Carbon source utilization during recombinant protein production in the methylotrophic yeast *Pichia pastoris*: novel strong promoters and specific transcriptional and translational responses

Dissertation

zur Erlangung des Doktorgrades an der Universität für Bodenkultur Wien

eingereicht von

Mag. Roland Prielhofer

Betreuer:

Univ. Prof. Dr. Diethard Mattanovich

Department für Biotechnologie

Universität für Bodenkultur Wien

Wien, Mai 2015

Acknowledgements - Danksagung

First I would like to thank my supervisors and mentors Diethard Mattanovich and Brigitte Gasser for their support and supervision during my work. I am very grateful for their help and guidance, and also for the freedom and chances I´ve got during my work!

I also thank all of my colleagues and coworkers. I really enjoyed the working atmosphere, the fun in the lab and all the nice parties we had. Special thanks for fruitful discussions go to Stefan Ertl, Nils Landes, Markus Buchetics, Marizela Delic, Fabian Schneider, Lars Töllner, Michael Maurer, Jonas Hohlweg, Minoska Valli and Alexandra Graf.

The bachelor and master students Christine Ruschitzka, Lisa-Sophie Handler and Christine Sanystra worked on related projects and were a great help. Thank you for that!

I would also like to thank our company partner LONZA for financial support, and especially our contacts Christoph Kiziak, Joachim Klein and Henriette Schneider for nice meetings and fruitful discussions.

Special thanks also go to our cooperation partners Roslyn Bill and Stephanie Cartwright. Thank you for supervision during my stay at Aston University in Birmingham and for the good cooperation!

Außerdem möchte ich mich bei all meinen Freunden bedanken, die immer für mich da waren und oft zur richtigen Zeit für Ablenkung gesorgt haben.

Ich danke insbesondere auch meinen Eltern. Vielen Dank für euren Glauben an mich und eure immerwährende Unterstützung!

Nicht zuletzt gilt mein größter Dank meiner großen Liebe Conny, die mir oft weitergeholfen hat und immer eine sehr wichtige Stütze für mich war!

Table of Contents

Abstract

The methylotrophic yeast *Pichia pastoris* (syn. *Komagataella* sp.) is a valuable expression host for recombinant protein production. Strain engineering aims to better understand cellular gene expression processes and to make production more efficient. Also, carbon source utilization is linked to production processes in *P. pastoris*: Methanol is often used to drive the expression of the gene of interest, although its use has disadvantageous implications.

The first part of this thesis focused on the identification and characterization of novel gene promoters in *P. pastoris* and both, inducible and constitutive promoter sequences suitable for protein production were discovered. Contrary to the already established strong methanol-driven promoters, these new inducible promoters are activated by glucose-limit, which is applied by glucose-based fed-batch cultivation. The gene controlled by the strongest of these promoters (P_{GTH1}) was shown to encode a high-affinity glucose transporter and named *GTH1*. An additional project was performed to study the functionality of P_{GTH1} in more detail and to generate engineered promoter variants with increased expression strength.

Inspired by these results, the second part of this thesis concentrated on a comprehensive gene expression and polysome analysis of *P. pastoris* in conditions typical for protein production. Translational regulation was found to be global rather than transcript-specific. Methanol induction could be connected to a general increase of translational activity, linking high protein productivity directly to the growth conditions in addition to promoter strength.

Zusammenfassung

Die methylotrophe Hefe *Pichia pastoris* (syn. *Komagataella* sp.) ist ein wertvoller Wirtsorganismus für rekombinante Proteinproduktion. Stammverbesserungsansätze zielen darauf ab, zelluläre Genexpressionsprozesse besser zu verstehen und nutzen zu können, um effizientere Produktionsprozesse zu ermöglichen. Außerdem spielt die Verwertung der Kohlenstoffquelle in *P. pastoris* eine Rolle bei Produktionsprozessen. Methanol wird oft verwendet um die Expression eines Zielgens zu forcieren, obwohl sein Einsatz Nachteile mit sich bringt.

Der erste Teil der vorliegenden Dissertation beschäftigt sich mit der Identifizierung und Charakterisierung neuer Gen-Promotoren in *P. pastoris*. Induzierbare und konstitutive Promotoren welche zur Proteinproduktion geeignet sind, konnten identifiziert werden. Im Gegensatz zu den bereits etablierten Methanol-induzierten Promotoren werden die neuen regulierbaren Promotoren im Glukose-Limit induziert, welches durch Glukose-basierte Fed-batch Kultivierung angewandt wird. Im Weiteren wurde die Funktion des Gens, welches unter Kontrolle des stärksten dieser neuen Promotoren (P_{GTH1}) liegt, analysiert, als hochaffiner Glucosetransporter identifiziert und *GTH1* benannt. Ein weiteres Projekt wurde durchgeführt um die detaillierte Funktionsweise von P_{GTH1} zu untersuchen und um Promoter-Varianten mit erhöhter Expressionsstärke zu generieren.

Diese Ergebnisse inspirierten dazu, im zweiten Teil dieser Arbeit eine umfassende Genexpressions- und Polysom- Analyse von *P. pastoris* in typischen Proteinproduktionsbedingungen durchzuführen. Es zeigte sich dass sich Translationsregulation eher global und kaum Transkript- spezifisch verhält. Die erhöhte Produktivität bei Methanol- Induktion, die bisher auf Promotorstärke zurückgeführt wurde, konnte aufgrund der allgemeinen Erhöhung der Translationsaktivität direkt zur Wachstumsbedingung verknüpft werden.

Aim of the study

At the beginning of this thesis, the objective was to identify and characterize novel gene promoters for recombinant protein production in *Pichia pastoris*. When the study was planned, only a few strong promoters were known and applied for production purposes in this host. Several *P. pastoris* genes such as *AOX1, AOX2, FDH, FLD1, DAS1 and DAS2* are essential for methanol utilization (MUT) and are characteristically highly expressed. The major MUT enzyme, alcohol oxidase 1 (Aox1), makes up one third of the cellular protein mass during methylotrophic growth and its promoter sequence (and those from other MUT genes) is often applied for protein production in *P. pastoris*.

On the other hand, industrial large-scale use of methanol has major drawbacks. First, methanol is highly flammable and toxic, and therefore requires health and safety considerations and could have implications for the final product and its purification. Second, methanol fed bioreactor cultivations have an increased oxygen demand and generate high heat evolution, making the process much more difficult to operate.

Therefore, the key criterion for our project was to identify novel promoters enabling strong expression without the use of methanol or any other foreign inducer substance. After successful completion of this goal, sequence-based promoter engineering was conducted for the strongest of the newly identified promoters (P_{GTH1}). Additionally, another novel promoter (P_{CS1}) with strong constitutive expression was identified and characterized.

The results of the first project raised questions regarding gene regulation due to growth conditions and carbon source utilization. Therefore we initiated a second project to investigate the responses of *P. pastoris* to different bioreactor-like conditions on both, transcriptional and translational level.

Background

This dissertation is a compilation of peer-reviewed articles and patent applications. Most of the background of this thesis is covered by our review entitled ´*Pichia pastoris*: protein production host and model organism for biomedical research´ (Gasser *et al.*, 2013 (Publication 2 of this thesis)). Therefore, this framework gives an overview on gene expression and its applications, highlighting new findings and topics which are relevant here.

Recombinant protein production capacities and capabilities are increasingly required for pharmaceutical, biotechnological and scientific applications. Production is often performed in the host *Escherichia coli*, but yeasts and mammalian cell lines are applied as well. Despite recent advances in *E. coli*, expression in higher organisms is beneficial due to efficient protein folding and post-translational modifications. Yeasts have advantages over mammalian cell lines for recombinant protein production because of the availability of broad molecular and genetic resources and their higher specific growth rates (Bill, 2014). Amongst others, the methylotrophic yeast *Pichia pastoris* is frequently used and under intense scientific investigation. The annual number of publications related to this yeast is constantly increasing since 20 years and it has reached about one tenth of the annual publications related to ´the yeast´ *Saccharomyces cerevisiae* (about 4500 annual publications).

Strain improvement uses –omics studies, directed evolution approaches and rational considerations to further improve the productivity of a certain strain. In this respect, the detailed understanding of gene expression processes and knowledge about related genes

and factors is of great value, especially because complex biological compounds such as immunoglobulins (Ig) cannot be produced in cell-free systems in considerable quantities.

Eukaryotic gene expression

The purpose of gene expression is the synthesis of a functional gene product, which can be a protein or a non-coding ribonucleic acid (ncRNA), such as ribosomal RNA (rRNA), transfer RNA (tRNA) and small nuclear RNA (snRNA). RNA is produced by RNA polymerases by transcription from genetic information. Proteins are synthesized by translation of messenger RNAs (mRNAs) into polypeptide chains by ribosomes. Further processes facilitate post-translational modifications (folding, structural changes, addition of functional groups to amino acids such as acetate, phosphate, glycans) and protein transport to assure proper functionality and localization of the protein.

In contrast to prokaryotes like *Escherichia coli*, eukaryotic cells are compartmentalized and possess specific organelles such as the endoplasmic reticulum (ER) and the Golgi apparatus, which enable highly complex gene expression processes and regulations. Most importantly, eukaryotes perform post-translational modifications and folding of complex proteins (e. g. immunoglobulins). Hence, gene expression steps are spatially separated in eukaryotes. Briefly, transcription takes place in the nucleus and goes along with or is followed by RNA splicing for intron-containing transcripts. Translation-associated RNAs undergo export through the nuclear pore, and might translocate to other cellular compartments (e. g. mitochondria) or stay in the cytoplasm and take part in active translation. The protein sequence is encoded by triplet nucleotides of the mRNA´s coding region, called codons. Each tRNA carries a certain amino acid and contains an anticodon which corresponds to the encoding codon. The ribosome catalyzes the binding of

anticodons to their respective codons, brings neighboring tRNAs closely together and synthesizes the polypeptide chain. Translation also takes place in the mitochondria or other compartments (e. g. synapses of neurons). Polypeptides are folded by proteins called chaperones and some are translocated into the ER through the Sec61 pore, and both of these processes can happen co- and post-translationally. Secretory proteins enter the secretion pathway, routing through the Golgi apparatus to vesicles excreting cargo to the exterior. More distant roles in gene expression are linked to cellular compartments dealing with degradation and storage processes, such as the vacuole and the lysosome.

Gene promoters and transcription factors

Transcription in eukaryotes is a thoroughly regulated and well organized process with a multiplicity of factors acting in concert. The key regulatory element of a gene is its promoter sequence (reviewed by Riethoven (2010)), which is usually located upstream (5´) of the coding sequence and comprises of core, proximal and distal promoter elements (see **Figure 1**).

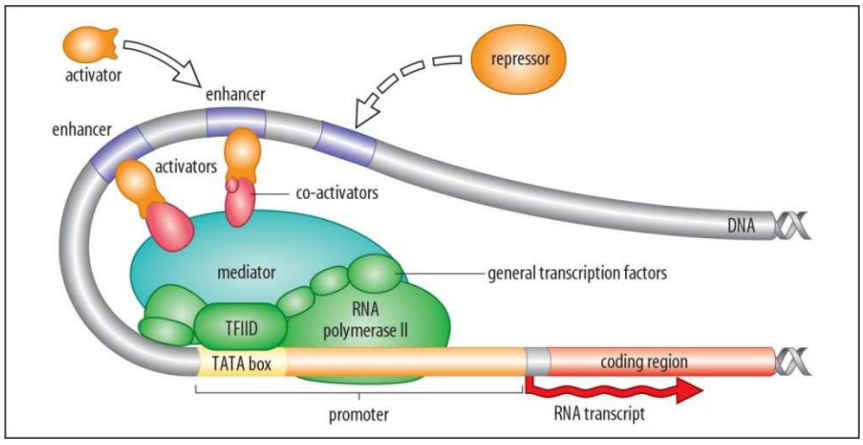

Figure 1: Schematic representation of transcription initiation at a eukaryotic promoter (figure taken from **Stavely (2014)**. Core promoter elements (such as TATA) act as binding site for TBP/TFIID complex and specific transcription factors (activators and repressors) bind to proximal/distal enhancer elements to regulate transcription.

Core promoter elements - which are best studied in vertebrates – serve as binding sites for several factors (general transcription factors). The TATA box, defined by the motif 5'-TATAAA-3', usually found 25 bp upstream of the start codon in eukaryotes, is the most ancient promoter element. TATA is bound by the TATA-binding protein (TBP) and TATA-associated factors (TAFs), forming the multi-subunit initiator complex TFIID (can also bind to the binding sites Inr (initiator element) and DPE (downstream promoter element)). The pre-initiation complex (PIC), which facilitates RNA polymerase II (RNAPII) positioning, DNA denaturation and DNA positioning to the active site of RNAPII, is built up by RNA polymerase II and basal transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH).

It became obvious that the concept of eukaryotic core promoters is more complex than initially thought. The mammalian PIC is often assembled independently of the TATA box. Most genes in higher eukaryotes possess dispersed core promoters and potential transcription starts sites (TSS) over a broad range of 50-100 bp. For yeast, this range was reported to be about 26 bp (Pelechano *et al.*, 2013).

Yeast core promoter elements are still poorly characterized and metazoan elements are thought to be generally missing, except for the TATA box, which is located between -40 and -120 bp of the TSS (Yang *et al.*, 2007). TATA boxes are be present in the vicinity of just ~20% of all *S. cerevisiae* genes (Basehoar *et al.*, 2004), but often also with improper consensus sequence (Lubliner *et al.*, 2013). It is assumed that transcription initiation in yeast depends on yet unidentified TAF-dependent core promoter motifs. TBP is complexed by 13 TAFs in yeast, forming the general transcription factor TFIID, and communication of TFIID to TFIIA was suggested to be bridged by TAF40 (Kraemer *et al.*, 2001).

A recent approach to understand where transcription starts and ends globally resulted in a unified model for yeast transcription (de Boer *et al.*, 2014). In agreement with experimental data, it indicates that promoters are largely defined by nucleosome-depleted regions (NDRs), which are formed due to sequence properties (G/C content, poly-A/T tracts and binding sites for chromatin-modifying factors).

Specific transcription factors

Gene-specific control is done via transcription factors (TFs) which bind to so-called enhancers and silencers (proximal and distal promoter elements). These DNA regions can be found up- or down-stream of the TSS, also very distantly and even on other chromosomes, and are often bound by more than one factor. Bound activators or repressors physically interact with the transcription machinery to execute their function. For this kind of interactions – which are important for promoters and transcription in general – so-called second order properties come into play: compared to non-promoter regions, promoters are characterized by lower stability, decreased bendability and enhanced curvature in pro- and eukaryotes (Kanhere & Bansal, 2005). Interestingly, previous work on DNA structure has revealed that different shapes other than the known double helix are also possible (Liu *et al.*, 2014), and this, in turn, could serve as another regulatory feature of DNA.

Transcription

When a stable and functional PIC is formed, the RNA polymerase initiates transcription and uses the template strand (non-coding) as template for RNA synthesis. Initiation can

break up, being abortive, or continue and result in productive transcription. Post-translational modifications of the C-terminal domain (CTD) of RNAPII (coined as the CTD code) is an important regulation mechanism for transcription and recruits different proteins (Shandilya & Roberts, 2012). Proximal histone modifications (known as histone code) also influence transcription initiation and elongation, for instance methylation at H3K4 and H3K27 are known for active and repressive chromatin states, respectively (Chiba *et al.*, 2010). The transcription factor P-TEFb is important for productive transcription, it phosphorylates the factors DSIF and NELFS (reverting their negative effect on transcription; both missing in yeast) and stimulates TFIIF for active transcription (Cheng & Price, 2007). During elongation, TFIIS, elongin, Rtr1/RPAP2 (yeast/human), Spt5, the histone chaperone complex (FACT) and other histone modifying factors (such as the SAGA complex) are crucial regulators (Shandilya & Roberts, 2012). Regulators often target CTD modifications of RNAPII, and these modifications are also modulated during its progress towards the 3´ end of the coding region. At completion of mRNA synthesis, termination can proceed in poly (A)- or Nrd1-Nab3-Sen1-dependent manner, though most protein coding genes in eukaryotes possess a highly conserved poly (A) signal. When transcribing the poly (A) signal, the RNAPII reduces processivity and pauses further downstream which results in endoribonucleolytic cleavage of the transcript and subsequent polyadenylation of the upstream- and degradation of the downstream product. Interestingly, *S. cerevisiae* lacks the RNAi pathway (involved in gene regulation in eukaryotes), but possesses unique 3´ end processing mechanisms for its CUTs and SUTs (cryptic unstable- and stable uncharacterized transcripts). After completion, the pre-mRNA and the RNAPII and factors of the transcription machinery are released and might be recycled for another round of transcription, which was described as a process termed ´gene looping´ (Wang *et al.*, 2010).

Two other RNA polymerases are known in eukaryotes. RNAPI transcribes ribosomal RNAs (except 5S rRNA) and its initiation depends on an upstream control sequence (UCS), bound by the upstream binding factor which recruits TBP and other associated factors (called selective factor 1, SL1), leading to Rrn3/TIF-IA phosphorylation and transcription start (Russell & Zomerdijk, 2006). RNAPIII synthesizes structural and catalytic RNAs (5S rRNA, tRNA and small RNAs) and three respective classes (I, II and III) of transcription initiation are known (Schramm & Hernandez, 2002).

