusätzlich macht das Vorhandensein von starken und gut regulierten Promotoren diesen Wirt sehr interessant. Jedoch ist die industrielle Verwendbarkeit von *Pichia pastoris* in manchen Fällen eingeschränkt, weil Methanol für die Protein Produktion mit einem der beliebtesten Promotoren (P(AOX1)) benötigt wird. Des Weiteren können starke Promotoren hohe Transkript Mengen induzieren, was zu einer Überlastung der zellulären post-translationalen Maschinerie führen kann und als Resultat falsch gefaltete, nicht prozessierte oder falsch lokalisierte Proteine zur Folge hat. Dadurch kann es auch zu einer Beeinträchtigung des Wachstums und der Produktivität kommen. Deswegen sind Alternativen gefragt. Kürzlich wurde, basierend auf der P(AOX1) Sequenz, eine Promoter Bibliothek generiert, die eine Feinabstimmung der Protein Expression in *Pichia pastoris* gewährleistet. Das Hauptziel der vorliegenden Studie war zu beweisen, dass die Protein Expression in *Pichia pastoris* verbessert werden kann, indem man diese neuen Promoter Varianten einsetzt. Des Weiteren wollten wir demonstrieren, dass *Pichia* für die Expression von Lignocellulose modifizierenden Enzymen eine interessante Alternative zu schwierigeren pilzlichen Expressionssystemen bietet. Deswegen wurden vier verschiedene Lignocellulose modifizierende Enzyme als Ziele ausgewählt. Ein Screeningverfahren wurde an Mikrogröße angepasst und hat ihre Anwendbarkeit als sensitive und verlässliche Nachweismethode für das Engineering von lignozellulolytischen Enzymen gezeigt. Anschließend wurden die Effekte von synonymem Codon Austausch, Promoter Wahl und Gen Dosierung auf Protein Produktion genauer evaluiert, um optimierte Bedingungen für die Protein Produktion zu bekommen. Es wurde demonstriert, dass *Pichia* in der Lage ist große Mengen an Lignocellulose modifizierenden Enzymen zu produzieren, die hohe spezifische Aktivitäten aufweisen. Außerdem wurde gezeigt, dass die neuen synthetischen Promotoren einen positiven Einfluss auf die Protein Produktion in *Pichia pastoris* haben. Es konnten *P. pastoris* Stämme hergestellt werden, die einen Titer von 1.142 g/l *Tr*bMan, 1.2 g/l *Tl*XynA und 18 g/l *Tr*CBH2 an Protein produziert haben.

Table of Content

Introduction

1. Bioethanol

The reservoir of fossil fuels is steadily decreasing and together with environmental issues, in terms of greenhouse gas emissions and unstable oil markets, this leads to the demand for alternative fuels. Plant biomass as the most abundant renewable resource on earth which also comprises energy- rich carbohydrates seems to be a perfect candidate for sustainable energy exploitation. Carbohydrates can be released from plant biomass by enzymatic hydrolysis and subsequently can be converted to bioethanol and many other value-added products. Bioethanol is a very promising fuel due to several desirable characteristics. These characteristics include good combustion efficiency, a reduction of particulate and NO_x emission from combustion, a high octane value, its CO_2 neutral production as well as its compatibility with the existing transport infrastructure [1-6]. Currently, the most common feedstocks for bioethanol production are either corn grain (starch) or sugar cane (sucrose) [7]. However, within the last years more and more concerns arose regarding the application of those C-sources. The so called “food versus fuel” concern is the most prominent one but there are also other socially, economically, environmentally and technically ones [1, 6]. Therefore, alternative biomass sources are wanted.

1.1. Lignocellulosic biomass composition

The so called second generation bioethanol is produced from lignocellulosic biomass and exhibits energetic, economic and environmental advantages in comparison to bioethanol from starch or sugar [6]. However, the recalcitrant nature of lignocellulosic biomass impedes its profitable utilization at industrial scale. Corn grain consists of mainly one carbohydrate (starch) thus comprises a far less complex structure than lignocellulosic biomass [8]. Lignocellulosic biomass is obtained from higher plants, softwood or hardwood and consists of three main components: cellulose (40–55%), hemicellulose (24–40%) and lignin (18–35%) [9].

1.1.1. Cellulose

In principle, starch from corn grain and cellulose consist of the same molecule, glucose. In both cases glucose is linked together in chains via glycosidic bonds. The main difference between starch and cellulose is the linkage between the individual glucose units itself. In starch α-1,4 and α-1,6 glycosidic linkages (amylose and amylopectin) occur while in cellulose glucose is linked via β-1,4 glycosidic bonds [6], see Figure 1. Due to the β-1,4 glycosidic bonds, long straight chains of glucose monomers can be arranged. The average chain length of cellulose is between 7000-15000 sugar units. These chains can be connected through the evenly distributed hydroxides on both sides of the monomers. By doing so several layers of cellulose can be created. This characteristic structure makes cellulose highly crystalline and compact and thus, very resistant for water or enzymes to penetrate the structure. Therefore, cellulose is a major structural compound of cell walls that is responsible for mechanical strength and chemical stability of plants [8-11].

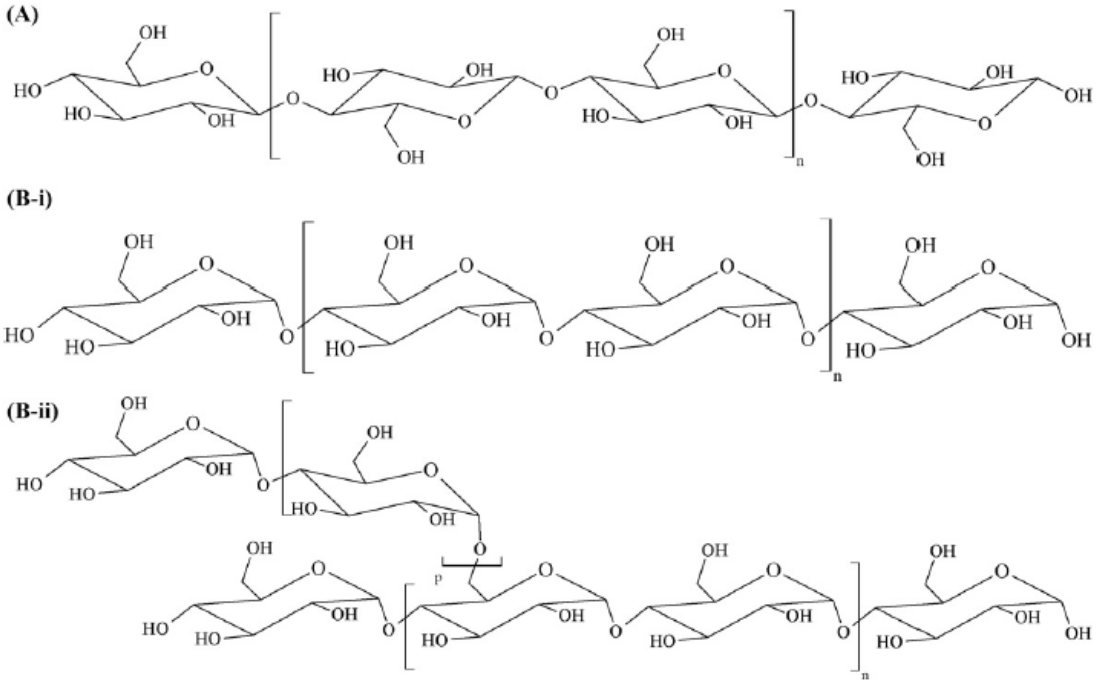


Figure 1: The structures of (A) cellulose, (B) starch consisting of (i) amylose and (ii) amylopectin as proposed by Nimtz, taken from [8]

1.1.2. Hemicellulose

Hemicellulose also exists in the plant cell wall. They are short, highly branched heteropolymers of different five-carbon (pentoses) and six-carbon (hexoses) sugars and a number of sugar acids. Their composition and structure is depending on the source and extraction method [11]. In softwood the dominant sugar of hemicellulose is mannose whereas in hardwood and agriculture residues the dominant sugar is xylose. In principle, the backbones of hemicellulose are β-1,4 linked polysaccharides with an equatorial configuration. From these backbones various short branches of mannose, arabinose, galactose, glucuronicacid, etc can be found. The average chain length of hemicellulose is between 500-3000 sugar units [11]. Depending on the side chain different chemical characteristics of hemicelluloses can be obtained. For example, there are hemicelluloses known which are soluble in water and some which are not. This effect can be explained by a higher percentage of acids in the side chains which lead to a slight charge of the hemicellulose. Contrary to cellulose hemicellulose is not crystalline. This is mainly due to the highly branched structure, and the presence of acetyl groups connected to the polymer chains. Besides this, hemicellulose can be also easier hydrolyzed because of its branched, amorphous nature. Moreover, in lignocellulosic materials, hemicellulose is the most thermal-chemically sensitive fraction. Hemicelluloses are a bonding agent between cellulose and lignin and therefore, limit the stretchiness of the cell wall. One biological function of hemicellulose is the controlling of cell enlargements [8-11].

1.1.3. Lignin

Lignin is a complex, amorphous, aromatic polymer of phenyl propanoid units with molecular masses in excess of 10,000 Da [11]. The exact chemical structure is difficult to ascertain since it is fragmented during extraction. Moreover, the molecule consists of various types of substructures that appear to repeat in a random manner and varies with plant and species. Basically, lignin contains of three aromatic alcohols coniferyl alcohol, sinapyl and p-coumaryl that are methoxylated to various degrees. Resulting units from the alcohols are incorporated into lignin and are further on referred to as guaiacyl (G), syringyl (S), and p-hydroxyphenyl (H) units [11]. Typically, the main components from softwood (gymnosperms) lignin are G-units and to a minor extent H-units. Unlike hardwood (angiosperm) dicots, here lignins are

composed of G- and S-units. In softwood compression wood and also in higher grasses the amount of H-units is slightly increased. The lignin content of softwoods is generally higher than that of hardwoods. Lignin composition can not only differ among taxa and species but can also differ among individual cell types and even at the level of individual cell wall layers. Besides, a variety of non hydrolysable C-C and C-O-C bonds this complex matrix also comprises a variety of functional groups [11]. These are methoxyl, phenolic hydroxyl and few terminal aldehyde groups and are supposed to have an impact on its reactivity toward the delignification process. Only a small percentage of the phenolic hydroxyl groups are free since most are occupied in bonds to adjoining phenylpropane linkages. This chemical constitution makes lignin hydrophobic and prohibits its penetration by water and enzymes. Because of its hydrophobic nature lignin also plays a crucial role in conduction water in plant stems. Additionally, lignin is has a stabilization effect of the cell walls. It is covalently linked to hemicellulose as well as to cellulose and forms a protective layer for the plant cell walls. Since lignin is the most recalcitrant component of the lignocellulosic biomass this component is the biggest obstacle for its efficient utilization [8-11].

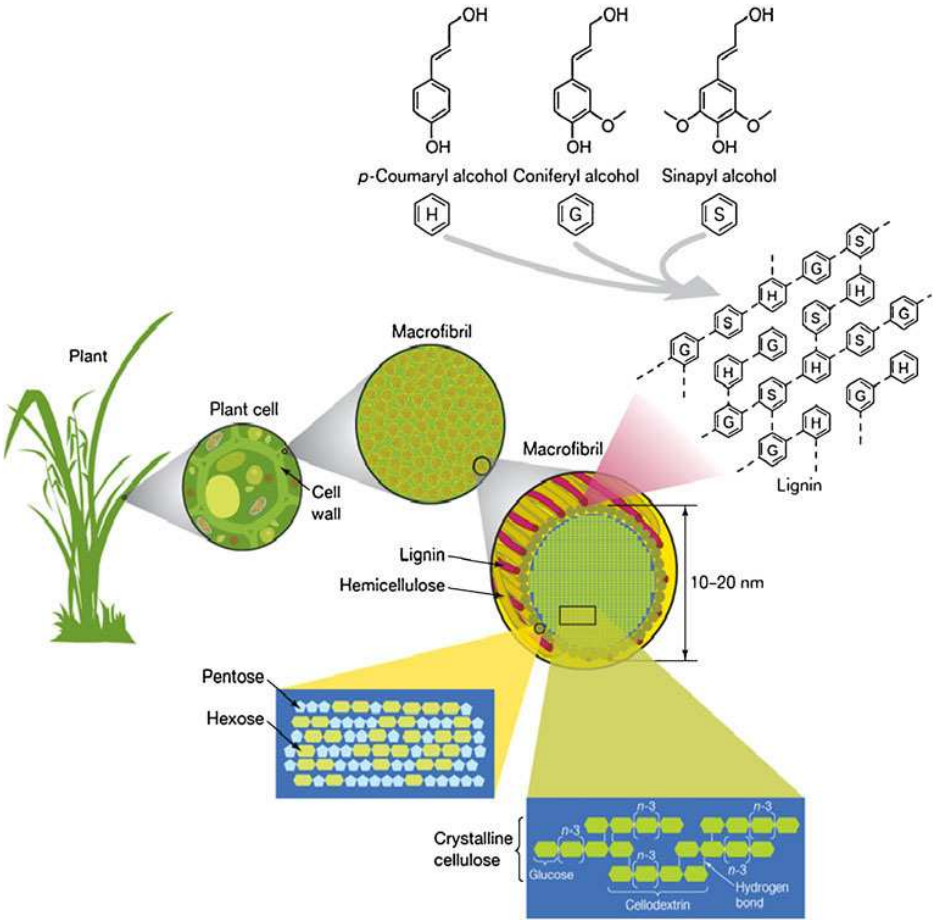


Figure 2: Molecular component of plant cell wall structure, taken from [10].

1.2. Pretreatment of lignocellulosic biomass

To facilitate lignocellulosic biomass conversion various pretreatment technologies are available. The main purpose of pretreatments is to alter the macro- and microscopic size and structure of the biomass, as well as its submicroscopic chemical composition to get an easier excess to the cellulose fraction. Each pretreatment method has specific impacts on the cellulose, hemicellulose and lignin parts. Therefore, different pretreatment methods and conditions should be chosen depending on the following application for hydrolysis and fermentation. Furthermore, pretreatments have a severe impact on an economical efficient production of bioethanol from lignocellulosic biomass, in terms of investment costs, energy consumption, time frame, etc. [9, 12-14].

Hemicellulose removal leads to an increase of the mean pore size of the substrate and thereby increases the accessibility and the probability of the cellulose to become hydrolyzed. Nevertheless, higher total fermentable sugar production can be obtained if hemicellulosic sugars can be recovered in the pretreated solids. Since hemicellulose can account for 37-48% [15] of a plant primary cell wall it represents an abundant and valuable source of pentose sugars. Besides its recalcitrant nature, lignin appears to reduce cellulose hydrolysis by non-productively binding cellulolytic enzymes and therefore has to be removed. However, lignin is an interesting starting source for the production of various chemicals, combined heat and power or other purposes thus makes recovery worth it [9, 12-15]. Since pretreatments are not on the scope of this thesis they will be just briefly summarized in the following subchapters (see also Table 1).

1.2.1. Physical pretreatment methods

These techniques utilize mechanical forces in order to reduce cellulose crystalline and/or increase the surface area of the material. The accessibility of the material is improved by a combination of processes such as chipping, grinding and milling. These treatments are very helpful tools to increase the yield of bioethanol from lignocellulosic biomass. However, they are usually found quite unattractive due to their high energy and capital costs [12-15].

1.2.2. Chemical pretreatment

These techniques are based on the disruption of the biomass structure through chemical reactions. Lignin and hemicelluloses are going to be removed by their interactions with different chemicals, such as ozone, diluted acids (H_2SO_4 and HCl), alkali, peroxides, and organic solvents. Unfortunately, most of these techniques are not cost competitive at large scale so far [12-15].

1.2.3. Physico-chemical pretreatment

These techniques combine mechanical as well as chemical actions. The most prominent pretreatment method in this group is steam explosion. It is a hydrothermal pretreatment in which the biomass is exposed to hot pressurized steam for a certain time ranging from seconds to several minutes, and then suddenly depressurized. On account of this the fibre structure breaks down, and the lignocellulose components get separated and partially solubilized. However, this method is also quite energy intensive [12-15].

1.2.4. Biological pretreatment

These techniques employ microorganisms such as brown, white and soft rot fungi to produce lignin degrading enzymes like lignin peroxidases, Mn-dependent peroxidises and laccases (monophenol oxidases). Advantageously, the energy requirement for this process is quite low and it also needs mild environmental conditions. Nevertheless, slow kinetics prohibits the application at industrial scale so far [12-15].

1.2.5. Example processes for separation of cellulose hemicellulose and lignin

1.2.5.1. Biofine process

Opposite to other dilute-acid lignocellulosic-fractionating technologies no free monomeric sugars are generated during the Biofine Process. Two distinct reactors are employed to break down the chemical structure into three main compounds furfural, levulinic acid and lignin. A high temperature tubular reactor into which steam is injected is employed followed by a second stage mixed reactor. The produced levulinic acid and furfural are concentrated,

purified and can be converted to a variety of other chemicals, including substitutions for petrochemicals (e.g. methyl tetrahydrofuran) [16].

1.2.5.2. CIMV process

The feedstock is treated at atmospheric pressure with a mixture of acetic acid and formic acid to dissolve the lignin and also to hydrolyze the hemicellulose into oligo- and monosaccharides harboring a high xylose content. After filtering of the raw pulp the solvent is removed and the residue is bleached by hydrogen peroxide. Organic acids are recycled by the evaporation from the organic solution. The remaining syrup is then treated with water to easily separate lignin from it by precipitation. Fortunately, the obtained lignin shows quite unusually linear structures and can be used as a reactant for chemical production of new polymers and composites synthesis. The raw syrup can be directly employed for numerous industrial applications without further purification steps needed [17].

Table 1: Pretreatment processes of lignocellulosic materials, adapted from [18]

Method	Processes	Mechanism changes on biomass
Physical pretreatments	Milling Hydrothermal High pressure steaming Extrusion Expansion Pyrolysis Irradiation	Increases in accessible surfaces area and size of pores Decrease of the cellulose crystallinity and its degrees of polymerization Partial hydrolysis of hemicelluloses Partial de-polymerization of lignin
Physicochemical & chemical pretreatments	*Explosion:* -Steam, ammonia fiber (AFEX), CO_2, SO_2 explosion *Alkali:* -Sodium hydroxide -Ammonia -Ammonium Sulfite *Gas:* -Chlorine and nitrogen dioxide *Acid:* -Sulfuric, hydrochloric, phosphoric acid -Sulfur dioxide *Oxidizing agents:* -Hydrogen peroxide -Wet oxidation -Ozone *Cellulose solvents:* -Cadoxen -CMCS *Solvent extraction of lignin:* -Ethanol-water, benzene-water, ethylene-glycol and butanol-water extraction -Swelling agents	Delignification Decrease of cellulose crystallinity and its degrees of polymerization Partial or complete hydrolysis of hemicelluloses
Biological pretreatments	Actinomycetes Fungi	Delignification Reduction in degree of polymerization of hemicellulose and cellulose

1.3. Hydrolysis of lignocellulosic biomass

Hydrolysis of cellulose and hemicellulose can be carried out either chemically by e.g. diluted sulfuric acid or enzymatically. Both methods exhibit several advantages and disadvantages. Acid hydrolysis requires high temperature and low pH which results in corrosive conditions, whereas, for enzymatic hydrolysis conversion can be performed under mild conditions. For enzymatic hydrolysis it is possible to obtain yields of close to 100% cellulose hydrolysis, while, it is rather difficult to achieve such high yields under acid conditions. Although in both processes inhibitory compounds can be found, this is far less severe for enzymatic hydrolysis [11-15, 18]. The biggest obstacles for an efficient utilization of enzymes for lignocellulosic biomass hydrolysis are the time which is needed for the hydrolysis and the costs for the enzymes. Therefore, to compete with grain derived ethanol, the enzymes used for lignocellulosic biomass hydrolysis must become more efficient and less expensive.

Enzymatic hydrolysis is carried out to release the monosaccharides from cellulose and hemicellulose and is mediated by glycoside hydrolases. Glycoside hydrolyses (GHs) are extremely common enzymes predominantly produced by microorganisms such as bacteria, archaea and fungi to utilize plant biomass. GHs harbor very broad as well as string substrate specificities and are classified in EC 3.2.1 [11-15, 18]. Currently it consists of 130 families that are arranged into 14 different clans (CAZy; http://www.cazy.org) based on their amino acid sequence and three dimensional folds. Along with the catalytic GH domain, enzymes involved in the hydrolysis of lignocellulosic biomass usually comprise also a carbohydrate-binding module (CBM). CBMs are small domains that have a specific affinity towards particular carbohydrate linkages and consequently are in charge to target the catalytic unit to its intended substrate [15].

1.3.1. Cellulase

Enzymes involved in the hydrolysis of cellulose are broadly referred to as cellulases. Indeed, three different kinds of cellulolytic enzyme activities have to be performed for its complete and efficient conversion. This includes endoglucanases (1,4-β-D-glucan glucohydrolase; EC 3.2.1.4) and exoglucanases (1,4-β-D-glucan cellobiohydrolase; EC 3.2.1.91) that operate at the liquid: solid interface and β-glucosidases (β-D-glucoside glucohydrolase; EC3.2.1.21) that operate on the soluble degradation products of cellulose [15]. Endoglucanases randomly break

down internal glycosidic linkages primarily in regions of low crystallinity. Hence, a rapid decrease in polymer length and a continuous increase in the number of released reducing ends can be observed. Exoglucanases attach to a free end of the cellulose chain and then creep down the strands and concomitantly release cellobiose units as they progress. This can be performed either from the reducing or non-reducing ends of the cellulose chains but not within them. Therefore, the amount of reducing ends is rapidly increasing but this process has almost no effect on the polymer length itself [15]. Cellobiose, the repeating structural element of cellulose, is a disaccharide that consist of two β-1,4-linked glucose molecules. This disaccharide can be further converted by β-glucosidases to release glucose. β-glucosidase activity is mandatory for an efficient and complete hydrolysis of cellulose due to a possible end-product inhibition of endo- and exoglucanases by cellobiose [11-15, 18]. Enzymes involved in the hydrolysis of cellulose are visualized in Figure 3

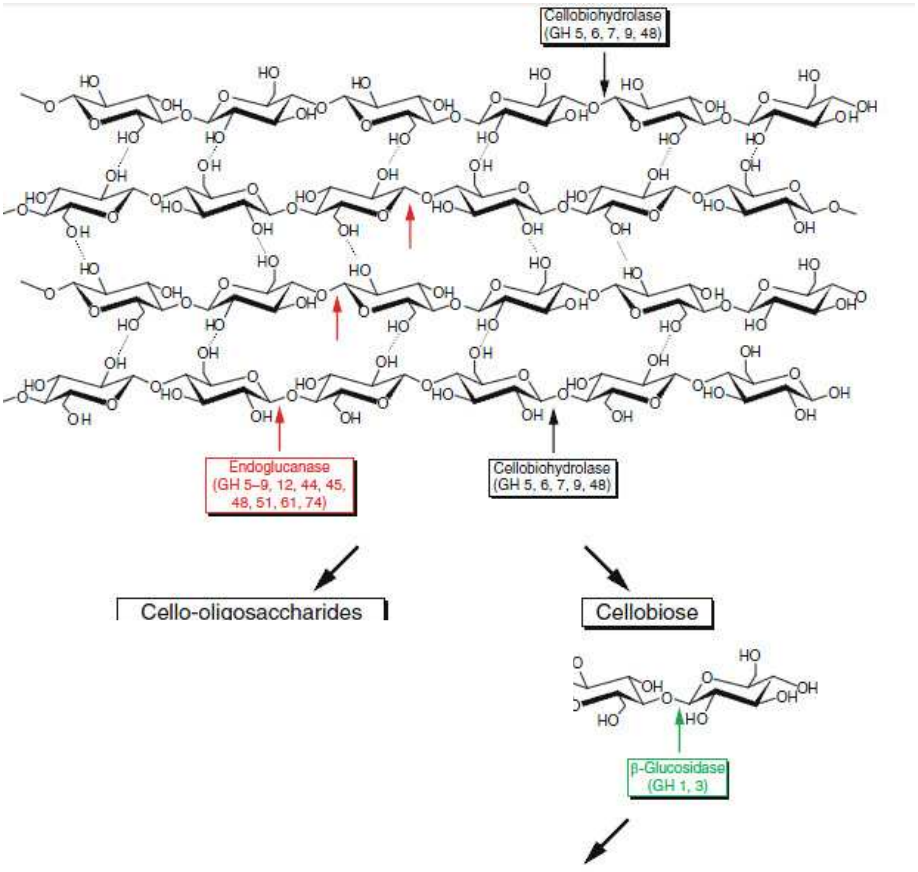


Figure 3: Enzymatic activities associated with cellulose deconstruction, adapted from [15].

1.3.2. Hemicellulase

Enzymes involved in the hydrolysis of hemicellulose are broadly referred to as hemicellulase. Since hemicellulose is far more complex than cellulose, a mixture of several enzymes has to be provided for its modification [11-15, 18]. Following enzyme classes are employed endo-β-1,4-xylanases (EC 3.2.1.8), xylan 1,4-β-xylosidases (EC 3.2.1.37), α-L-arabinofuranosidases (EC 3.2.1.55), α-glucuronidases (EC 3.2.1.139), acetylxylan esterases (EC 3.1.1.72]), feruloyl esterases (EC 3.1.1.73), mannan endo- 1,4-β-mannanases (EC 3.2.1.78), β-1,4-mannosidases (EC 3.2.1.25), and arabinan endo-1,5-α-L-arabinosidases (EC 3.2.1.99). In principle, the linear β-1,4 linked polysaccharide backbone, mainly xylan, is attacked by endo-β-1,4-xylanases and β-D-xylosidases whereas branches are attacked by enzymes such as α-L-arabinofuranosidases, α-glucuronidases, and esterases (see Figure 4). As a result six-carbon sugars as well as five-carbon sugars can get released [15].

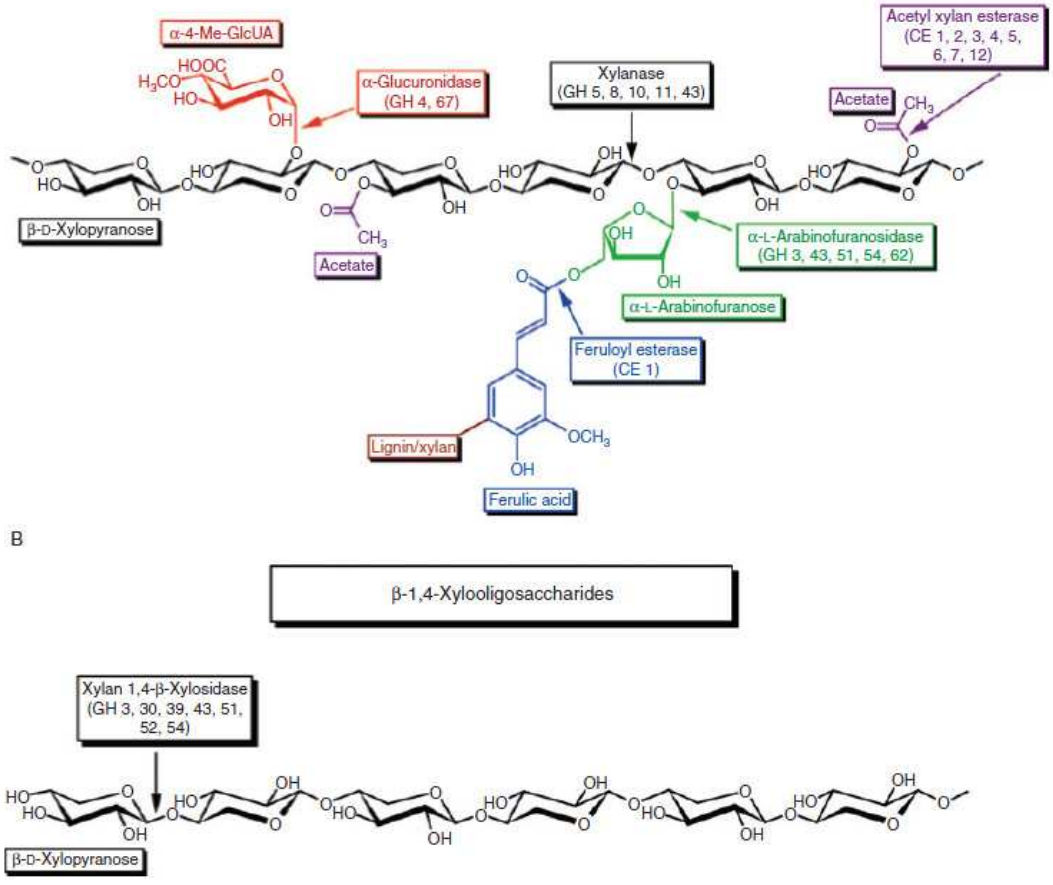


Figure 4: Enzymatic activities associated with xylan deconstruction. (A) Endoxylanases cleave the backbone of xylan chains to release shorter xylo-oligosaccharides which are further debranched by accessory enzymes. (B) b-Xylosidases release xylose monomers from the nonreducing ends of debranched xylo-oligosaccharides, adapted from [15].