Epigenetic regulation

DNA and histone proteins are packed into nucleosomes and, following higher order structures, into chromatin. Euchromatin is less tightly packed than heterochromatin, and is therefore known to be transcriptionally more active. Transcription factors and coactivators regulate chromatin organization, eventually leading to nucleosome depletion, which is a key event for regulatory functions and transcription (reviewed by Rando & Winston (2012)). It seems that the epigenetic regulation platform is more complex than initially anticipated and well interconnected with the whole transcription regulation machinery through chromatin remodelers and chaperones, ncRNA, TFs and others. Interestingly, stress genes are often regulated by many chromatin-regulating genes and their expression is noisier compared to housekeeping genes (Rando & Winston, 2012).

Transcriptional networks

Enormous research effort on yeast gene expression provides an advanced understanding of transcription at the genome scale today (reviewed by Hughes & de Boer (2013)). High-confidence binding site data is available for the majority of *S. cerevisiae* TFs and transcription initiation mechanisms of many promoters can be explained. On the other hand, TFs act differently at different sites – alone or cooperatively, through other factors, together with chromatin-modifying factors, and some even possess two distinct DNA binding motifs (Gordan *et al.*, 2011).

However, TFs were assigned to certain conditions such as oxidative stress (Toledano *et al.*, 2013), heat shock (Lee *et al.*, 2000), unfolded protein response (UPR) (Walter & Ron, 2011), osmotic regulation (Ni *et al.*, 2009) and cellular processes such as cell cycle (Iyer *et al.*, 2001), diauxic shift and stationary phase (Galdieri *et al.*, 2010).

Advanced methods, e. g. ChIP (chromatin immunoprecipitation) or the “calling-cards” (based on Sir4 fusion of TFs, which facilitates Ty5 retrotransposon integration at TF binding sites) method, were important for new discoveries and further methodical innovations will be essential for future research and applications.

The life cycle of mRNAs

After biogenesis by RNAPII, the 5´end of eukaryotic mRNA is capped (7-methylguanine nucleotide, 5’ to 5’ triphosphate linkage), introns are removed by the spliceosome (RNA-protein complex) and the 3´end is matured and polyadenylated. Numerous mRNA-maturing factors and heterogeneous nuclear ribonucleoproteins (hnRNPs) associate and finally pack into mature mRNPs – a process that has also been named ´mRNA imprinting´

(Choder, 2011). The formation of export-competent mRNPs is dependent on the transcription/export (TREX) complex (THO complex: Hpr1, Mft1, Tho2, Thp2, and mRNA export factor RNA helicase Sub2), RNA binding protein Yra1, export receptor Mex67:Mtr2, many hnRNPs (Gbp1, Hrb1, Tex1) and the poly(A)-binding protein Pab1 (Das & Das, 2013). The transition from the nucleolar to the cytoplasmic phase via the nuclear pore complex (NPC) is continued by the (single) pioneer round of translation, through which mRNPs undergo major structural remodeling and finally go into productive translation (Maquat *et al.*, 2010). Transport, specific localization or storage follows for some mRNPs. All processing steps are linked with and influence each other, and quality control mechanisms are present at all stages (reviewed by Das & Das (2013)).

However, the abundance of an mRNA is not only determined and regulated by its synthesis, but also by its degradation, and newest results showed that degradation processes are functionally linked to quality control pathways (Das & Das, 2013). Remarkably, gene promoters in yeast were shown to affect both, transcription and decay of its gene, likely due to imprinting (Dori-Bachash *et al.*, 2012, Bellofatto & Wilusz, 2011).

Eukaryotic mRNA structure

Sequence elements up- and downstream of the coding sequence such as the hairpin structure, internal ribosome entry site (IRES) and poly(A) tail are summarized as untranslated regions (UTRs). These sites are the determinants of the translational activity, subcellular localization and stability of the mRNA (see **Figure 2**).

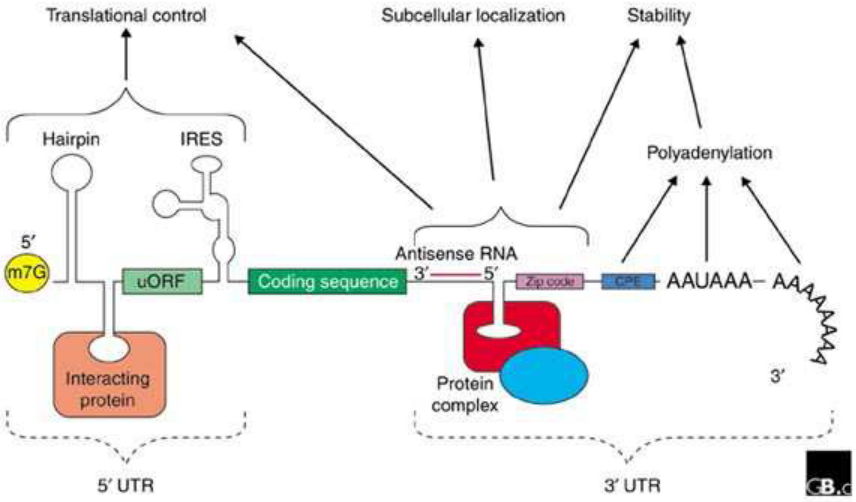

Figure 2: Eukaryotic mRNA structure (figure taken from Mignone *et al.* (2002)).

As mentioned before, transcriptional heterogeneity is quite common (~13% of all yeast genes possess uORFs (Lawless *et al.*, 2009)). Transcripts with alternative TSS, and hence alternative 5´UTRs, were reported to differ in translational activity, cap-independent translation (facilitated by internal ribosome entry site (IRES)) and even to act as on/off switch for certain genes (Rojas-Duran & Gilbert, 2012). Downstream elements (located in the 3´UTR) are significantly associated with stability and subcellular localization of mRNAs and interestingly, the presence or absence of certain motifs is correlated with different steady-state levels (Shalgi *et al.*, 2005). The poly(A) tail also makes a difference, but the question whether longer or shorter tails are better in terms of gene expression is not easily answered (Jalkanen *et al.*, 2014).

Protein synthesis

The process of protein biosynthesis, in which mRNA is translated into a polypeptide, can be divided into three steps: initiation, elongation and termination. Translation initiation is considered as the most important part, because of its regulatory features (reviewed by Jackson et al. (2010)).

At the start of the translation process, the small ribosomal subunit (40S) binds to the mRNA via proteins associated to the mRNA´s 5´UTR and to the 5´cap via eIF4G, which is a part of the eIF4F complex (consisting of eIF4A, eIF4E and eIF4G). Initiation might also start independently of the 5´ cap, mediated by the internal ribosome entry site (IRES). The eIF4G complex also interacts with eIF3 (associated to the 40S subunit, preventing the large subunit (60S) from binding) and poly(A)-binding protein (PABP), which was reported to be relevant for mRNA circularization during translation (Malys & McCarthy, 2011). The small subunit and methionine-charged tRNA (and eIF3, eIF1, eIF1A and probably eIF5) form the pre-initation complex (43S), and move in 5´ - 3´ direction until recognizing the start codon (usually AUG). Interfering secondary structures are resolved by eIF4A, eIF4B and eIF4F. At the start codon, eIF2-bound GTP is hydrolyzed (promoted by eIF5 and eIF5B), which in turn leads to unbinding of eIFs and to association of the 60S subunit, thereby forming the elongation-competent 80S ribosome.

Elongation, termination and recycling phases of translation in eukaryotes were reviewed by Dever *et al.* (2012). In short, the elongation phase starts with binding of aminoacyl-tRNA by the eukaryotic elongation factor eEF1A, which directs it to the A (aminoacyl-tRNA) site of the ribosome and hydrolyses GTP. When the codon is recognized by the tRNA, the peptide bond formation is facilitated by the peptidyl transferase center (PTC)

of the ribosome. Translocation to E and P (peptidyl-tRNA and exit) sites is done by eEF2 and the cycle starts again. In contrast to other eukaryotes, yeasts and fungi need another elongation factor, eEF3 – which is an ATPase required for the binding of the aminoacyl-tRNA-eEF1A-GTP ternary complex to the A site and likely facilitates clearance of deacyl-tRNA from the E site. When the stop codon is reached, translation termination occurs in dependence of eukaryotic release factors eRF1 and eRF3.

Translational control

Protein biosynthesis can either be controlled globally e. g. through initiation factors or transcript-specific via regulatory proteins or non-coding RNAs (Gebauer & Hentze, 2004). Because of its high energy demand, cells quickly down-regulate translation as response to stress (e. g. amino acid starvation, glucose depletion, and oxidative stress). Several initiation factors (eIFs) and associated proteins (eIF-binding proteins, protein kinases) are involved in global reactions, but the main cellular regulator(s) still need to be identified (reviewed by Simpson & Ashe (2012)). Another – possibly underestimated – global mechanism is ribosome heterogeneity: specialized ribosomes translate certain mRNA subsets and expand expression diversity (Filipovska & Rackham, 2013).

Protein folding and secretion

Nascent proteins are co- and post-translationally folded by chaperones to ensure that proteins can fulfill their biological function. Two important chaperones are SSB-RAC (stress 70 B–ribosome-associated complex) and NAC (nascent polypeptide–associated complex), and they are both ribosome-tethered, functionally connected and were found to control ribosome biogenesis (Koplin *et al.*, 2010). Depending on the target localization which is given by the signal peptide (5-30 N-terminal amino acids), proteins can be

translocated into the ER via co- or post- translational translocation, whereas the latter is depending on the signal recognition particle (SRP) in yeast. Products targeted to be localized within the secretory pathway, the cell wall, the periplasmic space (correctly termed inner wall space) or in the extracellular space enter the secretory pathway (reviewed by Delic *et al.* (2013)). Improper folding and accumulation of misfolded proteins induces the unfolded protein response (UPR), which in turn leads to the up-regulation of genes involved in protein folding and ER-mediated protein degradation (ERAD). Protein folding and secretion can be limiting processes for recombinant protein production (Delic *et al.*, 2014).

Recombinant protein production in *Pichia pastoris*

The methylotrophic yeast *Pichia pastoris* can grow to more than 150 g/L dry cell weight and has a high capacity for the production of secretory proteins. Glyco-engineered strains are available which can be used to generate human-like and tailor-made glycosylations (De Pourcq *et al.*, 2010, Hamilton & Gerngross, 2007, Bollok *et al.*, 2009). Its qualities and recent developments have made *P. pastoris* a highly valuable workhorse for biotechnological applications (Vogl *et al.*, 2013, Goncalves *et al.*, 2013).

Numerous different proteins for biopharmaceutical, biomedical and technical applications are produced in this host and very high product titers are possible (e. g. Mellitzer *et al.* (2014)and Hao *et al.* (2013)). Unfortunately, not all proteins can be produced easily.

Approaches to improve *P. pastoris* production processes are versatile and have brought substantial improvements. Different gene promoters (Gasser et al., 2013, Vogl *et al.*, 2014a), including libraries (Qin *et al.*, 2011) and synthetic promoter variants (Hartner *et*

al., 2008), can be used to enforce the transcription of the target gene at different levels. Amongst all available promoters in *P. pastoris*, those derived from genes involved in the methanol utilization pathway are of great interest for recombinant protein expression. Alcohol oxidase 1, the most prominent candidate, is strongly induced– and accounts for up to one third of the cellular protein during methylotrophic growth (Veenhuis *et al.*, 1983). On the other hand, methanol use has major drawbacks. Considering its industrial application, methanol needs safety precautions and leads to very high oxygen consumption and heat evolution during the fed batch process, which is problematic at large scale (Niu *et al.*, 2013). Another aspect is that methanol could be harmful or disruptive to the protein product and needs nearly complete clearance from the final product. In addition, methanol feeding is difficult to handle: overfeeding can deteriorate the production performance and even electrode-based control was reported to be complex to establish (Gao *et al.*, 2013).

Another key parameter for tuning recombinant expression in *P. pastoris* is the gene copy number (Hohenblum *et al.*, 2004). Post-transformational vector amplification can boost the expression of integrated genes and is therefore applied to generate improved production strains (Marx *et al.*, 2009, Aw & Polizzi, 2013). Furthermore, translation efficiency and proper protein secretion are dependent on coding sequence properties: codon usage and signal peptides potentially improve productivity in *P. pastoris* (Lin-Cereghino *et al.*, 2013, Mellitzer et al., 2014).

Heterologous gene expression largely affects and possibly exhausts gene expression processes, heat shock response, ERAD, UPR and other cellular functions (Young & Robinson, 2014). Overexpression of endogenous factors (sometimes called helper factor

genes) can be used to relieve this burden and thereby improve productivity of high yielding strains (Delic et al., 2014, Yu *et al.*, 2014, Gasser *et al.*, 2008). Potential interaction partners for antibody Fab fragments were identified in an interactome study (Pfeffer *et al.*, 2012).

Last but not least, bioreactor cultivation and downstream processing can be optimized to make production processes more efficient (e. g. Hao et al. (2013)).

Promoter regulation in *P. pastoris*

Inducible promoters are dependent on specific transcription factors (for a comprehensive review on promoter regulation in *P. pastoris* see Vogl & Glieder (2013)). *P. pastoris AOX1* and *AOX2* promoters are known to be repressed in non-growth limiting conditions by glycerol, glucose and ethanol and to be strongly induced by methanol, but just little is known about the responsible transcription factors, and the regulation of these genes is species dependent (Hartner & Glieder, 2006). Mxr1 (methanol expression regulator, homologuous to Adr1 in *S. cerevisiae*) regulates genes involved in methanol utilization, peroxisome biogenesis and likely also in the activation of β-oxidation. The regulatory protein 14-3-3 was suggested to be its regulator in a carbon-source dependent manner (Parua *et al.*, 2012). Rop (repressor of phosphoenolpyruvate carboxykinase) exhibits the same DNA binding specificity as Mxr1, acts antagonistically and is the first identified negative regulator of methanol metabolism in yeast (Kumar & Rangarajan, 2012). New evidence showed that the transcription factor Trm1 (homologous to Mpp1 in *Hansenula polymorpha*) is also essential for MUT gene induction (Sahu *et al.*, 2014).

Carbon source- and glucose repression mechanisms have been under investigation for a long time, and involved promoters have a big potential for recombinant protein

production in fed batch cultivation (Weinhandl *et al.*, 2014). Glucose signal transduction is mediated by hexose transporters and hexokinases, leading to increased intracellular AMP:ATP ratios which activate Snf1, a central kinase directly involved in gene regulation by carbon sources (Wilson *et al.*, 1996). Regarding glucose repression, Mig1 is the most important factor. It localizes to the nucleus at high glucose levels and recruits the Ssn6-Tup1 repressor complex which leads to conformational chromatin changes making transcription factor binding sites inaccessible (Bu & Schmidt, 1998).

However, transcription factor function is subject to fast evolutionary adaption: Aft1, which is known as regulator of ferrous transport in *S. cerevisiae*, lacks its iron- binding motif in *P. pastoris*. Instead, it was found to control the expression of genes involved in carbohydrate metabolism and many secretory genes, thereby offering itself for overexpression as helper factor (Ruth *et al.*, 2014). For a more detailed understanding of individual gene and promoter regulation in *P. pastoris*, further work on specific transcription factors, synthetic approaches (e. g. Vogl et al. (2014a)) and novel methods are required in future.

Results and Discussion

The first part of this thesis was a project in collaboration with LONZA (Visp, CH), intended to identify novel strong inducible promoters in *P. pastoris*. Since mainly methanol-inducible promoters such as P_{AOX1} were available for heterologous protein production prior to this project, the aim was to find alternative inducible promoters to avoid the need of methanol for strong induction. Methanol use is disadvantageous in industrial fermentation because it requires safety precautions and leads to strongly enhanced heat evolution and oxygen demand. As a follow-up, two other projects focused on sequence-based engineering of one regulated promoter and on a strong constitutive promoter.

For the second part of this thesis, inspired by the results of the first, a comprehensive gene expression study was conducted in cooperation with the group of Prof. Roslyn Bill. (Aston University, Birmingham, UK) Transcriptome- and polysome profiles, and translational regulation was analyzed in *P. pastoris* wildtype cells grown in four different bioreactor-like conditions. Despite transcriptional regulations, the aim was to compare translational activity and regulations. Therefore, the method of translational profiling was established for *P. pastoris*. Cultivations and sample preparations were done at Aston University in Birmingham (UK) and the sample processing, microarray and data analysis were performed at BOKU Vienna.

Novel promoters enable high-level expression of heterologous proteins in *Pichia pastoris*

Gene promoters, the first basis determining transcription, can simply be cloned upstream of a gene of interest. Depending on the application, a few characteristics need to be considered: how strong should the promoter be and should it enable constitutive or

regulatable transcription? For *P. pastoris*, promoters derived from genes of the methanol utilization pathway (*AOX1, AOX2, FLD1, DAS1, DAS2* and others) are often used for strong inducible expression (Tschopp *et al.*, 1987, Ohi *et al.*, 1994, Shen *et al.*, 1998), while strong constitutive expression is featured by the GAP promoter (Waterham *et al.*, 1997).

With this work we aimed to provide alternative promoters, basically to avoid the addition of inducer substances such as methanol. Novel inducible promoter candidates were identified by browsing for genes with good regulatory properties and superior induction strength in transcript profiles of *P. pastoris* cells grown in bioreactor cultivations (Prielhofer *et al.*, 2013 (Publication 1 of this thesis), Mattanovich *et al.*, 2013). These promoters are simply switched on by glucose limitation – a condition present during the fed batch in the bioreactor.

Initial screenings with clones expressing eGFP under control of these novel P_G promoters could verify the induction properties of the novel promoters. Besides, their induction behavior in decreasing glucose concentrations was shown. Further screenings and bioreactor cultivations with clones expressing human serum albumin, carboxypetidase B and Fab driven by the novel P_G promoters were successful and product titers exceeded those obtained under the control of P_{GAP} by more than two fold. Additionally, the gene under control of the strongest novel promoter (P_{GTH1}), was functionally analyzed and found to encode a high-affinity glucose transporter.

Further work focused on engineering of P_{GTH1} (unpublished manuscript Prielhofer *et al.*, 2015 (Publication 4 of this thesis). In a sequence-based approach, we could clearly identify a regulatory motif which is essential and characteristic for P_{GTH1} regulation: it stretches over 32 bp and contains four carbon source-related transcription factor binding sites

(F$RGT1.02, F$SIP4.01, F$CSRE.01 and F$RGT1.01) and likely involves the transcription factors Rgt1, Cat8-1 and/or Cat8-2. Most importantly, we could successfully generate promoter variants with improved induction whereas repression is unchanged.

Furthermore, a strong constitutive promoter called P_{CS1} which controls the gene with the highest transcript level (Mattanovich *et al.*, 2014) could be identified and characterized. P_{CS1}- driven product titers were twofold in screenings and more than threefold in bioreactor cultivations, compared to expression driven by the P_{GAP} promoter, respectively.

***Pichia pastoris* regulates its response to different carbon sources at the transcriptional, rather than the translational, level**

Transcriptional regulation has been studied for a long time in yeast, and since a decade also in *P. pastoris* (Liang *et al.*, 2012, Vogl *et al.*, 2014b, Sauer *et al.*, 2004, Graf *et al.*, 2008, Dragosits *et al.*, 2010, Baumann *et al.*, 2011). Compared to that, just a few approaches were dedicated to examine translational regulation in *S. cerevisiae*. Protein synthesis is essential for each and every cell and its proper function is important for recombinant protein production. Translational regulation is known to exist on a global and on a transcript-specific level. To investigate these regulations, the method of choice is polysome profiling (also called ribosome- or translational- profiling) which is based on the fixation of ribosomes on mRNAs using cycloheximide, followed by polysome isolation, profiling and optional purification and further analysis (for a comprehensive review see e. g. Ingolia (2014)).

A cooperation project with Prof. Roslyn Bill and Stephanie Cartwright from Aston University in Birmingham (UK) was initiated to perform translational profiling experiments with *P. pastoris* (Prielhofer *et al.*, 2015 (Publication 3 of this thesis)). Our main goal was to

clarify how different carbon sources, in particular methylotrophic growth and glucose limit, affect translation and protein production.

Translation profiles revealed that growth conditions have an impact on global translational activity, but that transcript-specific translational regulation is minimal. Transcription, on the other hand, was found to be regulated extensively in response to different carbon sources.

Superior promoters derived from methanol utilization genes are known to enable high level expression and that is why methylotrophic growth is often used for protein production. The present translation profiling experiments could reveal that translational activity is greatly enhanced during growth on methanol compared to excess glucose or glycerol, while the growth rate is less than half in the former condition (μ~0.1 h^{-1} compared to μ~0.23 h^{-1}). This suggests that the methanol condition itself provides superior prerequisites for protein production.

Glucose repression was described for genes related to glycerol uptake and metabolism, gluconeogenesis, metabolism of alternative carbon sources and glycolysis. Contrary to *S. cerevisiae*, respiratory processes and mitochondrial genes were found to be independent from glucose repression in *P. pastoris*. Carbon substrate suppression, meaning repression in surplus amounts of other carbon sources such as glycerol and glucose, was found for genes linked to methanol utilization and peroxisomes.

For *S. cerevisiae*, it was shown that UPR induction by DTT does not result in a major redistribution of polysome peaks (Payne *et al.*, 2008) (which means that global translational activity remains constant), on the contrary a shift from a fermentable to a non-fermentable carbon source does indeed (Kuhn *et al.*, 2001). Transcript-specific

translational regulation was found for both of these conditions and for oxidative stress (Shenton *et al.*, 2006). Polysome profiling is not only used to investigate gene regulation, it also adds to the understanding of the translation process itself. Transcript heterogeneity (Pelechano et al., 2013), ribosome diversity (Filipovska & Rackham, 2013) and other regulatory mechanisms such as upstream ORFs (Waern & Snyder, 2013) attract attention, and might possibly serve as engineering targets in future.

Global translation can be controlled by GAAC (general amino acid control) and TOR (target of rapamycin), both of them connecting to the inhibition of eukaryotic initiation factor eIF2 (mediates the binding of $tRNA^{met}$ to the ribosomal subunit 40S); other mechanisms are known and still under investigation (Simpson & Ashe, 2012). Yeast Gcn2 phosphorylates – and thereby reduces the activity – of eIF2α in response to amino acid starvation (Garcia-Barrio *et al.*, 2000). Interestingly, the transcriptional activator Gcn4 is translationally induced when eIF2 gets inactivated. This was proposed to rely on a mechanism involving four uORFs which prevent initiation at the Gcn4 start codon in non-starving conditions or high ternary complex (eIF2·GTP·Met-$tRNA_i^{Met}$) concentration (Hinnebusch, 2005). However, Shah and colleagues used a model to depict the rate-limiting step in yeast translation and concluded that protein production in healthy yeast cells depends on ribosome availability, while initiation and elongation rates are more important during stress (Shah *et al.*, 2013).

Conclusion and Outlook

Recombinant protein production in *P. pastoris* often takes advantage of its strong inducible promoters during methylotrophic growth. On the contrary, the use of methanol at industrial scale has drawbacks - it requires safety precautions and has technical

implications (addressed in the background´s section). Other promoters with beneficial regulation and strength were hardly available prior to this work.

The first objective of this thesis was to provide new promoters with superior properties, avoiding the use of methanol. Therefore, microarray analysis was successfully used to identify appropriate endogenous promoter sequences enabling both, strong regulatable (P_{G1}(=P_{GTH1}), P_{G3}, P_{G4}, P_{G6}, P_{G7} and P_{G8}) or strong constitutive (P_{CS1}) expression. The P_G promoters are derived from genes encoding hexose transporters (promoters P_{GTH1} and P_{G7}) and genes with possible roles in central metabolism (promoters P_{G3}, P_{G4}, P_{G6} and P_{G8}), and the protein encoded downstream of the constitutive P_{CS1} is predicted to be GPI-anchored on the cell surface. The P_G promoters are activated by limited glucose addition during fed batch in the bioreactor, and tightly repressed by carbon excess in the batch phase. Several proteins were successfully produced under control of these promoters and product titers exceeding those obtained in P_{GAP}-driven expression could be reached in small-scale and in bioreactor cultivation.

In another project, P_{GTH1} variants were generated aiming to better understand its regulation and to engineer the promoter sequence to further improve its strength. Therefore, promoter variants with altered length and deletions were cloned and screened. The main regulatory region was found to be located between -400 and -200 bp upstream of the ATG, improved promoter variants (unchanged repression) were obtained and its most important transcription factor binding sites could be identified.

In the second part of this thesis, a comprehensive gene expression study on transcription and translation in *P. pastoris* grown in different bioreactor-like conditions was used to better understand condition-specific differences. Thereby we could reveal that

methylotrophic growth strongly enhances translational activity. This indicates the fundamental influence of methanol utilization on the protein expression capacity and puts the idea of superior productivity on methanol in a different perspective.

Translation can also be controlled in a transcript-specific manner, especially in stress conditions. We showed that translational regulation in response to different carbon sources is minimal. Rather, the translational ratio (relative polysomal abundance) of individual transcripts strongly correlates with ORF length and/or transcript level. Genes become shortened and codon biased during evolutionary optimization, which enhances both, its transcript abundance and relative translational activity. Hence, it is important to understand and to take codon usage into consideration for heterologous genes.

The transcript profiles were used to make general observations and to define gene groups responding to glucose and carbon substrate repression. Most importantly, genes related to glycerol uptake and metabolism, glycolysis, gluconeogenesis and metabolism of alternative carbon sources are coordinated by glucose repression, while methanol utilization and peroxisomal genes underlie carbon substrate repression. Respiration and mitochondrial genes do not respond to glucose repression in the Crabtree-negative yeast *P. pastoris*, which is a fundamental difference to *S. cerevisiae*. The responsible transcription factors remain to be identified.

Methanol-dependent protein production processes for *P. pastoris* are predominant. Strain engineering approaches might enable to create *P. pastoris* strains with translation capacities similarly high on all carbon sources. However, it is questionable if it is possible to engineer changes like the ones described here or if this task would need a whole new

strategy. This directs towards the creation of designed cell factories, but the current knowledge of eukaryotic gene expression processes restricts rational strain engineering.

References

Aw, R. & K.M. Polizzi, (2013) Can too many copies spoil the broth? *Microb Cell Fact* **12**: 128.

Basehoar, A.D., S.J. Zanton & B.F. Pugh, (2004) Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **116**: 699-709.

Baumann, K., L. Dato, A.B. Graf, G. Frascotti, M. Dragosits, D. Porro, D. Mattanovich, P. Ferrer & P. Branduardi, (2011) The impact of oxygen on the transcriptome of recombinant *S. cerevisiae* and *P. pastoris* - a comparative analysis. *BMC Genomics* **12**: 218.

Bellofatto, V. & J. Wilusz, (2011) Transcription and mRNA stability: parental guidance suggested. *Cell* **147**: 1438-1439.

Bill, R.M., (2014) Playing catch-up with Escherichia coli: using yeast to increase success rates in recombinant protein production experiments. *Front Microbiol* **5**: 85.

Bollok, M., D. Resina, F. Valero & P. Ferrer, (2009) Recent patents on the *Pichia pastoris* expression system: expanding the toolbox for recombinant protein production. *Recent Pat Biotechnol* **3**: 192-201.

Bu, Y. & M.C. Schmidt, (1998) Identification of cis-acting elements in the SUC2 promoter of *Saccharomyces cerevisiae* required for activation of transcription. *Nucleic Acids Res* **26**: 1002-1009.

Cheng, B. & D.H. Price, (2007) Properties of RNA polymerase II elongation complexes before and after the P-TEFb-mediated transition into productive elongation. *J Biol Chem* **282**: 21901-21912.

Chiba, K., J. Yamamoto, Y. Yamaguchi & H. Handa, (2010) Promoter-proximal pausing and its release: molecular mechanisms and physiological functions. *Exp Cell Res* **316**: 2723-2730.

Choder, M., (2011) mRNA imprinting: Additional level in the regulation of gene expression. *Cell Logist* **1**: 37-40.

Das, S. & B. Das, (2013) mRNA quality control pathways in *Saccharomyces cerevisiae*. *J Biosci* **38**: 615-640.

de Boer, C.G., H. van Bakel, K. Tsui, J. Li, Q.D. Morris, C. Nislow, J.F. Greenblatt & T.R. Hughes, (2014) A unified model for yeast transcript definition. *Genome Res* **24**: 154-166.

De Pourcq, K., K. De Schutter & N. Callewaert, (2010) Engineering of glycosylation in yeast and other fungi: current state and perspectives. *Appl Microbiol Biotechnol* **87**: 1617-1631.

Delic, M., R. Gongrich, D. Mattanovich & B. Gasser, (2014) Engineering of protein folding and secretion-strategies to overcome bottlenecks for efficient production of recombinant proteins. *Antioxid Redox Signal* **21**: 414-437.

Delic, M., M. Valli, A.B. Graf, M. Pfeffer, D. Mattanovich & B. Gasser, (2013) The secretory pathway: exploring yeast diversity. *FEMS Microbiol Rev* **37**: 872-914.

Dever, T.E. & R. Green, (2012) The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harb Perspect Biol* **4**.

Dori-Bachash, M., O. Shalem, Y.S. Manor, Y. Pilpel & I. Tirosh, (2012) Widespread promoter-mediated coordination of transcription and mRNA degradation. *Genome Biol* **13**: R114.

Dragosits, M., J. Stadlmann, A. Graf, B. Gasser, M. Maurer, M. Sauer, D. Kreil, F. Altmann & D. Mattanovich, (2010) The response to unfolded protein is involved in osmotolerance of *Pichia pastoris*. *BMC Genomics* **11**: 207.

Filipovska, A. & O. Rackham, (2013) Specialization from synthesis: how ribosome diversity can customize protein function. *FEBS Lett* **587**: 1189-1197.

Galdieri, L., S. Mehrotra, S. Yu & A. Vancura, (2010) Transcriptional regulation in yeast during diauxic shift and stationary phase. *Omics* **14**: 629-638.

Gao, M.J., X.B. Zhan, Z.Y. Zheng, J.R. Wu, S.J. Dong, Z. Li, Z.P. Shi & C.C. Lin, (2013) Enhancing pIFN-alpha production and process stability in fed-batch culture of *Pichia pastoris* by controlling the methanol concentration and monitoring the responses of OUR/DO levels. *Appl Biochem Biotechnol* **171**: 1262-1275.

Garcia-Barrio, M., J. Dong, S. Ufano & A.G. Hinnebusch, (2000) Association of GCN1-GCN20 regulatory complex with the N-terminus of eIF2alpha kinase GCN2 is required for GCN2 activation. *Embo j* **19**: 1887-1899.

Gasser, B., R. Prielhofer, H. Marx, M. Maurer, J. Nocon, M. Steiger, V. Puxbaum, M. Sauer & D. Mattanovich, (2013) *Pichia pastoris*: protein production host and model organism for biomedical research. *Future Microbiol* **8**: 191-208.

Gasser, B., M. Saloheimo, U. Rinas, M. Dragosits, E. Rodríguez-Carmona, K. Baumann, M. Giuliani, E. Parrilli, P. Branduardi, C. Lang, D. Porro, P. Ferrer, M. Tutino, D. Mattanovich & A. Villaverde, (2008) Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microb Cell Fact* **7**: 11.

Gebauer, F. & M.W. Hentze, (2004) Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol* **5**: 827-835.

Goncalves, A.M., A.Q. Pedro, C. Maia, F. Sousa, J.A. Queiroz & L.A. Passarinha, (2013) *Pichia pastoris*: a recombinant microfactory for antibodies and human membrane proteins. *J Microbiol Biotechnol* **23**: 587-601.

Gordan, R., K.F. Murphy, R.P. McCord, C. Zhu, A. Vedenko & M.L. Bulyk, (2011) Curated collection of yeast transcription factor DNA binding specificity data reveals novel structural and gene regulatory insights. *Genome Biol* **12**: R125.

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

Hamilton, S. & T. Gerngross, (2007) Glycosylation engineering in yeast: the advent of fully humanized yeast. *Curr Opin Biotechnol* **18**: 387-392.

Hao, J., L. Xu, H. He, X. Du & L. Jia, (2013) High-level expression of Staphylococcal Protein A in *Pichia pastoris* and purification and characterization of the recombinant protein. *Protein Expr Purif* **90**: 178-185.

Hartner, F.S. & A. Glieder, (2006) Regulation of methanol utilisation pathway genes in yeasts. *Microb Cell Fact* **5**: 39.

Hartner, F.S., C. Ruth, D. Langenegger, S.N. Johnson, P. Hyka, G.P. Lin-Cereghino, J. Lin-Cereghino, K. Kovar, J.M. Cregg & A. Glieder, (2008) Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. *Nucleic Acids Res* **36**: e76.

Hinnebusch, A., (2005) Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol* **59**: 407-450.

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

Hughes, T.R. & C.G. de Boer, (2013) Mapping yeast transcriptional networks. *Genetics* **195**: 9-36.

Ingolia, N.T., (2014) Ribosome profiling: new views of translation, from single codons to genome scale. *Nat Rev Genet* **15**: 205-213.

Iyer, V.R., C.E. Horak, C.S. Scafe, D. Botstein, M. Snyder & P.O. Brown, (2001) Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409**: 533-538.

Jackson, R.J., C.U. Hellen & T.V. Pestova, (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11**: 113-127.

Jalkanen, A.L., S.J. Coleman & J. Wilusz, (2014) Determinants and implications of mRNA poly(A) tail size - Does this protein make my tail look big? *Semin Cell Dev Biol*.

Kanhere, A. & M. Bansal, (2005) Structural properties of promoters: similarities and differences between prokaryotes and eukaryotes. *Nucleic Acids Res* **33**: 3165-3175.

Koplin, A., S. Preissler, Y. Ilina, M. Koch, A. Scior, M. Erhardt & E. Deuerling, (2010) A dual function for chaperones SSB-RAC and the NAC nascent polypeptide-associated complex on ribosomes. *J Cell Biol* **189**: 57-68.

Kraemer, S.M., R.T. Ranallo, R.C. Ogg & L.A. Stargell, (2001) TFIIA interacts with TFIID via association with TATA-binding protein and TAF40. *Mol Cell Biol* **21**: 1737-1746.

Kuhn, K.M., J.L. DeRisi, P.O. Brown & P. Sarnow, (2001) Global and specific translational regulation in the genomic response of *Saccharomyces cerevisiae* to a rapid transfer from a fermentable to a nonfermentable carbon source. *Mol Cell Biol* **21**: 916-927.

Kumar, N.V. & P.N. Rangarajan, (2012) The zinc finger proteins Mxr1p and repressor of phosphoenolpyruvate carboxykinase (ROP) have the same DNA binding specificity but regulate methanol metabolism antagonistically in *Pichia pastoris*. *J Biol Chem* **287**: 34465-34473.

Lawless, C., R.D. Pearson, J.N. Selley, J.B. Smirnova, C.M. Grant, M.P. Ashe, G.D. Pavitt & S.J. Hubbard, (2009) Upstream sequence elements direct post-transcriptional regulation of gene expression under stress conditions in yeast. *BMC Genomics* **10**: 7.

Lee, S., T. Carlson, N. Christian, K. Lea, J. Kedzie, J.P. Reilly & J.J. Bonner, (2000) The Yeast Heat Shock Transcription Factor Changes Conformation in Response to Superoxide and Temperature. *Mol Biol Cell* **11**: 1753-1764.