1.4.Fermentation

During the fermentation process microorganism can convert the released monosaccharides to bioethanol and many other value-added products. For this purpose different methods are available that are influenced by several factors, such as reaction conditions, what kind of microorganism, contamination risk, performed pretreatments, construction of the reactors, etc. [19-21] Since fermentation is not on the scope of this thesis it will be just briefly summarized in the following subchapters.

1.4.1. Separate Enzymatic Hydrolysis and Fermentation (SHF)

Separate units are employed for the hydrolysis of pretreated lignocelluloses to glucose and for the subsequent fermentation to ethanol. The fermentation of hexoses and pentoses has to be performed by different microorganism. Hence, different vessels have to be employed. Advantageously, each step can be performed under its own optimal conditions. Usually, the optimum for the enzymes in the hydrolysis step is somewhere between 45 and 50°C, depending on the cellulase- producing microorganism. Unlike, the optimum for most of the ethanol- producing microorganisms, here favored temperatures are between 30 and 37°C [19-21]. The main drawback for SHF is the high inhibition possibility of cellulase activity by the released sugars, cellobiose and glucose. Moreover, due to the dilute sugar solution, microbial contamination risk is also a big problem in SHF. Since hydrolysis takes quite a long time, e.g. one to four days the risk is even further increased [19-21].

1.4.2. Simultaneous Saccharification and Fermentation (SSF)

This process, referred to as SSF, combines the enzymatic hydrolysis of pretreated lignocelluloses and the fermentation in one step, by employing lignocellulose modifying enzymes and microorganism for the subsequent fermentation in the same vessel [19-21]. Since hydrolysis is usually the rate limiting step in SSF, produced glucose can be consumed immediately by the fermenting microorganism present in the culture. Hence, the inhibition effect of cellobiose and glucose to the enzymes is minimized by keeping a low concentration of the sugars in the media. Moreover, there is less risk for a contamination in SSF processes compared to SHF processes since the presence of ethanol reduces the chance of contamination

[19-21]. Although the fermentation of pentoses has still to be performed in a different vessel, lower capital costs are obtained for SSF due to a reduced number of vessels required for this process compared to SHF. Nevertheless, a main drawback in this process is the different optimum temperature for the hydrolyzing enzymes and the fermenting microorganism. As already mentioned above, favored temperatures for enzymatic hydrolysis are between 45 and 50°C, whereas, preferred temperatures for the microorganism are between 30 and 37° [19-21]. The strategy for SSF processes is to have the optimum conditions for both processes as close as possible by e.g. employing thermo-tolerant bacteria and yeasts.

1.4.3. Consolidated Bioprocessing (CBP)

Up to now, all of the considered processes employ a separate enzyme production unit or the enzymes have to be provided externally. In consolidated Bioprocessing (CBP) just one vessel is required in which ethanol along with all of the essential enzymes is produced by a single microorganism´s community. Mono- or co-cultures of microorganism are employed with the ability to ferment cellulose (and hemicellulose) to ethanol. Therefore, no additional costs arise regarding the purchase of enzymes or its production. So far, there is no (native) microorganism available which suits this application satisfyingly. Hence, strains are genetically modified to increase bioethanol yield. Therefore, two strategies are common. Either excellent ethanol producers are genetically engineered to become also efficient cellulase producers or vice versa, excellent cellulase producers are engineered in such way that they become also efficient ethanol producers [19-21].

2. Expression of lignocellulolytic enzymes

2.1. Homologous protein expression

Lignocellulolytic enzymes can be produced by various fungi species such as ascomycetes (e.g. *Trichoderma reesei* and *Aspergillus niger*), basidiomycetes (e.g. *Phanerochaete chrysosporium* and *Fomitopsis palustris*) and a few anaerobic species (e.g. *Orpinomyces sp.)*. Additionally, there are also bacteria known which are capable to produce lignocellulolytic enzymes, e.g. *Clostridium thermocellum.* Fungi mainly produce large amounts of extracellular lignocellulolytic enzymes. Whereas anaerobic bacteria and a few anaerobic fungal strains primarily produce cellulolytic enzymes in a complex called cellulosome, which is associated with the cell wall. Besides hydrolytic enzymes the cellulosome also consist of non-catalytic scaffolding proteins that are responsible for its structure [22-25].

Although fungi can produce large amounts of particular extracellular lignocellulolytic enzymes, efficient biomass conversion usually cannot be performed due to missing enzymes. For example *T. reesei*, that excellently secretes cellobiohydrolase 1 and 2, and several endoglucanases, only produces small amounts of β-glucosidases. Since β-glucosidases is mandatory for the degradation of cellobiose to glucose, cellobiose is accumulating and thus inhibits further cellulose hydrolysis. Moreover, fungi systems offer also other drawbacks. First of all, compared to other systems, fungi are quite complex and therefore more difficult to handle. Second, strain or protein engineering is hampered due to insufficient tools or time consuming screening methods. Last but not least, homologous proteins can be expressed in large quantities however heterologously expression of interesting proteins might be not that good. Also bacteria cannot be employed for an efficient lignocellulosic biomass conversion at industrial scale so far, due to unique substrate requirements, low hydrolytic activities and/or low protein concentrations. Therefore, alternative expression hosts are wanted [22-25].

2.2. Heterologous protein expression in *Pichia pastoris*

The Yeast *Pichia pastoris* is a relatively simple, single-celled microorganism that is mostly faster, cheaper and easier to handle compared to more complex expression systems like mammalian and insect cell cultures. Nevertheless, many posttranslational modifications such as glycosylation, disulfide bond formation and proteolytic processing can be still performed because it is a eukaryotic [26]. Typically, lignocellulolytic enzymes require post-translational modifications for correct folding, stability and activity; therefore, *P. pastoris* is an interesting alternative to other more challenging fungal secretory expression systems. Moreover, on account of strong promoters high levels of heterologous proteins have been successfully expressed in *Pichia*. One of the most popular promoters is the strong and tightly regulated methanol inducible promoter AOX1. However, the application of P(AOX1) comprises also disadvantages. On the one hand, not every protein requires strong promoters. There is also the possibility that a high transcript level of the desired gene caused by strong promoters leads to an overload of the cellular post-translational machinery. Consequently, misfolded, unprocessed, or mislocalized proteins are produced [27]. On the other hand, the application of methanol at large scale is not favorable since its dangerous character, in terms of e.g. explosion hazard and toxicity. Therefore, a library of promoter variants based on the wild-type P(AOX1) was previously generated [28]. This library facilitates heterologous protein expression because the promoters span a broad range of different activities and can be regulated by available carbon sources, kind of similar to the P(AOX1). Advantageously, this library also contains promoters that need no methanol for induction, since they are active under derepressed conditions.

At first, to prove *Pichia´*s capability for the production of high level and active lignocellulolytic enzymes a high- throughput screening assay had to be established. Therefore, a reducing sugar assay was adapted to micro-scale and compared to other commonly available assays. We have demonstrated that our assay is a sensitive and reliable assay which is suitable for the activity based detection of different kinds of (hemi-) cellulases. These results are summarized in **Chapter 1** of the thesis. In **Chapter 2** we demonstrated the importance of gene optimization and strain characterization for successfully improving (secreted) protein levels. Codon-optimized genes together with particular promoters and verified numbers of integrated expression cassettes allowed us to develop *P. pastoris* strains producing high levels of lignocellulolytic enzymes. Moreover, our obtained specific enzymatic activities were

compared to the same enzyme expressed in different hosts. We concluded that *Pichia* seems to be a good choice for the heterologous expression of individual lignocellulolytic enzymes.
In Chapter 2 we already got an idea that differently optimized coding sequences can have a huge effect on the amount of produced protein. To further evaluate the effect of synonymous codon substitution we decided to test 57 differently optimized variants of the lignocellulolytic enzyme *Tr*CBH2. Additionally, we also wanted to test the influence of altering expression conditions. Therefore, three different promoters were employed. The constitutive wild type promoter P(GAP) was tested in presence of glucose, the derepressed synthetic promoter P(DeS) in presence of sorbitol and the inducible synthetic promoter P(En) in presence of methanol. Due to the sensitive and reliable high- throughput screening assay 171 putative single copy landscapes were generated and compared. Thus, a model for the prediction of heterologous protein expression in *Pichia pastoris* was established. Fed-batch bioreactor cultivations were performed to verify some of these results also at large scale. By combining synthetic genes and synthetic promoters we were able to generate *Pichia* strains capable of producing 20 g/L of secreted *Tr*CBH2 even without methanol but still quite tight regulated. These results are summarized in **Chapter 3**.

Reference list

1. Von Blottnitz, H. and M.A. Curran, *A review of assessments conducted on bio-ethanol as a transportation fuel from a net energy, greenhouse gas, and environmental life cycle perspective*, in *Journal of cleaner production*. 2007. p. 607-619.
2. Zinoviev, S., et al., *Next-generation biofuels: Survey of emerging technologies and sustainability issues.* ChemSusChem, 2010 **3**(10): p. 1106-33.
3. Fang, X., et al., *Status and prospect of lignocellulosic bioethanol production in China.* Bioresour Technol, 2010. **101**(13): p. 4814-9.
4. Kim, J.S., et al., *Production of bioethanol from lignocellulose: Status and perspectives in Korea.* Bioresour Technol, 2010. **101**(13): p. 4801-5.
5. Mabee, W.E. and J.N. Saddler, *Bioethanol from lignocellulosics: Status and perspectives in Canada.* Bioresour Technol, 2010. **101**(13): p. 4806-13.
6. Demain, A.L., *Biosolutions to the energy problem.* J Ind Microbiol Biotechnol, 2009. **36**(3): p. 319-32.
7. Gray, K.A., L. Zhao, and M. Emptage, *Bioethanol.* Curr Opin Chem Biol, 2006. **10**(2): p. 141-6.
8. Collinson, S.R. and W. Thielemans, *The catalytic oxidation of biomass to new materials focusing on starch, cellulose and lignin.* Coordination Chemistry Reviews, 2010. **254**(15): p. 1854-1870.
9. Harmsen, P., et al., *Literature review of physical and chemical pretreatment processes for lignocellulosic biomass*. 2010: Wageningen UR, Food & Biobased Research.
10. Chandel, A.K. and O.V. Singh, *Weedy lignocellulosic feedstock and microbial metabolic engineering: advancing the generation of 'Biofuel'.* Appl Microbiol Biotechnol, 2011. **89**(5): p. 1289-303.
11. Laine, C., *Structures of hemicelluloses and pectins in wood and pulp*. 2005, Keskuslaboratorio.
12. Alvira, P., et al., *Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review.* Bioresour Technol, 2010. **101**(13): p. 4851-61.
13. Talebnia, F., D. Karakashev, and I. Angelidaki, *Production of bioethanol from wheat straw: An overview on pretreatment, hydrolysis and fermentation.* Bioresour Technol, 2010. **101**(13): p. 4744-53.
14. Zhao, X., K. Cheng, and D. Liu, *Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis.* Appl Microbiol Biotechnol, 2009. **82**(5): p. 815-27.
15. Yeoman, C.J., et al., *Thermostable enzymes as biocatalysts in the biofuel industry.* Advances in applied microbiology, 2010. **70**: p. 1-55.

16. Hayes, D.J., et al., *Chapter 7: The Biofine Process – Production of Levulinic Acid, Furfural, and Formic Acid from Lignocellulosic Feedstocks.* Biorefineries-Industrial Processes and Products: , 2006: p. 139-164.
17. Delmas, M., *Vegetal Refining and Agrochemistry.* Chemical engineering & technology, 2008. **31**(5): p. 792-797.
18. Taherzadeh, M.J. and K. Karimi, *Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: a review.* Int J Mol Sci, 2008. **9**(9): p. 1621-51.
19. Hasunuma, T. and A. Kondo, *Development of yeast cell factories for consolidated bioprocessing of lignocellulose to bioethanol through cell surface engineering.* Biotechnol Adv, 2012.
20. Lynd, L.R., et al., *Consolidated bioprocessing of cellulosic biomass: an update.* Curr Opin Biotechnol, 2005. **16**(5): p. 577-83.
21. Xu, Q., A. Singh, and M.E. Himmel, *Perspectives and new directions for the production of bioethanol using consolidated bioprocessing of lignocellulose.* Curr Opin Biotechnol, 2009. **20**(3): p. 364-71.
22. Bayer, E.A., R. Lamed, and M.E. Himmel, *The potential of cellulases and cellulosomes for cellulosic waste management.* Curr Opin Biotechnol, 2007. **18**(3): p. 237-45.
23. Dashtban, M., H. Schraft, and W. Qin, *Fungal bioconversion of lignocellulosic residues; opportunities & perspectives.* Int J Biol Sci, 2009. **5**(6): p. 578-95.
24. Sanchez, C., *Lignocellulosic residues: biodegradation and bioconversion by fungi.* Biotechnol Adv, 2009. **27**(2): p. 185-94.
25. Schwarz, W.H., *The cellulosome and cellulose degradation by anaerobic bacteria.* Appl Microbiol Biotechnol, 2001. **56**(5-6): p. 634-49.
26. Cereghino, J.L. and J.M. Cregg, *Heterologous protein expression in the methylotrophic yeast Pichia pastoris.* FEMS Microbiol Rev, 2000. **24**(1): p. 45-66.
27. Hohenblum, H., et al., *Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant Pichia pastoris.* Biotechnol Bioeng, 2004. **85**(4): p. 367-75.
28. Hartner, F.S., et al., *Promoter library designed for fine-tuned gene expression in Pichia pastoris.* Nucleic Acids Res, 2008. **36**(12): p. e76.

CHAPTER 1

Technical Report

Sensitive high-throughput screening for the detection of reducing sugars

Andrea Mellitzer[1], Anton Glieder[2], Roland Weis[3], Christoph Reisinger[4] and Karlheinz Flicker[2]

1 Institute of Molecular Biotechnology, Graz University of Technology, Graz, Austria
2 ACIB GmbH, Austrian Centre of Industrial Biotechnology, Graz, Austria
3 VTU Technology GmbH, Grambach, Austria
4 Süd-Chemie AG, Munich, Germany

Biotechnol. J. 2012, 7, 155–162 DOI 10.1002/biot.201100001

Chapter 1

Technical Report

Sensitive high-throughput screening for the detection of reducing sugars

Andrea Mellitzer[1], Anton Glieder[2], Roland Weis[3], Christoph Reisinger[4] and Karlheinz Flicker[2]

[1] Institute of Molecular Biotechnology, Graz University of Technology, Graz, Austria
[2] ACIB GmbH, Austrian Centre of Industrial Biotechnology, Graz, Austria
[3] VTU Technology GmbH, Grambach, Austria
[4] Süd-Chemie AG, Munich, Germany

Received 5 January 2011
Revised 4 February 2011
Accepted 20 February 2011

The exploitation of renewable resources for the production of biofuels relies on efficient processes for the enzymatic hydrolysis of lignocellulosic materials. The development of enzymes and strains for these processes requires reliable and fast activity-based screening assays. Additionally, these assays are also required to operate on the microscale and on the high-throughput level. Herein, we report the development of a highly sensitive reducing-sugar assay in a 96-well microplate screening format. The assay is based on the formation of osazones from reducing sugars and *para*-hydroxybenzoic acid hydrazide. By using this sensitive assay, the enzyme loads and conversion times during lignocellulose hydrolysis can be reduced, thus allowing higher throughput. The assay is about five times more sensitive than the widely applied dinitrosalicylic acid based assay and can reliably detect reducing sugars down to 10 μM. The assay-specific variation over one microplate was determined for three different lignocellulolytic enzymes and ranges from 2 to 8%. Furthermore, the assay was combined with a microscale cultivation procedure for the activity-based screening of *Pichia pastoris* strains expressing functional *Thermomyces lanuginosus* xylanase A, *Trichoderma reesei* β-mannanase, or *T. reesei* cellobiohydrolase 2.

Keywords: Cellulase · High-throughput screening · *para*-Hydroxybenzoic acid hydrazide · *Pichia pastoris* · Reducing sugar

1 Introduction

The development and the production of efficient and cheap lignocellulolytic enzymes is one of the key points for second-generation biofuel production from renewable resources. Typical biotechnological methods for enzyme development, such as directed evolution or strain engineering, usually require large numbers of samples to be evaluated by activity-based assays. Microassays are preferred for this purpose because they facilitate rapid screening of a large number of samples and substantially reduce reagent consumption. However, the use of natural, recalcitrant, and often insoluble lignocellulosic substrates imposes some distinct problems on activity-based screening, particularly on the microscale. The insolubility of these substrates causes problems with substrate handling and dosing. Moreover, enzymatic hydrolysis of natural cellulosic substrates can essentially only be monitored by the detection of reducing sugars or by enzyme-coupled assays [1]. Direct (physical) methods based on nephelometry, turbidimetry [2], and viscosimetry [3] have been developed, however, these methods rely on unnatural, physically modi-

Correspondence: Dr. Karlheinz Flicker, ACIB GmbH, Petersgasse 14, 8010 Graz, Austria
E-mail: karlheinz.flicker@acib.at

Abbreviations: ANOVA, analysis of variation; **AOX1**, alcohol oxidase 1; **BCA**, bicinchoninic acid; **BMD**, buffered minimal medium; **CBHI**, cellobiohydrolase I; **CV**, coefficient of variation; **DNS**, 3,5-dinitrosalicylic acid; **DWP**, deep-well plate; **FPU**, filter paper units; **GAP**, glyceraldehyde-3-phosphate dehydrogenase; **IUPAC**, International Union of Pure and Applied Chemistry; **pHBAH**, *p*-hydroxybenzoic acid hydrazide; ***Tl*XynA**, *Thermomyces lanuginosus* xylanase A; ***Tr*bMan**, *Trichoderma reesei* β-mannanase; ***Tr*CBH2**, *Trichoderma reesei* cellobiohydrolase 2

fied (amorphous) cellulosic substrates. Additionally, their adaption to the microscale has not yet been reported. More convenient alternatives to these natural substrates are chemically modified derivatives of cellulose, such as carboxymethyl cellulose, covalently dyed cellulose (e.g., Cellulose Azure) [4], soluble oligosaccharides [5], and soluble chromogenic and/or fluorogenic small-molecule substrates [6] (e.g., *para*-nitrophenyl and 4-methylumbelliferyl glycosides). Heavily modified or small substrates, however, are likely to introduce a bias in enzyme engineering and should be applied thoughtfully.

The International Union of Pure and Applied Chemistry (IUPAC) recommends that cellulolytic activities are measured in "filter paper units" (FPU) [7].The FPU assay is based on the 3,5-dinitrosalicylic acid (DNS) method reported by Miller [8]. It is defined exclusively for the release of glucose and requires a dilution series of each enzyme solution to be assayed so that exactly 2 mg of glucose is released during 60 min of substrate conversion. Clear disadvantages of the FPU assay in high-throughput applications are the difficult substrate handling (e.g., paper strips) and the requirement for a dilution series for each enzyme, especially if large differences in protein expression are observed within a screening experiment. These characteristics ultimately result in time-consuming measurements and consequently reduce throughput. Moreover, although the FPU [9] and DNS assays [10] have been adapted to the microscale level, there are more sensitive microscale assays available [11], including the *p*-hydroxybenzoic acid hydrazide (pHBAH) assay. More sensitive assays are especially useful for the realization of short hydrolysis times at low conversion levels with lignocellulosic substrates. Other sensitive reducing-sugar assays, such as the Nelson–Somogyi assay, have also been established at the microscale. However, the throughput of the described methods was limited by the substrate conversion time (24 h) and the requirement for partially purified enzymes [12] or, in a reported screening of bacterial lysates, by the time-consuming (2 h) substrate conversion and color development [13]. A microscale modification of the ferricyanide-based method (Park–Johnson assay) was recently also reported, however, the authors again report sensitivity to interfering compounds [14]. Generally, it can be assumed that, for fast and efficient activity-based screening of lignocellulolytic activities in crude samples, a certain trade-off between sensitivity and interference of the applied assay has to be accepted.

Lignocellulose degradation is very heterogeneous and usually slows down at higher conversion rates, for reasons that are still not completely understood [15]. One explanation relates these effects to a limited number of accessible enzyme binding sites on the substrate [15, 16]. Also, product inhibitory effects, as observed in the case of cellobiohydrolases [17], are minimized when the overall conversion (conversion time) of substrate is limited.

Herein, we present a sensitive and broadly applicable reducing-sugar assay in a 96-well microplate format. The assay is based on a previously described method described by Lever for detecting reducing sugars [18] and is combined with microscale cultivation of *P. pastoris* in deep-well plates (DWPs) [19]. The reducing sugars chemically react with pHBAH to form strongly absorbing osazones that can be photometrically detected. Compared with the widely applied DNS assay developed by Miller [8], the pHBAH assay is much more sensitive and less toxic. In the established procedure, 20 µl of cultivation supernatant from microcultures of secreting yeast strains are used for substrate conversions in a total assay volume of 170 µL. Due to the small amount of enzyme in the assay, only a limited number of the available enzyme binding sites on the substrate are occupied by enzymes. The pHBAH assay shows high sensitivity and only short substrate conversion times can be realized. Additionally, the assay is easily tuned to fit different detection ranges by changing the incubation temperature during osazone formation.

In combination with the well-established high-throughput microscale cultivation of *P. pastoris* in DWP [19], we have developed an activity-based high-throughput screening method for (hemi-)cellulase-expressing *P. pastoris* strains. To demonstrate the value of this assay for fast and reliable expression screening, we have applied it to *Thermomyces lanuginosus* xylanase A (*Tl*XynA), *Trichoderma reesei* β-mannanase (*Tr*bMan), and *Trichoderma reesei* cellobiohydrolase 2 (*Tr*CBH2) expressing strains of *P. pastoris*. We were able to show that the complete assay procedure offers a sensitive, quantitative, and rapid screening method for (hemi-)cellulases in *P. pastoris* on a functional and high-throughput level.

2 Materials and methods

2.1 Materials

All disposable plastic materials and self-adhesive aluminum foil for sealing the 96-well microtiter and PCR plates were from Greiner bio-one (Frick-

enhausen, Germany). Semi-skirted 96-well PCR plates were from VWR (Vienna, Austria).

2.2 Solutions and reagents

All reagents were obtained from Carl Roth GmbH (Karlsruhe, Germany) unless otherwise stated. 3,5-Dinitrosalicylic acid (order no. D0550) was from Sigma. Avicel PH-101 (order no. 11365) and pHBAH (order no. 54600) were obtained from Fluka. D-(+)-mannose and D-(+)-cellobiose were from Fluka, D-(+)-xylose was from Sigma, and D-(+)-glucose monohydrate was from Carl Roth (Karlsruhe, Germany). BSA was from Thermo Scientific's bicinchoninic acid (BCA) protein assay kit (order no. 23225), *Thermomyces lanuginosus* cellobiohydrolase I (CBHI) was from Megazyme (Bray, Co. Wicklow, Ireland). For the BCA assay Thermo Scientific's BCA protein assay kit (order no. 23225) was used.

2.3 Construction of *P. pastoris* strains

The coding sequences of *Tl*XynA (UniProtKB/Swiss-Prot: O43097), *Tr*bMan (UniProtKB/TrEMBL: Q99036), and *Tr*CBH2 (UniProtKB/Swiss-Prot: P07987) were codon-optimized for *P. pastoris* expression by applying the Gene Designer software (DNA2.0, Menlo Park, CA, USA). Synthetic genes were cloned into the multiple cloning site of the *E. coli*/*P. pastoris* shuttle vector pPpB1 [20] through *Eco*RI/*Not*I restriction sites. *Tr*bMan and *Tl*XynA were cloned downstream of the glyceraldehyde-3-phosphate dehydrogenase (GAP) and *Tr*CBH2 downstream of the alcohol oxidase 1 (AOX1) promoters. Electro-competent *P. pastoris* CBS 7435 Mut^S cells were prepared and transformed according to ref. [21]. 1 to 3 μg of the *Bgl*II-linearized pPpB1 vector construct were used for each transformation. Transformants were plated on YPD-zeocin (100 μg/mL zeocin) agar plates and grown at 30°C for 72 h.

2.4 Microscale cultivation of *P. pastoris*

P. pastoris strains expressing *Tl*XynA, *Tr*bMan, and *Tr*CBH2 were cultivated on the microscale in DWPs [19]. DWPs containing appropriate media were inoculated with fresh transformants from agar plates with sterile toothpicks and were then cultivated in shakers (INFORS Multitron, Bottmingen, Switzerland) at 28°C, 320 rpm, and 80% relative humidity. The conditions for AOX1- and GAP-promotor-driven expression were the same as those described previously [22].

2.5 Reducing-sugar assays

A stock solution of 5% w/v pHBAH in 0.5% v/v HCl was prepared as described by Lever [18]. Insoluble matter was removed by filtering through an 8 μm syringe filter. The working solution was prepared freshly for each measurement (max. 24 h storage at 4°C) by diluting the stock with 0.5 M NaOH 1:4 v/v. The reagent solution for the DNS assay was prepared as described by Miller [8]. Solutions for the BCA assay were prepared as described in the manual of the BCA protein assay kit (Thermo Fisher Scientific Inc., Waltham, MA, USA). Standard dilutions of the reducing sugars were prepared in 50 mM sodium acetate buffer at pH 4.8.

For the lignocellulosic substrate conversion, 150 μL of appropriate substrate (stirred suspension of 1% Avicel; solutions of 0.5% xylan or 0.2% locust bean gum) in 50 mM citrate buffer (pH 4.8) were transferred to a 96-well plate. 20 μL of the enzyme sample (culture supernatant) was added to each well. The plates were sealed with adhesive aluminum foil and were incubated at 50°C for 30 min and 300 rpm (unless otherwise stated) followed by substrate pelletizing at 4000 *g* for 5 min at 4°C. For the subsequent reducing-sugar assay, 50 μL of the substrate conversion reaction (or, in the case of the standard sugars, appropriate dilutions of the reducing sugar) were pipetted into 150 μL of the pHBAH working solution in a 96-well PCR plate. After sealing with adhesive aluminum foil, the PCR plates were incubated for 5 min at 95°C and then cooled to 4°C in an Applied Biosystems 2720 thermocycler. 150 μL of the assay samples were transferred to a new polystyrol microplate and the absorption was read at 410 (pHBAH assay), 540 (DNS assay), and 562 nm (BCA assay) in a SPECTRA MAX Plus384 plate reader (Molecular Devices Corp., Sunnyvale, CA, USA). All 96-well pipetting steps were done using a Quadra Tower (Tomtec, Hamden, CT, USA) 96-tip pipetting robot.

2.6 Determination of the assay-specific variation

For these experiments, sterile-filtered, shake-flask culture supernatants of *P. pastoris* strains with GAP-promotor-regulated expression of *Tl*XynA and *Tr*bMan were used (250 mL baffled shake flasks containing 50 mL buffered minimal medium (BMD) [19] with 5% glucose, 60 h at 28°C, 150 rpm). *Tr*CBH2 was a rehydrated lyophilisate of a *P. pastoris* fermentation supernatant. The substrate solutions were 0.5% xylan (*Tl*XynA), 0.2% locust bean gum (*Tr*bMan), and 1% Avicel (*Tr*CBH2) in 50 mM citrate buffer (pH 4.8). The pHBAH assay procedure was the same as described before (see sec-

tion 2.5 Reducing-sugar assays) except that substrate conversion was 5, 10, and 30 min for *Tl*XynA, *Tr*bMan, and *Tr*CBH2, respectively. Different conversion times were chosen to compensate for the different hydrolysis rates of the tested enzymes and to avoid absorbance readings in the nonlinear range.

2.7 Data evaluation

Linear standard curves were obtained for each assay by linear least-squares fitting. The slopes of these linear standard curves of different reducing sugars were used as "response" parameters to compare the sensitivity of the assays. The lower linear detection-range limits for each assay were determined from the mean of the buffer blank plus three times the standard deviation of the blank plus three times the standard deviation of a low concentration sample. The upper limits were defined as the concentration value corresponding to a maximum absorption value of 2.5 (i.e., this value represents a parameter specific for the plate reader).

2.8 Statistics

Standard errors were determined from ≥2 individual experiments of 4 technical repeats unless otherwise stated. The optimal assay conditions for the pHBAH assay were selected by a multifactorial experimental setup using D-(+)-glucose standard dilutions at three concentration levels (low: 0.015 mg/mL, middle: 0.125 mg/mL, high: 0.4 mg/mL) and applying variations in incubation time (1, 5, and 10 min at 95°C) and strength and sort of incubation buffer (25, 50, and 100 mM sodium citrate or sodium acetate buffer at pH 4.8). The parameters were tested for significant contributions to the final absorption values by using a one-way analysis of variance (ANOVA) with a sample number of 4 and p=0.05. The coefficient of variation (CV) was calculated from the percent ratio of standard deviation and mean of a measurement.

3 Results and discussion

Fast and reliable activity-based screening for the functional expression of lignocellulolytic enzymes in high-throughput systems is required for strain selection, directed evolution, and the engineering of protein expression and secretion. Ideally, reliable microassays that serve this purpose also cope with crude samples (culture supernatants) and are able to detect small differences in enzymatic activities. The adapted pHBAH assay presented herein can be used to screen for (hemi-) cellulolytic activity of heterologous enzymes produced by *P. pastoris* strains in cultures of a few hundred microliters. In a single step, 96 strains can be screened in parallel and the short assay time allows screening of hundreds to thousands of strains per day.

Compared with other assay procedures, the inhomogeneity of heat distribution and inefficient heat transfer during reducing-sugar detection are eliminated in the present assay by using a PCR thermocycler as a heating device. Additionally, errors derived from liquid handling are overcome by using a liquid-handling system capable of performing simultaneous 96-well pipetting steps. This is especially important when insoluble particulate substrates are used because it is extremely difficult to avoid substrate carryover during manual pipetting. To evaluate and benchmark the pHBAH assay, we compared this assay with the BCA and DNS reducing-sugar assays. To identify the best pHBAH assay conditions, a multifactorial assay setup combined with one-way ANOVA tests was used. According to the data, incubation for 5 min at 95°C is enough to achieve full signal development. The assay-specific background reaction level for acetate and citrate buffer (pH 4.8) is negligible (<0.1 absorbance units) for the pHBAH and DNS assays for buffer concentrations from 0 to 150 mM. In contrast, the BCA assay showed absolute background absorption in the range of 0.15 ± 0.02 and 0.23 ± 0.01 absorbance units for citrate and acetate buffer, respectively.

The linear detection ranges of the three assays were compared based on a standard dilution series of reducing sugars. Table 1 and Fig. 1 show that only pHBAH and BCA are capable of reliably detecting different reducing sugars below a concentration of 1 mM. Regarding sensitivity (compare Fig.1 D and Table 1) the BCA and pHBAH assays show an approx. 13- and 5-fold increase of sensitivity relative to the DNS assay (averaged over the four sugars tested). The variability of the pHBAH assay in terms of linear detection range and response changes was investigated by using D-(+)-glucose as the reducing sugar. According to Table 1, the sensitivity of the pHBAH assay is reduced 2- and 10-fold when the incubation temperature is lowered to 80 and 70°C, respectively. As exemplified for the detection of D-(+)-glucose, the concomitant extension of the linear detection range (Table 1) enables direct measurement of glucose concentrations ranging from 0.01 (lower limit of linear detection range) to 10.76 mM (upper limit of linear detection range) of D-(+)-glucose, thus adding additional flexibility to the pHBAH assay. In summary, the range of detection for the pHBAH assay can be adapted as re-

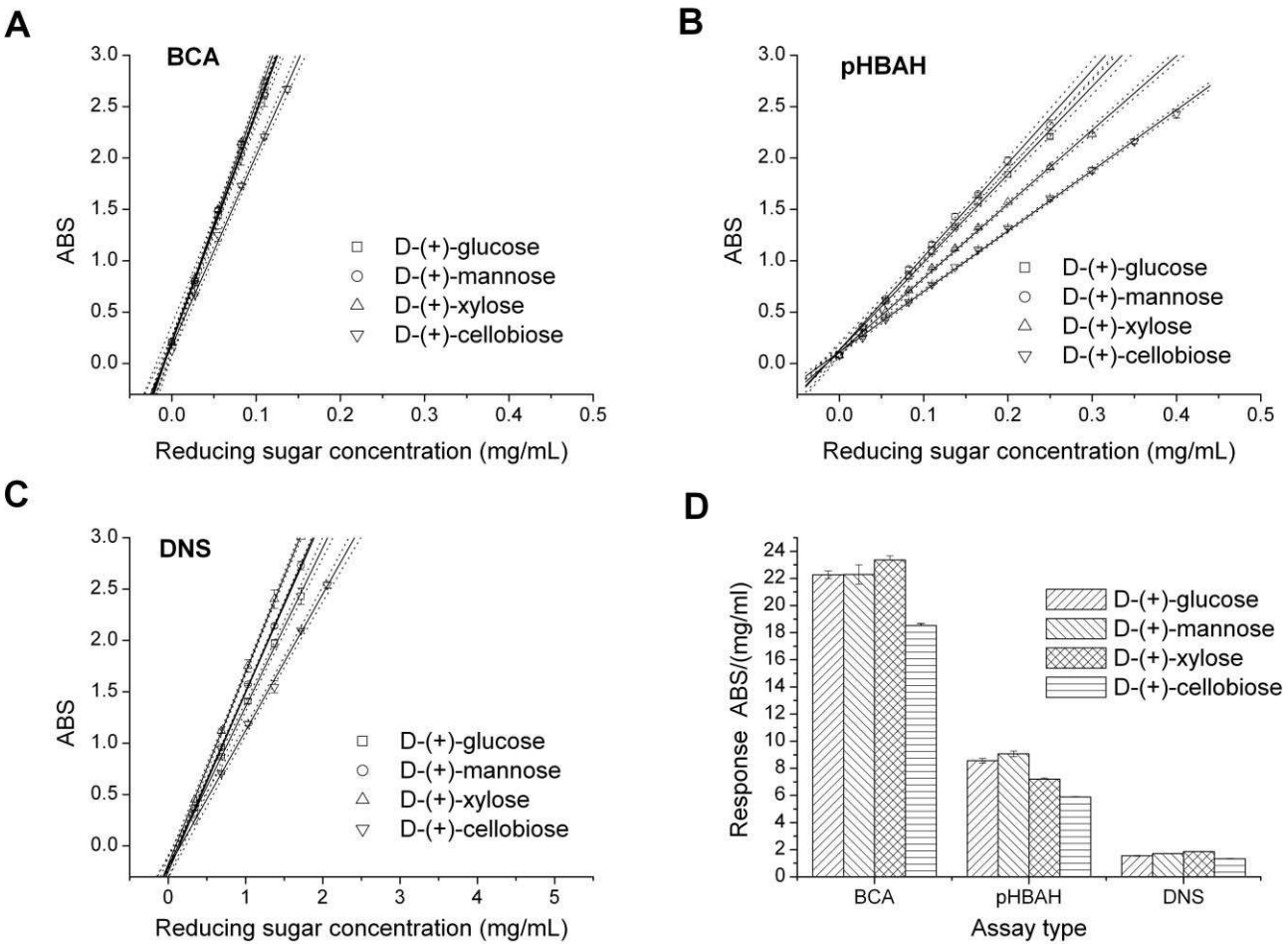


Figure 1. Linear standard curves for four different reducing sugars using (A) BCA, (B) pHBAH, and (C) DNS assays. Confidence intervals (95%) are displayed for each fit as dotted lines. (D) The different sensitivities (response slopes) of the compared assays for the four different reducing sugars are given in absorbance (ABS) units per reducing-sugar concentration in mg/mL. All data were determined in four independent measurements and error bars indicate ± standard deviation.

quired by a simple change of the incubation temperature during osazone formation.

The protein-specific background reaction of the reducing-sugar assays was investigated by using dilution series of 0 to 2 mg/ml solutions of BSA and *Thermomyces lanuginosus* CBH1, which represented examples of non-glycosylated and glycosylated proteins, respectively. The BCA assay showed a linear increase in the background signal with increasing protein concentration, whereas the pHBAH and DNS assays produced only a constant, low background of approximately 0.15 absorbance units over the investigated concentration range. Therefore, although the BCA assay has been described as a sensitive assay for the determination of reducing sugars [11, 23], it may not be suitable for assaying protein-rich samples (e.g., culture supernatants or lysates) due to its high cross-reactivity with proteins. Further descriptions of cross-reaction of similar highly sensitive redox assays, such as the Nelson–Somogyi or the Park–Johnson assays, with buffer components or proteins have been reported by Zhang et al. [24] and by Moretti and Thorson [11]. Unlike the BCA assay, the pHBAH assay is only sensitive to very high protein concentrations [18]. In fact, it has been used to determine glucose levels in blood serum [25]. Therefore, selective quantification of reducing sugars in culture supernatants containing proteins is more reliably done with the pHBAH assay than with metal-ion-based redox assays.

The pHBAH assay performance (including substrate conversion and reducing-sugar detection) was tested with culture supernatants from flask cultures (*Tl*XynA, *Tr*bMan) or with a rehydrated lyophilisate of a concentrated fermentation supernatant (*Tr*CBH2). To test the assay variability, the enzyme samples were uniformly distributed over

Table 1. Comparison of characteristic assay parameters

		Incubation temperature (°C)					
		95°C		80°C		70°C	
	Assay	min	max	min	max	min	max
Linear detection range (mM)[a]	BCA	0.04[b]	0.52[b]	n.d.	n.d.	n.d.	n.d.
		0.04[c]	0.57[c]				
		0.06[d]	0.66[d]				
		0.05[e]	0.37[e]				
	pHBAH	0.01[b]	1.40[b]	0.12[b]	2.45[b]	0.57[b]	10.76[b]
		0.01[c]	1.45[c]				
		0.05[d]	2.21[d]				
		0.03[e]	1.18[e]				
	DNS	1.04[b]	8.78[b]	n.d.	n.d.	n.d.	n.d.
		1.15[c]	8.77[c]				
		1.13[d]	9.55[d]				
		0.78[e]	5.94[e]				
	Nelson-Somogyi[h]			0.55[b]	8.33[b]		
Response (ABS/mM)[f]	pHBAH	2.02 ± 0.08		0.97 ± 0.02		0.22 ± 0.01	
Relative responses[g]	BCA	13.44 ± 0.50					
	pHBAH	4.77 ± 0.15					
	DNS	1.00 ± 0.02					

a) Linear detection ranges were determined from linear standard curves by using 50 µL standard dilutions of the respective reducing sugar and a reaction time of 5 min at the indicated temperature.
b) Reducing sugar: D-(+)-glucose.
c) Reducing sugar: D-(+)-mannose.
d) Reducing sugar: D-(+)-xylose.
e) Reducing sugar: D-(+)-cellobiose.
f) Responses determined for D-(+)-glucose at the indicated temperatures, ABS: absorbance.
g) Determined at 95°C, normalized to the response of the DNS assay and averaged for each of the reducing sugars; errors correspond to standard deviation.
h) Calculated linear range of the standard curve from the original publication (molecular weight glucose: 180.16) [12].
n.d.: not determined.

one microplate. Based on the detected reducing-sugar levels, the pHBAH assay showed only a small variation over a complete microplate (Fig. 2 A and B). For all three enzymes tested, the middle 50% of all (mean-normalized) data points determined for one microplate were within ± 5% of the plate mean (Fig. 2 A). The assay-specific CV values were 2.30, 3.80, and 7.95% for *Tr*CBH2, *Tl*XynA, and *Tr*bMan, respectively.

To further verify the applicability of the pHBAH assay, we generated example activity landscapes by screening microscale DWP cultures of *P. pastoris* strains producing *Tl*XynA, *Tr*bMan, and *Tr*CBH2 (Fig. 2 C and D). For this purpose, we transformed *P. pastoris* CBS 7435 Mut^S with expression constructs containing the constitutive GAP promoter (*Tl*XynA, *Tr*bMan) and the methanol-inducible AOX1 promoter (*Tr*CBH2). The activity landscapes from these transformants were determined directly from the harvested DWP culture supernatants. This way we were also able to show the applicability of the assay for different expression conditions. By using the mean and the standard deviation of each landscape, CV values of 23.57, 32.81, and 20.42% were determined for *Tr*CBH2, *Tl*XynA, and *Tr*bMan, respectively. These values are approx 10-fold higher for *Tr*CBH2 than the assay specific variation (compare Fig. 2 B and C). A similar value can be observed for *Tl*XynA (approx. 9-fold), whereas a lower value of 3-fold is observed for *Tr*bMan. Thus, the pHBAH assay proved to be a very reliable reducing-sugar screening assay, which generates activity landscapes that are statistically robust towards assay-specific variations. As a result, the total variation in the activity landscapes of the transformants can be attributed mostly to growth- or expression-related variations or to differences in gene copy number. However, these activity variations are only partly host-related and are also specific for the expressed enzyme. Due to the high sensitivity of the pHBAH assay, weakly expressing strains, which typically comprise strains containing single copies of the expression cassettes, can also be screened and subjected to engineering for improved performance. Also, only single-copy clones allow discrimi-

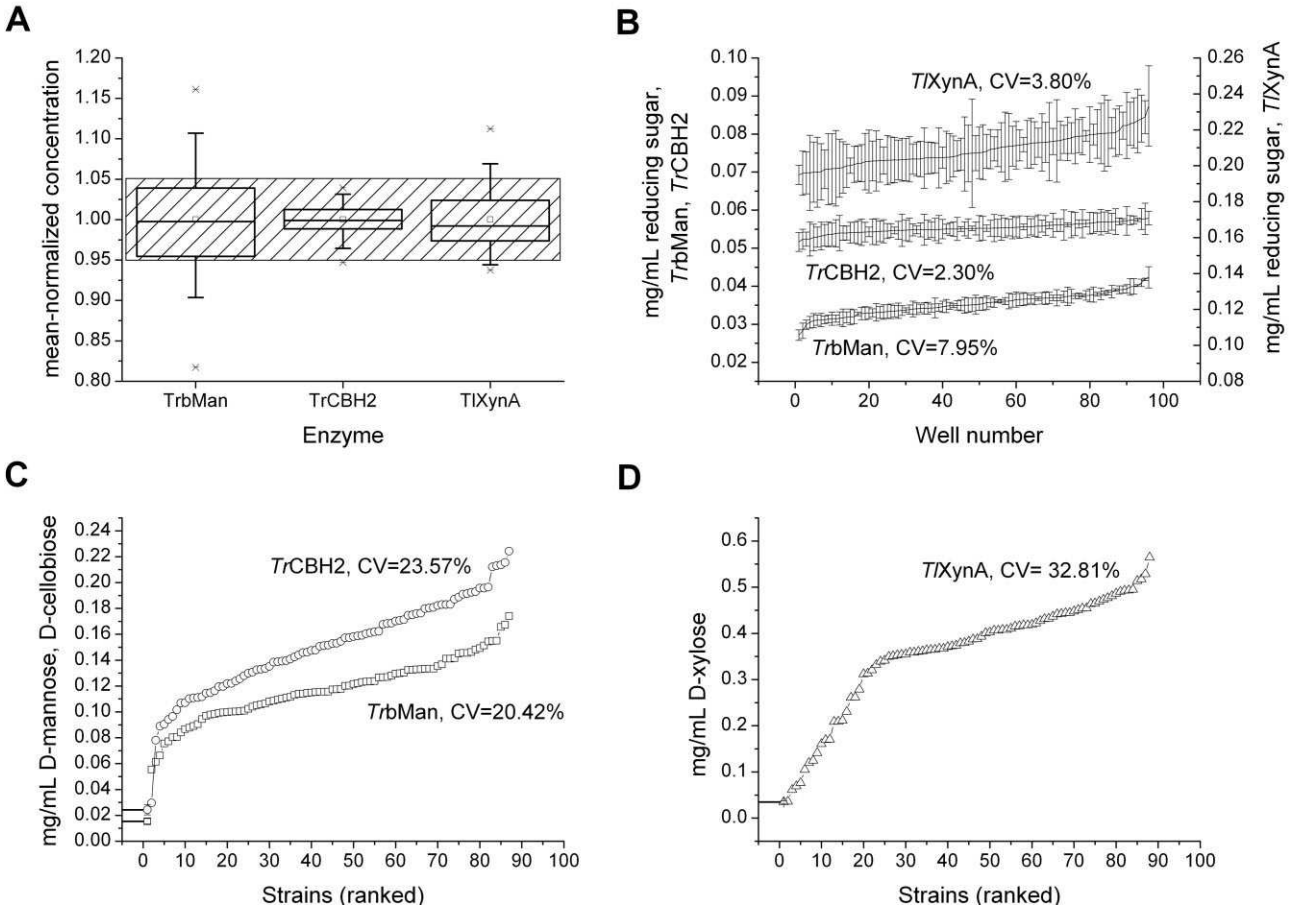


Figure 2. Plate variability of the established microscale pHBAH screening assay for three different expressed enzymes represented as (A) box plots of plate-mean normalized data and (B) ranked mean data of three independent measurements (error bars represent ± standard deviation). The shaded region in (A) indicates the ± 5% region. Data were corrected for negative controls (*P. pastoris* CBS 7435 Mut^S transformed with empty vector). Total protein concentrations are (69 ± 2), (177 ± 4), and (60 ± 1) µg/mL for *Tr*bMan, *Tr*CBH2, and *Tl*XynA, respectively. The activity landscapes of 88 individual *P. pastoris* transformants expressing *Tr*bMan and *Tr*CBH2 (C) and *Tl*XynA (D) are shown, as determined with the pHBAH assay. The activity levels of the vector control strains are indicated by a horizontal line.

nation between effects from mutations versus multi-copy integration in the case of screenings for improved activity. Based on the pHBAH assay's adaptability of the linear detection range (Table 1), this assay can be used throughout a complete engineering experiment in which the absolute activities of strains or enzyme variants usually increase considerably.

4 Concluding remarks

The pHBAH assay is a sensitive reducing-sugar assay that is suitable for implementation in high-throughput protein-expression screening on the microscale. Its functionality was demonstrated by the activity-based high-throughput expression analysis of *P. pastoris* strains producing (hemi-)cellulases in DWP microscale cultures. Ideally, the combination of the pHBAH assay with microscale cultivation facilitates parallel screening of 96 clones in one single experiment.

Compared with bioreactor cultivations, lower protein concentrations are observed in *P. pastoris* DWP culture supernatants. Thus, the use of small volumes of culture supernatant in combination with the sensitive pHBAH assay allows conversion of lignocellulosic substrates at low enzyme loads and short conversion times; this ultimately allows higher throughput. Furthermore, we have shown, in principle, that the sensitivity of the pHBAH assay can be reduced by lowering the incubation temperature during osazone formation. This adaptability fits the requirements of the higher detection range associated with the selection of improved strains or enzyme variants.

The clear advantages of the presented high-throughput pHBAH microassay over other sensitive assays are its sensitivity combined with its suitability for assaying crude-culture supernatants. Therefore, this simple and reliable assay can be universally applied in strain engineering and directed evolution for the activity-based screening of enzymes that degrade natural sugar polymers.

This study was performed within the COMET K-Project MacroFun (FFG project number 815076). We acknowledge the FFG, the SFG, and the Province of Styria, Austria, for financial support.

The authors have declared no conflict of interest.

5 References

[1] Bethke, P. C., Busse, J. C., Validation of a simple, colorimetric, microplate assay using Amplex red for the determination of glucose and sucrose in potato tubers and other vegetables. *Am. J. Potato Res.* 2008, *85*, 414–421.

[2] Nummi, M., Fox, P. C., Niku-Paavola, M. L., Enari, T. M., Nephelometric and turbidometric assays of cellulase activity. *Anal. Biochem.* 1981, *116*, 133–136.

[3] Pala, H., Mota, M., Gama, F. M., Enzymatic depolymerisation of cellulose. *Carbohydr. Polym.* 2007, *68*, 101–108.

[4] Lai, T. E., Pullammanappallil, P. C., Clarke, W. P., Quantification of cellulase activity using cellulose-azure. *Talanta* 2006, *69*, 68–72.

[5] Nidetzky, B., Zachariae, W., Gercken, G., Hayn, M. et al., Hydrolysis of cellooligosaccharides by Trichoderma reesei cellobiohydrolases: Experimental data and kinetic modeling. *Enzyme Microb. Technol.* 1994, *16*, 43–52.

[6] Van Tilbeurgh, H., Claeyssens, M., de Bruyne, C. K., The use of 4-methylumbelliferyl and other chromophoric glycosides in the study of cellulolytic enzymes. *FEBS Lett.* 1982, *149*, 152–156.

[7] Ghose, T. K., Measurement of cellulase activities. *Pure Appl. Chem.* 1987, *59*, 257–268.

[8] Miller, G. L., Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 1959, *31*, 426–428.

[9] Xiao, Z., Storms, R., Tsang, A., Microplate-based filter paper assay to measure total cellulase activity. *Biotechnol. Bioeng.* 2004, *88*, 832–837.

[10] King, B. C., Donnelly, M. K., Bergstrom, G. C., Walker, L. P. et al., An optimized microplate assay system for quantitative evaluation of plant cell wall-degrading enzyme activity of fungal culture extracts. *Biotechnol. Bioeng.* 2009, *102*, 1033–1044.

[11] Moretti, R., Thorson, J. S., A comparison of sugar indicators enables a universal high-throughput sugar-1-phosphate nucleotidyltransferase assay. *Anal. Biochem.* 2008, *377*, 251–258.

[12] Green III, F., Clausen, C. A., Highley, T. L., Adaptation of the Nelson-Somogyi reducing-sugar assay to a microassay using microtiter plates. *Anal. Biochem.* 1989, *182*, 197–199.

[13] Leemhuis, H., Euverink, G. J. W., Dijkhuizen, L., High-throughput screening for gene libraries expressing carbohydrate hydrolase activity. *Biotechnol. Lett.* 2003, *25*, 1643–1645.

[14] Rigouin, C., Ladrat, C. D., Sinquin, C., Colliec-Jouault, S. et al., Assessment of biochemical methods to detect enzymatic depolymerization of polysaccharides. *Carbohydr. Polym.* 2009, *76*, 279–284.

[15] Bansal, P., Hall, M., Realff, M. J., Lee, J. H. et al., Modeling cellulase kinetics on lignocellulosic substrates. *Biotechnol. Adv.* 2009, *27*, 833–848.

[16] Kongruang, S., Han, M. J., Breton, C. I. G., Penner, M. H., Quantitative analysis of cellulose-reducing ends. *Appl. Biochem. Biotechnol.* 2004, *113*, 213–231.

[17] Gruno, M., Väljamäe, P., Pettersson, G., Johansson, G., Inhibition of the Trichoderma reesei cellulases by cellobiose is strongly dependent on the nature of the substrate. *Biotechnol. Bioeng.* 2004, *86*, 503–511.

[18] Lever, M., A new reaction for colorimetric determination of carbohydrates. *Anal. Biochem.* 1972, *47*, 273–279.

[19] Weis, R., Luiten, R., Skranc, W., Schwab, H. et al., Reliable high-throughput screening with Pichia pastoris by limiting yeast cell death phenomena. *FEMS Yeast Res.* 2004, *5*, 179–189.

[20] Abad, S., Nahalka, J., Bergler, G., Arnold, S. A. et al., Stepwise engineering of a Pichia pastoris D-amino acid oxidase whole cell catalyst. *Microb. Cell Fact.* 2010, *9*, 24.

[21] Lin-Cereghino, J., Wong, W. W., Xiong, S., Giang, W. et al., Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast Pichia pastoris. *Biotechniques* 2005, *38*, 44–48.

[22] Pscheidt, B., Liu, Z., Gaisberger, R., Avi, M. et al., Efficient biocatalytic synthesis of (R)-pantolactone. *Adv. Synth. Catal.* 2008, *350*, 1943–1948.

[23] Waffenschmidt, S., Jaenicke, L., Assay of reducing sugars in the nanomole range with 2,2′-bicinchoninate. *Anal. Biochem.* 1987, *165*, 337–340.

[24] Zhang, Y. H. P., Himmel, M. E., Mielenz, J. R., Outlook for cellulase improvement: Screening and selection strategies. *Biotechnol. Adv.* 2006, *24*, 452–481.

[25] Fingerhut, B., Automated method for serum glucose with p-hydroxybenzoic acid hydrazide. *Clin. Chem.* 1973, *19*, 1022–1026.

CHAPTER 2

Expression of lignocellulolytic enzymes in *Pichia pastoris*

Andrea Mellitzer[1]
Email: andrea.mellitzer@tugraz.at

Roland Weis[3]
Email: roland.weis@vtu.com

Anton Glieder[2]
Email: anton.glieder@acib.at

Karlheinz Flicker[2*]
* Corresponding author
Email: karlheinz.flicker@tugraz.at

1 Institute of Molecular Biotechnology, Graz University of Technology, Graz, Austria
2 ACIB GmbH, Austrian Centre of Industrial Biotechnology, Graz, Austria
3 VTU Technology GmbH, Grambach, Austria

Microbial Cell Factories 2012, **11**:61 doi:10.1186/1475-2859-11-61

Microbial Cell Factories

BioMed Central
The Open Access Publisher

Expression of lignocellulolytic enzymes in Pichia pastoris

Microbial Cell Factories 2012, **11**:61 doi:10.1186/1475-2859-11-61

Andrea Mellitzer (andrea.mellitzer@tugraz.at)
Roland Weis (roland.weis@vtu.com)
Anton Glieder (anton.glieder@acib.at)
Karlheinz Flicker (karlheinz.flicker@acib.at)

ISSN 1475-2859

Article type Research

Submission date 9 November 2011

Acceptance date 14 May 2012

Publication date 14 May 2012

Article URL http://www.microbialcellfactories.com/content/11/1/61

Expression of lignocellulolytic enzymes in *Pichia pastoris*

Andrea Mellitzer[1]
Email: andrea.mellitzer@tugraz.at

Roland Weis[3]
Email: roland.weis@vtu.com

Anton Glieder[2]
Email: anton.glieder@acib.at

Karlheinz Flicker[2*]
[*] Corresponding author
Email: karlheinz.flicker@tugraz.at

[1] Institute of Molecular Biotechnology, Graz University of Technology, Graz, Austria

[2] ACIB GmbH, Austrian Centre of Industrial Biotechnology, Graz, Austria

[3] VTU Technology GmbH, Grambach, Austria

Abstract

Background

Sustainable utilization of plant biomass as renewable source for fuels and chemical building blocks requires a complex mixture of diverse enzymes, including hydrolases which comprise the largest class of lignocellulolytic enzymes. These enzymes need to be available in large amounts at a low price to allow sustainable and economic biotechnological processes.

Over the past years *Pichia pastoris* has become an attractive host for the cost-efficient production and engineering of heterologous (eukaryotic) proteins due to several advantages.

Results

In this paper codon optimized genes and synthetic alcohol oxidase 1 promoter variants were used to generate *Pichia pastoris* strains which individually expressed cellobiohydrolase 1, cellobiohydrolase 2 and beta-mannanase from *Trichoderma reesei* and xylanase A from *Thermomyces lanuginosus*. For three of these enzymes even gram quantities of enzyme per liter were obtained by fed-batch cultivation. Additionally, we compared our achieved yields of secreted enzymes and the corresponding activities to literature data.

Conclusion

In our experiments we could clearly see the importance of gene optimization and strain characterization for successfully improving secretion levels. We also give a basic guideline for understanding the interplay of promoter strength and gene dosage for a successful improvement of the secretory production of lignocellulolytic enzymes in *Pichia pastoris*.