Liang, S., B. Wang, L. Pan, Y. Ye, M. He, S. Han, S. Zheng, X. Wang & Y. Lin, (2012) Comprehensive structural annotation of *Pichia pastoris* transcriptome and the response to various carbon sources using deep paired-end RNA sequencing. *BMC Genomics* **13**: 738.

Lin-Cereghino, G.P., C.M. Stark, D. Kim, J. Chang, N. Shaheen, H. Poerwanto, K. Agari, P. Moua, L.K. Low, N. Tran, A.D. Huang, M. Nattestad, K.T. Oshiro, J.W. Chang, A. Chavan, J.W. Tsai & J. Lin-Cereghino, (2013) The effect of alpha-mating factor secretion signal mutations on recombinant protein expression in *Pichia pastoris*. *Gene* **519**: 311-317.

Liu, J., J. Huang, T. Su, K. Bertoldi & D.R. Clarke, (2014) Structural transition from helices to hemihelices. *PLoS One* **9**: e93183.

Lubliner, S., L. Keren & E. Segal, (2013) Sequence features of yeast and human core promoters that are predictive of maximal promoter activity. *Nucleic Acids Res* **41**: 5569-5581.

Malys, N. & J.E. McCarthy, (2011) Translation initiation: variations in the mechanism can be anticipated. *Cell Mol Life Sci* **68**: 991-1003.

Maquat, L.E., W.Y. Tarn & O. Isken, (2010) The Pioneer Round of Translation: Features and Functions. *Cell* **142**: 368-374.

Marx, H., A. Mecklenbräuker, B. Gasser, M. Sauer & D. Mattanovich, (2009) Directed gene copy number amplification in Pichia pastoris by vector integration into the ribosomal DNA locus. *FEMS Yeast Res* **9**: 1260-1270.

Mattanovich, D., B. Gasser, M. Maurer, R. Prielhofer, J. Klein & J. Wenger, (2014) Regulatable promoter. **US Patent 20140242636**.

Mattanovich, D., B. Gasser & R. Prielhofer, (2013) Constitutive promoter. **US Patent US20140274761**.

Mellitzer, A., C. Ruth, C. Gustafsson, M. Welch, R. Birner-Grunberger, R. Weis, T. Purkarthofer & A. Glieder, (2014) Synergistic modular promoter and gene

optimization to push cellulase secretion by *Pichia pastoris* beyond existing benchmarks. *J Biotechnol* **191**: 187-195.

Mignone, F., C. Gissi, S. Liuni & G. Pesole, (2002) Untranslated regions of mRNAs. *Genome Biol* **3**: reviews0004-reviews0004.0010.

Ni, L., C. Bruce, C. Hart, J. Leigh-Bell, D. Gelperin, L. Umansky, M.B. Gerstein & M. Snyder, (2009) Dynamic and complex transcription factor binding during an inducible response in yeast. *Genes Dev* **23**: 1351-1363.

Niu, H., L. Jost, N. Pirlot, H. Sassi, M. Daukandt, C. Rodriguez & P. Fickers, (2013) A quantitative study of methanol/sorbitol co-feeding process of a *Pichia pastoris* Mut+/pAOX1-lacZ strain. *Microbial Cell Factories* **12**: 33.

Ohi, H., M. Miura, R. Hiramatsu & T. Ohmura, (1994) The positive and negative cis-acting elements for methanol regulation in the *Pichia pastoris* AOX2 gene. *Mol Gen Genet* **243**: 489-499.

Parua, P.K., P.M. Ryan, K. Trang & E.T. Young, (2012) Pichia pastoris 14-3-3 regulates transcriptional activity of the methanol inducible transcription factor Mxr1 by direct interaction. *Mol Microbiol* **85**: 282-298.

Payne, T., C. Hanfrey, A.L. Bishop, A.J. Michael, S.V. Avery & D.B. Archer, (2008) Transcript-specific translational regulation in the unfolded protein response of *Saccharomyces cerevisiae*. *FEBS Lett* **582**: 503-509.

Pelechano, V., W. Wei & L.M. Steinmetz, (2013) Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature* **497**: 127-131.

Pfeffer, M., M. Maurer, J. Stadlmann, J. Grass, M. Delic, F. Altmann & D. Mattanovich, (2012) Intracellular interactome of secreted antibody Fab fragment in *Pichia pastoris* reveals its routes of secretion and degradation. *Appl Microbiol Biotechnol* **93**: 2503-2512.

Prielhofer, R., S.P. Cartwright, A.B. Graf, M. Valli, R.M. Bill, D. Mattanovich & B. Gasser, (2015) *Pichia pastoris* regulates its gene-specific response to different carbon

sources at the transcriptional, rather than the translational, level. *BMC Genomics* **16**: 1393.

Prielhofer, R., M. Maurer, J. Klein, J. Wenger, C. Kiziak, B. Gasser & D. Mattanovich, (2013) Induction without methanol: novel regulated promoters enable high-level expression in *Pichia pastoris*. *Microb Cell Fact* **12**: 5.

Qin, X., J. Qian, G. Yao, Y. Zhuang, S. Zhang & J. Chu, (2011) GAP promoter library for fine-tuning of gene expression in *Pichia pastoris*. *Appl Environ Microbiol* **77**: 3600-3608.

Rando, O.J. & F. Winston, (2012) Chromatin and transcription in yeast. *Genetics* **190**: 351-387.

Riethoven, J.J., (2010) Regulatory regions in DNA: promoters, enhancers, silencers, and insulators. *Methods Mol Biol* **674**: 33-42.

Rojas-Duran, M.F. & W.V. Gilbert, (2012) Alternative transcription start site selection leads to large differences in translation activity in yeast. *RNA* **18**: 2299-2305.

Russell, J. & J.C. Zomerdijk, (2006) The RNA polymerase I transcription machinery. *Biochem Soc Symp*: 203-216.

Ruth, C., M. Buchetics, V. Vidimce, D. Kotz, S. Naschberger, D. Mattanovich, H. Pichler & B. Gasser, (2014) *Pichia pastoris* Aft1 - a novel transcription factor, enhancing recombinant protein secretion. *Microb Cell Fact* **13**: 120.

Sahu, U., K. Krishna Rao & P.N. Rangarajan, (2014) Trm1p, a Zn(II)(2)Cys(6)-type transcription factor, is essential for the transcriptional activation of genes of methanol utilization pathway, in *Pichia pastoris*. *Biochem Biophys Res Commun* **451**: 158-164.

Sauer, M., P. Branduardi, B. Gasser, M. Valli, M. Maurer, D. Porro & D. Mattanovich, (2004) Differential gene expression in recombinant *Pichia pastoris* analysed by heterologous DNA microarray hybridisation. *Microb Cell Fact* **3**: 17.

Schramm, L. & N. Hernandez, (2002) Recruitment of RNA polymerase III to its target promoters. *Genes Dev* **16**: 2593-2620.

Shah, P., Y. Ding, M. Niemczyk, G. Kudla & J.B. Plotkin, (2013) Rate-limiting steps in yeast protein translation. *Cell* **153**: 1589-1601.

Shalgi, R., M. Lapidot, R. Shamir & Y. Pilpel, (2005) A catalog of stability-associated sequence elements in 3' UTRs of yeast mRNAs. *Genome Biol* **6**: R86.

Shandilya, J. & S.G. Roberts, (2012) The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. *Biochim Biophys Acta* **1819**: 391-400.

Shen, S., G. Sulter, T.W. Jeffries & J.M. Cregg, (1998) A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene* **216**: 93-102.

Shenton, D., J.B. Smirnova, J.N. Selley, K. Carroll, S.J. Hubbard, G.D. Pavitt, M.P. Ashe & C.M. Grant, (2006) Global translational responses to oxidative stress impact upon multiple levels of protein synthesis. *J Biol Chem* **281**: 29011-29021.

Simpson, C.E. & M.P. Ashe, (2012) Adaptation to stress in yeast: to translate or not? *Biochem Soc Trans* **40**: 794-799.

Stavely, B., (2014) Homepage for Molecular & Developmental Biology, http://www.mun.ca/biology/desmid/brian/BIOL3530/.

Toledano, M.B., A. Delaunay-Moisan, C.E. Outten & A. Igbaria, (2013) Functions and cellular compartmentation of the thioredoxin and glutathione pathways in yeast. *Antioxid Redox Signal* **18**: 1699-1711.

Tschopp, J.F., P.F. Brust, J.M. Cregg, C.A. Stillman & T.R. Gingeras, (1987) Expression of the lacZ gene from two methanol-regulated promoters in *Pichia pastoris*. *Nucleic Acids Res* **15**: 3859-3876.

Veenhuis, M., J.P. Van Dijken & W. Harder, (1983) The significance of peroxisomes in the metabolism of one-carbon compounds in yeasts. *Adv Microb Physiol* **24**: 1-82.

Vogl, T. & A. Glieder, (2013) Regulation of *Pichia pastoris* promoters and its consequences for protein production. *N Biotechnol* **30**: 385-404.

Vogl, T., F.S. Hartner & A. Glieder, (2013) New opportunities by synthetic biology for biopharmaceutical production in Pichia pastoris. *Curr Opin Biotechnol* **24**: 1094-1101.

Vogl, T., C. Ruth, J. Pitzer, T. Kickenweiz & A. Glieder, (2014a) Synthetic core promoters for Pichia pastoris. *ACS Synth Biol* **3**: 188-191.

Vogl, T., G.G. Thallinger, G. Zellnig, D. Drew, J.M. Cregg, A. Glieder & M. Freigassner, (2014b) Towards improved membrane protein production in *Pichia pastoris*: general and specific transcriptional response to membrane protein overexpression. *N Biotechnol* **31**: 538-552.

Waern, K. & M. Snyder, (2013) Extensive transcript diversity and novel upstream open reading frame regulation in yeast. *G3 (Bethesda)* **3**: 343-352.

Walter, P. & D. Ron, (2011) The unfolded protein response: from stress pathway to homeostatic regulation. *Science* **334**: 1081-1086.

Wang, Y., J.A. Fairley & S.G. Roberts, (2010) Phosphorylation of TFIIB links transcription initiation and termination. *Curr Biol* **20**: 548-553.

Waterham, H.R., M.E. Digan, P.J. Koutz, S.V. Lair & J.M. Cregg, (1997) Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. *Gene* **186**: 37-44.

Weinhandl, K., M. Winkler, A. Glieder & A. Camattari, (2014) Carbon source dependent promoters in yeasts. *Microb Cell Fact* **13**: 5.

Wilson, W.A., S.A. Hawley & D.G. Hardie, (1996) Glucose repression/derepression in budding yeast: SNF1 protein kinase is activated by phosphorylation under derepressing conditions, and this correlates with a high AMP:ATP ratio. *Curr Biol* **6**: 1426-1434.

Yang, C., E. Bolotin, T. Jiang, F.M. Sladek & E. Martinez, (2007) Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene* **389**: 52-65.

Young, C.L. & A.S. Robinson, (2014) Protein folding and secretion: mechanistic insights advancing recombinant protein production in *S. cerevisiae*. *Curr Opin Biotechnol* **30c**: 168-177.

Yu, P., Q. Zhu, K. Chen & X. Lv, (2014) Improving the Secretory Production of the Heterologous Protein in *Pichia pastoris* by Focusing on Protein Folding. *Appl Biochem Biotechnol*.

Publications

Prielhofer, R., M. Maurer, J. Klein, J. Wenger, C. Kiziak, B. Gasser & D. Mattanovich, (**2013**) Induction without methanol: novel regulated promoters enable high-level expression in *Pichia pastoris*. Microb Cell Fact 12: 5.

Gasser, B., **R. Prielhofer**, H. Marx, M. Maurer, J. Nocon, M. Steiger, V. Puxbaum, M. Sauer & D. Mattanovich, (**2013**) *Pichia pastoris*: protein production host and model organism for biomedical research. Future Microbiol 8: 191-208.

Prielhofer, R., S.P. Cartwright, A.B. Graf, M. Valli, R.M. Bill, D. Mattanovich & B. Gasser, (**2015**) *Pichia pastoris* regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. BMC Genomics 16: 167.

Prielhofer, R., D. Mattanovich & B. Gasser, (**2015**) P_{GTH1} promoter variants reveal key regulatory elements and enable improved protein production in *Pichia pastoris*. *Unpublished manuscript.*

Patent applications

1 Mattanovich D, Gasser B, Maurer M, Prielhofer R, Klein J, Wenger J: Regulatable promoter. US Patent 20140242636.

2 Mattanovich D, Gasser B, Prielhofer R, Ltd L: Constitutive promoter. 2014, WO 2014139608 A1 and US 20140274761 A1.

MICROBIAL CELL FACTORIES

RESEARCH Open Access

Induction without methanol: novel regulated promoters enable high-level expression in *Pichia pastoris*

Roland Prielhofer[1], Michael Maurer[2,3], Joachim Klein[4], Jana Wenger[4], Christoph Kiziak[4], Brigitte Gasser[1,2*] and Diethard Mattanovich[1,2]

Abstract

Background: Inducible high-level expression is favoured for recombinant protein production in *Pichia pastoris*. Therefore, novel regulated promoters are desired, ideally repressing heterologous gene expression during initial growth and enabling it in the production phase. In a typical large scale fed-batch culture repression is desired during the batch phase where cells grow on a surplus of e.g. glycerol, while heterologous gene expression should be active in the feed phase under carbon (e.g. glucose) limitation.

Results: DNA microarray analysis of *P. pastoris* wild type cells growing in glycerol-based batch and glucose-based fed batch was used for the identification of genes with both, strong repression on glycerol and high-level expression in the feed phase. Six novel glucose-limit inducible promoters were successfully applied to express the intracellular reporter eGFP. The highest expression levels together with strong repression in pre-culture were achieved with the novel promoters P_{G1} and P_{G6}.
Human serum albumin (HSA) was used to characterize the promoters with an industrially relevant secreted protein. A P_{G1} clone with two gene copies reached about 230% of the biomass specific HSA titer in glucose-based fed batch fermentation compared to a P_{GAP} clone with identical gene copy number, while P_{G6} only achieved 39%. Two clones each carrying eleven gene copies, expressing HSA under control of P_{G1} and P_{G6} respectively were generated by post-transformational vector amplification. They produced about 1.0 and 0.7 g L^{-1} HSA respectively in equal fed batch processes. The suitability in production processes was also verified with HyHEL antibody Fab fragment for P_{G1} and with porcine carboxypeptidase B for P_{G6}. Moreover, the molecular function of the gene under the control of P_{G1} was determined to encode a high-affinity glucose transporter and named *GTH1*.

Conclusions: A set of novel regulated promoters, enabling induction without methanol, was successfully identified by using DNA microarrays and shown to be suitable for high level expression of recombinant proteins in glucose-based protein production processes.

Keywords: *Pichia pastoris*, Heterologous protein production, Glucose-limited fed batch cultivation, Inducible promoter, High-affinity glucose transporter

* Correspondence: brigitte.gasser@boku.ac.at
[1]University of Natural Resources and Life Sciences, Department of Biotechnology, Muthgasse 18, Vienna 1190, Austria
[2]Austrian Centre of Industrial Biotechnology (ACIB GmbH), Muthgasse 11, Vienna 1190, Austria
Full list of author information is available at the end of the article

Background

The methylotrophic yeast *Pichia pastoris* is widely used as a production platform for heterologous proteins. Latest developments in strain engineering for improved protein folding and secretion and glyco-engineering have recently been reviewed by Damasceno et al. [1].

Another important target for strain development is the promoter driving expression of the heterologous gene. A summary of the most important promoters of non-methylotrophic and methylotrophic yeasts is provided by Mattanovich et al. [2]. While production of recombinant proteins in *P. pastoris* has been successfully achieved under control of the constitutive glyceraldehyde-3-phosphate dehydrogenase promoter (P_{GAP}), regulated promoters have several advantages: they enable initial biomass gain without product formation and allow tuning of the production process. Additionally, a potential impact of product accumulation on growth or viability of the cells can be prevented by decoupling growth from the production phase.

However, today's available regulated promoters of *P. pastoris* have drawbacks. Many of them derive from methanol utilization pathway genes, which are generally repressed by glucose and/or ethanol and strongly induced by methanol. P_{AOX1} induces high-level expression of its encoded alcohol oxidase 1, which catalyzes the oxidation of methanol to formaldehyde [3]. Its weaker homolog P_{AOX2} has been used for protein production as well [4]. Another strong promoter of this pathway is P_{FLD1}, its gene formaldehyde dehydrogenase is either induced by methylamine or methanol [5]. The promoter of dihydroxyacetone synthase, P_{DAS}, was reported to be similarly regulated and even stronger than P_{AOX1} [3], however it is not commonly used for protein production. Methanol and methylamine are both highly flammable and hazardous to health, so safety precautions are required for their industrial use. In addition to that, methanol consumption is technically disadvantageous because it causes high heat evolution and an increased oxygen demand during the fed batch phase [6].

The P_{ICL1} promoter controls the expression of isocitrate lyase and is regulated by the carbon source used for cell growth. No detectable promoter activity is present when cells are growing on glucose, while it gets turned on when cells are stationary or growing on ethanol [7]. Hence, this promoter might be an alternative for some applications, but its regulatory properties are poor. *PHO89* is a regulated sodium phosphate symporter and its promoter was investigated and shown to produce reasonable amounts of protein [8]. Cells must be phosphate-limited for the full activation of P_{PHO89}, and an increase in product titer was even shown in phosphate-limited stationary phase. Additionally, an impact on growth was reported and reduced cellular fitness can be assumed in these conditions.