Keywords

Xylanase, Mannanase, Cellobiohydrolase, Synthetic gene, Synthetic promoter, Quantitative real time PCR, *Pichia Pastoris*, Fermentation, Strain development

Background

Although *Pichia pastoris* is a relatively simple eukaryotic organism it can perform many posttranslational modifications such as glycosylation, disulfide bond formation, and proteolytic processing [1]. Therefore, *Pichia* serves as an interesting alternative to other (more difficult to handle) fungal secretory expression systems that are used to produce lignocellulolytic enzymes and other eukaryotic proteins which typically require post-translational modifications for correct folding, stability and activity. The recalcitrant and complex nature of lignocellulosics [2] affords the application of complex enzyme mixtures for efficient hydrolysis of these renewable sources. Consequently, for a sustainable production of fuels, chemical building blocks, and functional macromolecules from plant biomass a multitude of different enzymes is needed. To produce all these enzymes and variants thereof, production strains which can be handled and engineered in a simple way need to be generated. Therefore, being a well- described and widely applied expression host [3] *P. pastoris* was the first choice for the heterologous expression of the selected target proteins. Furthermore, in contrast to many other eukaryotic expression systems *P. pastoris* secretes no endogenous lignocellulolytic enzymes in significant amounts [4]. Therefore, recombinant *Pichia* strains can provide almost pure heterologous enzyme preparations without the need of extensive and costly downstream processing. In addition, simple media requirements and relative easy handling in bioreactors enable inexpensive large-scale cultivations of *Pichia* [5]. All these characteristic features of *Pichia* contribute to its high potential for cost reduction during the production of lignocellulolytic enzymes, particularly for application studies when only low- and medium-scale enzyme productions are required.

However, even though *Pichia pastoris* is a good host for the expression of heterologous proteins [3] there is still room for improvements on transcriptional [6,7] and (post-) translational level [8,9]. In this work we exemplify the impact of gene optimization on the overall expression level of lignocellulolytic enzymes in *Pichia pastoris*. Most genomes are heterogeneous in codon usage [10] and, accordingly, the codon bias of small subsets of genes may differ clearly from the average codon usage of the genome. To optimize protein coding sequences for enhanced protein expression in *Pichia pastoris* we use an in-house developed biased codon usage table [11]. This codon usage is biased towards the codons of selected, highly expressed [12,13] endogenous and heterologous genes when the AOX1 promoter and methanol were used for induction in *Pichia pastoris*. In addition to gene optimization, enzyme expression can be influenced on a transcriptional level by varying copy numbers of the integrated expression cassettes and by the choice of the promoter. So far the wildtype

AOX1 promoter (P(AOX1)) and, to a certain extent, the GAP promoter (P(GAP)) were mostly used for heterologous protein production in *Pichia pastoris* [3]. However, since the P(GAP) is strong and constitutive it is not a good choice for production of physiologically problematic or cytotoxic proteins [14]. In contrast, the P(AOX1) is even stronger but also tightly regulated. Nonetheless, for some heterologous proteins the high transcript level generated by P(AOX1) can overload the cellular post-translational machinery, resulting in misfolded, unprocessed, or mislocalized proteins that can trigger a complex cellular response known as the unfolded protein response [15,16]. To overcome these disadvantages of the wild-type GAP and AOX1 promoter a library of promoters based on the wild-type P(AOX1) was previously generated [6]. The distinct properties of these novel promoters regulate the transcript level of target mRNA in response to the available carbon source level and type and concomitantly achieve a fine-tuned protein expression in *Pichia pastoris*.

The aim of the present study was to show the functional expression of lignocellulolytic enzymes in *Pichia pastoris* at high quantities and investigating the effect of gene optimization and of alternate promoters on the expression level of these enzymes. Our expression studies highlight basic principles for designing suitable expression constructs and for successful strain development for different cellulolytic enzymes. For this study *Trichoderma reesei* cellobiohydrolase 1 and 2 (*Tr*CBH1 and *Tr*CBH2) and beta-mannanase (*Tr*bMan), and *Thermomyces lanuginosus* xylanase A (*Tl*XynA) were chosen as target enzymes.

Results and discussion

The goal of this study was to evaluate the potential of *Pichia pastoris* to express lignocellulolytic enzymes. In particular, we improved the expression of selected (hemi-) cellulases by codon optimization of the target genes, investigated the effect of promoter choice, and characterized the performance of selected producer strains in small-scale bioreactors. This characterization also included the effects of multi-copy integration on the productivity for the selected target enzymes.

To investigate the effect of different methods for codon optimization three different gene variants of *Trichoderma reesei* cellobiohydrolase 2 (*Tr*CBH2) were employed; the native gene variant (*Tr*CBH2-wt), a gene variant with optimized codon pairs by a commercial supplier (*Tr*CBH2-CP) and an in-house-optimized variant (*Tr*CBH2-HM). For the in-house design a codon usage table [17] derived from genes which are highly expressed in *Pichia pastoris* in methanol containing media was used.

The effects of gene optimization and promoter type were characterized by comparing activity landscapes of different strains (Figure 1). For this purpose *P. pastoris* strains were cultivated in 96 deep-well plates according to [18] and subsequently screened for lignocellulolytic activities using a reducing sugar assay that was recently adapted to high-throughput [19]. Owing to the low standard deviation of this assay, the detected changes in the activity landscapes mainly reflect actual changes in the expression level [19]. These differences can either be due to the number of integrated expression cassettes or caused by specific effects of the individual gene variants. Figure 1 shows enzyme activity landscapes of *Tr*CBH2-wt and the two differently optimized *Tr*CBH2 gene variants which have all been separately incorporated into the same expression vector and host. Stable integration of expression cassettes into the *Pichia pastoris* genome is generally based on homologous recombination

but can also be an effect of non-homologous end-joining (NHEJ). Depending on the length, type and structure of the homologous flanking regions, untargeted (random) genome integration mediated by NHEJ becomes prevalent over locus-specific targeting (own observation for our vector system). Therefore, expression levels may be influenced not only by the number of integrated gene copies [20] but also by the integration locus which influences the transcript levels of the integrated genes. Our results demonstrate a clear effect of gene optimization on expression level. This is corroborated by the fact that our interpretation of expression level does not rely on a single observation but is averaged over a whole activity landscape of many individual transformants (Figure 1). This could be substantiated by reliably proving low copy numbers among differently optimized genes, in order to get a decent comparability of the influence. The 2-fold increase in expression level of *Tr*CBH2-HM compared to *Tr*CBH2-wt suggests a more efficient transcription and/or translation of this variant in *P. pastoris*. Contrary to this, the gene optimized by the commercial service using codon pair optimization, *Tr*CBH2-CP, showed a 2-fold lower expression level than *Tr*CBH2-wt. *B*eing originally designed to assist co-translational protein folding [21] of multi-domain proteins we expected the optimization based on codon pair signaling to show improved expression for the two-domain enzyme *Tr*CBH2. However, as we observed the opposite effect for *Tr*CBH2-CP we speculate that the bottleneck of *Tr*CBH2 expression is rather on transcriptional level than on the posttranslational level of protein folding. Summarizing, the optimized gene variant *Tr*CBH2-HM was superior to all other variants under the tested methanol-inducing conditions. This suggests that preferring codons with a high codon adaptation index (CAI) for highly expressed proteins under methanol inducing conditions is a good choice for *Tr*CBH2.

Figure 1 Activity landscapes of individual *P. pastoris* transformants expressing three different *Tr*CBH2 variants controlled by P(AOX1). Codon pair optimized sequence in black (CP), wt sequence in grey, high CAI codons for methanol induced gene expression in white (HM) [11]. Released cellobiose concentration is represented in bars. The untransformed strain *P. pastoris* CBS7435 Mut^S was used as negative control

Especially for secreted proteins the level of expression strongly depends on the number of integrated expression cassettes. Therefore, often the production efficiency of a strain can be predicted by quantifying the number of genome-integrated expression cassettes (copy number, CN) [15,20,22,23]. In *P. pastoris* an initial (linear) positive correlation between copy number and productivity that stagnates at a defined upper limit can be observed [20,22]. Furthermore, in some cases also a loss of productivity above a certain number of integrated copies has been described [15,23]. In fact, high mRNA-levels caused by strong promoters or by high numbers of the expression cassettes can overload the folding and secretion machinery of the host. Depending on the protein this can entail an accumulation of unfolded proteins which triggers dedicated signaling pathways, commonly known as the unfolded protein response [24]. For comparative studies it is, therefore, essential to characterize the strains with regard to their copy numbers. Doing so will also allow the separation of promoter and/or copy number related effects on expression levels.

To determine the individual expression levels of *Tr*CBH2 expressing *P. pastoris* strains under bioreactor conditions we selected suitable strains based on initial micro-scale screenings in 96 deep-well plates and by quantitative gene copy number determination using qRT-PCR. Gene expression was driven either by the wild- type promoters P(AOX1) and P(GAP) or by synthetic promoter variants. These synthetic promoter variants are part of a newly generated promoter library based on P(AOX1) which was designed to fine-tune protein expression in

Pichia pastoris [6]. With regard to their particular regulatory features two of these synthetic promoters were chosen to be tested in this study, namely P(En) and P(De). In the original publication by Hartner *et al.* [6] P(En) showed similar low expression of the reporter protein green fluorescent protein (GFP) under derepressed conditions but increased expression up to 166%, when compared to the wild-type promoter P(AOX1) on single copy level after 0 h and 72 h of methanol induction, respectively. P(De) showed more than 4-fold higher GFP fluorescence intensity under derepressed conditions but decreased expression down to 55%, if compared to the wild-type promoter P(AOX1) on single copy level after 0 h and 72 h of methanol induction, respectively. Even though, GFP expression driven by P(De) resulted in decreased protein production it was shown that this promoter was favorable for difficult to secrete proteins such as horseradish peroxidase (HRP). The overall productivity in fed-batch cultivations of HRP expression driven by P(De) was significantly higher than compared to the overall productivity of HRP expression driven by P(AOX1) [6]. In Figure 2A the time-courses of protein concentrations in the supernatants during fed-batch cultivations of *Tr*CBH2 are compared. Figure 2A shows that the strain P(De)-*Tr*CBH2-CP-CN25 ± 7 which harbors about 25 expression cassettes achieved around 4 g/l of *Tr*CBH2. This is comparable to the expression of P(De)-*Tr*CBH2-HM-CN7 ± 1 which is optimized using our in-house HM method. It can also be seen from Figure 2B that our in-house gene optimization method HM outperforms that of the commercial supplier (compare also Figures 1 and 3B). Although different promoters were used the expression of P(De)-*Tr*CBH2-CP-CN25 ± 7 and P(AOX)-*Tr*CBH2-CP-CN7 ± 1 normalized to the same level suggesting that a linear correlation of expression independent of promoter type up to a CN of 25 exists for the CP optimization. Based on these data and on data from literature [20,22] we observed two properties for *Tr*CBH2 expression. Firstly, using the AOX1 promoter variant P(De) we observed a positive initial (linear) correlation up to at least 7 copies between copy number and productivity. Secondly, gene optimization with our in-house method results in higher expression level at low copy numbers.

Figure 2 Time-course of protein concentration during fed-batch cultivations. Gene sequences were optimized either by codon pair optimization (CP) [21] or by applying the high CAI codons for methanol induced genes (HM) [11]. Copy numbers (CN) are specified in the legend. Panel A: Expression of differently codon optimized *Tr*CBH2 variants under the control of P(AOX1), P(GAP) or the synthetic promoter P(De) [6]. HM optimized variants (closed symbol), CP optimized variants (open symbol). Panel B: Time-course of *Tr*CBH2 expression normalized to copy number. Legend labeling see panel A. Panel C: Expression of *Tr*bMan (open symbol) and *Tl*XynA (closed symbol) under the control of P(En) [6]. For virtual gels of the protein yields during the fermentation runs please refer to Additional file 1 and the raw data can be found in Additional file 2

Figure 3 Comparison of *Tr*CBH2 gene/ promoter variants normalized by gene copy numbers (CN). Panel A: Protein concentration of T*r*CBH2 fed-batch cultivations normalized by CN after 90 h of induction. Panel B: Relative ratios of the expression levels of the different gene optimization variants (HM/CP) under the control of P(AOX1) (white bar) and P(De) (black bar). Panel C: Relative ratios of the normalized expression levels of the different methanol inducible promoters (P(AOX1)/P(De)) expressing either *Tr*CBH2-CP (black bars) or *Tr*CBH2-HM (white bar) after 90 h of induction. Please refer to Additional file 2 for the raw data and Results and discussion section for more detailed information

As observed in the micro-scale screening (Figure 1) *Tr*CBH2-HM led to a higher expression level than *Tr*CBH2-CP (Figure 3) also in fed-batch cultivations. This effect was even more

pronounced for the expression regulated by P(AOX1) with a 5-fold improvement of *Tr*CBH2-HM over *Tr*CBH2-CP (Figure 3B). In contrast, for the P(De)-*Tr*CBH2 variants we only observed a 3-fold improvement (Figure 3B) which can be explained by the lower promoter strength of P(De) as described previously [6]. In addition, the relative ratios of the different gene optimization variants (Figure 3B) and the relative ratios of the different gene promoter variants (Figure 3C) allow also a better comparison between the methanol inducible promoters P(AOX1) and P(De). (Figure 3C) shows that the strong methanol-inducible P(AOX1) increases expression of *Tr*CBH2-HM around 1.7-fold compared to expression under the control of P(De). For *Tr*CBH2-CP expression under the control of different promoters only a 1.2-fold improvement can be seen (Figure 3C). These results clearly indicate for *Tr*CBH2 expression that the codon optimization which is based on the codon bias of highly transcribed genes under methanol-inducing conditions gives even higher expression when a strong methanol-inducible promoter is employed.

After 90 h of induction the single copy expression level of P(De)-*Tr*CBH2-HM-CN1 is approximately 0.43 g/l (Figure 2B) whereas a higher single copy expression level of about 0.930 g/l (normalized to CN) can be calculated for P(AOX1)-*Tr*CBH2-HM-CN3. Based on these data, P(AOX1) gives an around 2-fold higher expression level than P(De). Although the strain P(De)-*Tr*CBH2-HM-CN8 ± 2 performed best under the tested MeOH-inducing conditions our results, based on the normalized data, indicate that strong methanol-inducible promoters such as P(AOX1) or the even stronger methanol-inducible P(En) [6] can further increase the expression of *Tr*CBH2. To verify this hypothesis on fermenter scale we decided to screen for higher copy number strains expressing *Tr*CBH2 under the control of P(AOX1) and P(En). As seen in Figure 4/A the selected strains with increased copy numbers P(AOX1)-*Tr*CBH2-HM-CN5 ± 1 and P(En)-*Tr*CBH2-HM-CN6 ± 1 indeed produced significantly more protein over the whole induction period than the best strain of the first fermentation P(De)-*Tr*CBH2-HM-CN7 ± 2. Within the first 70 h of induction the productivity of P(En)-*Tr*CBH2-HM-CN6 ± 1 was higher than the productivity of P(AOX1)-*Tr*CBH2-HM-CN5 ± 1. This confirms the results of the previously reported GFP expression experiments [6] using an improved synthetic AOX1 promoter variant. The final protein yield of both strains was comparable at around 6 g/l. Summarizing, using strong methanol inducible promoters in combination with high copy numbers of genes that are optimized to a high CAI for highly expressed proteins under methanol induction can further increase the yield of *Tr*CBH2. Moreover, we showed that pre-selection of strains using micro-scale screenings and further strain characterization using qRT-PCR for copy number determination is a useful tool to reduce bioreactor cultivations to a reasonable number.

Figure 4 Time-course of protein concentration during fed-batch cultivations. Gene sequences were optimized by applying the high CAI codons for methanol induced genes (HM) [11]. Copy numbers (CN) are specified in the legend. Dotted lines indicate previously obtained results (compare Figure 2). Panel A: *Tr*CBH2 expression under the control of P(AOX1) or the synthetic promoters En or De [6]. Panel B: *Tr*bMan expression under the control of P(De) and P(En). For a visual representation of the protein yields during the fermentation runs please refer to the virtual gels presented in Additional file 1 and the raw data can be found in Additional file 2

Although *Trichoderma reesei* typically can produce more than 100 g/l of cellulases [25], individual enzymes such as *Tr*CBH2 are expressed in much lower quantities (10-15%) [26]. Table 1 gives an overview of published expression yields and activities of the different lignocellulolytic enzymes in different host systems. So far, Miettinen-Oinonen *et al*. achieved

the highest protein concentration of 0.7 g/l *Tr*CBH2 in *T. reesei* strains cultivated in shake flasks [27] which is around 9-fold less than compared to our highest cellobiohydrolase concentration in *Pichia pastoris* bioreactor cultures. For other heterologous host systems such as *S. cerevisiae* [28] and *S. pombe* [29] even lower *Tr*CBH2 concentrations in the range of 0.1 g/l have been reported. Regarding the specific activities of *Tr*CBH2, we obtained 3.04 U/mg on Avicel, 5.30 U/mg on PASC and 1.51 U/mg on CMC whereas 2.52 U/mg on PASC and 0.09 U/mg on CMC have been reported for the *S. pombe* system [29].

Table 1 Protein yields and enzymatic activities of expressed lignocellulolytic enzymes

Enzyme	Host	Yield (g/l)	Activity (U/ml)	Spec. Activity (U/mg)	Reference
			18.21[1]	3.04	this study
*Tr*CBH2	*P. pastoris* [B]	5.984	31.70[2]	5.30	
			9.05[3]	1.51	
*Tr*CBH2	*S. cerevisiae* [B]	0.1	n.d.	n.d.	[28]
*Tr*CBH2	*T. reesei* [A]	0.7	n.d.	n.d.	[27]
*Tr*CBH2	*S. pombe* [A]	0.115	0.29[2]	2.52	[29]
			0.01[3]	0.09	
*Tr*bMan	*P. pastoris* [B]	1.142	109[4]	95.45	this study
*Tr*bMan	*S. cerevisiae* [A]	0.000150	0.01[4]	66.67	[30]
*Tr*bMan	*T. reesei* [A]	n.d.	85.85[4]	n.d.	[30]
*Tr*bMan	*T. reesei* [A]	n.d.	1.8[4]	n.d.	[31]
*Tl*XynA	*P. pastoris* [B]	1.2	138[5]	115.00	this study
*Tl*XynA	*P. pastoris* [A]	0.148	40.2[5]	271.62	[32]
*Tl*XynA	*T. lanuginosus* [A]	0.270	88.5[5]	327.78	[32]
*Tl*XynA	*P. pastoris* [A]	0.236	26.8[5]	113.56	[33]

[A] Shake flask cultivation
[B] fed-batch cultivation
[n.d] not determined
[1] avicel
[2] phosphoric-acid-swollen-cellulose (PASC)
[3] carboxymethylcellulose (CMC)
[4] locust bean gum
[5] birchwood xylan

To evaluate *P. pastoris's* capability for expressing various other lignocellulolytic enzymes we also expressed xylanase A from *Thermomyces lanuginosus* (*Tl*XynA), beta-mannase from *Trichoderma reesei* (*Tr*bMan) and cellobiohydrolase 1 from *Trichoderma reesei* (*Tr*CBH1). All genes were optimized using the in-house codon usage table and subsequently cloned downstream of the synthetic promoter P(En) [6]. *Tr*bMan was also cloned downstream of the constitutive P(GAP) promoter. For *Tl*XynA only the strong promoter P(En) was selected to push its already proven high expression in *Pichia pastoris* with the native AOX1 promoter [34]. Similar to the experiments for *Tr*CBH2 high-throughput deep-well plate screenings were performed with the adapted pHBAH-assay and the CNs were determined by qRT-PCR. Five *Tr*bMan P(En) strains harboring 1, 4, 6 ± 1, 16 ± 4 and 39 ± 9 copies, one single copy strain under the control of P(GAP) and three *Tl*XynA P(En) strains with 6 ± 1, 10 ± 3, and 18 ± 4 copies were fed-batch cultivated. Although we successfully expressed *Tr*CBH1 in the micro-scale screening bioreactor fermentations yielded similar low protein concentrations for heterologously protein expression in the range of a few mg per liter as previously reported in

literature [28,35,36]. Therefore, those strains were not characterized in more detail for this paper.

Although the initial micro-scale screening revealed expression of *Tr*bMan under the control of P(AOX1) and P(GAP) over a broad range of CNs only the single copy strain of *Tr*bMan P(En) successfully produced *Tr*bMan in the bioreactor (Figure 2). All other P(En) regulated strains with more than one copy had major growth problems shortly after induction resulting in attenuated growth (data not shown) when grown under our standard cultivation conditions. In contrast, the P(En)-*Tr*bMan-HM-CN1 strain showed normal growth after recovering from an initial cessation of growth post methanol-induction (data not shown). Under constitutive expression of *Tr*bMan using P(GAP) the growth rate was slowed down and no *Tr*bMan was produced even though just a single expression cassette was integrated into the *Pichia* genome (data not shown). This could be a further example of the potentially cyto-toxic effects [14] of constitutive heterologous protein expression with P(GAP) in *P. pastoris*. In addition, *Tr*bMan seems to be generally difficult to express in yeasts under constitutive promoters. As an example, *Tr*bMan was only produced at a level of 0.150 mg/l in *Saccharomyces cerevisiae* [30] under the control of the constitutive phosphoglycerate kinase (PGK) promoter. Our experiments revealed that the molecular weight ratio of glycosylated and deglycosylated *Tr*bMan was about 4 (determined by capillary electrophoresis, data not shown). Therefore, hyper-glycosylation of *Tr*bMan in *P. pastoris* might be another problem for expression.

To test if a weaker methanol-inducible promoter can increase the productivity we also tested the synthetic promoter P(De) for *Tr*bMan. As seen in Figure 4B, over the whole induction period the strain P(De)-*Tr*bMan-HM-CN5 ± 1 achieved significantly more protein than compared to P(En)-*Tr*bMan-HM-CN1 (Figure 4B). As previously seen in deep well experiments [6] P(De) has a weak onset of expression during the glucose depletion-derepression phase which could also be presumed for bioreactor cultivations. Based on that, we further assume that a weaker onset leads to a better adaptation of the *Pichia* system for the production of *Tr*bMan. In addition, it was recently shown that expression under the control of P(De) can result in positive effects on cell physiology compared to expression under the control of P(AOX1) [37]. Consequently, P(De)-*Tr*bMan-HM-CN5 ± 1 was capable of producing 1.142 g/l of protein (Figure 4B) devoid of any directly observable growth problems during fermentation (data not shown). The obtained yield of 1.142 g/l of *Tr*bMan is ~7600-fold higher than the so far highest reported heterologous yield of 0.150 mg/l expressed in *Saccharomyces cerevisiae* [30] (Table 1). The activity of *Tr*bMan expressed in our study was 109 U/ml using locust bean gum as substrate. This was similar to the results of Stalbrand *et al*. [30] who obtained ~86 U/ml in *Trichoderma reesei* shake flask cultures. However, *T. reesei* shake flask cultivations performed by Hagglund *et al*. [31] showed an activity of 1.8 U/ml. This value is about 60 times less than compared to our results. Comparing heterologously expressed *Tr*bMan our obtained activity of 109 U/ml is approximately 11000-fold higher than the 0.01 U/ml expressed in *S. cerevisiae* [30]. Regarding the specific activities for *Tr*bMan, we achieved 95.4 U/mg compared to 66.7 U/mg for *Tr*bMan expressed in *S. cerevisiae* [30] (see Table 1) which usually shows an even higher tendency for hyper mannosylations that could limit the activity of *Tr*bMan.

For the fourth target, *Tl*XynA, fed-batch bioreactor cultivation of *Pichia* strains regulated by P(En) and harboring 6 or 10 integrated expression cassettes produced around 1 g/l whereas one strain with 18 integrated expression cassettes showed a reduced protein concentration of 0.25 g/l (Figure 2). As mentioned before, such negative correlation at higher copy numbers

and productivity had already been described in literature for other proteins expressed in *Pichia* [15,23].

Our yield of 1.2 g/l *Tl*XynA represents a 5 to 8-fold increase in yield compared to earlier expression studies in *Pichia pastoris* by Damaso *et. al.* [32] and Gaffney *et. al.* [33], respectively. Compared to *T. lanuginosus* shake flask cultures we achieved about 4 times more protein than reported before in [32] (Table 1). Our obtained specific activity of 115.00 U/mg was similar to the specific activities of 113.56 U/mg and 271.62 U/mg that were obtained by *Pichia pastoris*, [33] and [32] respectively. The specific activity of homologously expressed *Tl*XynA of 327.8 U/mg [32] was approximately 3-fold higher than compared to our obtained values. The comparison to homologously expressed *Tl*XynA indicates that the enzyme produced in *P. pastoris* showed lower specific activity although total volumetric yields were higher.

Generally, we speculate that the variation in specific activities of all enzymes could predominantly be attributed to the different glycosylation pattern that is produced by *P. pastoris* [38]. This phenomenon has already been described in literature *e.g.* by Macauly-Patrick *et.al.* [14].

Unfortunately, there are only limited bioreactor cultivations reported for *Tr*CBH2, *Tr*bMan, and *Tl*XynA, therefore, above made direct comparison of bioreactor results to published shake flask expression experiments are biased. However, we can still conclude that homologous expression yielded the highest specific activities but not necessarily the highest total protein yields. Although *P. pastoris* is an excellent host for achieving high protein concentrations heterologous expression can also influence the activity of the expressed enzymes. Nevertheless, comparing the calculated specific activities from Table 1 there is a general trend that the specific activities of the enzymes produced by *P. pastoris* are in the range or even higher than the specific activities of the same enzymes expressed in other heterologous hosts. This makes *Pichia* a good compromise for the expression of high quantities of enzymes with relatively high specific activities. Furthermore, it also shows the possible relevance of host strain glyco-engineering for industrial enzyme production as it already has for the production of biologically active pharmaceutical proteins.

Conclusions

We have successfully constructed *P. pastoris* strains capable of producing maximum protein concentrations of 1.142 g/l *Tr*bMan, 6.55 g/l *Tr*CBH2, and 1.2 g/l *Tl*XynA in fed-batch bioreactor cultivations. Moreover, we showed that suitable codon optimization of the target genes helps to increase heterologous protein production by *P. pastoris*, thus providing a simple way of increasing heterologous protein production for individual enzymes.

Furthermore, we emphasize the importance of transcript level optimization by alternative promoters and gene dosage (numbers of integrated gene copies) for expression optimization. This was particularly evident for the functional expression of *Tr*bMan. The strong constitutive and methanol inducible promoters P(GAP) and, P(AOX1) respectively, secreted no or less protein than the weaker synthetic promoter P(De).

Basically there are three classes of genes (A,B,C) with varying dependence of yields of active proteins in relation to copy numbers: For class A genes an increase in copy number to more

than 10 copies has a positive effect on protein expression, as seen in the case of *Tr*CBH2. For class B genes the yield of active protein increases within the number of integrated copies up to a copy number of 2–10 and decreases with higher copy numbers, as seen in the case of *Tl*XynA. Finally, class C genes where yields of active protein get worse with increasing copy numbers, as seen in the case of *Tr*bMan. However, these effects definitely depend on the strength of the employed promoter as well as the gene encoding the respective target protein.

Our conclusions are based on a better understanding of promoter and/or copy number-related effects. Codon-optimized genes together with optimized promoters and numbers of integrated expression cassettes allowed us to develop *P. pastoris* strains producing high levels of lignocellulolytic enzymes. In combination with the high specific activities compared to the same enzymes expressed in other hosts, *Pichia* seems to be a good choice for the heterologous expression of individual lignocellulolytic enzymes.

Methods

Chemicals and Materials

Oligonucleotide primers were obtained from Integrated DNA Technologies (Leuven, Belgium). For plasmid isolation the GeneJET™ Plasmid Miniprep Kit of Fermentas (Burlington, Ontario, Canada) was used. All DNA-modifying enzymes were obtained from Fermentas GmbH (Burlington, Ontario, Canada). Chemicals were purchased if not stated otherwise from Becton, Dickinson and Company (Franklin Lakes, NJ, USA), Fresenius Kabi Austria (Graz, Austria), Carl Roth (Karlsruhe, Germany), and Sigma- Aldrich (St Louis, MO, USA). p-hydroxybenzoic acid hydrazide (order no. 54600) were obtained from Fluka (Hamburg, Deutschland). D-(+)-mannose and D-(+)-cellobiose were from Fluka, D-(+)-xylose from Sigma, D-(+)-glucose monohydrate from Carl Roth (Karlsruhe, Germany).

Media

For *E. coli* standard LB-medium containing 25 μg/ml zeocin was used. YPD for *P. pastoris* contained 10 g/l yeast extract, 20 g/l peptone and 20 g/l glucose. For antibiotic selection 100 μg/ml zeocin were used. 15 g/l agar was added for plate media. Buffered minimal media BMD (1%), BMM2 and BMM10 consisted per liter of 200 ml 1 M potassium phosphate buffer (pH 6), 13.4 g yeast nitrogen base without amino acids, 0.0004 g/l biotin and 11 g/l glucose or 1 or 5% (v/v) methanol, respectively. All pre-cultures were prepared using YPhyD medium containing 20 g/l Phytone-Peptone, 10 g/l Bacto-Yeast Extract and 20 g/l glucose. BSM medium contained per liter $CaSO_4$_$2H_2O$ 0.47 g, K_2SO_4 9.1 g, KOH 2.07 g, $MgSO_4$_$7H_2O$ 7.5 g, EDTA 0.6 g, H_3PO_4 (85%) 13.4 ml, Glycerol 40.0 g, NaCl 0.22 g and 4.35 ml PTM1. PTM1 Trace elements solution contained per liter 0.2 g Biotin, 6.0 g $CuS0_4$_$5H_2O$, 0.09 g KI, 3.0 g $MnSO_4$_H_2O, 0.2 g Na_2MoO_4_$2H_2O$, 0.02 g H_3BO_3, 0.5 g $CoCl_2$, 42,2 g $ZnSO_4$_7H2O, 65 g Fe(II)SO_4_$7H_2O$ and 5 ml H_2SO_4. The fed-batch feed media were either 60% (w/w) Glycerol or concentrated MeOH and were supplemented with 12 ml/l PTM1 mineral salts solution.

Construction of *P. pastoris* strains

The coding sequences of xylanase A from *Thermomyces lanuginosus* (*Tl*XynA) [UniProtKB/Swiss-Prot: O43097], beta-mannanase (*Tr*bMan) [UniProtKB/TrEMBL: Q99036], cellobiohydrolase 1 (*Tr*CBH1) [UniProtKB/Swiss-Prot: P62694] and cellobiohydrolase 2 (*Tr*CBH2) [UniProtKB/Swiss-Prot: P07987] from *Trichoderma reesei* inclusive of their natural secretion leaders were codon optimized for *P. pastoris* expression applying the Gene Designer software (DNA2.0, Menlo Park, CA, USA) based on an in-house developed codon bias [11]. The GC content was set to be between 40 and 60% without local peaks and restriction sites for cloning were avoided. In addition, one other variant of *Tr*CBH2 was ordered from a commercial supplier (CODA Genomics, Laguna Hills, CA) by which optimization was done with the method of codon pair signaling [21]. The native DNA sequence was kindly provided by Frances H. Arnold. To further optimize translation all genes were cloned after a defined Kozak consensus sequence (gaaacg) [39]. The synthetic genes were cloned into the multiple cloning site of the *E. coli*/*P. pastoris* shuttle vector pPpB1 [11] via *EcoR*I/*Not*I restriction sites. The *Tr*CBH2 variants were cloned downstream of the wild type promoters P(GAP) and P(AOX1) and synthetic promoter variants with distinctly different regulation patterns were also included, namely P(En) and P(De). P(En) can be induced by methanol and showed increased GFP expression up to 166%, if compared to the wild-type promoter P(AOX1). P(De) can either be induced by methanol or under derepressed conditions as described by Hartner *et al*. [6]. *Tr*bMan and *Tl*XynA were cloned downstream of the synthetic promoter P(En) [6] and in addition *Tr*bMan was cloned downstream of P(GAP) and P(De). Plasmids were linearized with *Bgl*II, subsequently purified and concentrated using the Wizard_ SV Gel and PCR Cleanup System (Promega Corp.). Electro-competent *P. pastoris* CBS 7435 mut^S cells were prepared and transformed with 1- to 2 μg of the *Bgl*II-linearized pPpB1 vector construct according to Lin-Cereghino [40]. Transformants were plated on YPD-Zeocin (100 μg/ml Zeocin) agar plates and grown at 28°C for 48 h.

Micro-scale cultivation and high-throughput screening

P. pastoris strains expressing *Tl*XynA, *Tr*bMan, and *Tr*CBH2 were cultivated in 96-deep well plates as described by Weis *et al* [18]. Incubation was done in shakers (INFORS Multitron, Bottmingen, Switzerland) at 28°C, 320 rpm, and 80% relative humidity. After an initial batch phase for 60 h on 1% glucose the cultures were induced with 0.5% of methanol for a total of 72 h (additional supplementations to 0.5% methanol were added after 12, 36 and 60 h of the first induction with methanol). After induction the cells were pelleted at 4000 rpm and enzymatic activities were determined in the supernatants using the pHBAH-assay as previously described by Mellitzer *et al*. [19]. Substrate conversions were performed in 50 mM citrate buffer containing appropriate substrate for each enzyme (either suspensions of 1% Avicel®, 0.25% PASC or solutions of 0.25% CMC, 0.5% xylan or 0.2% locust bean gum) at 50°C (*Tr*CBH2, *Tr*CBH1 and *Tr*bMan) or at 59°C (*Tl*XynA). The incubation time was 2 h for the cellobiohydrolases and 20 min for *Tl*XynA and *Tr*bMan. For the subsequent reducing sugar assay 50 μL of the substrate reaction (or, in the case of the standard sugars, appropriate dilutions of the reducing sugars) were pipetted into 150 μl of the pHBAH working solution in a 96-well PCR plate. The plate was sealed and incubated at 95°C for 5 min and then cooled to 4°C. 150 μl of the assay samples were transferred to a new micro-titer-plate and the absorption measured at 410 nm in a SPECTRA MAX Plus384 plate reader (Molecular Devices Corp., Sunnyvale, CA, USA). For exact quantification of reducing sugars a standard curve of the respective reducing sugar (0–1 mg/ml) was included on each plate. Activity units

for the expressed enzymes refer to the amount of released reducing sugar over time and correspond to the standard IUPAC definition μM/min.

Copy number determination by quantitative real-time PCR

Copy numbers of integrated expression cassettes in the *Pichia* genome were determined using quantitative real-time PCR (qRT-PCR) as described by Abad *et al.* [17].

Fed-batch cultivations of *Pichia pastoris* strains

Pre-cultures of individual strains were grown in 50 and 200 ml YPhyD in wide-necked, baffled shake flasks at 120 rpm at 28°C. Each fermenter (Infors Multifors system (Infors AG, Bottmingen, Switzerland)) containing 450 ml BSM-media (pH 5.0) was inoculated from the pre-culture to an OD600 of 2.0. During the batch phase *P. pastoris* was grown on glycerol (4%) at 28°C. At the beginning of the glycerol feeding phase the temperature was decreased to 24°C. For methanol-fed cultures, the fed-batch phase was started upon depletion of initial glycerol with 16 g/(l*h) glycerol feed solution followed by methanol induction. In the early stages, the methanol-feed was set to 2 g/(l*h) and was gradually increased within the next 70 h to 6 g/(l*h). Likewise, the glycerol-feed was phased down during the first hour of methanol induction to 0 g/(l*h). Dissolved oxygen was set to 30% throughout the whole process. After 91.5 h of methanol induction the fermentations were stopped. For glycerol-fed strains, the batch phase was directly followed by a constant glycerol-feed with 6 g/(l*h). Protein concentrations were determined by micro-fluidic capillary electrophoresis (CE) using fluorescence detection (Caliper GXII System, Hopkinton, USA). Standard deviations of this robust system are usually below 10%, even at high protein loads (exemplified in Additional file 3). Therefore, just single measurements of every sample were performed. More specifically, proteins were quantified by calibrating the integrated areas of the protein-specific peaks in the electropherograms to an external reference protein standard (BSA) of known concentration. For glycosylated proteins, peak areas of diluted deglycosylated samples were compared to those of untreated samples to compensate for glycosylation-related differences in quantification (a comparison of glycosylated and non-glycosylated enzyme samples is exemplified for *Tr*bMan in Additional file 4). The dilutions of samples were in a range to give peak areas of the samples that were comparable to those of the reference protein standard. Importantly, the absence of comparable protein peaks in the vector-only control strains further validates the quantification of the secreted enzymes (see Additional file 1).

Abbreviations

pHBAH, p-hydroxybenzoic acid hydrazide; DNS, 3,5-dinitrosalicylic acid; *Tl*XynA, *Thermomyces lanuginosus* xylanase A; *Tr*bMan, *Trichoderma reesei* beta-mannanase; *Tr*CBH2, *Trichoderma reesei* cellobiohydrolase 2; qRT-PCR, Quantitative real time polymerase chain reaction; P(GAP), Glyceraldehyde-3-phosphate dehydrogenase promoter; P(AOX1), Alcohol oxidase 1 promoter; PASC, Phosphoric acid-swollen cellulose; CMC, Carboxymethylcellulose

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

AG conceived the project and KF managed the project and helped with the experimental implementation. AG and KF both helped to draft and improve the manuscript. AM carried out all experiments except of the bioreactor cultivations and drafted the manuscript. Bioreactor cultivations were performed and analyzed by RW. All authors have read and approved the final manuscript.

Acknowledgements

We would like to thank VTU Technology for scientific and financial support. Furthermore, the authors acknowledge the FFG, the Province of Styria the SFG and the Austrian COMET program for financial support.

References

1. Cereghino GP, Cregg JM: **Applications of yeast in biotechnology: protein production and genetic analysis.** *Curr Opin Biotechnol* 1999, **10**(5):422–427.

2. Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS: **Microbial cellulose utilization: fundamentals and biotechnology.** *Microbiol Mol Biol Rev* 2002, **66**(3):506–577. table of contents.

3. Cereghino JL, Cregg JM: **Heterologous protein expression in the methylotrophic yeast Pichia pastoris.** *FEMS Microbiol Rev* 2000, **24**(1):45–66.

4. Cregg JM, Vedvick TS, Raschke WC: **Recent advances in the expression of foreign genes in Pichia pastoris.** *Biotechnology (N Y)* 1993, **11**(8):905–910.

5. Tolner B, Smith L, Begent RH, Chester KA: **Production of recombinant protein in Pichia pastoris by fermentation.** *Nat Protoc* 2006, **1**(2):1006–1021.

6. Hartner FS, Ruth C, Langenegger D, Johnson SN, Hyka P, Lin-Cereghino GP, Lin-Cereghino J, Kovar K, Cregg JM, Glieder A: **Promoter library designed for fine-tuned gene expression in Pichia pastoris.** *Nucleic Acids Res* 2008, **36**(12):e76.

7. Welch M, Govindarajan S, Ness JE, Villalobos A, Gurney A, Minshull J, Gustafsson C: **Design parameters to control synthetic gene expression in Escherichia coli.** *PLoS One* 2009, **4**(9):e7002.

8. Gasser B, Sauer M, Maurer M, Stadlmayr G, Mattanovich D: **Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts.** *Appl Environ Microbiol* 2007, **73**(20):6499–6507.

9. Inan M, Aryasomayajula D, Sinha J, Meagher MM: **Enhancement of protein secretion in Pichia pastoris by overexpression of protein disulfide isomerase.** *Biotechnol Bioeng* 2006, **93**(4):771–778.

10. Davis JJ, Olsen GJ: **Modal codon usage: assessing the typical codon usage of a genome.** *Mol Biol Evol* 2009, **27**(4):800–810.

11. Abad S, Nahalka J, Bergler G, Arnold SA, Speight R, Fotheringham I, Nidetzky B, Glieder A: **Stepwise engineering of a Pichia pastoris D-amino acid oxidase whole cell catalyst.** *Microb Cell Fact* 2010, **9**(26):24.

12. 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* 1985, **5**(5):1111–1121.

13. Koutz P, Davis GR, Stillman C, Barringer K, Cregg J, Thill G: **Structural comparison of the Pichia pastoris alcohol oxidase genes.** *Yeast* 1989, **5**(3):167–177.

14. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM: **Heterologous protein production using the Pichia pastoris expression system.** *Yeast* 2005, **22**(4):249–270.

15. Hohenblum H, Gasser B, Maurer M, Borth N, Mattanovich D: **Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant Pichia pastoris.** *Biotechnol Bioeng* 2004, **85**(4):367–375.

16. Mattanovich D, Gasser B, Hohenblum H, Sauer M: **Stress in recombinant protein producing yeasts.** *J Biotechnol* 2004, **113**(1–3):121–135.

17. Abad S, Kitz K, Hormann A, Schreiner U, Hartner FS, Glieder A: **Real-time PCR-based determination of gene copy numbers in Pichia pastoris.** *Biotechnol J* 2010, **5**(4):413–420.

18. 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):179–189.

19. Mellitzer A: **Sensitive high throughput screening for detecting reducing sugars.** *Biotechnol J* 2011 **:manuscript accepted**(6).

20. Vassileva A, Chugh DA, Swaminathan S, Khanna N: **Effect of copy number on the expression levels of hepatitis B surface antigen in the methylotrophic yeast Pichia pastoris.** *Protein Expr Purif* 2001, **21**(1):71–80.

21. Irwin B, Heck JD, Hatfield GW: **Codon pair utilization biases influence translational elongation step times.** *J Biol Chem* 1995, **270**(39):22801–22806.

22. Clare JJ, Romanos MA, Rayment FB, Rowedder JE, Smith MA, Payne MM, Sreekrishna K, Henwood CA: **Production of mouse epidermal growth factor in yeast: high-level secretion using Pichia pastoris strains containing multiple gene copies.** *Gene* 1991, **105**(2):205–212.

23. Zhu T, Guo M, Zhuang Y, Chu J, Zhang S: **Understanding the effect of foreign gene dosage on the physiology of Pichia pastoris by transcriptional analysis of key genes.** *Appl Microbiol Biotechnol* 2011, **89**(4):1127–1135.

24. Graf A, Gasser B, Dragosits M, Sauer M, Leparc GG, Tuchler T, Kreil DP, Mattanovich D: **Novel insights into the unfolded protein response using Pichia pastoris specific DNA microarrays.** *BMC Genomics* 2008, **9**(19):390.

25. Xu Q, Singh A, Himmel ME: **Perspectives and new directions for the production of bioethanol using consolidated bioprocessing of lignocellulose.** *Curr Opin Biotechnol* 2009, **20**(3):364–371.

26. Kubicek-Pranz EM, Gruber F, Kubicek CP: **Transformation of Trichoderma reesei with the cellobiohydrolase II gene as a means for obtaining strains with increased cellulase production and specific activity.** *J Biotechnol* 1991, **20**(1):83–94.

27. Miettinen-Oinonen A, Paloheimo M, Lantto R, Suominen P: **Enhanced production of cellobiohydrolases in Trichoderma reesei and evaluation of the new preparations in biofinishing of cotton.** *J Biotechnol* 2005, **116(3):**305–317.

28. Penttila ME, Andre L, Lehtovaara P, Bailey M, Teeri TT, Knowles JK: **Efficient secretion of two fungal cellobiohydrolases by Saccharomyces cerevisiae.** *Gene* 1988, **63(1):**103–112.

29. Okada H, Sekiya T, Yokoyama K, Tohda H, Kumagai H, Morikawa Y: **Efficient secretion of Trichoderma reesei cellobiohydrolase II in Schizosaccharomyces pombe and characterization of its products.** *Appl Microbiol Biotechnol* 1998, **49**(3):301–308.

30. Stalbrand H, Saloheimo A, Vehmaanpera J, Henrissat B, Penttila M: **Cloning and expression in Saccharomyces cerevisiae of a Trichoderma reesei beta-mannanase gene containing a cellulose binding domain.** *Appl Environ Microbiol* 1995, **61**(3):1090–1097.

31. Hagglund P, Eriksson T, Collen A, Nerinckx W, Claeyssens M, Stalbrand H: **A cellulose-binding module of the Trichoderma reesei beta-mannanase Man5A increases the mannan-hydrolysis of complex substrates.** *J Biotechnol* 2003, **101**(1):37–48.

32. Damaso MC, Almeida MS, Kurtenbach E, Martins OB, Pereira N Jr, Andrade CM, Albano RM: **Optimized expression of a thermostable xylanase from Thermomyces lanuginosus in Pichia pastoris.** *Appl Environ Microbiol* 2003, **69**(10):6064–6072.

33. Gaffney M, Carberry S, Doyle S, Murphy R: **Purification and characterisation of a xylanase from Thermomyces lanuginosus and its functional expression by Pichia pastoris.** *Enzyme Microb Technol* 2009, **45**(5):348–354.

34. Schlacher A, Holzmann K, Hayn M, Steiner W, Schwab H: **Cloning and characterization of the gene for the thermostable xylanase XynA from Thermomyces lanuginosus.** *J Biotechnol* 1996, **49**(1–3):211–218.

35. Boer H, Teeri TT, Koivula A: **Characterization of Trichoderma reesei cellobiohydrolase Cel7A secreted from Pichia pastoris using two different promoters.** *Biotechnol Bioeng* 2000, **69**(5):486–494.

36. Den Haan R, McBride JE, Grange DCL, Lynd LR, Van Zyl WH: **Functional expression of cellobiohydrolases in Saccharomyces cerevisiae towards one-step conversion of cellulose to ethanol.** *Enzyme Microb Technol* 2007, **40**(5):1291–1299.

37. Hyka P, Züllig T, Ruth C, Kovar K: **Combined Use of Fluorescent Dyes and Flow Cytometry To Quantify the Physiological State of Pichia pastoris during the Production of Heterologous Proteins in High-Cell-Density Fed-Batch Cultures.** *Appl Environ Microbiol* 2010, **76**(13):4486–4496.

38. 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, **140**(Pt 1):159–166.

39. Kozak M: **An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs.** *Nucleic Acids Res* 1987, **15**(20):8125–8148.

40. Lin-Cereghino J, Wong WW, Xiong S, Giang W, Luong LT, Vu J, Johnson SD, Lin-Cereghino GP: **Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast Pichia pastoris.** *Biotechniques* 2005, **38**(1):44. 46, 48.

Additional files

Additional file_1_as PDF
Additional file 1 Supplementary figures. These figures provide virtual protein gels for each of the expressed enzymes at different time points during the fermentation runs and give a relative estimation of the purity of the expressed enzymes.

Additional file_2_as PDF
Additional file 2 These tables provide the individual data points of the measured target protein concentrations for the Figures 2, 3 and 4.

Additional file_3_as PDF
Additional file 3 This figure provides a comparison of a triplicate measurement of a *Tr*CBH2 sample to exemplify the accurateness of the detection method.

Additional file_4_as PDF
Additional file 4 Deglycosylation of *Tr*bMan. These figures provide a comparison of glycysylated and EndoH-deglycosylated protein samples of *Tr*bMan.

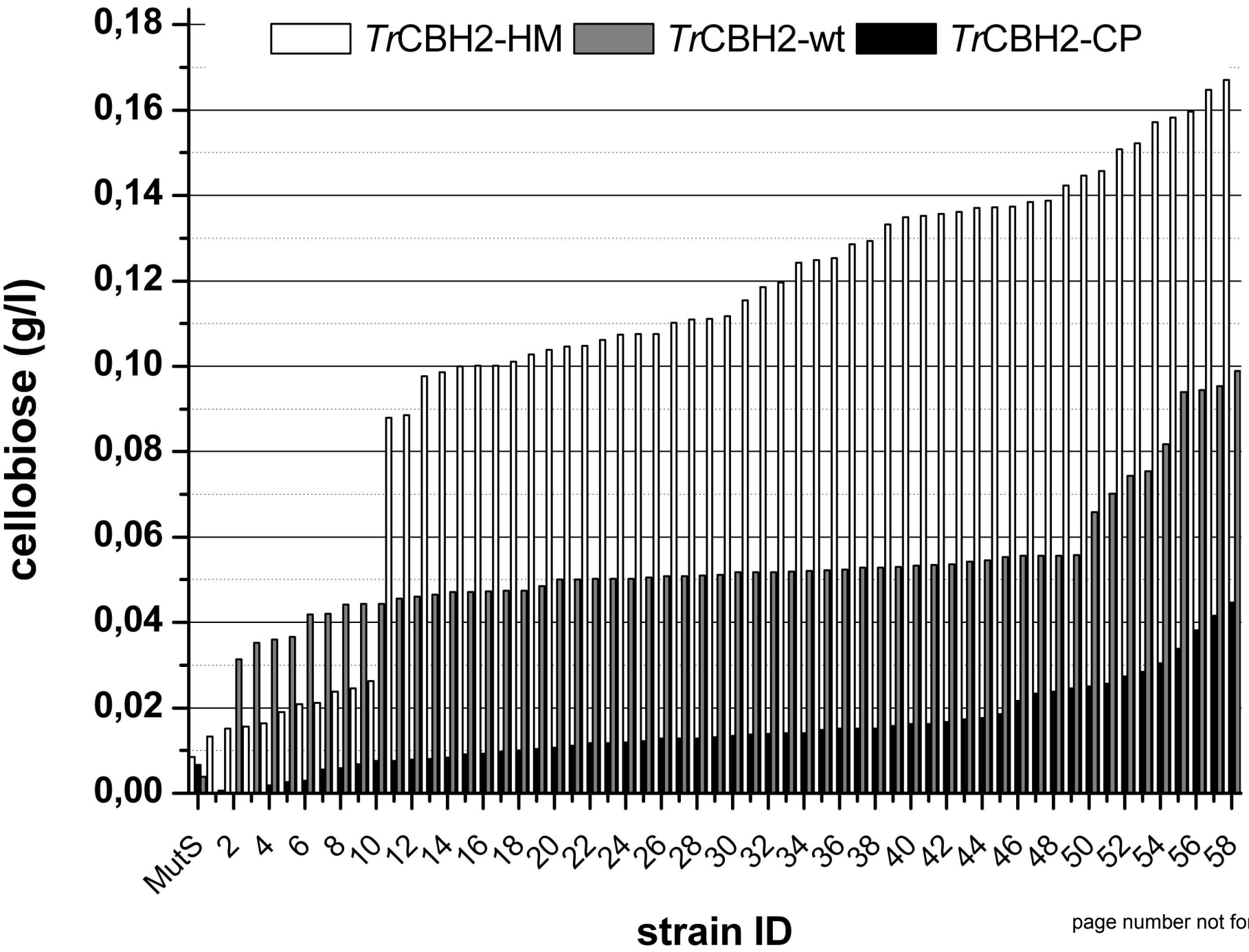

TrCBH2-HM
TrCBH2-wt
TrCBH2-CP
cellobiose (g/l)
0,18
0,16
0,14
0,12
0,10
0,08
0,06
0,04
0,02
0,00
MutS
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
40
42
44
46
48
50
52
54
56
58
strain ID

Figure 1

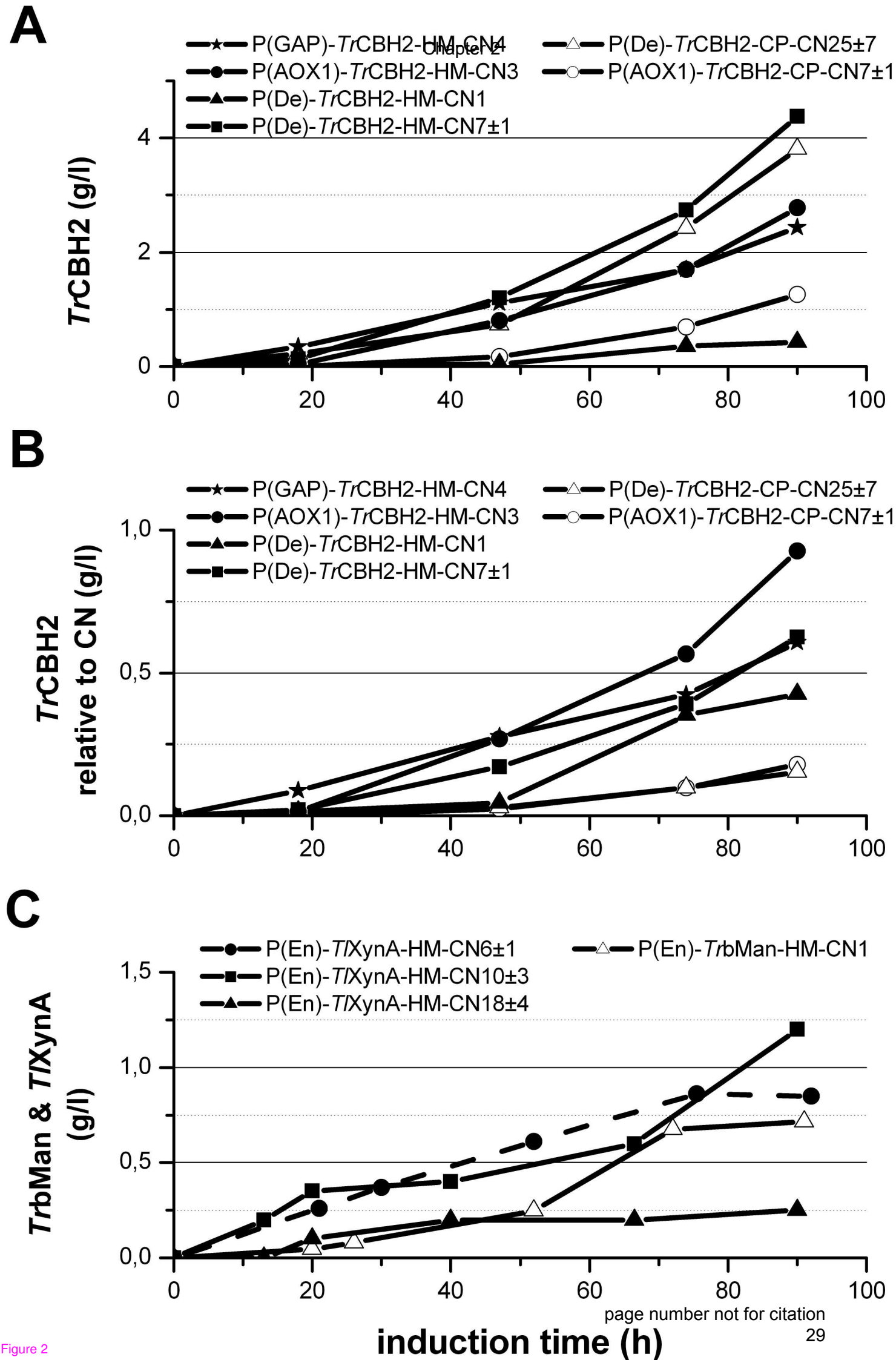
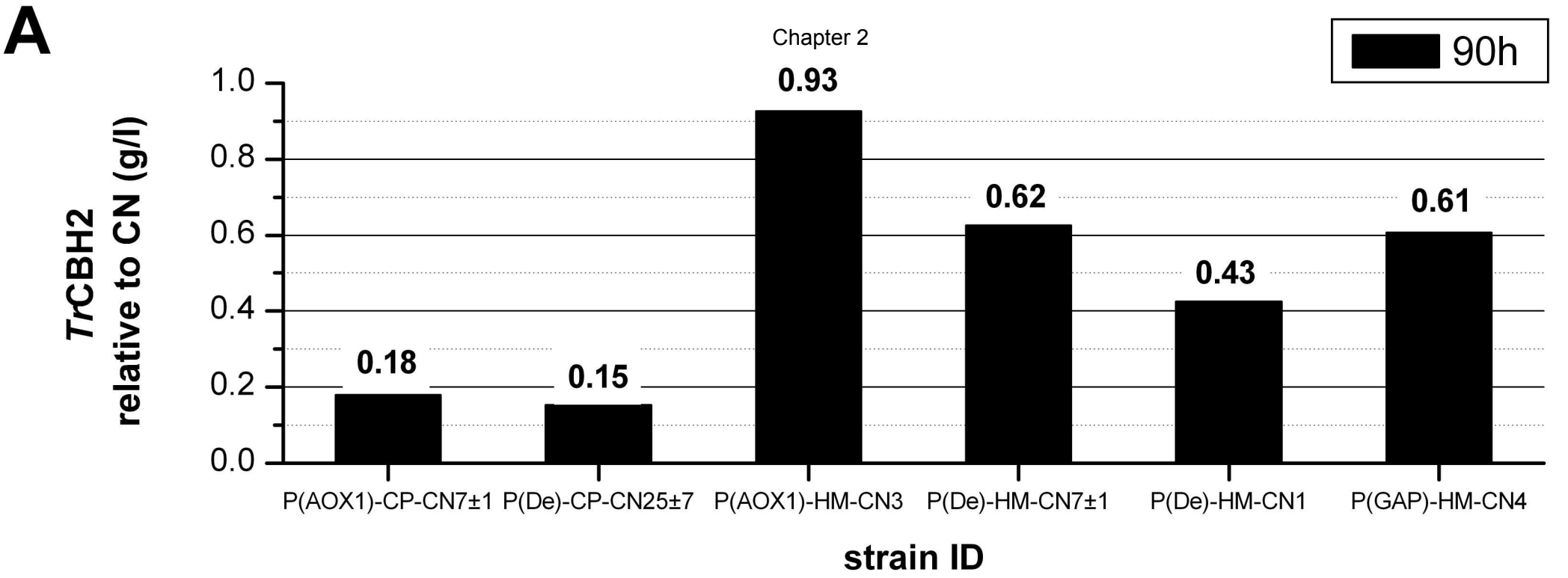
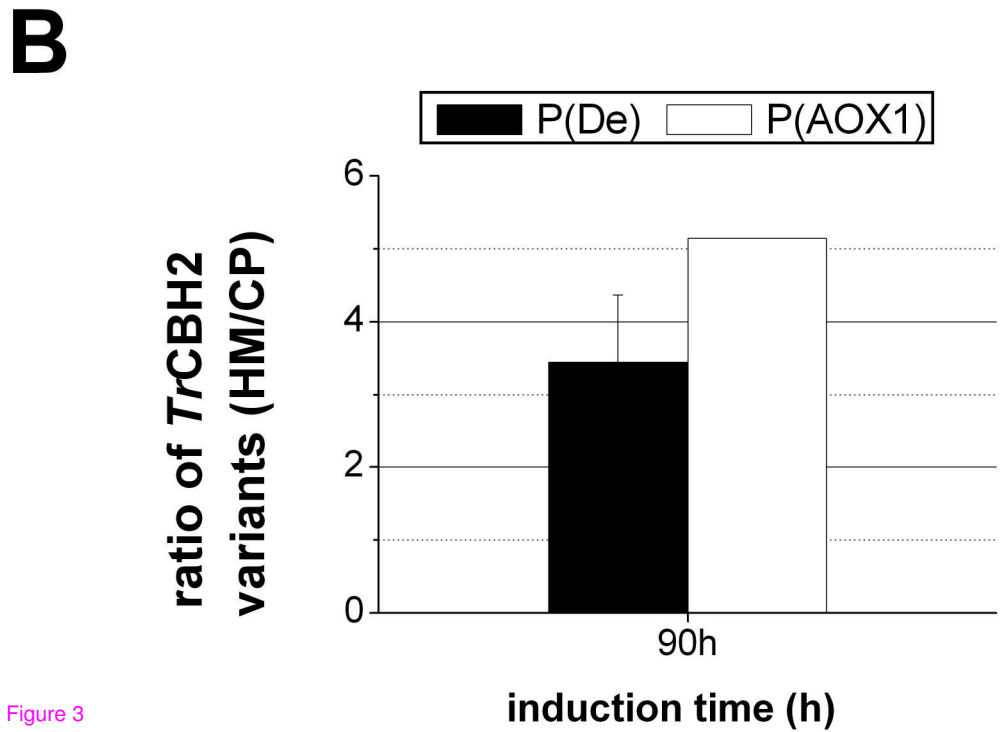