On the other hand, constitutive promoters might be advantageous for the over-expression of genes or to co-express helper factors and marker genes. The widely-used P_{GAP} controls the expression of glyceraldehyde-3-phosphate dehydrogenase at a high basal level [9]. Its productivity can be influenced by controlling the growth rate at the optimal activity of P_{GAP} [10], and by a decrease of available O_2 levels [11]. The promoter of the translational elongation factor EF-1 alpha gene, P_{TEF1}, was analyzed and showed a tighter growth-associated regulation than P_{GAP} [12].

A promoter library of P_{AOX1} was generated, leading to a few variants that were slightly stronger than wild type P_{AOX1}, and a number of variants with altered regulatory properties, some of them being active without methanol [13]. Another library approach was done for P_{GAP} by mutation and clones expressing yeast-enhanced green fluorescent protein (yEGFP) under the control of obtained variants produced 8 to 218% of fluorescence intensity compared to the wild type promoter [14]. Potential promoter libraries can also be deduced from microarray data and rational considerations. Focussing on highly transcribed genes in general, 15 promoters were selected for characterization and the promoter of the thiamine biosynthesis gene P_{THI11}, which is regulated by the availability of thiamine in the growth medium, was discovered [15].

As described above, the number of strong promoters with advantageous properties for protein production is limited in *P. pastoris*. This work was designated to identify novel promoters with both, high expression and an optimal regulation in production process conditions. Equally important, the addition of inducers was to be avoided, because their use is often associated with extra costs and safety precautions in large scale fermentation processes.

A typical production process under the control of P_{GAP} uses glycerol in the batch phase, and a constant glucose fed batch for 100 hours to reach more than 100 g L^{-1} cell dry weight [16]. In order to identify potential inducible promoters in the course of this process, we used DNA microarray analysis to compare gene expression patterns of glycerol-excess (=batch growth phase) and glucose-limited (=fed batch production phase) conditions. The expression capacity of selected promoter targets was characterized with model proteins and verified in fed batch processes.

Results and discussion

Identification of novel promoters with desired induction properties

A typical *P. pastoris* protein production process avoiding methanol induction starts with a glycerol batch (surplus of carbon source) which is followed by a glucose fed

batch (limit of carbon source) [10]. DNA microarrays were used to analyze gene expression patterns and to identify potential promoters for this cultivation strategy. In order to eliminate growth rate related effects, glucose-limited conditions were analyzed in chemostat cultivation where the growth rate, similar to that in the batch phase, was fixed by controlling the dilution rate at 0.1 h^{-1}.

The microarray data was mined for genes with both, high difference in expression level between repressed and induced state (fold change) as well as high signal intensity in the induced state to identify potent promoters for inducible high-level protein production in *P. pastoris*. Six potential promoters (abbreviated as P_{G1}, P_{G3}, P_{G4}, P_{G6}, P_{G7} and P_{G8}, see Figure 1 and Table 1) were considered for further characterization.

Verification of promoter strength and regulation

At first, the strength and regulation of the novel promoters were assayed with the intracellular reporter protein eGFP in small scale screening cultures. Both, repressive conditions in pre-culture (glycerol excess) and induced ones during main culture (glucose limit) were analyzed during the screening. In order to simulate fed batch like conditions in screenings, we had to adapt the screening strategy. Instead of usual feedings with certain amounts of glucose which lead to repeated batch phases, we used slow glucose releasing polymer particles (12mm feed beads, Kuhner, CH), liberating glucose at a non-linear rate of $1.63 \cdot t^{0.74}$ mg per disc (t = time [h]), which equals to 28.6 mg per disc after 48 hours.

As shown in Figure 2, P_{G1} and P_{G6} had superior properties in terms of both, regulation and induction strength. In order to visualize gene dosage effects, genomic DNA of several clones was isolated and analyzed by real-time PCR to determine the gene copy number (GCN) of eGFP. Compared to a P_{GAP} clone with one gene copy, the specific fluorescence of P_{G1} and P_{G6} controlled expression of eGFP (normalized to GCN) were induced from almost zero in batch phase to about 150% and 100% after 48 h screening culture, respectively. The other promoters P_{G3}, P_{G4}, P_{G7}, and P_{G8} still showed a good regulation and induction strength suitable for inducible protein expression, with expression strengths spanning a spectrum of about 20% to 120% relative to P_{GAP} (Figure 2A). The next step was to investigate the induction behaviour of the novel promoters in more detail.

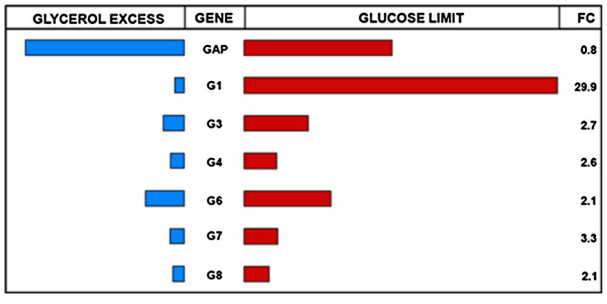

Figure 1 Microarray data (red channel) of identified target genes in comparison to GAP. Bars represent relative expression levels in glycerol excess (batch phase, blue bars on the left side) and in glucose limit (chemostat cultivation, red bars on the right side). Numbers in the right column represent the fold change of signal intensity between glucose limit and glycerol excess conditions.

Table 1 Identified promoter candidates

Promoter	Gene	*P. pastoris* gene identifier (strain GS115)
P_{GAP}	GAP	PAS_chr2-1_0437
P_{G1}	G1	PAS_chr1-3_0011
P_{G3}	G3	PAS_chr4_0550
P_{G4}	G4	PAS_chr4_0043
P_{G6}	G6	PAS_chr2-1_0853
P_{G7}	G7	PAS_chr1-4_0570
P_{G8}	G8	PAS_chr1-3_0165

Analysis of the glucose dependent regulation

The induction behaviour of the novel promoters was characterized in screenings with eGFP producing clones in YP media containing different amounts of glucose (ranging from 20 to 0.002 g L^{-1}). The cells were cultivated for 5–6 hours and eGFP expression was analyzed by flow cytometry.

Promoters P_{G1} and P_{G7} showed a flat induction course leading to full activity only with less than 0.05 g L^{-1} glucose. That is clearly different to P_{G3}, P_{G4} and P_{G6}´s steeper regulation pattern which reach their top activity already at around 4 g L^{-1} glucose (Figure 3). In other words, P_{G1} is not only the strongest but also most tightly repressed by glucose among the promoters tested here.

Based on these regulatory features we intended to characterize the functions of the genes under control of the P_G promoters. At the time of their identification, no or only putative functions were assigned to the underlying genes aside from G1. Therefore, we used NCBI Conserved Domain search to analyze the protein sequences in order to identify putative gene functions.

The gene under the control of P_{G1} was previously functionally clustered with *K. lactis* high-affinity glucose transporter *HGT1* [17]. It contains two major facilitator superfamily domains, same as the G7 gene, which is therefore assumed to be a glucose transporter too. For *Saccharomyces cerevisiae*, it was reported that hexose transporters underlie complex regulation patterns and are expressed in dependence of glucose concentration [18]. The regulation pattern exhibited by the promoters

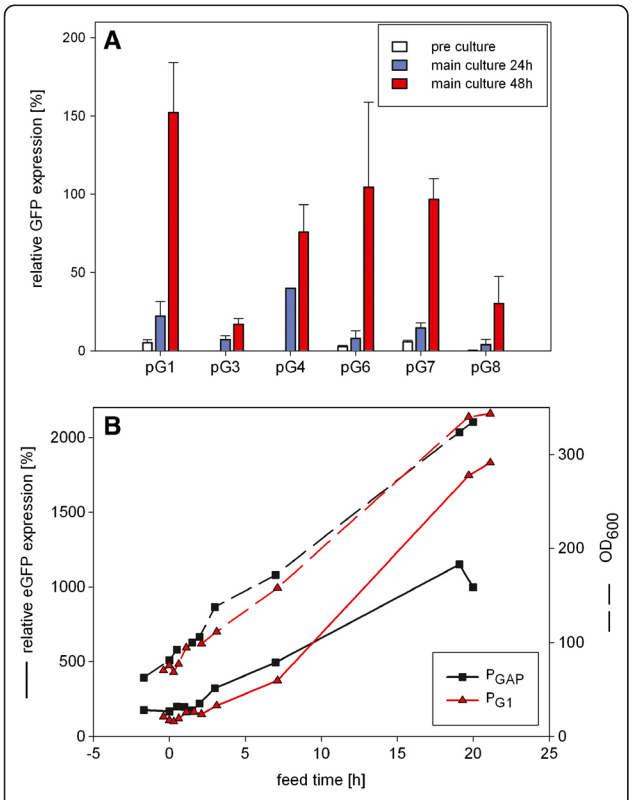

Figure 2 Expression of eGFP under control of the novel promoters P_{G1}, P_{G3}, P_{G4}, P_{G6}, P_{G7} and P_{G8}. (A) Specific eGFP fluorescence in shake flask screenings related to P_{GAP} and to eGFP gene copy number. **(B)** Fed batch cultivations of single gene copy clones expressing eGFP under the control of P_{GAP} and P_{G1}. Relative eGFP expression (solid lines) and OD_{600} (dashed lines) are shown over the feed time.

P_{G3}, P_{G4}, P_{G6} and P_{G8} might be associated with a role in central metabolism. An AKR (aldo keto reductase) domain was found in the gene expressed under P_{G3}. The genes controlled by P_{G4} and P_{G6} are both putative aldehyde dehydrogenases, predicted to be localized in the cytosol and in the mitochondria, respectively. The Gti1/Pac2 family domain found in G8 plays a role in gluconate uptake upon glucose starvation and in sexual development in *Schizosaccharomyces pombe*.

We could show the need of an explicit glucose limit for full activity of the novel promoters, which is most pronounced for P_{G1}. This demonstrates that the microarray data-based promoter selection is excellently suited to select for promoters with features relevant for bioprocesses, and secondly indicates the novel promoter´s advantages in fed batch fermentation.

To prove this statement, the application of the strongest and most promising promoter P_{G1} was tested in a fed batch fermentation where truly glucose-limited conditions are present [10]. A single gene copy clone expressing eGFP was chosen for comparison to an equivalent single gene copy clone of eGFP under the control of P_{GAP}. Thereby we could show that the P_{G1} promoter remains repressed during the batch phase and that its induction during fed batch clearly exceeds the strength of P_{GAP}. Relative eGFP expression (fluorescence related to the culture volume) and OD_{600} over the feed time are shown in Figure 2B.

Expression of secreted human serum albumin

P_{G1} and P_{G6} were further selected to assay the production of a secreted protein under their control. Human serum albumin (HSA) is efficiently produced in *P. pastoris*. Therefore it can be regarded as an industrially relevant secreted reporter protein. Its expression under the control of P_{G1}, P_{G6} and P_{GAP} was screened in shake flasks with glucose-limited conditions (through the use of feed beads) in the main culture. During the glycerol (batch) pre-culture, both P_{G1} and P_{G6} promoters remained well repressed. As seen before, during main culture expression under P_{G1} was stronger than under P_{G6} and, in relation to gene copy number, reached around 77% of the biomass specific HSA yield of cells expressing under P_{GAP}, while P_{G6} produced about 22% of P_{GAP} (Figure 4A). However, the novel promoters might not show their full potential in shake flask screenings, since these conditions are not strictly glucose limited during the entire production phase.

To exploit the full potential of the novel promoters, glucose limited fed batch cultivations were performed. Based

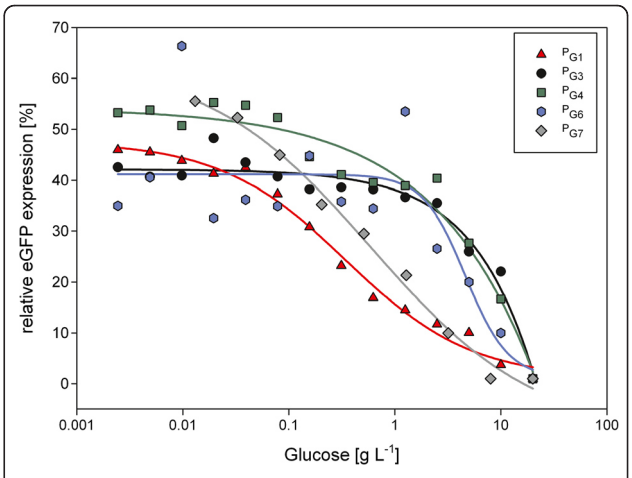

Figure 3 Induction behaviour of the novel promoters. Specific eGFP fluorescence of clones expressing eGFP under the control of P_{G1}, P_{G3}, P_{G4}, P_{G6} and P_{G7} in media containing different amounts of glucose. Data is related to P_{GAP}, normalized to 1.0 at the highest glucose concentration of 20 g L^{-1} and plotted against the logarithmic glucose concentration (trend line calculation: four parameter logistic curve). The glucose concentration given on the x-axis refers to the glucose set point at the beginning of the cultivation (serial dilutions ranging from 20 to 0.002 g L^{-1}). This screening setup and data processing points out relative promoter activities, thereby showing the kinetics of induction, but does not allow comparison of promoter strength.

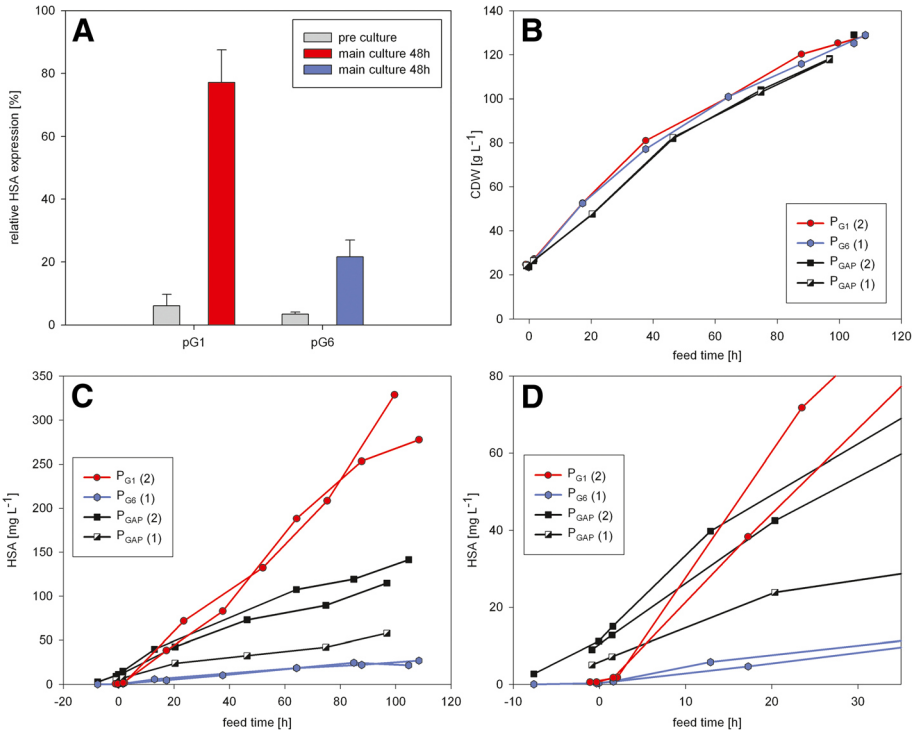

Figure 4 Expression of secreted HSA using the novel promoters P_{G1} and P_{G6} in shake flask and fed batch cultivations. (A) HSA expression in shake flask screenings related to P_{GAP} and to the gene copy number. **(B)** Dry cell weight and **(C)** HSA titer in fed batch cultivations of double and single gene copy clones expressing under the control of P_{G1} (circle, two copies), P_{G6} (diamond, one copy) and P_{GAP} (black square, two copies and black-and-white, one copy). **(D)** Detail of (**C**) showing late batch and early fed batch phase, highlighting the different regulation properties of the promoters. Except for the single gene copy P_{GAP} clone, all fermentations were performed in duplicates.

on the screening results, one clone each expressing sufficient amounts of HSA under the control of P_{G1} and P_{G6} was selected. Those clones, harbouring two and one gene copies respectively, were compared with their respective gene copy equivalent P_{GAP} clones (Figure 4B, 4C and 4D). Dry cell weight (DCW) and HSA titers are summarized in Table 2. HSA titers of P_{GAP} clones correlate with their respective gene copy number. Again, the P_{G1} clone showed superior properties - it clearly outperformed the P_{GAP} clone with the same gene copy number and produced about 230% of the biomass specific product yield compared to P_{GAP}. P_{G6} produced about 39% of the biomass specific HSA yield of its gene copy equivalent P_{GAP} clone. Besides, PGAP-driven expression was already active in the batch phase, and more than 5% of the final HSA amount was already present at the batch end for both clones expressing under its control. While this is not an issue in case of HSA, it is a clear disadvantage compared to inducible promoters such as P_{G1} in case of toxic or difficult to express products. Figure 4C shows HSA titer over the feed time, and the unique repression/induction efficiency of P_{G1} is clearly pointed out in the first hours (Figure 4D). Both, P_{G1} and P_{G6} showed good repression in the batch phase and induction by the glucose limited feed.

To verify that the novel promoters also exhibit their superior regulatory properties and expression capacity

Table 2 Summary of fed batch cultivations of *P. pastoris* expressing HSA under the control of P_{G1}, P_{G6} and P_{GAP}

Promoter	GCN	Batch end			Fed batch end			
		DCW [g L^{-1}]	HSA [mg L^{-1}]	HSA/DCW [mg g^{-1}]	DCW [g L^{-1}]	HSA [mg L^{-1}]	HSA/DCW [mg g^{-1}]	% HSA/DCW of P_{GAP}
P_{G1}*	2	24.3	0.5	0.0	126.9	303.1	2.4	231.2
P_{G6}*	1	23.9	0.3	0.0	127.1	24.3	0.2	38.9
P_{GAP}*	2	23.8	10.1	0.4	123.7	128.2	1.0	
P_{GAP}	1	24.2	5.0	0.2	117.7	57.8	0.5	

* performed in duplicates.

Table 3 Summary of fed batch cultivations of GCN amplified HSA expressing clones under the control of P_{G1} and P_{G6}

Promoter	GCN	Batch end			Fed batch end		
		DCW [g L^{-1}]	HSA [mg L^{-1}]	HSA/DCW [mg g^{-1}]	DCW [g L^{-1}]	HSA [mg L^{-1}]	HSA/DCW [mg g^{-1}]
P_{G1}	11	18.9	0.2	0.0	114.0	1060.8	9.3
P_{G6*}	11	22.5	0.3	0.0	110.8	728.7	6.6

* performed in duplicates.

in industrially relevant conditions, we elevated HSA gene copy number by post-transformational vector amplification as described previously [19]. Thereby, we were able to produce more than 1 g L^{-1} HSA under the control of P_{G1} with a clone harbouring 11 gene copies, which corresponds to the 3.4-fold titer of its two copy clone (Table 3). Again, P_{G1} outperformed a comparable clone with the same gene copy number under the control of the P_{GAP} promoter, which produced 607 mg L^{-1} HSA in a similar fermentation [19]. High level HSA production was also achieved with an amplified clone expressing HSA (11 gene copies as well) under the control of the weaker P_{G6} promoter, which produced more than 720 mg L^{-1} HSA (Table 3). This titer is approximately 30-fold higher than the titer reached with the P_{G6} single copy clone (24 mg L^{-1}), thus indicating that multiple copies of expression cassettes under control of a weaker promoter can also lead to high productivities. One possible explanation of this effect could be that the ratio of a repressing protein to promoter copy number and thus repressor binding sites is decreased in the amplified clones, therefore leading to higher transcription.

Additionally, the porcine enzyme carboxypeptidase B (CpB) that is used for human insulin production, and an antibody Fab fragment were produced under the control of P_{G1} and P_{G6} respectively, and exceeded the production levels compared to P_{GAP}. Thereby we verified again the suitability of these promoters in standard glucose based production processes.

P_{G1} activity depends on specific growth rate

We elucidated the expression activity of P_{G1} at different growth rates using an HSA clone with two gene copies under its control. It was cultivated in chemostat with different dilution rates and the highest specific product formation was found at a growth rate of about 0.07 h^{-1} (Figure 5). This clearly differs to the profile obtained with P_{GAP} in [10], where the highest specific product formation was obtained only at higher growth rates. Growth rate dependency may be utilized to optimize space-time yield or other parameters in the production processes [10].

Knock out of G1

Furthermore, we decided to clarify the function of the gene PAS_chr1-3_0011, which underlies the control of the promoter P_{G1}. It contains 12 transmembrane domains (predicted by TMHMM Server v. 2.0), two Major Facilitator Superfamily (MFS) and other transporter domains. Based on the sequence homology to the *K. lactis* high-affinity glucose transporter *HGT1*, the gene controlled by P_{G1} was expected to have a function in glucose transport [17], [20]. Strong activity of its promoter at very low glucose concentrations further strengthened this assumption. For further verification, the gene was disrupted using the split marker cassette technique (primers given in Additional file 1: Table S1) as described by Heiss et al. [21]. Similar as described by Jørgensen and his colleagues [22], we compared the glucose uptake of the wild type and a G1 knock out clone in glucose-limited chemostat cultivations at different growth rates. The glucose saturation constants were calculated from the residual glucose concentrations (Table 4) and a K_S of 9.7, 23.1 and 69.3 μM was obtained for three different dilution rates (μ=0.14, 0.1 and 0.05 h^{-1}) for the wild type. Changing K_S values are observed for the whole cell in different conditions, which is due to the differential regulation of its several transporters. A reduced

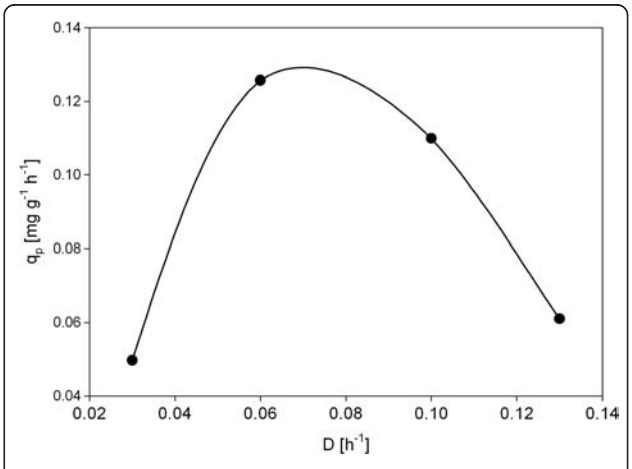

Figure 5 Correlation of specific productivity to specific growth rate using P_{G1}. Specific product formation rate (q_p) observed in chemostat cultivation at different dilution rates of a clone expressing HSA under the control of P_{G1} as well as the respective trend curve (spline curve).

Table 4 Glucose substrate saturation constants of a wildtype and a G1 knock out clone

	G1 k. o., μ_{max} = 0.18 h^{-1}			Wildtype, μ_{max} = 0.18 h^{-1}		
D [h^{-1}]	S [μM]	X [g L^{-1}]	K_S [μM]	S [μM]	X [g L^{-1}]	K_S [μM]
0.14	316.4	29.5	90.4	33.9	30.5	9.7
0.10	123.8	31.3	99.0	28.9	32.1	23.1
0.05	79.9	31.3	207.8	26.6	30.8	69.3

Residual glucose (S), dry cell weight (X) and substrate saturation constant (K_S) of a wild type and a G1 knock out (k. o.) clone at different dilution rates in chemostat cultivations.

capacity of glucose uptake at low specific growth rates has been reported before [22]. The G1 knock out clone appeared to have much higher saturation constants of 90.4, 99.0 and 207.8 μM at the same dilution rates, which was also described for the high-affinity glucose transporter disruption in *Aspergillus niger* [22]. The knock out clone does not display the low K_S values of the wild type, so that the gene PAS_chr1-3_0011 was clearly identified as a high-affinity glucose transporter. As the short name *HGT1* is used as an alias for a peptide transporter in *S. cerevisiae* we suggest the short name *GTH1* (glucose transporter with high affinity) for this *P. pastoris* gene.

Conclusions

Efficient regulated promoters cannot necessarily be found by classical batch screening approaches. Simulating production conditions in lab scale and searching the promoter space offers a new target oriented approach. We could show here that the cultivation of *P. pastoris* in conditions where repression or induction are desired, followed by the analysis of transcript levels with DNA microarrays offers a potent opportunity to find new, strong and regulated promoters.

Six novel promoters were identified and further characterized. All of them are activated by carbon source depletion. The new promoters provide a tool box for expression of recombinant genes and are thus well suitable for protein production processes. P_{G1} had the most favourable repression kinetics and exceeded the expression levels of the well-established constitutive GAP promoter in glucose limited fed batch cultures by more than twofold. The molecular function of the gene under its control was identified as high-affinity glucose transporter and named *GTH1*.

Materials and methods

Strains and cultivation

Escherichia coli DH10B (Invitrogen) was used for subcloning. It was routinely cultivated in petri dishes or shake flasks using LB media supplemented with 25 μg mL^{-1} Zeocin. A wild type *Pichia pastoris* strain CBS2612 which can grow on minimal media supplemented with biotin, was used for protein expression in this work.

The main culture for screenings was either done with YP or BM media and glucose feed beads (12 mm, Kuhner, CH) which provided the carbon source.

YP media contained 20 g L^{-1} peptone and 10 g L^{-1} yeast extract, which can be supplemented with 12.6 g glycerol or 20 g glucose to obtain YPG and YPD, respectively. For cultivation on plates, 5 g L^{-1} agar-agar was added to the liquid medium. BM media was based on YP, supplemented with 13.4 g L^{-1} yeast nitrogen base (Cat.No. 291940, Becton Dickinson, FR) with ammonium sulfate, 0.4 mg L^{-1} biotin and 100 mM potassium phosphate buffer pH 6.0.

Identification of novel inducible promoters

a) Bioreactor cultivations

Fermentations for the identification of promoter candidates were done in 3.5 L working volume bioreactors (Minifors, Infors, CH) in three biological replicates. Cells were grown for about 24 h in batch on glycerol medium, followed by an exponential feed phase on glycerol fed batch medium calculated as described by Resina et al. [23] with a specific growth rate of μ= 0.1 h^{-1} and a substrate yield coefficient of $Y_{X/S}$ of 0.5 g g^{-1}. Sequentially, chemostat cultivation (D = μ = 0.1 h^{-1}) with high density glucose medium was performed.

Glycerol batch medium contained per liter: 2 g citric acid monohydrate, 39.2 g glycerol, 20.8 g $NH_4H_2PO_4$, 0.5 g $MgSO_4 \cdot 7H_2O$, 1.6 g KCl, 0.022 g $CaCl_2 \cdot 2H_2O$, 0.8 mg biotin and 4.6 mL PTM1 trace salts stock solution. HCl was added to set the pH to 5.0.

Glycerol fed-batch medium contained per liter: 632 g glycerol, 8 g $MgSO_4 \cdot 7H_2O$, 22 g KCl, and 0.058 g $CaCl_2 \cdot 2H_2O$.

High-density chemostat medium contained per liter: 2 g citric acid monohydrate, 99.42 g glucose monohydrate, 22 g $NH_4H_2PO_4$, 1.3 g $MgSO_4 \cdot 7H_2O$, 3.4 g KCl, 0.02 g $CaCl_2 \cdot 2H_2O$, 0.4 mg biotin and 3.2 mL PTM1 trace salts stock solution. HCl was added to set the pH to 5.0.

PTM_1 trace salts stock solution contained per liter: 6.0 g $CuSO_4 \cdot 5H_2O$, 0.08 g NaI, 3.36 g $MnSO_4 \cdot H_2O$, 0.2 g $Na_2MoO_4 \cdot 2H_2O$, 0.02 g H_3BO_3, 0.82 g $CoCl_2$, 20.0 g $ZnCl_2$, 65.0 g $FeSO_4 \cdot 7H_2O$, 0.2 g biotin and 5.0 mL H_2SO_4 (95%-98%).

b) Microarray hybridization

RNA purification and sample preparation as well as microarray hybridization (in-house designed *P. pastoris* specific oligonucleotide arrays, AMAD-ID: 034821, 8x15K custom arrays, Agilent) and data analysis were done as described by Graf et al. [24].

Characterization of promoter strength and regulation

a) Cloning

Cloning and transformation was done using the in-house vector pPuzzle [15], which contains a Zeocin resistance cassette for selection in both *E. coli* and yeast, an expression cassette for the gene of interest (GOI) consisting of a multiple cloning site and the *S. cerevisiae* CYC1 transcription terminator, and a locus for integration into the *P. pastoris* genome (3´ AOX1 region or rDNA locus). Promoter sequences (up to 1000 bps upstream of the start codon of their respective genes) were PCR-amplified from *P. pastoris* genomic DNA (primer sequences see Additional file 1: Table S1). The promoters were ligated into pPuzzle in front of the start codons of the model proteins, using the ApaI and the SbfI restriction sites of the multiple cloning site of the vector. Vectors expressing the respective model protein under control of P_{GAP} were used as controls throughout the study. For the expression of heterodimeric HyHEL antibody Fab fragment (HyHEL Fab), the expression cassettes of light chain and Fab heavy chain (each under control of P_{G1}) were combined into one vector (using the strategy described in [27]).

HSA was secreted by its native secretion leader, while for CpB and HyHEL Fab the *S. cerevisiae* alpha mating factor signal sequence was used. To avoid positional effects on reporter gene expression levels, genome integration of the expression plasmids was targeted to either the 3´flanking region of the *AOX1* gene or the ribosomal DNA locus (rDNA, for multicopy integration) of *P. pastoris,* respectively.

Plasmids were linearized within the genome integration region prior to electroporation (2 kV, 4 ms, GenePulser, BioRad) into electrocompetent *P. pastoris*. Multicopy integration of HSA expressing clones was done as described by Marx et al. [19] and selected at higher Zeocin concentrations (up to 1000 μg mL^{-1}).

P. pastoris cells were first selected and cultivated in petri dishes on YPD agar and then inoculated in an YPG medium as pre-culture for screenings and fermentations. Antibiotic selection by Zeocin was applied on plates and in pre-culture at a concentration of 25 μg mL^{-1} or higher.

b) Expression screening

Expression of intracellular eGFP and the secreted proteins HSA, CpB and HyHEL Fab with the novel promoters in comparison to P_{GAP} was evaluated in shake flask screenings. All screenings were performed at 25°C and with shaking at 180 rpm. Single colonies were inoculated in YPG medium with selection pressure (Zeocin) for pre-culture. After approximately 24 hours, the pre-culture was used to inoculate the main culture with an optical density (OD_{600}) of 0.1 (for eGFP) or 1 (for HSA, CpB and HyHEL Fab) in 10 mL YP or BM medium, respectively. Glucose feed beads (12 mm, Kuhner, CH) were used to generate glucose-limiting growth conditions. Expression of eGFP was measured at the end of pre-culture and at 24 and 48 hours of the main culture. Culture supernatant of clones expressing secreted protein was harvested from the pre-culture and after 48 hours and cell density was determined by measuring wet cell weight or OD_{600}.

c) Comparative analysis of *P. pastoris* promoter activity

In order to analyze relative transcription strength of the P_G promoters at different glucose concentrations, a comparative promoter activity study using various glucose concentrations (ranging from 20 to 0.002 g L^{-1} glucose) was performed with eGFP expressing clones in 24-well plates (Cat. No. 7701–5110, Whatman, UK) covered with breath seal membranes (Cat. No. B-100, Excel Scientific, CA). Glucose concentrations of 20, 10, 5, 2.5, 1.25, 0.63, 0.31, 0.16, 0.08, 0.04, 0.02, 0.01, 0.005, 0.002 g L^{-1} were obtained by serial dilution in YP media, and represent the inital setpoints. The main culture was inoculated from YPG-Zeocin pre-culture with an OD_{600} of 0.01 and samples were taken after 5–6 hours and analyzed by flow cytometry.

d) Fed batch cultivation

All fed batch fermentations were done in 1.0 L working volume bioreactors (SR0700ODLS, DASGIP, DE). The dissolved oxygen was controlled at DO = 20% with the stirrer speed (400 – 1200 rpm). Aeration rate was 18 L h^{-1} air, the temperature was controlled at 25°C and the pH was controlled at 5.85 for HSA [25] or pH 5.0 for the other proteins [10] with addition of ammonium hydroxide (25%). To start the fermentation, 300 mL batch medium was sterile filtered into the fermenter and a *P. pastoris* clone was inoculated from an overnight pre-culture with a starting optical density (OD_{600}) of 1. For the cultivation of clones expressing eGFP, the batch phase of approximately 25 h was followed by a fed batch phase with a feeding rate optimized according to [10]. HSA expressing strains were cultivated as described by Marx et al. [19], where the batch phase was followed by a constant feed of 2 g h^{-1} fed batch medium for 100 h, Carboxypeptidase B and HyHEL Fab expressing clones were cultivated similarly. Samples were taken during batch and fed batch phase, and analyzed for expression.

Glycerol batch and glucose fed batch media for eGFP, HyHEL Fab and Carboxypeptidase B expressing clones were exactly as described in [10], while for the production of HSA the media was described in [19].

e) Chemostat cultivation
A strain expressing HSA (2 GCN) under control of P_{G1} was tested for its growth rate dependent expression behaviour in chemostat at different dilution rates (D = 0.03, 0.06, 0.10, 0.13) [10].
For characterization of glucose uptake characteristics, the *P. pastoris* wild type strain and the strain deleted for PAS_chr1-3_0011 were cultivated in glucose limited chemostats at D = 0.05, 0.10, and 0.14 h^{-1}. Samples were taken rapidly as described below.

Analytical methods

a) Copy number determination with real-time PCR
Genomic DNA was isolated using the DNeasy Blood&Tissue Kit (Cat. No. 69504, Quiagen, DE). Gene copy numbers were determined with quantitative PCR using the SensiMix SYBR Kit (QT605-05, Bioline reagents, UK). The primers (supplementary Additional file 1: Table S1) and sample were mixed with the SensiMix and applied for real time analysis in a real-time PCR cycler (Rotor Gene, Qiagen, DE). All samples were analyzed in tri- or quadruplicates. Data analysis was performed with the two standard curve method of the Rotor Gene software. The actin gene *ACT1* was used as calibrator.

b) Determination of protein expression levels
A plate reader (Infinite 200, Tecan, CH) was used to determine eGFP fluorescence in fermentation samples. Therefore, samples were diluted to an OD_{600} of 5 and fluorescence intensity was then related to the culture volume.
Expression of eGFP in screenings was analyzed by flow cytometry as described before [15]. Specific eGFP fluorescence referred to in this study is the fluorescence intensity related to the cell volume for each data point as described by Hohenblum et al. [26]. Then the geometric mean of the population´s specific fluorescence was normalized by subtracting background signal (of non-producing *P. pastoris* wild type cells) and related to expression under the control of P_{GAP}.
For quantification of HSA in shake flask and fermentation supernatants, the Human Albumin ELISA Quantitation Set (Cat. No. E80-129, Bethyl Laboratories, TX) was used. The HSA standard was applied with a starting concentration of 400 ng mL^{-1}. Dilution-, Blocking- and Washing buffer were based on TBS (50 mM Tris–HCl, 140 mM NaCl, pH 8.0) and completed with BSA (1% (w/v)) and/or Tween20 (0.05% (v/v)) accordingly.
HyHEL Fab was determined with ELISA as described previously [27].
CpB was quantified using an enzymatic assay based on the cleavage of hippuryl-L-arginine (Cat. No. H2508, Sigma, MO). Generation of hippuric acid was monitored at its absorbance maximum of 254 nm. Prior to the measurement, the samples were desalted with Zeba Spin columns (Thermo Fisher Scientific, IL) and activated with trypsin (Cat. No. T8345, Sigma, MO).

c) Determination of residual glucose
The D-Glucose Assay - GOPOD-Format (Megazymes, IE) was used to determine residual glucose of chemostat samples. Supernatant sampling was done by pumping culture broth out of the bioreactor by producing an overpressure, and its direct sterile filtration using a vacuum filter unit (Cat. No. 5141178, Whatman, UK). Glucose-limited cultivations usually go along with very low residual glucose concentrations in the supernatant, so the manufacturer's protocol was adapted for glucose concentrations from 10 to 100 mg L^{-1}. Briefly, the ratio of reaction buffer to sample was changed from 30:1 to 3:1.