A
P(GAP)-TrCBH2-HM-CN4
P(De)-TrCBH2-CP-CN25±7
P(AOX1)-TrCBH2-HM-CN3
P(AOX1)-TrCBH2-CP-CN7±1
P(De)-TrCBH2-HM-CN1
P(De)-TrCBH2-HM-CN7±1
TrCBH2 (g/l)
0
2
4
0
20
40
60
80
100
B
P(GAP)-TrCBH2-HM-CN4
P(De)-TrCBH2-CP-CN25±7
P(AOX1)-TrCBH2-HM-CN3
P(AOX1)-TrCBH2-CP-CN7±1
P(De)-TrCBH2-HM-CN1
P(De)-TrCBH2-HM-CN7±1
TrCBH2 relative to CN (g/l)
0,0
0,5
1,0
0
20
40
60
80
100
C
P(En)-TlXynA-HM-CN6±1
P(En)-TrbMan-HM-CN1
P(En)-TlXynA-HM-CN10±3
P(En)-TlXynA-HM-CN18±4
TrbMan & TlXynA (g/l)
0,0
0,5
1,0
1,5
0
20
40
60
80
100
induction time (h)

Figure 2


A
90h
TrCBH2 relative to CN (g/l)
1.0
0.8
0.6
0.4
0.2
0.0
0.18
0.15
0.93
0.62
0.43
0.61
P(AOX1)-CP-CN7±1
P(De)-CP-CN25±7
P(AOX1)-HM-CN3
P(De)-HM-CN7±1
P(De)-HM-CN1
P(GAP)-HM-CN4
strain ID


B
P(De)
P(AOX1)
ratio of TrCBH2 variants (HM/CP)
6
4
2
0
90h
induction time (h)

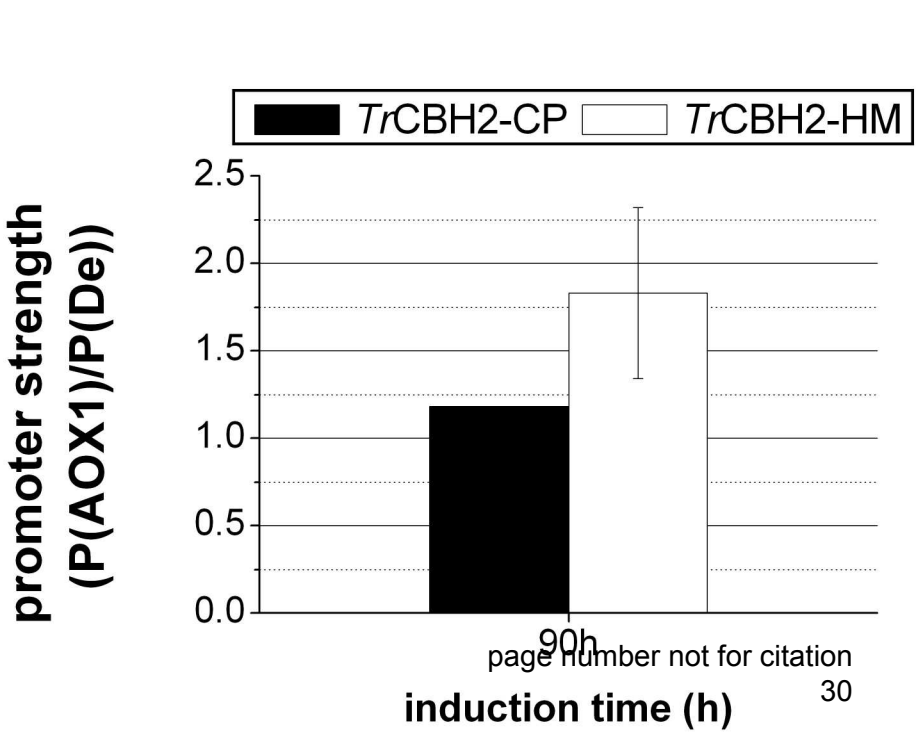

C
TrCBH2-CP
TrCBH2-HM
promoter strength (P(AOX1)/P(De))
2.5
2.0
1.5
1.0
0.5
0.0
90h
induction time (h)

Figure 3

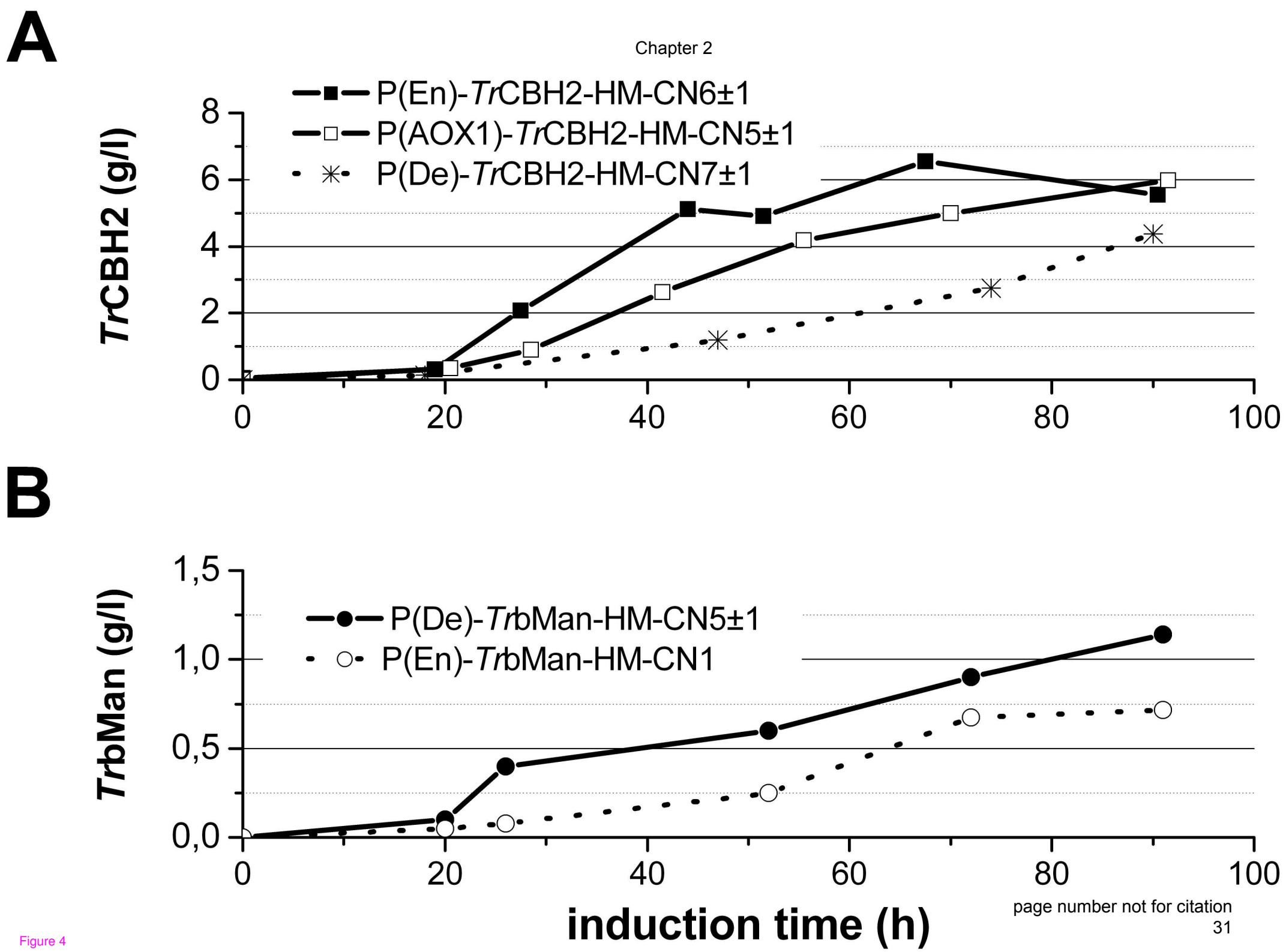

A
P(En)-TrCBH2-HM-CN6±1
P(AOX1)-TrCBH2-HM-CN5±1
P(De)-TrCBH2-HM-CN7±1
TrCBH2 (g/l)
0
2
4
6
8
0
20
40
60
80
100
B
P(De)-TrbMan-HM-CN5±1
P(En)-TrbMan-HM-CN1
TrbMan (g/l)
0,0
0,5
1,0
1,5
0
20
40
60
80
100
induction time (h)

Figure 4

Additional files provided with this submission:

Additional file 1: Additional file 1.pdf, 306K
http://www.microbialcellfactories.com/imedia/1785819385673073/supp1.pdf
Additional file 2: Additional file 2.pdf, 188K
http://www.microbialcellfactories.com/imedia/1000325836673073/supp2.pdf
Additional file 3: Additional file 3.pdf, 115K
http://www.microbialcellfactories.com/imedia/1509000764705726/supp3.pdf
Additional file 4: Additional File 4.pdf, 133K
http://www.microbialcellfactories.com/imedia/1347726795705727/supp4.pdf

CHAPTER 3

High level protein expression in *Pichia pastoris* combining synthetic promoters and synthetic genes

Mellitzer A. (1), Ruth C. (1), Gustafsson C. (2), Welch M. (2), Birner-Grünberger R. (3), Weis R. (4), Glieder A*. (1)

*Corresponding author
Email: anton.glieder@acib.at

[1]*ACIB GmbH c/o IMBT TUGraz, Austrian Centre of Industrial Biotechnology, Graz, Austria*
[2]*DNA2.0 Inc., Menlo Park, CA 94025, USA*
[3]*Center For Medical Research, Graz, Austria;*
[4]*VTU Technology GmbH, Grambach, Austria*

High level protein expression in *Pichia pastoris* combining synthetic promoters and synthetic genes

Mellitzer A. (1), Ruth C. (1), Gustafsson C. (2), Welch M. (2), Birner-Grünberger R. (3), Weis R. (4), Glieder A. (1)*

*Corresponding author

Email: anton.glieder@acib.at

[1]*ACIB GmbH c/o IMBT TUGraz, Austrian Centre of Industrial Biotechnology, Graz, Austria*

[2]*DNA2.0 Inc., Menlo Park, CA 94025, USA*

[3]*Center For Medical Research, Graz, Austria;*

[4]*VTU Technology GmbH, Grambach, Austria*

Keywords

Pichia Pastoris, synonymous codon substitution, synthetic promoters, gene dosage, landscape

ABSTRACT

Within the last years we developed a new toolbox to further enhance and facilitate protein expression by *Pichia pastoris*. Herein we demonstrate the beneficial combination of several of these tools for protein expression. A set of 48 lignocellulolytic genes was synthesized distinguished by synonymous codon substitution. This set was expressed under the control of the wild type P(GAP) as well as newly available synthetic promoters DeS and En. Putative single copy landscapes were generated by the released cellobiose concentration and varied from undetectable to ~600% of the initial activity. Transcript levels and protein yields revealed the superior performance of P(DeS) in fed-batch cultivations. We demonstrated that P(DeS) can be tightly regulated by the available carbon source type and level. Finally the individual results were combined to engineer high *P. pastoris* productivity strains. Therefore, strains with defined copy numbers expressing an improved gene variant under the control of P(DeS) were employed. Although no methanol was used for induction we obtained more than 18 g/L of secreted protein. This is the highest concentration of secreted protein reported in *P. pastoris* so far. Concluding, we were able to improve extracellular cellulose production by *Pichia pastoris* with the help of synthetic promoters and synthetic codon optimized genes.

Introduction

The yeast *Pichia pastoris* is a unicellular eukaryotic microorganism which harbors several advantageous compared to other systems. Since it is a eukaryote many post- translational modifications such as proteolytic processing, correct folding, disulfide bond formation and glycosylation can be still performed. Nevertheless, it can be quite easily genetically manipulated to get relative stable production strains without the urge of applying permanent selection pressure. Moreover, it can be cultivated to high cell densities in chemically defined media and does not need complex cultivation conditions [1, 2]. So far more than 600 different proteins of various organisms were successfully expressed in *Pichia pastoris* [3]. Nevertheless, to further increase its value as host systems for heterologous protein production new tools and techniques are required.

There are several factors known to strongly influence protein production in *Pichia pastoris*, like synonymous codon substitution, promoter choice and gene dosage [2, 4-6]. Regarding synonymous codon substitution, studies performed in the simpler microorganism *E. coli* obtained partly contradictory results. [7] speculated that variations of expression caused by synonymous codon substitution depend on the limiting step of tRNA recharging. Whereas [8] speculated that variations of expression caused by synonymous codon substitution depend on mRNA stability. However, both studies demonstrated that in *E. coli* variations of expression caused by synonymous codon substitution is principally regulated at transcriptional and/ or translational level. Whereas for *Pichia pastoris* different results were obtained, indicating that posttranslational processes (e.g. folding and secretion) and protein degradation of misfolded or unfolded proteins are the limiting factor for protein production [6]. It is limited due to a stress induced extra energy requirement needed for recombinant protein production which was summarized as metabolic burden (Glick, 1995). As a result Inhibition of growth or low level of product accumulation, decreased cell lysis, plasmid instability can influence the productivity of the expression system.

The most popular promoters in *Pichia pastoris* are the wild- type promoters glyceraldehyde-3-phosphate dehydrogenase (GAP) [9] and alcohol oxidase 1 (AOX1) [10] since they are strong promoters and easy to induce. However, due to the fact that P(GAP) is a constitutive promoter, it is not suited for the expression of physiologically problematic or cytotoxic proteins [2]. Unlike the P(AOX1), which can be tightly regulated by the available C-source [10, 11]. Unfortunately, the use of methanol in industrial applications is not favorable considering safety issues and methanol induced cell lysis and proteolysis [2, 3, 6, 12]. Meanwhile, alternatives to the wild- type promoters exist. A set of synthetic promoters was designed by deletion and duplication of putative transcription factor- binding sites within the

AOX1 promoter region [5]. Thereby, promoters were generated with an activity range between 6 % and 166 % of the wild- type promoter and exhibiting different regulatory features. For the P(GAP) library, error-prone PCR was used to introduce mutations in the promoter sequence. 33 Mutants were chosen for further characterization and spanned an activity range between 0.6 % and 19.6-fold of the wild- type promoter activity using yEGFP as reporter [13].

The aim of the present work was to independently characterize the effect of synonymous codon substitution, promoter choice and gene dosage in *Pichia pastoris* to find optimal conditions for heterologous protein production. Therefore, a set of differently optimized *Tr*CBH2 gene variants were expressed under the control of distinct promoters. Moreover, well defined strains were employed to further characterize the promoter performances on transcriptional level in fed- batch bio reactor cultivations. Finally the individual results were combined to engineer high *P. pastoris* productivity strains.

MATERIAL AND METHODS

Chemicals and Materials

Oligonucleotide primers were obtained from Integrated DNA Technologies (Leuven, Belgium). For plasmid isolation the GeneJET™ Plasmid Miniprep Kit of Fermentas (Burlington, Ontario, Canada) was used. All DNA-modifying enzymes were obtained from Fermentas GmbH (Burlington, Ontario, Canada). Chemicals were purchased if not stated otherwise from Becton, Dickinson and Company (Franklin Lakes, NJ, USA), Fresenius Kabi Austria (Graz, Austria) and Carl Roth (Karlsruhe, Germany). p-hydroxybenzoic acid hydrazide (order no. 54600) and D-(+)-cellobiose were obtained from Fluka (Hamburg, Deutschland). Cellobiase from *Aspergillus niger* (C6105) and cellulase from *Trichoderma reesei* ATCC 26921 (C2730) were obtained from Sigma- Aldrich (St Louis, MO, USA).

Media

For *E. coli* standard LB-medium containing 25 µg/ml zeocin was used. YPD for *P. pastoris* contained 10 g/l yeast extract, 20 g/l peptone and 20 g/l glucose. For antibiotic selection 100 µg/ml zeocin were used. 15 g/l agar was added for plate media. Buffered minimal media BMD (1%), BMM2 and BMM10 consisted per liter of 200 ml 1 M potassium phosphate buffer (pH 6), 13.4 g yeast nitrogen base without amino acids, 0.0004 g/l biotin and 11 g/l glucose or 1 or 5% (v/v) methanol, respectively. All pre-cultures were prepared using YPhyD medium containing 20 g/l Phytone-Peptone, 10 g/l Bacto-Yeast Extract and 20 g/l glucose. BSM medium contained per liter CaSO4_2H2O 0.47 g, K2SO4 9.1 g, KOH 2.07 g, MgSO4_7H2O 7.5 g, EDTA 0.6 g, H3PO4 (85%) 13.4 ml, Glycerol 40.0 g, NaCl 0.22 g and 4.35 ml PTM1. PTM1 Trace elements solution contained per liter 0.2 g Biotin, 6.0 g CuS04_5H2O, 0.09 g KI, 3.0 g MnSO4_H2O, 0.2 g Na2MoO4_2H2O, 0.02 g H3BO3, 0.5 g CoCl2, 42,2 g ZnSO4_7H2O, 65 g Fe(II)SO4_7H2O and 5 ml H2SO4. The fed-batch feed media were either 60% (w/w) Glycerol or concentrated MeOH and were supplemented with 12 ml/l PTM1 mineral salts solution.

Construction of *P. pastoris* strains

The codon sequence of cellobiohydrolase 2 (*Tr*CBH2) from *Trichoderma reesei* inclusive of the natural secretion leader was codon optimized for *P. pastoris* expression applying the Gene Designer software (DNA2.0, Menlo Park, CA, USA). The GC content was set to be between 40 and 60% without local peaks and restriction sites for cloning were avoided. To further optimize translation all genes were cloned after a defined Kozak consensus sequence (gaaacg). The synthetic genes were cloned into the multiple cloning site of the E. coli/P. pastoris shuttle vector pPpB1 [4] via EcoRI/NotI restriction sites of pPpT4, Näätsaari *et al.*. The *Tr*CBH2 variants were cloned downstream of the wild type promoters P(GAP) and

P(AOX1) and synthetic promoter variants with distinctly different regulation patterns were also included, namely P(En) and P(DeS). P(En) can be induced by methanol and showed increased GFP expression up to 166%, if compared to the wild-type promoter P(AOX1). P(DeS) can either be induced by methanol or be active under derepressed conditions. Plasmids were linearized with BglII, subsequently purified and concentrated using the Wizard_ SV Gel and PCR Cleanup System (Promega Corp.). Electro-competent *P. pastoris* CBS 7435 mut^S cells were prepared and transformed with 1- to 2 µg of the BglII-linearized pPpB1 or pPpT4 vector construct according to Lin-Cereghino [14]. Transformants were plated on YPD-Zeocin (100 µg/ml Zeocin) agar plates and grown at 28°C for 48 h.

Micro-scale cultivation and high-throughput screening

P. pastoris strains expressing *Tr*CBH2 were cultivated in 96-deep well plates as described by Weis et al [15]. Incubation was done in shakers (INFORS Multitron, Bottmingen, Switzerland) at 28°C, 320 rpm, and 80% relative humidity. After an initial batch phase for 60 h on 1% glucose the cultures were induced with 0.5% of methanol for a total of 72 h for expression under the control of P(De). Glucose was used instead of methanol for P(GAP) and for P(DeS) sorbitol was used instead of methanol, similar to the methanol induction. After induction the cells were pelleted at 4000 rpm and enzymatic activities were determined in the supernatants using the pHBAH -assay as previously described by [16]. Substrate conversions were performed in 50 mM citrate buffer containing 1% Avicel® at 50°C for 2 h. For the subsequent reducing sugar assay 50 µL of the substrate reaction (or, in the case of the standard sugars, appropriate dilutions of the reducing sugars) were pipetted into 150 µl of the pHBAH working solution in a 96-well PCR plate. The plate was sealed and incubated at 95°C for 5 min and then cooled to 4°C. 150 µl of the assay samples were transferred to a new micro-titer-plate and the absorption measured at 410 nm in a SPECTRA MAX Plus384 plate reader (Molecular Devices Corp., Sunnyvale, CA, USA). For exact quantification of reducing sugars a standard curve of the reducing sugar (0–1 mg/ml) was included on each plate. Activity units refer to the amount of released reducing sugar over time and correspond to the standard IUPAC definition µM/min.

Copy number determination by quantitative real-time PCR

Copy numbers of integrated expression cassettes in the Pichia genome were determined using quantitative real-time PCR (qRT-PCR) as described by [17].

Fed-batch cultivations of *Pichia pastoris* strains

Pre-cultures of individual strains were grown in 50 and 200 ml YPhyD in wide-necked, baffled shake flasks at 120 rpm at 28°C. Each fermenter (Infors Multifors system (Infors AG,

Bottmingen, Switzerland)) containing 450 ml BSM-media (pH 5.0) was inoculated from the pre-culture to an OD600 of 2.0. During the batch phase *P. pastoris* was grown on glycerol (4%) at 28°C. At the beginning of the glycerol feeding phase the temperature was decreased to 24°C. For methanol-fed cultures, the fed-batch phase was started upon depletion of initial glycerol with 16 g/(l*h) glycerol feed solution followed by methanol induction. In the early stages, the methanol-feed was set to 2 g/(l*h) and was gradually increased within the next 70 h to 6 g/(l*h). Likewise, the glycerol-feed was phased down during the first hour of methanol induction to 0 g/(l*h). Dissolved oxygen was set to 30% throughout the whole process. After 91.5 h of methanol induction the fermentations were stopped. For glycerol-fed strains, the batch phase was directly followed by a constant glycerol-feed with 5 g/(l*h). Protein concentrations were determined by micro-fluidic capillary electrophoresis (CE) using fluorescence detection (Caliper GXII System, Hopkinton, USA). Standard deviations of this robust system are usually below 10%, even at high protein loads (exemplified in Additional file 3). Therefore, just single measurements of every sample were performed. More specifically, proteins were quantified by calibrating the integrated areas of the protein-specific peaks in the electropherograms to an external reference protein standard (BSA) of known concentration. For glycosylated proteins, peak areas of diluted deglycosylated samples were compared to those of untreated samples to compensate for glycosylation-related differences in quantification. The dilutions of samples were in a range to give peak areas of the samples that were comparable to those of the reference protein standard.

Results

Effects on heterologous protein expression and vector design

The effects of synonymous codon substitution and promoter choice were studied in more detail to establish a model to predict heterologous protein expression in *Pichia pastoris*. An initial set of 48 *Tr*CBH2 genes was synthesized by DNA2.0. DNA2.0 used an in- silico design (DoE) setup called “48-run Plackett-Burman” design, which applied distinct biases to the gene features. This design allowed covering a broad range of variation while minimizing co-variation between the features. This was important to be able to distinguish independent effects of the features. For all genes known splicing and polyA motifs as well as large GC, AT and homonucleotide runs of 6 or more were avoided. This was to avoid local extremes, which have been associated with various local effects, such as translational frameshifts and transcriptional termination. DNA2.0 did not include mRNA structure near the translational start or elsewhere in the design scheme. Furthermore, CAI or the tAI as variables were not included in the DoE. The sequences were verified and cloned in a newly available *Pichia pastoris* shuttle vector system (pPpT4), Näätsaari *et al*.

In addition to synonymous codon substitution, effects on expression under altering conditions were also examined by employing three different promoters. Therefore, *Tr*CBH2 was cloned under the control of the somehow constitutive wild type promoter P(GAP) or under the control of the synthetic promoter P(DeS) or under the control of the methanol inducible synthetic promoter P(En).

Putative single copy landscapes

Putative single copy activity landscapes were generated by cultivating 88 *P. pastoris* strains per construct in 96 deep-well plates. For further information see Material and Methods section. Subsequent, screening for lignocellulolytic activity was performed employing the high- throughput reducing sugar assay established by [16]. Each 96- deep well plate included positive as well as negative controls. The positive control was a verified single copy strain under the control of respective promoter expressing the initial variant *Tr*CBH2-HM (see [18]). The negative control, a *Pichia pastoris* MutS strain, was defined as background and its activity abstracted of each single clone. Random strains of different landscapes were chosen to be tested using quantitative real- time PCR (qRT-PCR) and approved single copy integration of expression cassettes.

Figure 1 summarizes obtained single copy activities of the 48 gene variants expressed under the control of P(GAP) in box- and whisker plots (for P(En) and P(DeS) see supplementary). Released sugar concentrations varied from undetectable (<0.003 mg/ml ([16])) to 0.06 mg/ml for P(GAP), to 0.07 mg/ml for P(DeS) and to 0.16 mg/ml for P(En). Compared to the initial variant *Tr*CBH2-HM a two-fold, six-fold or four-fold increase was obtained for P(GAP), P(DeS) and P(En), respectively. Obtained activities of *Tr*CBH2 variants expressed under the control of P(DeS) were not as much scattered as variants expressed under the control of P(GAP) and P(En). Most likely this effect is due to a not optimized induction protocol for P(DeS) since so far only protocols for glucose or methanol driven expression are established at our institute. Nevertheless, there is an obvious trend for some of the *Tr*CBH2 variants regardless of which cultivation protocol. There are *Tr*CBH2 variants which are showing no or almost no activity, variants which are showing similar activities as the initial *Tr*CBH2-HM variant as well as *Tr*CBH2 variants which are showing improved activities. Out of the 48 active *Tr*CBH2 variants synonymous codon changes caused 16-fold differences in expression.

Initial Partial Least Square (PLS) model

Models to predict expression as a function of codon usage were constructed using Partial Least Squares (PLS) regression [7]. For each cultivation condition measured expression level was plotted against the expression predicted from a PLS model. The predictive models were built as described in Welch *et al*. The correlation between individual codon frequencies (59 are variable) and regarding expression were looked in more detail. Variable such as mRNA structure and 5`-AT content were also included in the initial modeling but have not seen strong contributions, so for simplicity the analysis was confined to codon usage.