Additional file

Additional file 1: Table S1. Primer Sequences. Sequences of oligonucleotides used for amplification of promoters, determination of gene copy numbers of the model protein expression cassettes, and generation of G1 disruption cassette (including verification of positive knock-outs).

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
RP performed the experimental work, data analysis, contributed to study design and drafted the manuscript. MM contributed to study design and planning of bioreactor cultivations. JK and JW conceived and provided industrially relevant screening conditions. CK supported data interpretation and planning of promoter characterization. BG planned and supervised the experimental work, and contributed to data analysis and drafting the manuscript. DM coordinated the project and contributed to drafting the manuscript. BG, MM and DM conceived of the study. All authors read and approved the final manuscript.

Acknowledgements
The authors thank Alexandra B. Graf for the processing of the microarray data. Special thanks go to Lisa-Sophie Handler, Christine Ruschitzka and Christine Sanystra for their help with promoter characterization during their bachelor theses or internship. This project was financially supported by LONZA (Visp, CH).

Author details
[1]University of Natural Resources and Life Sciences, Department of Biotechnology, Muthgasse 18, Vienna 1190, Austria. [2]Austrian Centre of Industrial Biotechnology (ACIB GmbH), Muthgasse 11, Vienna 1190, Austria. [3]University of Applied Sciences FH-Campus Vienna, School of Bioengineering, Muthgasse 18, Vienna 1190, Austria. [4]Lonza AG, Rottenstraße 6, Visp 3930, Switzerland.

Received: 10 December 2012 Accepted: 23 January 2013
Published: 24 January 2013

References

1. Damasceno LM, Huang CJ, Batt CA: **Protein secretion in Pichia pastoris and advances in protein production.** *Appl Microbiol Biotechnol* 2012, **93:**31–39.
2. Mattanovich D, Branduardi P, Dato L, Gasser B, Sauer M, Porro D: **Recombinant protein production in yeasts.** *Methods Mol Biol* 2012, **824:**329–358.
3. Tschopp JF, Brust PF, Cregg JM, Stillman CA, Gingeras TR: **Expression of the lacZ gene from two methanol-regulated promoters in Pichia pastoris.** *Nucleic Acids Res* 1987, **15:**3859–3876.
4. Ohi H, Miura M, Hiramatsu R, Ohmura T: **The positive and negative cis-acting elements for methanol regulation in the Pichia pastoris AOX2 gene.** *Mol Gen Genet* 1994, **243:**489–499.
5. Shen S, Sulter G, Jeffries TW, Cregg JM: **A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast Pichia pastoris.** *Gene* 1998, **216:**93–102.
6. Jungo C, Marison I, von Stockar U: **Mixed feeds of glycerol and methanol can improve the performance of Pichia pastoris cultures: A quantitative study based on concentration gradients in transient continuous cultures.** *J Biotechnol* 2007, **128:**824–837.
7. Menendez J, Valdes I, Cabrera N: **The ICL1 gene of Pichia pastoris, transcriptional regulation and use of its promoter.** *Yeast* 2003, **20:**1097–1108.
8. Ahn J, Hong J, Park M, Lee H, Lee E, Kim C, Lee J, Choi E, Jung J: **Phosphate-responsive promoter of a Pichia pastoris sodium phosphate symporter.** *Appl Environ Microbiol* 2009, **75:**3528–3534.
9. Waterham H, Digan M, Koutz P, Lair S, Cregg J: **Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter.** *Gene* 1997, **186:**37–44.
10. Maurer M, Kühleitner M, Gasser B, Mattanovich D: **Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with Pichia pastoris.** *Microb Cell Fact* 2006, **5:**37.
11. Baumann K, Maurer M, Dragosits M, Cos O, Ferrer P, Mattanovich D: **Hypoxic fed-batch cultivation of Pichia pastoris increases specific and volumetric productivity of recombinant proteins.** *Biotechnol Bioeng* 2008, **100:**177–183.
12. Ahn J, Hong J, Lee H, Park M, Lee E, Kim C, Choi E, Jung J: **Translation elongation factor 1-alpha gene from Pichia pastoris: molecular cloning, sequence, and use of its promoter.** *Appl Microbiol Biotechnol* 2007, **74:**601–608.
13. 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:**e76.
14. Qin X, Qian J, Yao G, Zhuang Y, Zhang S, Chu J: **GAP promoter library for fine-tuning of gene expression in Pichia pastoris.** *Appl Environ Microbiol* 2011, **77:**3600–3608.
15. Stadlmayr G, Mecklenbrauker A, Rothmuller M, Maurer M, Sauer M, Mattanovich D, Gasser B: **Identification and characterisation of novel Pichia pastoris promoters for heterologous protein production.** *J Biotechnol* 2010, **150:**519–529.
16. Cos O, Ramón R, Montesinos J, Valero F: **Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast Pichia pastoris under different promoters: a review.** *Microb Cell Fact* 2006, **5:**17.
17. Mattanovich D, Graf A, Stadlmann J, Dragosits M, Redl A, Maurer M, Kleinheinz M, Sauer M, Altmann F, Gasser B: **Genome, secretome and glucose transport highlight unique features of the protein production host Pichia pastoris.** *Microb Cell Fact* 2009, **8:**29.
18. du Preez JC, de Kock SH, Kilian SG, Litthauer D: **The relationship between transport kinetics and glucose uptake by Saccharomyces cerevisiae in aerobic chemostat cultures.** *Antonie Van Leeuwenhoek* 2000, **77:**379–388.
19. Marx H, Mecklenbräuker A, Gasser B, Sauer M, Mattanovich D: **Directed gene copy number amplification in Pichia pastoris by vector integration into the ribosomal DNA locus.** *FEMS Yeast Res* 2009, **9:**1260–1270.
20. Baruffini E, Goffrini P, Donnini C, Lodi T: **Galactose transport in Kluyveromyces lactis: major role of the glucose permease Hgt1.** *FEMS Yeast Res* 2006, **6:**1235–1242.
21. Heiss S, Maurer M, Hahn R, Mattanovich D, Gasser B: **Identification and deletion of the major secreted protein of Pichia pastoris.** *Appl Microbiol Biotechnol* 2012, doi:10.1007/s00253-00012-04260-00254.
22. Jørgensen TR, VanKuyk PA, Poulsen BR, Ruijter GJ, Visser J, Iversen JJ: **Glucose uptake and growth of glucose-limited chemostat cultures of Aspergillus niger and a disruptant lacking MstA, a high-affinity glucose transporter.** *Microbiology* 2007, **153:**1963–1973.
23. Resina D, Cos O, Ferrer P, Valero F: **Developing high cell density fed-batch cultivation strategies for heterologous protein production in Pichia pastoris using the nitrogen source-regulated FLD1 Promoter.** *Biotechnol Bioeng* 2005, **91:**760–767.
24. Graf A, Gasser B, Dragosits M, Sauer M, Leparc G, Tüchler T, Kreil D, Mattanovich D: **Novel insights into the unfolded protein response using Pichia pastoris specific DNA microarrays.** *BMC Genomics* 2008, **9:**390.
25. Kobayashi K, Kuwae S, Ohya T, Ohda T, Ohyama M, Tomomitsu K: **High level secretion of recombinant human serum albumin by fed-batch fermentation of the methylotrophic yeast, Pichia pastoris, based on optimal methanol feeding strategy.** *J Biosci Bioeng* 2000, **90:**280–288.
26. Hohenblum H, Borth N, Mattanovich D: **Assessing viability and cell-associated product of recombinant protein producing Pichia pastoris with flow cytometry.** *J Biotechnol* 2003, **102:**281–290.
27. Gasser B, Maurer M, Gach J, Kunert R, Mattanovich D: **Engineering of Pichia pastoris for improved production of antibody fragments.** *Biotechnol Bioeng* 2006, **94:**353–361.

doi:10.1186/1475-2859-12-5
Cite this article as: Prielhofer *et al.*: Induction without methanol: novel regulated promoters enable high-level expression in *Pichia pastoris*. *Microbial Cell Factories* 2013 **12**:5.

Review
Future Microbiology

Pichia pastoris: protein production host and model organism for biomedical research

Brigitte Gasser[1,2], Roland Prielhofer[1], Hans Marx[1], Michael Maurer[2,3], Justyna Nocon[1], Matthias Steiger[1,2], Verena Puxbaum[1,2], Michael Sauer[1,2] & Diethard Mattanovich*[1,2]

[1]University of Natural Resources & Life Sciences (BOKU), Department of Biotechnology, 1190 Vienna, Austria
[2]Austrian Centre of Industrial Biotechnology (ACIB GmbH), 1190 Vienna, Austria
[3]University of Applied Sciences FH-Campus Vienna, School of Bioengineering, 1190 Vienna, Austria
*Author for correspondence: Tel.: +43 1 47654 6569 ■ Fax: +43 1 3697615 ■ diethard.mattanovich@boku.ac.at

Pichia pastoris is the most frequently used yeast system for heterologous protein production today. The last few years have seen several products based on this platform reach approval as biopharmaceutical drugs. Successful glycoengineering to humanize *N*-glycans is further fuelling this development. However, detailed understanding of the yeast's physiology, genetics and regulation has only developed rapidly in the last few years since published genome sequences have become available. An expanding toolbox of genetic elements and strains for the improvement of protein production is being generated, including promoters, gene copy-number enhancement, gene knockout and high-throughput methods. Protein folding and secretion have been identified as significant bottlenecks in yeast expression systems, pinpointing a major target for strain optimization. At the same time, it has become obvious that *P. pastoris*, as an evolutionarily more 'ancient' yeast, may in some cases be a better model for human cell biology and disease than *Saccharomyces cerevisiae*.

The methylotrophic yeast *Pichia pastoris* is widely used as a production platform for heterologous proteins and as a model organism for peroxisome and secretory organelle proliferation [1–4]. Although *P. pastoris* has been established as a production host over more than two decades, the lack of genetic information has handicapped basic research significantly, so that most biotechnological improvements focused on culture media and fermentation strategies. The genome sequence of *P. pastoris* strains has been published only recently [5–7], boosting genetic and physiological research into this yeast. This review summarizes recent developments of *P. pastoris* research, highlighting novel methods enabling both rapid development of protein production strains and fundamental understanding of physiological processes. It will be emphasized where and why *P. pastoris* may serve as a better model for medical research than the traditional yeast model, *Saccharomyces cerevisiae*.

P. pastoris as a host for protein production

Background

The development and application of *P. pastoris* as production host for recombinant proteins has been reviewed elsewhere, and these fundamentals will not be repeated here. A growing number of protein products are reaching the market (Box 1). However, it has become obvious that a number of methodological as well as cellular limitations reduce the efficiency of the production strain and process development. Efficient DNA transfer is a prerequisite for many high-throughput methods that will be summarized here. The early protein production steps require efficient promoters and translation signals. It has become obvious that folding and secretion of complex proteins (e.g., the typical biopharmaceutical drug candidates of human origin) constitute major bottlenecks of productivity, which are discussed here together with potential solutions. Efficient production of proteins requires balanced culture media. A set of well-established media is described below.

Transformation of *P. pastoris*

For the overexpression of a heterologous protein in *P. pastoris*, the most commonly used method is the stable integration of an expression vector into the *P. pastoris* genome via homologous recombination (HR). Therefore, several different transformation methods [8] and protocols can be applied. The most commonly used transformation method is electroporation, yielding up to 10^5 transformants/µg of DNA [8]. A protocol is provided in Box 2.

Keywords
■ gene copy number ■ peroxisome ■ promoter ■ protein folding ■ protein secretion ■ proteomics ■ transcriptomics ■ transformation

Future Medicine part of fsg

Box 1. Biopharmaceuticals and research materials produced with *Pichia pastoris*.

- Several products from *Pichia pastoris*, such as human serum albumin, insulin, IFN-α and hepatitis B vaccine, are marketed in India and/or Japan [121]. In 2009, the US FDA approved the recombinant kallikrein inhibitor ecallantide (Kalbitor®, Dyax, MA, USA) for the treatment of hereditary angioedema and in the prevention of blood loss in cardiothoracic surgery. Ecallantide is a synthetic peptide produced with *P. pastoris*. Several more therapeutic product candidates that are produced with *P. pastoris* are in the clinical pipeline. ALD518, a humanized monoclonal antibody directed against human IL-6 (Alder Biopharmaceuticals, WA, USA), is already clinically validated for the treatment of rheumatoid arthritis, and currently in Phase II trials for the treatment of cancer. A fusion protein consisting of an anticarcinoembryonic antigen single-chain fragment and the enzyme carboxypeptidase G2 is in clinical investigation for antibody-directed enzyme prodrug therapy against breast cancer (MFECP1 [122], Cancer Research Technology, UK). Furthermore, a malaria vaccine candidate [123] is in clinical investigation, and a number of other drug candidates are in development [124].
- Apart from biopharmaceutical products, a number of reagents for biomedical research are produced with *P. pastoris*, for example, recombinant HEGF, GM-CSF, interleukins, endostatin and human albumin (e.g., marketed by Sigma [MO, USA], Merck [NJ, USA] or Aviva [CA, USA]).

Gene copy-number amplification

For high-level production of recombinant pharmaceutical proteins, it can be favorable to use a production clone with a higher gene dosage of the corresponding gene. High gene dosage can overcome certain limits of the cellular protein production machinery, such as the suboptimal 5′-untranslated region, mRNA secondary structure and protein stability. There are several possible ways to gain such clones harboring a higher gene dosage. One possibility is to construct a multimer of expression cassettes *in vitro* prior to transformation, but the construction of such a multimer can be quite laborious and, furthermore, it is limited to a number of approximately eight expression cassettes. A more convenient method to gain so-called 'jackpot' clones is making use of a drug-selectable marker (e.g., antibiotic-resistance genes) on the expression plasmid. Transformants, which are able to grow on enhanced concentrations of the selective drug (such as Zeocin™ [InvivoGen, CA, USA], G418 [9] and hygromycin B [10]), are likely to harbor a higher copy number of expression plasmids. Nevertheless, subjecting transformants to an enhanced antibiotic selection pressure directly after transformation often results in very low transformation efficiencies. To circumvent this effect, Sunga *et al.* established a post-transformational vector-amplification

Box 2. Protocol for efficient transformation of *Pichia pastoris*.

- This is a protocol derived from a previously described procedure [8]. High transformation frequencies of 10^5 per microgram DNA can be reached reproducibly, so that even library sizes of 10^7 can be achieved with reasonable effort.

***Preparation of electrocompetent* Pichia pastoris**

- Inoculate 200 ml of fresh YPD medium with an overnight YPD culture and incubate at 28°C with vigorous shaking up to an OD_{600} of two. Transfer the culture into sterile centrifuge tubes and centrifuge at 1.500 × *g* and 4°C for 5 min. Discard the supernatant and resuspend in 100 ml of prewarmed YPD medium containing 2 ml of 1 M HEPES pH 8.0 and 2.5 ml of 1 M dithiothreitol. Incubate for 15 min at 28°C at 170 rpm. Add 400 ml of ice-cold water and harvest the cell by centrifugation as described above. Wash the pellet in 250 ml of sterile and ice-cold 1 mM HEPES, with centrifugation as above. Resuspend the pellet in 50 ml of sterile ice-cold 1 M sorbitol and centrifuge as above. Finally, resuspend cells in 0.5 ml of sterile ice-cold 1 M sorbitol. Aliquot the cells into 80 µl vessels, and keep them on ice until transformation. For longer-term storage, freezing at -70°C is appropriate.

***Electroporation of* P. pastoris**

- An 80-µl aliquot of the electrocompetent *P. pastoris* cells is mixed gently with the DNA (0.5–10 µg in a maximum of 30 µl of water), and the mixture is transferred into a chilled electroporation cuvette (2 mm) and incubate on ice for 5 min. Electroporation is performed at the following parameters: 2000 V, 25 µF and 200 Ω. Immediately after transformation, 1 ml of ice-cold YPD (or 1 M sorbitol for auxotrophy selection) is added, and the mixture is transferred to a sterile microcentrifuge tube. The yeast cells are allowed to regenerate for at least 1.5–3 h at 28°C before they are plated in aliquots on selective agar.

OD_{600}: Optical density of a sample measured at a wavelength of 600 nm.

method for *P. pastoris*, where transformants are preselected on a lower/normal concentration of Zeocin (100 µg/ml) and transferred to higher concentrations of 500 and 2000 µg/ml Zeocin consecutively [9]. With this method, they were able to select for 'jackpot' clones harboring more than ten expression cassettes with a frequency of 5–6%. By Southern blot analysis, it was shown that these copies are inserted in the same locus in a head-to-tail configuration [9].