Initial regression analysis with all sense codons suggested that frequencies of only a subset of codons could explain most expression variation. A genetic algorithm was used to evolve 888 highly-predictive unique PLS models, each with a reduced set of codons (average of 14.2 codons per model). The predictions of the best models are shown in Figure 2. This initial model was refined by nine hybrid genes derived from the first set, where coding sequences of two genes were recombined. With the help of these chimeras region specific effects can be more easily determined. The chimeras were chosen to sample variation along trajectories that appeared informative. Basically, pairs of genes that were either different in expressed activity or similar in activity but very different in overall codon usage were selected.

Characterization of the individual promoters

Fed-batch bioreactor cultivations were carried out to further characterize the performance of the individual promoters. Consequently, for this comparative study well defined strains had to be engineered to eliminate locus and copy number dependent effects on expression. *Tr*CBH2-HM was cloned into pPpT4 under the control of either P(GAP), P(AOX1), P(En) or P(DeS), subsequently linearised in the AOX1 terminator region and transformed to the *Pichia ku*70 strain, Näätsaari *et al*. The *ku*70 strain harbors a knock-out in the *KU70* locus thus results in increased frequencies of homologous integration. Due to this effect a higher percentage of a correctly inserted expression cassette was obtained and fewer strains had to be screened. Correctly integrated expression cassettes were verified by control colony PCRs. For this reason, primers binding either on the promoter region of the expression cassette or on the coding sequence of *Tr*CBH2-HM and primers binding outside of the expression cassette downstream in the AOXTT region [19] were employed (data not shown). Additionally, for each promoter *Tr*CBH2-HM construct strains were analyzed for single copy integration by qRT-PCR. In this way four single copy strains were generated expressing *Tr*CBH2-HM under the control of either P(GAP), P(AOX1), P(En) or P(DeS) with verified integration of the expression cassette in the AOX1 locus. Fed batch cultivations of these strains were performed and *Tr*CBH2 expressed in presence of either methanol or glycerol. Samples were taken during the whole cultivation process and subsequently analyzed.

Since the main disadvantage of P(GAP) is its constitutive behavior we wanted to proof that the P(DeS) can be tightly regulated by the available carbon source type and level. Therefore, the promoter performance of P(AOX1) was compared to P(DeS) in presence of methanol or glycerol by transcript level analysis. Samples were analyzed for *Tr*CBH2 mRNA shortly before the fed-batch start, during fed-batch and after induction start. As seen in Figure 3, in our experiments P(AOX1) exhibits a low basal promoter activity in presence of glycerol which is increased by a factor of 5 in presence of methanol, on average 2.3 and 11.7 respectively. These values are less than expected [10] and might be related to the deficiency strain *ku*70 nevertheless, the transcript level was significantly increased after addition of methanol. Comparable results were obtain by P(DeS) regulated expression for both cultivation strategies. Indeed, *Tr*CBH2 mRNA levels were increased by a factor of 6 shortly after addition of methanol or after the reduction of glycerol from $16g*l^{-1}*h^{-1}$ to $5g*l^{-1}*h^{-1}$ (Figure 3), on average 2.0 and 12.6. These results indicate that glycerol can either repress or derepress the activity of P(DeS) depending on its concentration. Furthermore, P(DeS) can also be induced by methanol.

In Addition to transcript level analysis we also wanted to compare the produced protein titers of the fed-batch cultivated single copy strains. As already mentioned, the cultivation protocol

for the micro-scale screening was not optimized for P(DeS), hence diverge results were expected. The protein titers of the fed-batch cultivations clearly demonstrated that the synthetic P(DeS) outperformed the wild-type promoters P(AOX1) and P(GAP), see Figure 4. Within the first 45 h of process time roughly the same amount of *Tr*CBH2 was expressed by all strains. Afterwards the strain under the control of P(DeS) and induced by methanol was superior over all other strains. In fact, after 90 h around 40 % more protein was produced by P(DeS) than compared to the wild-type promoter AOX1, 0.73 g/l and 0.51 g/l respectively. Moreover, in the case of methanol free induction P(DeS) achieved a 12-fold higher end-titer than compared to the wild-type promoter P(GAP), 0.12 g/l and 0.01 g/l respectively. Although the strain harboring the other synthetic promoter P(En) achieved slightly higher protein concentrations till 45 h of process time the end titer was around 30 % less than compared to the wild type promoter AOX1, 0.37 g/l and 0.51 g/l respectively.

Summarizing, equal conditions were chosen to compare the performance of the synthetic promoters to the wild-type promoters. As demonstrated in these experiments, P(DeS) was tightly regulated on transcriptional level and outperformed the other promoters in terms of protein titer.

Combining synthetic promoters and synthetic genes

So far micro-scale experiments and s.c. fed-batch cultivations revealed that synonymous codon substitution and promoter choice have a huge effect on *Tr*CBH2 protein production. To test if these results are also valid in up-scale experiments and to evaluate if they can be combined to further increase *Tr*CBH2 production additional strains were generated. Fed-batch cultivations were performed and either *Tr*CBH2-HM or *Tr*CBH2-V04 was expressed by varying gene dosages under the control of P(AOX1) and P(DeS) in presence of either methanol or glycerol, see Figure 5.

Regarding the effect of synonymous codon substitution, although the same promoter and the same gene dosage was used to express *Tr*CBH2 the strain P(DeS)_*Tr*CBH2-V04_CN4±1 clearly outperformed the strain P(DeS)_*Tr*CBH2-HM_CN4±1 in terms of protein titers Figure 5. The last two time points 74.5 and 95.5 h of the cultivation process revealed that the protein titer is 3 to 2 fold higher if the codon optimized *Tr*CBH2-V04 variant is expressed instead of *Tr*CBH2-HM, 16.18 and 15.71 g/l compared to 5.50 and 6.55 g/l respectively. Based on the micro-scale screening results we even expected an up to 5-fold improvement of the protein titers. This discrepancy can be related either to the not optimized micro-scale assay or to a protein saturation effect (see also following paragraphs).

Regarding the promoter performances, strains expressing an identical *Tr*CBH2 variant and harboring comparable amounts of expression cassettes were used for the comparison,

namely P(DeS)_*Tr*CBH2-HM_CN4±1 and P(AOX1)_*Tr*CBH2-HM_CN3. As seen in Figure 5, P(DeS)_*Tr*CBH2-HM_CN4±1 clearly outperformed the strain P(AOX1)_*Tr*CBH2-HM_CN4±1 with and even without methanol in terms of protein titers over the whole induction period. The last two time points 74.5 and 91.5 h of the cultivation process revealed that the protein titer is 3 to 2 fold higher if *Tr*CBH2 expression is induced by methanol under the control of P(DeS) compared to P(AOX1), 5.50 and 6.55 g/l compared to 1.70 and 2.78 g/l respectively. This enhancement is better than the results obtained for the single copy fed-batch cultivations. Most likely this is due to the replacement of the deficiency strain *ku*70 by the producer strain mut^S and might be also related to the gene dosage.

Regarding the gene dosage, *Tr*CBH2-HM was expressed under the control of P(DeS) by varying copy numbers either in presence of methanol or glycerol, namely P(DeS)_*Tr*CBH2-HM_CN4±1 or P(DeS)_*Tr*CBH2-HM_CN36±9. Although methanol induction achieved higher protein titers the difference was less pronounced if the gene dosage was further increased. In the case of the before mentioned s.c. strains the differences was around 600 % (see Figure 4) at the last time point whereas for P(DeS)_*Tr*CBH2-HM_CN4±1 around 20 % and for P(DeS)_*Tr*CBH2-HM_CN36±9 less than 0.01% was obtained at the last time point, 0.73 compared to 0.12 g/l, 6.55 compared to 5.31 g/l and 13 compared to 12.96 g/l respectively see Figure 5. This saturation effect indicates that there is no linear relation of gene dosage and *Tr*CBH2 titer for higher copy number strains.

Additionally to the strains seen in Figure 5 other strains were also fed-batch cultivated. In these strains the gene dosage was even further increased. They expressed either *Tr*CBH2-V04 or *Tr*CBH2-V09 under the control of P(DeS) in presence of methanol or glycerol. To accurately quantify obtained protein titers after 90 h different methods were applied and compared to enzymatic activities. Table 1 is summarizing the results of all fed-batch cultivated strains in the study. The values in Table 1 indicate that if the gene dosage is too high protein production will drop down, as already seen for *Tl*XynA production [18]. Most likely, due to the strong promoter and the optimized gene sequence the increased transcript level overloaded folding and secretion machinery of the cells [6, 20, 21]. This is even more evident in the case of methanol induced *Tr*CBH2 expression.

Although lower copy number strains induced without methanol were not as good as strains induced with methanol the highest titer was obtained by P(DeS)_*Tr*CBH2-HM_CN40±9 without methanol. Less than 1g/l of *Tr*CBH2 was achieved if methanol was used for expression due to major growth related problems. However, also for *Tr*CBH2 expression under the control of P(DeS) in presence of glycerol an upper limit of gene dosage was observed. In the case of P(DeS)_*Tr*CBH2-HM_CN54±13 less protein was achieved than with the strain P(DeS)_*Tr*CBH2-HM_CN40±9.

To additionally check the quantity and quality of the produced protein, the supernatants were loaded on a SDS gel and compared to commercial available lignocellulolytic enzymes produced either by *Trichoderma reesei* or by *Aspergillus niger*. Moreover, different concentrations of BSA were loaded on the SDS gel to better estimate obtained protein concentrations. As seen in Figure 6, in the supernatants of *T. reesei* and *A. niger* several proteins can be detected whereas in *P. pastoris* mainly the heterologously expressed protein can be detected. Three different supernatants of *P. pastoris* strains were loaded on the SDS gel, namely the strain P(DeS)_*Tr*CBH2-V04_CN4±1 cultivated with or without methanol and P(DeS)_*Tr*CBH2-HM_CN40±9 cultivated without methanol. As seen in Figure 6, the highest protein titer was achieved by P(DeS)_*Tr*CBH2-HM_CN40±9 without methanol induction, ~18 g/l. Furthermore, SDS gel pictures and LC-MS analysis revealed that in the supernatant of methanol induced strains more degradations products can be detected compared to supernatants of strains induced without methanol.

Summarizing, optimization of the gene sequence in combination with a strong synthetic promoter induced just by low levels of glycerol led to the highest protein secretion reported in *Pichia pastoris* so far.

DISCUSSION

One aim of this study was to further characterize the effect of synonymous codon substitution on protein production. The lignocellulolytic enzyme *Tr*CBH2 was chosen as target protein since a reliable and sensitive assay was available for the screening. Moreover, we have already demonstrated that the amount of produced protein of this class A gene can be almost linearly increased in accordance with at least 7 integrated expression cassettes [18]. Since we also expect a variation of gene expression in the order of several magnitudes for synonymous codon substitution, *Tr*CBH2 was an excellent target for its evaluation.

The in- silico design DeO was applied to reduce the incredible high number of possible targets to a rational number that can be dealt with. Nevertheless, more than 15000 *Pichia pastoris* transformants were screened and resulting putative single copy (s.c.) landscapes were analyzed. Particularly interesting were those *Tr*CBH2 variants which showed no or reduced activity compared to the initial *Tr*CBH2 variant. Apparently, these sequences were containing negatively influencing elements and/or factors. This offers valuable clues of *Pichia pastoris´* gene expression and which factors should be better avoided.

First models to predict expression as a function of codon usage were constructed using Partial Least Squares (PLS) regression. These models showed that similar results can be obtained for *Tr*CBH2 production in *Pichia pastoris,* independently from the employed promoter and expression conditions. In *E. coil,* Welch *et al*. [7] speculated that variations of expression levels aroused by synonymous codon substitution depended on the limitation of tRNA recharging. Since the tRNA pool should be influenced by varying C- sources it is rather unlikely that the basis of the *E. coli* results apply also for *P. pastoris*. In *E. coli* energy depletion can be caused by high transcript levels due to limitations of precursor availabilities [22]. Whereas for *Pichia pastoris* a translational limitation is not very plausible due to moderate specific productivities at least at low to medium expression levels [6]. More likely protein production is influenced by posttranslational processes (e.g. folding and secretion) and/or protein degradation of misfolded or unfolded proteins. This was also evident in the experiments performed by [6, 23]. In the case of [23] a temperature decrease led to a reduction of the metabolic burden however, protein production rate was up to threefold increased. Our results indicated too, that *Tr*CBH2 production was not limited in the first place at transcriptional level. Increasing of gene dosage led to an improved protein production till a certain upper limit of integrated expression cassettes. We assume that afterwards the metabolic burden produced by high transcript levels was getting too high. Thus, growth related problems of the cell and loss of productivity was observed. This effect was already reported in other studies [6, 20, 24].

From industrial point of few, lower copy number strains are more interesting for biotechnology applications since strain stability is improved. [25] demonstrated that if the amount of integrated expression cassettes was too high loss of copies was observed during methanol induction. Thus long term productivity can be affected.

Additionally, analysis of the putative s.c. landscape data were performed regarding the relative impact the codon frequency had on expression in the PLS model and regarding of its variable importance. Hence, particular amino acids and the frequency of certain synonymous codons appeared to have significant influence on the protein productivity.

To test the PLS model, hybrid genes derived from the first *Tr*CBH2 set were employed. Even though the model could not correctly predict the enzymatic activities of all of the chimeras they still helped to further improve the model since region specific effects could be more easily determined. In the end, we have shown that synonymous codon substitution had a huge effect on protein production. Thereby, a model was generated to predict the expression of heterologous proteins in *P. pastoris*. However, this is just an initial model and does not directly provide an optimal codon usage. Rather it suggests which codons should be used more and less often than their average use. Therefore, it might be worth it to test several variants of your gene of interest to be sure to get maximized protein titers. Moreover, for class B and especially for class C genes we emphasize that the gene variants should be optimized in a non optimal way. By doing so, the transcript level will get reduced but still more protein might be produced since the metabolic burden will be reduced too [21, 23, 26].

The second aim of this study was to further characterize the newly available synthetic promoters. Transcriptional analysis and protein titer determination revealed, P(DeS) is a strong promoter that can be induced by methanol and additionally is active under derepressed conditions. Contrary to P(GAP), P(DeS) can be still tightly regulated by the available carbon source type and level. Moreover, studies performed with another synthetic derepressed promoter P(De) revealed that there is a positive effect on cell physiology compared to expression under the control of P(AOX1) [27] which might also apply for P(DeS).

The highest titer of produced *Tr*CBH2 was approximately 18 g/l which were achieved under the control of P(DeS) without methanol induction. We did not want to relay on just one determination method therefore, we confirmed this value by four different methods. On account of this we can definitely declare that this value is the highest concentration of secreted protein in *P. pastoris* reported so far [28, 29]. Additionally, almost pure supernatants of proteins were produced by *Pichia* compared to commercial available supernatants of lignocellulolytic enzymes produced in other fungi host systems. Moreover, methanol induced

protein production exhibited more degradation products than compared to methanol free induction. It is already known that presence of methanol can lead to cell lysis which further on negatively effects protein quality [21, 30]. Besides this glycerol (or glucose)- based fermentations are highly interesting for industrial applications since a 10-fold decreased heat production, significantly reduced oxygen consumption and higher protein synthesis rates as compared to methanol-based processes can be obtained [21].

Summarizing, protein production in *Pichia pastoris* is strongly influenced by several factors. Therefore, to get maximal protein titers without optimizing expression conditions is very unlikely. Here we demonstrated that certain elements of the expression cassette were independently optimized and obtained results subsequently combined. These results confirm that synthetic biology is a helpful tool to further increase protein production.

Figures

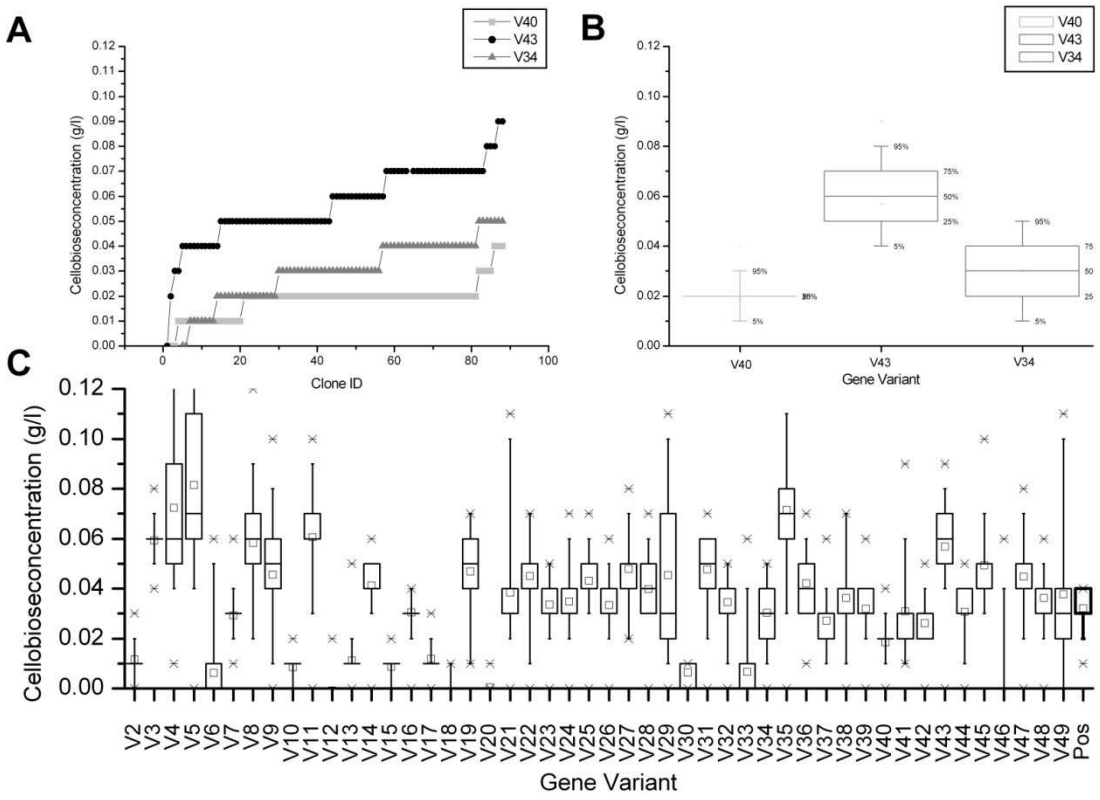


Figure 1: Released Cellobiose concentration of the *Tr*CBH2 set expressed under the control of P(GAP) determined by the pHBAH-Assay. Panel A) Putative single copy landscapes of three representative gene *Tr*CBH2 gene variants. Panel B) This panel shows how the values of the putative s.c. landscapes of panel A are interpreted. Panel C) Summarizes all results obtained with the 48 TrCBH2 gene variants. The initial TrCBH2-HM variant is represented in bold at the left side.

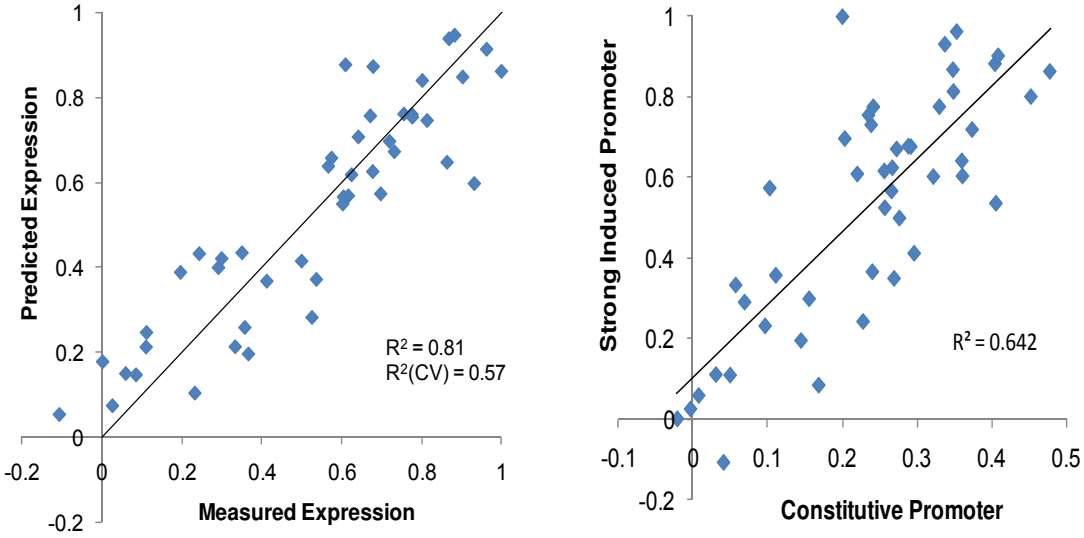


Figure 2: PLS codon frequency models. For each variant the measured expression level was plotted against the expression predicted from a PLS model using genetic algorithm-selected codons [7]. Panel A) Model fit for *Tr*CBH2 set expressed under the control of P(GAP). Panel B) Model fit for a comparison between the constitutive and induced expression of the *Tr*CBH2 gene set. Expression in each set was normalized to the highest expression level in that set. R2(CV) indicates the correlation coefficient for the fit of the model in cross- validation.

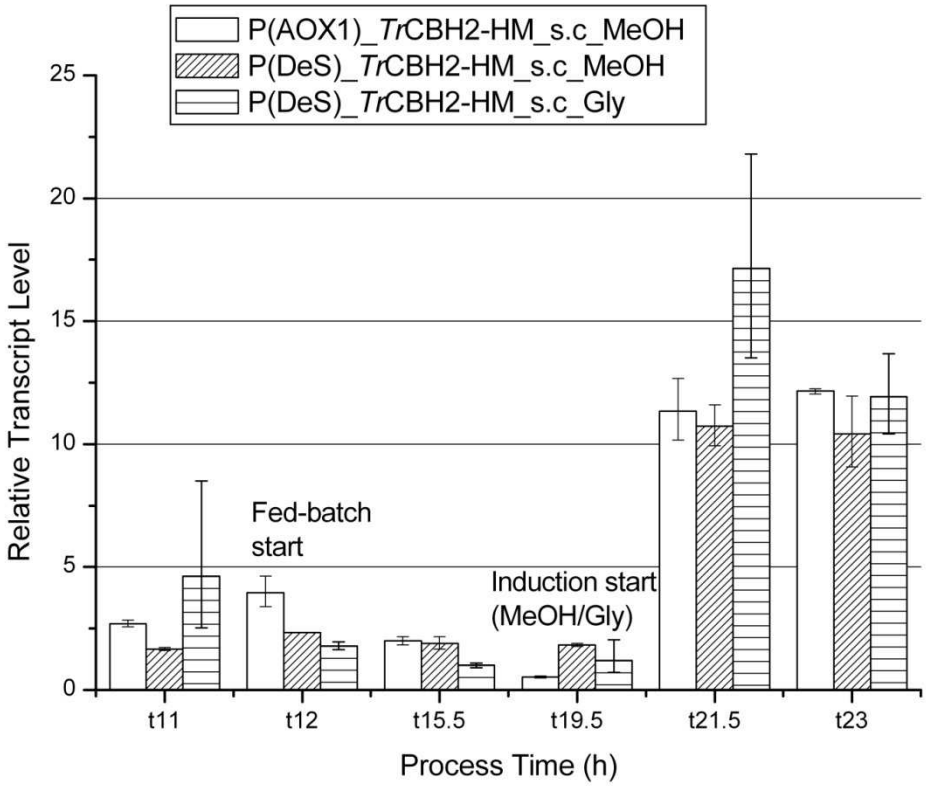


Figure 3: Transcript level analysis of fed-batch cultivated s.c. strains. Samples were taken at the end of the batch phase, at the start of the fed-batch phase and at the beginning of the induction phase, as indicated in the legend. *Tr*-CBH2-HM was expressed under the control of P(AOX1) in presence of methanol (white bar) or under the control of P(DeS) in presence of methanol (dense bar) or under the control of P(DeS) in presence of glycerol (sparse bar). All values are normalized to the value of the strain P(DeS)-*Tr*CBH2-HM-s.c. cultivated in presence of glycerol after 15.5 h of process time.

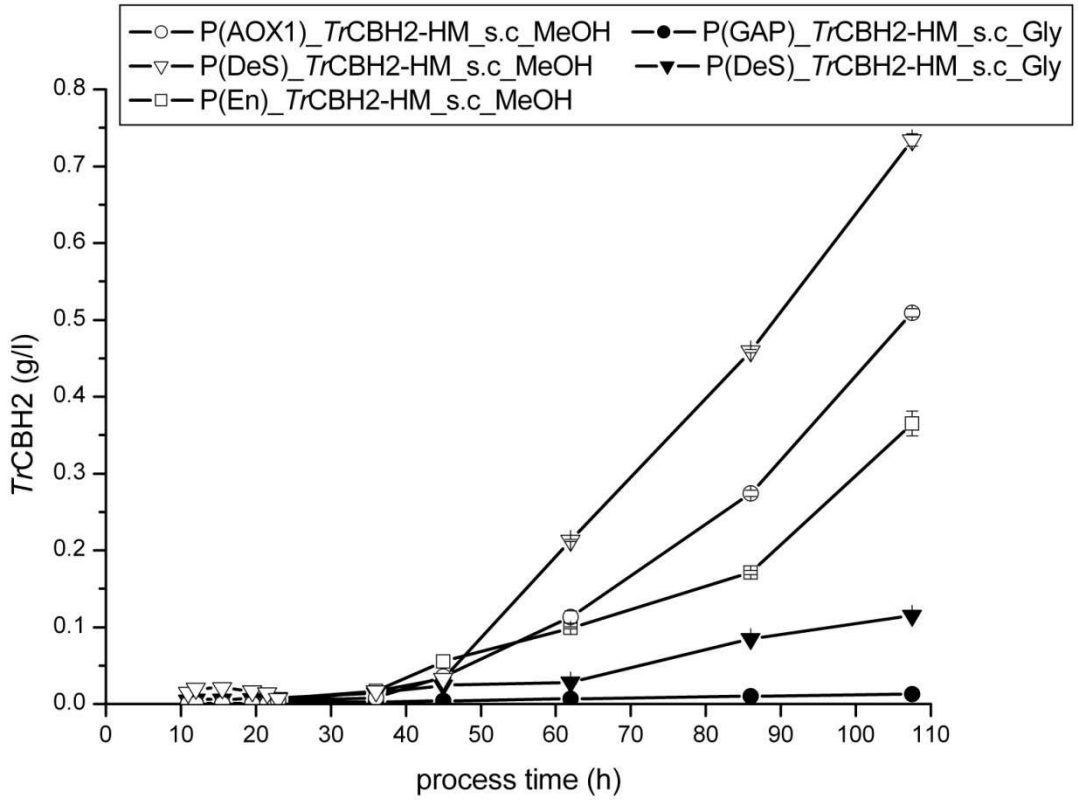


Figure 4: Time- course of protein concentration during fed-batch cultivation of s.c. strains. *Tr*-CBH2-HM was expressed under the control of P(AOX1) in presence of methanol (open circle), *Tr*-CBH2-HM was expressed under the control of P(DeS) in presence of methanol (open triangle), *Tr*-CBH2-HM was expressed under the control of P(En) in presence of methanol (open square), *Tr*-CBH2-HM was expressed under the control of P(GAP) in presence of glycerol (closed circle), *Tr*-CBH2-HM was expressed under the control of P(DeS) in presence of glycerol (closed triangle).

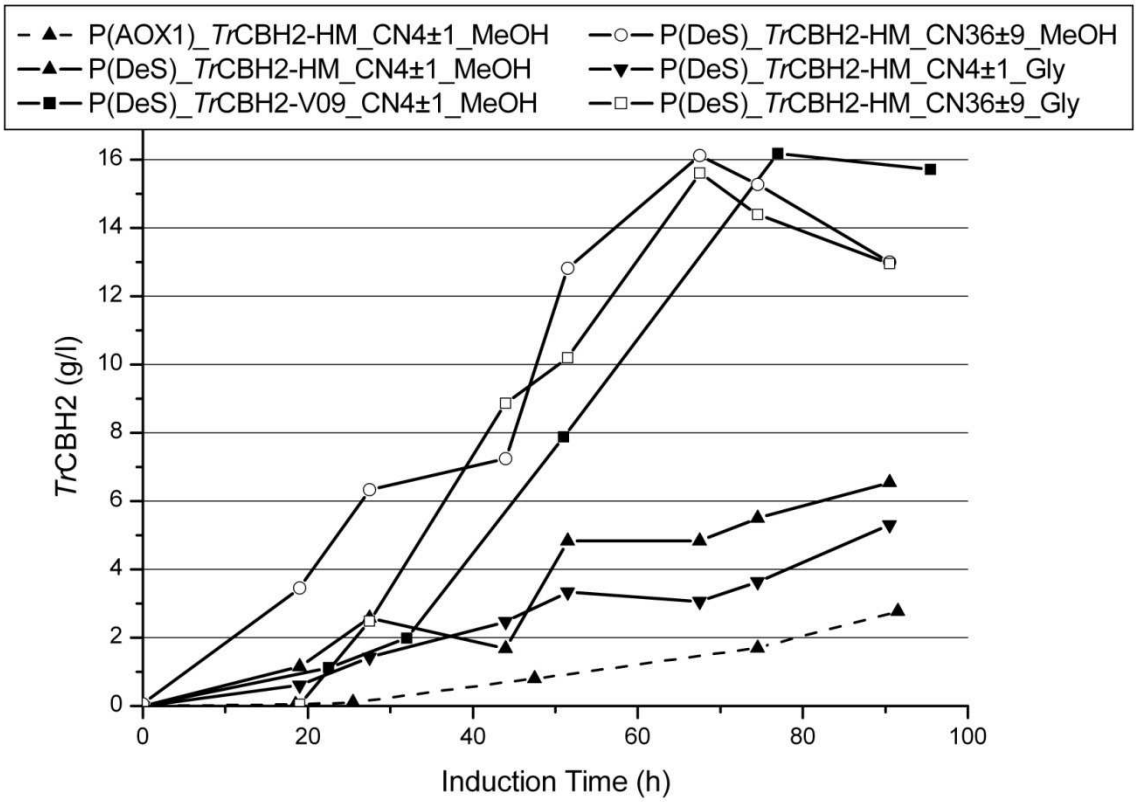


Figure 5: Time- course of protein concentration during fed-batch cultivation of m.c. strains. Differently codon optimized *Tr*CBH2 variants were expressed under the control of P(AOX1) or P(DeS). Copy numbers (CN) are specified in the legend. *Tr*-CBH2-HM was expressed under the control of P(AOX1) in presence of methanol harboring 4 expression cassettes (dotted line, closed triangle), *Tr*-CBH2-HM was expressed under the control of P(DeS) in presence of methanol harboring 4 expression cassettes (closed upright triangle), *Tr*-CBH2-V09 was expressed under the control of P(DeS) in presence of methanol harboring 4 expression cassettes (closed square), *Tr*-CBH2-HM was expressed under the control of P(DeS) in presence of methanol harboring 36 expression cassettes (open circle), *Tr*-CBH2-HM was expressed under the control of P(DeS) in presence of glycerol harboring 4 expression cassettes (closed triangle) and *Tr*-CBH2-HM was expressed under the control of P(DeS) in presence of glycerol harboring 36 expression cassettes (open square).

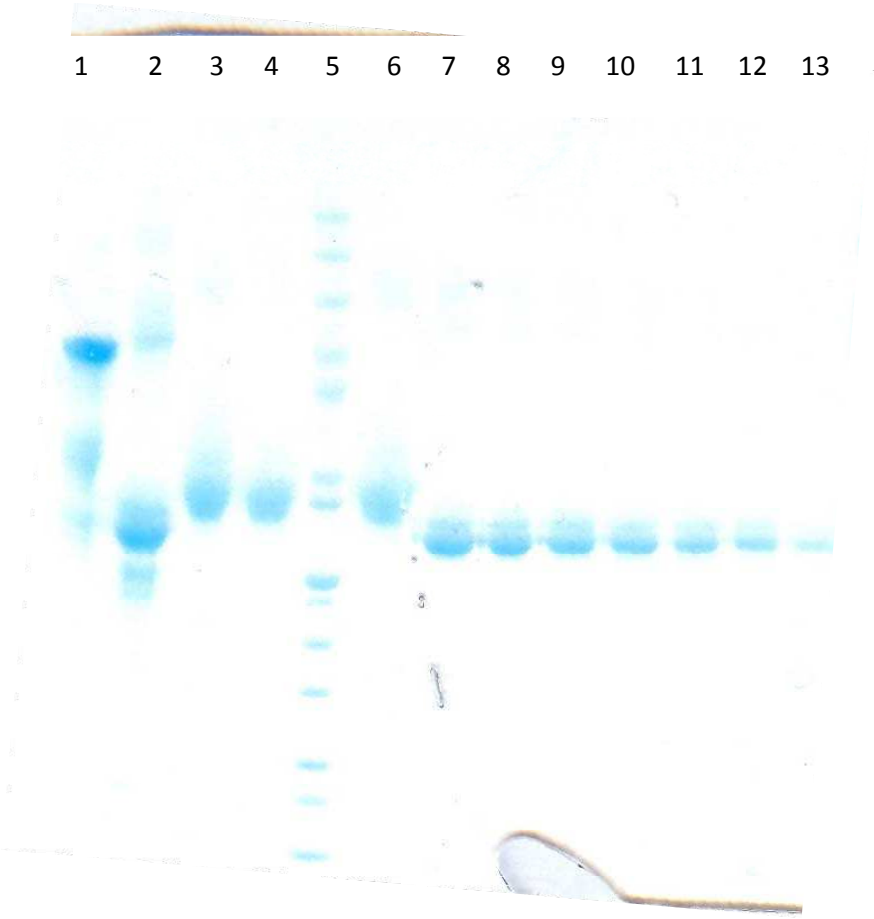


Figure 6: SDS gel picture of lignocellulolytic enzymes expressed in different hosts. Lane 1: Cellobiase from *Aspergillus niger* (C6105) 3µl of 1:20 diluted supernatant; Lane 2: Cellulase from Trichoderma reesei ATCC 26921 (C2730), 3µl of 1:20 diluted supernatant; Lane 3: *Pichia pastoris* strain P(DeS)_*Tr*CBH2-HM_CN36±9 in presence of methanol, 4µl of a 1:20 diluted supernatant; Lane 4: *Pichia pastoris* strain P(DeS)_*Tr*CBH2-HM_CN36±9 in presence of glycerol, 4µl of a 1:20 diluted supernatant; Lane 5: PageRuler™ Protein Ladder (SM0661); Lane 6: *Pichia pastoris* strain P(DeS)_*Tr*CBH2-HM_CN40±9 in presence of glycerol, 2µl of a 1:20 diluted supernatant; Lane 7: 0.7g/l of BSA, 3µl loaded; Lane 8: 0.6g/l of BSA, 3µl loaded; Lane 9: 0.5g/l of BSA, 3µl loaded; Lane 10: 0.4g/l of BSA, 3µl loaded; Lane 11: 0.3g/l of BSA, 3µl loaded; Lane 12: 0.2g/l of BSA, 3µl loaded; Lane 13: 0.1g/l of BSA, 3µl loaded.

References

1. Cregg, J.M., et al., *Expression in the yeast Pichia pastoris.* Methods Enzymol, 2009. **463**: p. 169-89.
2. Macauley-Patrick, S., et al., *Heterologous protein production using the Pichia pastoris expression system.* Yeast, 2005. **22**(4): p. 249-70.
3. Zhang, A.L., et al., *Recent advances on the GAP promoter derived expression system of Pichia pastoris.* Mol Biol Rep, 2009. **36**(6): p. 1611-9.
4. Abad, S., et al., *Stepwise engineering of a Pichia pastoris D-amino acid oxidase whole cell catalyst.* Microb Cell Fact, 2010. **9**: p. 24.
5. Hartner, F.S., et al., *Promoter library designed for fine-tuned gene expression in Pichia pastoris.* Nucleic Acids Res, 2008. **36**(12): p. e76.
6. Mattanovich, D., et al., *Stress in recombinant protein producing yeasts.* J Biotechnol, 2004. **113**(1-3): p. 121-35.
7. Welch, M., et al., *Design parameters to control synthetic gene expression in Escherichia coli.* PLoS One, 2009. **4**(9): p. e7002.
8. Kudla, G., et al., *Coding-sequence determinants of gene expression in Escherichia coli.* Science, 2009. **324**(5924): p. 255-8.
9. Waterham, H.R., et al., *Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter.* Gene, 1997. **186**(1): p. 37-44.
10. Tschopp, J.F., et al., *Expression of the lacZ gene from two methanol-regulated promoters in Pichia pastoris.* Nucleic Acids Res, 1987. **15**(9): p. 3859-76.
11. Hartner, F.S. and A. Glieder, *Regulation of methanol utilisation pathway genes in yeasts.* Microb Cell Fact, 2006. **5**: p. 39.
12. Menendez, C., et al., *Functional production and secretion of the Gluconacetobacter diazotrophicus fructose-releasing exo-levanase (LsdB) in Pichia pastoris.* Enzyme and microbial technology, 2004. **34**(5): p. 446-452.
13. Qin, X., et al., *GAP promoter library for fine-tuning of gene expression in Pichia pastoris.* Appl Environ Microbiol, 2011. **77**(11): p. 3600-8.
14. Lin-Cereghino, J., et al., *Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast Pichia pastoris.* Biotechniques, 2005. **38**(1): p. 44, 46, 48.
15. Weis, R., et al., *Reliable high-throughput screening with Pichia pastoris by limiting yeast cell death phenomena.* FEMS Yeast Res, 2004. **5**(2): p. 179-89.
16. Mellitzer, A., et al., *Sensitive high-throughput screening for the detection of reducing sugars.* Biotechnol J, 2011. **7**(1): p. 155-62.

17. Abad, S., et al., *Real-time PCR-based determination of gene copy numbers in Pichia pastoris.* Biotechnol J, 2010. **5**(4): p. 413-20.
18. Mellitzer, A., et al., *Expression of lignocellulolytic enzymes in Pichia pastoris.* Microb Cell Fact, 2012. **11**(1): p. 61.
19. Kuberl, A., et al., *High-quality genome sequence of Pichia pastoris CBS7435.* J Biotechnol, 2011. **154**(4): p. 312-20.
20. Glick, B.R., *Metabolic load and heterologous gene expression.* Biotechnol Adv, 1995. **13**(2): p. 247-61.
21. Heyland, J., et al., *Quantitative physiology of Pichia pastoris during glucose-limited high-cell density fed-batch cultivation for recombinant protein production.* Biotechnol Bioeng, 2010. **107**(2): p. 357-68.
22. Sanden, A.M., et al., *Limiting factors in Escherichia coli fed-batch production of recombinant proteins.* Biotechnol Bioeng, 2003. **81**(2): p. 158-66.
23. Heyland, J., et al., *Carbon metabolism limits recombinant protein production in Pichia pastoris.* Biotechnol Bioeng, 2011. **108**(8): p. 1942-53.
24. Ramon, R., P. Ferrer, and F. Valero, *Sorbitol co-feeding reduces metabolic burden caused by the overexpression of a Rhizopus oryzae lipase in Pichia pastoris.* J Biotechnol, 2007. **130**(1): p. 39-46.
25. Zhu, T., et al., *A systematical investigation on the genetic stability of multi-copy Pichia pastoris strains.* Biotechnol Lett, 2009. **31**(5): p. 679-84.
26. Gasser, B., et al., *Monitoring of transcriptional regulation in Pichia pastoris under protein production conditions.* BMC Genomics, 2007. **8**: p. 179.
27. Hyka, P., et al., *Combined use of fluorescent dyes and flow cytometry to quantify the physiological state of Pichia pastoris during the production of heterologous proteins in high-cell-density fed-batch cultures.* Appl Environ Microbiol, 2010. **76**(13): p. 4486-96.
28. Werten, M.W., et al., *High-yield secretion of recombinant gelatins by Pichia pastoris.* Yeast, 1999. **15**(11): p. 1087-96.
29. Xiong, A.S., et al., *High level expression of a synthetic gene encoding Peniophora lycii phytase in methylotrophic yeast Pichia pastoris.* Appl Microbiol Biotechnol, 2006. **72**(5): p. 1039-47.
30. Sinha, J., et al., *Causes of proteolytic degradation of secreted recombinant proteins produced in methylotrophic yeast Pichia pastoris: case study with recombinant ovine interferon-tau.* Biotechnol Bioeng, 2005. **89**(1): p. 102-12.

CHAPTER 4

ACIB Method

Sensitive high-throughput screening for the detection of reducing sugars

ACIB Protocol

Title

Sensitive high-throughput screening for the detection of reducing sugars

Author: **Andrea Mellitzer**
Andrea.mellitzer@acib.at

Creation date: 27.01.2012 Validity: until revoked/cancelled This ACIB protocol replaces the version from: Developed in project/working group: MacroFun P4		
Notice of modification:		
Written by: Andrea Mellitzer and Karlheinz Flicker		
Date/signature: 26.06.2012		
Checked:	Yes ☐ No ☐	Date/Signature
Approved by: Karlheinz Flicker and Aleksandra Mitrovic. This method was also already published in [2].		

1. Purpose and Field of Application

This assay was established for the detection of lignocellulolytic enzymes in the crude supernatants of *Pichia pastoris* strains.

This high- through micro-scale assay allows screening of up to several thousand strains per day. So far, four different types of lignocellulolytic enzymes were verified by this method. Released sugar concentration down to 10µM can be reliable detected. For exact quantification of reducing sugars a standard curve of the respective reducing sugar (0-1mg/ml) is included on each plate.

2. Principle

The reducing sugars chemically react with pHBAH to form strongly absorbing osazones that can be photometrical detected at 410nm and correlate with the released sugar concentration.

3. Key Words, Definitions & Abbreviations

Abbreviations: pHBAH, p-hydroxybenzoic acid hydrazide;

Key words: Cellulase, high-throughput screening, para-Hydroxybenzoic acid hydrazide. Pichia pastoris, reducing sugar and lignocellulolytic enzymes

Unit: Activity units for the expressed enzymes refer to the amount of released reducing sugars over time and correspond to the standard IUPAC definition µM/min.

4. Methodology

4.1.Reagents

List of the chemicals needed and precautions if needed (storage conditions and time, risks)

Name	Formula	MW	Purity	Supplier	Order No.	Comments
pHBAH	HOC6H4CONHNH2	152.15	≥97.0%	Fluka	54600	
Salzsäure	HCl	36,46		Any supplier		Acid
Natronlauge	NaOH	40,01		Any supplier		Base
Trinatriumcit-rat-Dihydrat	Na3C6H5O7*H2O	294.10		Any supplier		
Zitronensäure	C6H8O7	192,13		Any supplier		Acid
Reducing sug-ars (e.g. cello-biose)	C12H22O11	342.3		Any supplier		
Substrate (e.g. Avicel)	(C6H10O5)n	(162)n		Any supplier		

4.2.Solutions

Description of the solutions and buffers needed to execute the method and precautions if needed (storage conditions and time, risks).

- 50 mM citrate buffer pH 5.5 (Trinatriumcitrat-Dihydrat, Zitronensäure) to dilute the respective reducing sugars and for background control
- pHBAH stock preparation: p-hydroxybenzoic acid hydrazide (4-Hydroxybenzhydrazide, Fluka). 5% in 0.5 M HCl. For 50 mL stock :
    - 2.5 g pHBAH, +40 ml ddH20, add 2 ml conc. HCl (37%=12M); ad 50 ml with ddH20→ pHBAH not completely dissolved after vortexing and 15 at RT
    - Store +/- one month at 4°C
- pHBAH working solution (WS):
    - Mix one volume pHBAH stock with 4 volumes 0.5 N NaOH. Every day fresh!

  Here: solution was made by adding 3.125 ml 4M NaOH to 5 ml of pHBAH stock, then filled up to 25 ml with ddH2O

4.3.Materials

Description of materials needed to perform the method

Name	Supplier	Order No.	Comments
50 ml Tubes	Greiner bio-one	227261	Any supplier ok
96 well microplates, white, PCR clean	Eppendorf		
self-adhesive aluminium foil	Any supplier		
96-well PCR plates	Eppendorf		
Semi-skirted 96-well PCR plates	VWR		
Multi-channel tips, 1 ml	Any supplier		
96-tips for pipetting robot	Any supplier		
Silicone Sealing Mat for PCR plates	VWR		

4.4.Apparatus

Description of the instruments needed to perform the test and specifications.

Name	Supplier	Comments
2720 thermocycler.	Applied Biosystems	
Quadra Tower, 96-tip pipetting robot,	Tomtec,	
Microplate reader Tecan Sunrise	Tecan	
Shaking tower		

4.5.Procedure

Substrate conversion

Use 20 µL of culture supernatant for each well. Add 150 µL of 50 mM citrate buffer (Trinatriumcitrat-Dihydrat und Zitronensäure) with the appropriate substrate:

- 150 µL 50 mM citrate buffer pH5.0 containing substrate (e.g. 1% Avicel)
- Add 20 µL of culture supernatant containing the desired lignocellulolytic enzyme
- seal plate with adhesive aluminium foil (and put a lid onto the plate)
- mix on a vibro skaker for 1 min
- incubate under shaking at 50°C for 2h (300 rpm)
- let cool to rt and centrifuge plate

Reducing sugar detection

- pellet insoluble substrate by centrifugation
- add 150 µL of pHBAH WS in a 96-well PCR plate to stop the reaction
- pipet 50 µL supernatant (and appropriate standard dilutions of reducing sugars; 0, 0.125; 0.25; 0.5, 1 mg/mL cellobiose) from the substrate conversion reaction
- seal plate with silicon mat
- Mix on a vibro-shaker.
- Heat to 95°C for 5 min in a PCR thermocycler.
- Let cool to RT
- centrifuge plate
- transfer 150 µL from the reaction to a 96-well plate for measurement (use 96-well plate pipetting robot if possible)

Read absorbance at 410 nm on a plate reader.

4.6.Calculations

From the absorbance readings the corresponding difference between the sample and the blank is calculated"

$\Delta A410$ = A410 (sample) –A410 (blank)

A standard curve is plotted from the results of the standard solutions. The scale of abscissa represents the amount of released sugar in mg/ml determined at A410 and the scale of the ordinate represents the employed sugar concentration.

5. Safety Precautions

Please follow instructions described in "acib-Mitarbeiterleitfaden Gefahrstoff- und Laborordnung"

6. Documentation

Besides the lab book documentation of Flicker and Mellitzer this method is also described in the published article "Sensitive high throughput screening for detecting reducing sugars".

7. References

1. Lever, M., *A new reaction for colorimetric determination of carbohydrates.* Analytical Biochemistry, 1972. **47**(1): p. 273-279.
2. Mellitzer A, *Sensitive high throughput screening for detecting reducing sugars.* Biotechnol J 2011

Summary and Conclusion

Summary and Conclusion

The main aim of the present study was to demonstrate that newly available synthetic promoter variants are a beneficial tool for the heterologously protein expression in *Pichia pastoris*. Moreover, we wanted to show that *Pichia pastoris* is a suitable alternative to other more difficult to handle fungal systems for the expression of lignocellulolytic enzymes. Therefore, *Trichodermareesei*cellobiohydrolase 1 and 2 (*Tr*CBH1 and *Tr*CBH2) and beta-mannanase(*Tr*bMan), and *Thermomyceslanuginosus*xylanase A (*Tl*XynA) were chosen as target and modelenzymes.

The first step to proof *Pichia´*s capability as convenient host for the expression and engineering of lignocellulolytic enzymes was to establish a screening assay. Therefore, a reducing- sugar assay was adapted to a highthroughput micro scale protein expression screening method. We have shown that all of our target enzymes could be actively expressed, characterized and identified by the adapted assay. Advantageously, this sensitive and reliable screen can be applied for assaying crude-culture supernatants. Hence, this simple assay can be universally applied in strain engineering and directed evolution for the activity-based screening of enzymes that degrade natural sugar polymers.

To establish *Pichia* as host for the expression and engineering of lignocellulolytic enzymes it was necessary to generate good producer strains of the target enzymes and to evaluate their potential as efficient enzyme producers. Therefore, coding sequences of the genes were optimized for expression in *Pichia pastoris*. Additionally, different promoters were employed to further influence protein production. Based on our results we introduced a new possible classification for genes to be expressed in *Pichia*. In these classes (A, B, C) protein production depends on gene dosage and the strength of the employed promoter. Finally, strains were constructed producing maximum protein concentrations of 1.142 g/l *Tr*bMan, 6.55 g/l *Tr*CBH2 and 1.2 g/l *Tl*XynA in fed-batch bioreactor cultivations. Additionally, since high specific activities compared to the same enzymes expressed in other host systems can be obtained we concluded that *Pichia* seems to be a good host for the expression of individual lignocellulolytic enzymes.

In the last part of the thesis we demonstrated the beneficial combination of synonymous codon substitution, promoter choice and gene dosage for protein production in *Pichia pastoris*. Therefore, to further characterize each element, individually experiments were performed. Finally, strains were engineered producing ~18 g/l of *Tr*CBH2, the highest

concentration of secreted protein reported in *P. pastoris* so far. Moreover, we have shown that compared to other fungal host systems,*Pichia* secrets almost no endogenous enzymes in significant amounts hence, downstream processes are facilitated. For this experiment anew synthetic promoter was employed, expressing the protein under tight control simple by glycerol feed reduction.

Within the frame of the present work we demonstrated that protein production in *Pichia pastoris*can be further improved by new techniques and tools. This increases its value as host for the expression of heterologous proteins and makes it even more interesting for industrial applications.

Appendix

Table 1: *Pichia pastoris* strains generated during the thesis

CC Nummer	Name
3587	B1-GAP-CBH2 (CI)-CN3/4
3585	B1-d6*-CBH2 (CI)-D7_6-8 Kopien
3586	B1-d6*-CBH2 (CI)-B6_s.c.
3588	B1-d6*-CBH2 (Coda)_F6_21-29 Kopien
3591	B1-AOX-CBH2 (Coda)_C3_7 Kopien
3590	B1-AOX-CBH2 (Coda)_C9_4 Kopien
3870	B1-d1+-bman_E8_CN=1
3872	B1-d6*-bman_A3_CN=5
3876	B1-Gap-bman_F1_CN=1
3877	B1-d1+-XynA_A11_CN=18
3878	B1-d1+-XynA_D8_CN=9/10
3879	B1-d1+-XynA_C12_CN=6
3869	pPpB1-P(d1+)-bman_G2_CN=39
3871	pPpB1-P(d1+)-bman_F1_CN=16
3873	pPpB1-P(d6*)-bman_A11_CN=12
6249	pPpT4_S_2x2x_DNA20-V04-P3-C9-CN40
6294	pPpT4_S_2x2x_DNA20-V31-D9-CN1
6227	pPpT4_S_2x2x_DNA20-V27-CN=0.7-A4
6229	pPpT4_S_2x2x_DNA20-V43-CN=3.0-F11
6230	pPpT4_S_2x2x_DNA20-V43-CN=0.6-B8
6301	pPpT4_S_2x2x_DNA20-V11-CN=0.5-C9
3921	pPpT4_S_2x2x_DNA20-V09-CN4/5
6295	pPpT4_S_2x2x_DNA20-V19-F9-CN1.3
6296	pPpT4_S_2x2x_DNA20-V19-F12-CN0.5
6291	pPpT4_S_2x2x_DNA20-V04-s.c.
6245	pPpB1_S_2x2x_DNA20-V09-W-3-CN23
6246	pPpB1_S_2x2x_DNA20-V09-M-2-CN18
6185	pPpB1_S_2x2x_DNA20-V09-CN11/12
6292	pPpB1_S_2x2x_DNA20-V09-Mittel-10-CN4/5
3950	pPpB1_S_2x2x_V04-CN59
6247	pPpB1_S_2x2x_DNA20-V04-P3-F5-CN54
6248	pPpB1_S_2x2x_DNA20-V04-P3-F1-CN43
3968	pPpB1_S_2x2x_V04-CN21-29-B4

6184	pPpB1_S_2x2x_DNA20-V04-CN15/19
6293	pPpB1_S_2x2x_DNA20-V04-P3-G9-CN10
6251	pPpB1_S_2x2x_DNA20-V04-P1-E12-CN7
6290	pPpB1_S_2x2x_DNA20-V04-B12-CN3/4
3856	pPpT4_2x2x_S_CBH2-Claudia-S.c.
3857	pPpT4_d1+_smi_CBH2-Claudia-S.c.
3858	pPpT4_AOX_smi_CBH2-Claudia-S.c.
3859	pPpT4_GAP_smi_CBH2-Claudia-S.c.
3919	pPpB1_S_2x2x_CL-CN36+2x2x_PDI-CN13/15-G10
3920	pPpB1_S_2x2x_CL-CN36+2x2x_PDI-CN10/12-F4
3881	B1-d1+-CBH2-Claudia_s.c._ref_DNA2.0
3882	B1-2x2x-CBH2-Claudia_s.c._ref_DNA2.0
3883	B1-GAP-CBH2-Claudia_s.c._ref_DNA2.0

Table 2: *Escherichia coli* strains generated during the thesis

CC Nummer	Organismus	Name
3986	E.coli	B1_2xmat2x201_CBH2-V09
3892	E.coli	pADH_ADHTT_B1_2xmat2x201_CBH2-V04
3893	E.coli	pADH_ADHTT_B1_2xmat2x201_CBH2-V27
3855	E.coli	pKan3,2-8-synPDI_2x2x
3854	E.coli	pKan3,2-8-synPDI_d1+
3853	E.coli	pPpT4_2x2x_S_CBH2-Cl
3852	E.coli	pPpT4_d1+_smi_CBH2-Cl
3851	E.coli	pPpT4_GAP_smi_CBH2-Cl
3850	E.coli	pPpT4_AOX_smi_CBH2-Cl
3849	E.coli	pADH_ADHTT_B1_Δ6*_b-man
3848	E.coli	pADH_ADHTT_B1_GAP_b-man
3847	E.coli	pADH_ADHTT_B1_GAP_XynA
3846	E.coli	pADH_ADHTT_B1_d1+_XynA
3845	E.coli	pADH_ADHTT_B1_Δ6*_XynA
3916	E.coli	pKan3,2-8-Smi-synPDI_d1+-Strain-25
3884	E.coli	pKan3,2-8-Smi-synPDI_2x2x-Strain-27
3885	E.coli	pKan3,2-8-Smi-TKL_2x2x-Strain-130
3886	E.coli	pKan3,2-8-Smi-synPDI_d1+-Strain-26
3326	E.coli	pKan3,2-8-Smi-synPDI_2x2x-Strain-28
3328	E.coli	pKan3,2-8-Smi-TKL_d1+-Strain-135
3329	E.coli	pKan3,2-8-Smi-TKL_d1+-Strain-134
3330	E.coli	pKan3,2-8-Smi-TKL_2x2x-Strain-133

CURRICULUM VITAE

PERSONAL DATA

Name:	Andrea Mellitzer
Date of Birth:	03.06.1983
Place of Birth:	Leoben, Austria
Nationality:	Austria

EDUCATION & RESEARCH EXPERIENCE

1993 – 2001	Bundesrealgymnasium Kapfenberg
October 2001 – July 2008	Diploma studies at Graz University of Technology Technical Chemistry – Biotechnology, Biochemistry and Food Chemistry Diploma thesis "*Development of promoter and reporter system for transient gene expression analysis*" Supervisor: Ao.Univ.-Prof. Mag. Dr. Anton Glieder Institute of Molecular Biotechnology, TU Graz
Since July 2008	Doctoral studies at Graz University of Technology PhD thesis "*Expression and engineering of eukaryotic enzymes for polymer modification using Pichia pastoris*" Supervisor: Ao.Univ.-Prof. Mag. Dr. Anton Glieder

Institute of Molecular Biotechnology, TU Graz

PUBLICATIONS

PAPERS:

- Mellitzer, A., et al., *Expression of lignocellulolytic enzymes in Pichia pastoris.* Microb Cell Fact, 2012. **11**(1): p. 61.

- Mellitzer, A., et al., *Sensitive high-throughput screening for the detection of reducing sugars.* Biotechnol J, 2011. **7**(1): p. 155-62.

- Ruth, C., et al., *Variable production windows for porcine trypsinogen employing synthetic inducible promoter variants in Pichia pastoris.* Syst Synth Biol, 2010. **4**(3): p. 181-91.

POSTERS & PRIZES

- *Pichia,* 2012, Alpbach, Austria (Poster)

- *Recombinant Protein Production, 2011, Wien (3. Poster prize)*

- 14th International Biotechnology Symposium and Exhibition, 2010, Rimini (Poster)