On the basis of post-transformational vector amplification work in other yeasts, Marx *et al.* established an efficient and directed gene copy-number amplification protocol for *P. pastoris* by integrating the expression vector into the rDNA locus of *P. pastoris* and repeated restreaking of the clones on increasing antibiotic concentrations [11].

Gene promoters in *P. pastoris*

Gene promoters, which act as key regulators of gene expression, are essential elements of expression cassettes. Versatile promoters covering a wide range of strengths with various regulation patterns are used for the expression of genes for strain tuning and desired products. Constitutive promoters are active on a basal expression level with low potential of variation, whereas inducible promoters vary the expression level in a regulated way, depending on certain conditions.

For production purposes, inducible expression is favored: it enables initial biomass accumulation without product formation and a better control of the whole production process. Accumulating product can have an impact on the growth or vitality of the cells; regulated promoters can also help to overcome these problems.

Most of the currently available regulated promoters of *P. pastoris* derive from its methanol utilization pathway genes, which are generally repressed by glucose and ethanol and strongly induced by methanol [12]. The most frequently used promoter for production purposes is P_{AOX1}. Its strong methanol-inducible gene encodes Aox1, which catalyzes the oxidation of methanol to formaldehyde, and is naturally expressed at tremendous levels, leading to it constituting up to 30% of the total cell protein.

Its weaker homolog P_{AOX2} has also been used for protein production [13]. Other strong promoters of this pathway are the methylamine – or methanol – inducible P_{FLD1} [14] (formaldehyde dehydrogenase) and P_{DAS1} (dihydroxyacetone synthase), which was reported to be similarly regulated and even more active than P_{AOX1} [15]. The industrial use of methanol and methylamine is compromised to some extent by their high flammability and health hazards, requiring appropriate safety precautions, and by the high oxygen demand and heat production of the methanol-utilization pathway, leading to technical limitations in the fermentation processes. Other regulated promoters of *P. pastoris* with lower expression levels than those mentioned previously include P_{ICL1} (isocitrate lyase) [16] and phosphate-regulated P_{PHO89} [17].

High-throughput approaches are an obvious choice for the analysis of promoters. A library of P_{AOX1} mutants was generated that displays between 6 and 160% of the wild-type promoter activity and features different regulatory properties, partially activated upon derepression by glucose [18]. Based on microarray analysis, a set of novel promoters, including the promoter of the thiamine biosynthesis gene P_{THI11}, which is activated upon depletion of thiamine in the growth medium, were identified [19].

Constitutive promoters have also been employed for efficient protein production, and they may be advantageous for the steady overexpression of genes encoding cellular factors supporting heterologous protein production. The promoter of the *P. pastoris* glyceraldehyde-3-phosphate dehydrogenase gene, P_{GAP}, is expressed at a high basal level and commonly used [20]. Its performance can be increased in hypoxic conditions [21] and by controlling the growth rate at the optimum for productivity [22]. Promoter engineering of P_{GAP} yielded variants with 8–218% expression strength compared with wild-type P_{GAP} [23]. The promoter of the translational elongation factor EF1-α gene, P_{TEF1}, has a tighter growth-associated regulation than P_{GAP} [24], while a large set of additional constitutive promoters with expression levels between 5 and 50% of P_{GAP} are available for strain engineering purposes [25,26].

The power to use highly potent promoter sequences to influence a complex process such as transcription is a great tool for biotechnological and scientific purposes. Various promoters should be tested for new products or applications and a deliberate choice is crucial for success. In contrast to that, almost every parameter can have an influence on gene expression: strain, media, cultivation conditions, fermentation strategy and the promoter–product combination [27].

Promoter divergence in yeast

For some organisms, such as *S. cerevisiae*, a promoter database is even available [28], but

P. pastoris has not been studied to this level of detail. Prediction of regulatory sequences based on analogy to *S. cerevisiae* is limited, as promoter regions, and therefore transcription factor binding sites, tend to evolve much faster than coding genes.

Limitations of protein production–protein folding & secretion

It has been shown in several cases that a higher gene copy number leads to an increase in product yield [1,29]. However, this correlation cannot be assumed *a priori*, and strongly depends on the produced protein. Increasing the gene dosage of recombinant secretory proteins over a certain threshold was shown to overburden host cell physiology, leading to an overload of the secretory pathway [11,30,31].

Translation of a secretory protein is followed by translocation into the endoplasmic reticulum (ER), which is mediated by a secretion signal peptide. For *P. pastoris*, in most cases, the use of the *S. cerevisiae* MFα pre-pro-leader proved to be the most effective [29,32]. However, as product quality, especially N-terminal uniformity, is not always satisfactory when using MFα, alternative secretion leaders have been investigated and are summarized in Table 1.

During secretion, nascent proteins intended for secretion are translocated into the ER lumen, where folding and formation of disulfide bonds take place and post-translational modifications such as glycosylation are initiated [33–36]. If protein folding and secretion exceed the capacity of the ER, or if perturbations alter ER homeostasis, unfolded proteins accumulate and form aggregates in the ER. This triggers the activation of the unfolded protein response (UPR) pathway, which aims to reduce ER stress conditions by the induction of genes involved in protein folding and the ER-associated degradation (ERAD) pathway [37–40].

During recombinant protein production, a high rate of polypeptide accumulation is not always accompanied by a corresponding folding rate. A substantial fraction of the heterologous protein intended for secretion is retained within the cells due to limitations in folding and/or secretion (reviewed in [29,41]). As most therapeutic proteins are secretory proteins, this is regarded as the major bottleneck limiting product yield. In order to alleviate the potential secretory limitations, strains were engineered for co-overexpression of chaperones such as Kar2/BiP [42], folding catalysts such as protein disulfide isomerase (*PDI1*) from *S. cerevisiae* [43,44], humans [45], *Plasmodium* species [46] and *P. pastoris* [30]; or the UPR transcription factor *HAC1* [43,47,48]. Formation of disulfide bonds in the ER and folding itself are described as the main rate-limiting steps during secretion, but vesicular transport or excretion at the plasma membrane/cell wall may also pose a bottleneck [48–51]. Other factors influencing secretion levels are the thermodynamic stability of the recombinant protein [52], or enhanced solubility upon fusion to maltose binding protein [53]. By modeling and measuring the fluxes of recombinant protein towards intracellular degradation and secretion, Pfeffer *et al.* showed that protein synthesis rate is the first rate-limiting step, until folding limitation leads to a plateau of achievable secretion rates [54]. This was verified by applying this model to single-cell analysis of GFP-secreting cells [55]. Additionally, rational literature-based approaches as well as high-throughput methods, such as library sorting and DNA microarrays, were successfully applied for strain engineering for increased secretion [26,49,56].

The tremendous achievements regarding humanization of the *P. pastoris N*-glycosylation patterns obtained by glycoengineering, as well as their impact on production of human therapeutics, have been recently summarized in excellent reviews [57–61] and thus will not be repeated here.

Cultivation media for shake-flask & fed-batch cultivation

P. pastoris has a rather broad substrate spectrum, utilizing D-glucose, L-rhamnose, trehalose, glycerol, D-glucitol, D-mannitol, DL-lactate, succinate, methanol, ethanol, propane 1,2-diol, ethylamine and cadaverine [62]. For the production of therapeutic proteins, it is most important that *P. pastoris* can grow on chemically defined media, so that the addition of any complex, less defined components, especially those of animal origin, can be avoided. It is possible to design cultivation media using the published elementary composition of *P. pastoris*, as shown in Table 2. Biotin is the only essential vitamin that has to be added in the range of 40 µg of biotin per gram of yeast cell dry weight, if an engineered biotin-prototrophic strain is not used [25]. Typical media and fermentation protocols are provided in Box 3.

Novel methods in *P. pastoris* research

Gene deletions in *P. pastoris*

Gene deletion is a powerful molecular–genetic technique that is widely used to investigate and

Table 1. Secretion leaders tested for secretion of recombinant proteins in *Pichia pastoris*.

Secretion leader	Advantages	Concerns	Ref.
MFα with EAEA	High-level secretion	Nonhomogeneous N-terminus possible (EAEA repeats attached to secreted protein)	[32,125]
MFα without EAEA	Secretion levels as high as with EAEA	Nonhomogeneous N-terminus possible	[32,125,126]
Human serum albumin	Homogeneous N-terminus, high-level secretion of HSA and HSA fusion proteins	Lower secretion levels for other heterologous proteins	[32]
Lysozyme	Homogeneous N-terminus	Secretion levels lower than with MFα	[32]
Fungal leaders	Homogeneous N-terminus, high-level secretion of native fungal enzymes	Not tested for heterologous proteins	[127,128]
Human leaders	–	Less efficient than MFα, unpredictable if they function	[43,125]
Hydrophobins	Homologous N-terminus, secretion levels of GFP comparable to MFα, very short	GFP targeted to vacuole, not tested for other heterologous proteins yet	[129]
Toxin	Homologous N-terminus, secretion levels of GFP comparable to MFα	Not tested for other heterologous proteins yet	[130]
Pichia pastoris PIR1 pre-pro	Secretion of eGFP and antitrypsin	No information about secretion levels	[131]
P. pastoris PHO1	Homologous N-terminus	Low secretion levels	[32]

EAEA: Tetrapeptide preceding the Kex2 leader sequence cleavage site.

analyze the functionality of genes. Gene deletions also have an important application in industrial processes, because by deleting genes, unwanted byproducts can be eliminated from a microbial product. In order to delete a gene, two fundamental prerequisites need to be fulfilled in an experiment: introduction of exogenous DNA into the cell; and a HR event between the exogenous DNA and the targeted genetic locus. At least one HR event is required to integrate DNA at a certain genetic locus. The delivery of exogenous DNA into *P. pastoris* cells is not a major obstacle and can be achieved by standard methods such as electroporation of competent cells (see '*P. pastoris* as a host for protein production' section and Box 2). A gene deletion can be obtained by two different integration strategies (Figure 1). The first, also referred to as gene disruption or knock in, requires only one HR event. On a circular vector, a homologous fragment of the targeted coding sequence is located together with a selection marker. After integration, the cassette disrupts the coding sequence and thus leads to a defective genetic locus. The advantage of this approach is the relatively easy construction of the transformation cassette, because only one homologous fragment from the target host needs to be assembled. In such a strain, the coding sequence of the gene is only disrupted but not eliminated. Therefore, these strains can easily restore the original genotype by another HR eliminating the selection marker cassette. The second strategy is referred to as gene replacement, cassette exchange or knockout, and requires two HR events. Two homologous regions, up- and down-stream of the targeted locus (flanking regions), with a selection marker cassette in between, constitute the transformation cassette. After successful integration, the genomic region between the two flanking sequences is deleted from the genome and a stable genotype is accomplished.

Table 2. Summary of the elemental composition of *Pichia pastoris* grown on the most commonly used C substrates.

Substrate	Elementary composition	Ref.
Methanol/ammonia	$CH_{1.691}O_{0.502}N_{0.176}$	[132]
Sorbitol/ammonia	$CH_{1.757}O_{0.629}N_{0.105}$	[132]
Glycerol/ammonia	$CH_{1.87}O_{0.56}N_{0.18}S_{0.008}$	[133]
Glucose/ammonia	$CH_{1.761}N_{0.143}O_{0.636}S_{0.0018}$	[134]

Box 3. Commonly used shake-flask and fermentation media for the cultivation of *Pichia pastoris*.

- Typical shake-flask media were published in the *Pichia* protocols [135], including buffered complex medium, which is composed of (per liter): 10 g yeast extract, 20 g peptone, 13.4 g yeast nitrogen base with ammonium sulfate and without amino acids (yeast nitrogen base), 0.4 mg biotin, 100 mM potassium phosphate (pH 6.0) and a suitable carbon source (either 20 g glucose, 10 g glycerol or 1% methanol). Alternatively, there are synthetic shake-flask culture media, such as the buffered M2 minimal media (per liter: 20 g citric acid, 3.15 g $[NH_4]_2HPO_4$, 0.03 g $CaCl_2*2H_2O$, 0.8 g KCl, 0.5 g $MgSO_4*7H_2O$, 2 ml biotin [0.2 g/l], 1.5 ml of PTM_1 without biotin trace salts stock solution, and a suitable carbon source [either 20 g glucose, 10 g glycerol or 1% methanol]; pH set to 3.0–6.5 with 5 M KOH) [136].
- Bioreactor cultivations are normally performed in a fed-batch mode with two phases for constitutive expression (e.g., when using the GADPH promoter) or three phases for inducible expression (e.g., for the alcohol oxidase promoter). Glycerol is the preferred carbon source for the batch phase, because even though *P. pastoris* is known to be Crabtree negative, the yeast produces reasonable amounts of ethanol ($q_{ethanol}$ = 0.08 g/g/h [82]) when grown on excess glucose. Ethanol and glucose repress the *AOX1* promoter [137] so that ethanol formation should be avoided by using glycerol as a nonfermentable carbon source. The most commonly used batch medium is the basal salt medium (per liter: 26.7 ml H_3PO_4, 0.93 g $CaSO_4$, 18.2 g K_2SO_4, 14.9 g $MgSO_4*7H_2O$, 4.13 g KOH, 4.35 ml PTM_1 trace salt solution and 40.0 g glycerol, set to the decided pH with 28% NH_4OH) [135], but owing to the high salt concentration and extensive salt precipitation below pH 5.0, the recipe requires further optimization. Prior to methanol fed-batch, a glycerol fed-batch for biomass propagation is usually performed with a simple glycerol feed stock consisting of 630 g glycerol and 12 ml PTM_1 trace salt solution per liter. The feed solution for the induction phase is composed of 988 ml 100% methanol and 12 ml PTM_1 trace salt stock [135]. A reasonable glucose medium for the fed-batch phase consists of 550 g glucose*$1H_2O$, 10 g KCl, 6.45 g $MgSO_4*7H_2O$, 0.35 g $CaCl_2*2H_2O$ and 12 ml PTM_1 stock solution per liter [22]. The recipe for the most commonly used trace salt solution PTM_1 is (per liter): 6.0 g $CuSO_4*5H_2O$, 0.08 g NaI, 3.0 g $MnSO_4*H_2O$, 0.2 g $Na_2MoO_4*2H_2O$, 0.02 g H_3BO_3, 0.5 g $CoCl_2$, 20.0 g $ZnCl_2$, 65.0 g $FeSO_4*7H_2O$, 0.2 g biotin and 5.0 ml H_2SO_4 (95–98%).

Nonhomologous end-joining

The HR event is a crucial step, because the cell has competing pathways to integrate exogenous DNA at random loci. This process is referred to as nonhomologous end-joining (NHEJ) and leads to a reduced probability of HR events. Whereas *S. cerevisiae* preferentially uses the HR route, many other organisms, including filamentous fungi, plants and insects, have substantial activity of the NHEJ pathway [63]. As a consequence, random integration of exogenous DNA sequences, even in the presence of long homologous DNA regions, is observed in *P. pastoris*, which is highly unappreciated in gene-targeting experiments. The NHEJ pathway was first discovered in mammalian cells and needs several proteins for its activity, including the KU70/80 complex, which is the first component that binds to the unprotected and broken DNA ends. In 2004, it was shown that by deleting either *mus51* or *mus52*, encoding the homologs of mammalian KU70/80, gene-targeting efficiency was dramatically increased in *Neurospora crassa* [63]. Since then, gene deletions of the KU70 or KU80 homologs were published in a plethora of organisms, including the recent deletion of the KU70 homolog in *P. pastoris* [64]. On the background of a *ku70*-deletion strain, the targeting efficiency was increased from approximately 12 to 100% for a *HIS4* deletion using 650-bp homologous flanking regions, and enabled high efficiencies with flanking regions as short as 200 bp. As the NHEJ pathway is one of the main pathways for restoring double-strand breaks, it is not surprising that strains lacking this repair mechanism are sensitive to several environmental stresses, including ionizing radiation [65,66], methyl methanesulfonate, bleomycin [63] and ultraviolet irradiation [67]. A higher sensitivity towards ultraviolet rays was also observed for the *ku70 P. pastoris* strain [64]. The reduced fitness of such a strain does not make it an optimal production strain and, therefore, the gene should be reintroduced into the cell before it is used for industrial purposes. Unfortunately, this decreases the overall applicability of the system, because a *ku70*-deletion mutant is needed as the recipient strain, and after the strain construction process, another transformation step is required to restore the *ku70* locus.

Split-marker system

Another effective option to increase the gene-targeting efficiency in *P. pastoris* is the split-marker

system. A gene-replacement cassette is constructed as described above, but the selection marker gene is split into two overlapping DNA fragments (Figure 2). The two parts of the cassette are then transformed simultaneously and only after a successful HR of the two fragments does the selection marker become active. The proximity of the other recombination sites increases the probability that these recombination events are also mediated by HR. The system was originally developed for *S. cerevisiae* in 1996 [68] and transferred to many other fungi, including *P. pastoris* [69]. Virtually any selection marker can be used for this system; however, the division needs to be introduced in a way that both truncated fragments alone are not active any more. In *P. pastoris*, the system was used with common markers like the KanMX cassette [69], as well as the hygromycin and Zeocin resistance cassettes [Mattanovich D *et al.*, Unpublished Data]. No special recipient strain is needed to apply the split-marker technology, which makes it a universal, applicable tool.

Transcriptomics & proteomics

With the advent of genome sequencing, the analysis of biological functions at the genome scale has become possible, usually summarized as systems biology. The core of such analyses is the availability of a genomic sequence, which was published for *P. pastoris* in 2009 [5,6]. The application of yeast systems biology to recombinant protein production [70] has, so far, been mainly based on the quantitative and comparative measurement of transcript and protein levels. Compared with other yeasts, the genome sequence of *P. pastoris* has been published quite late; however, a commercial

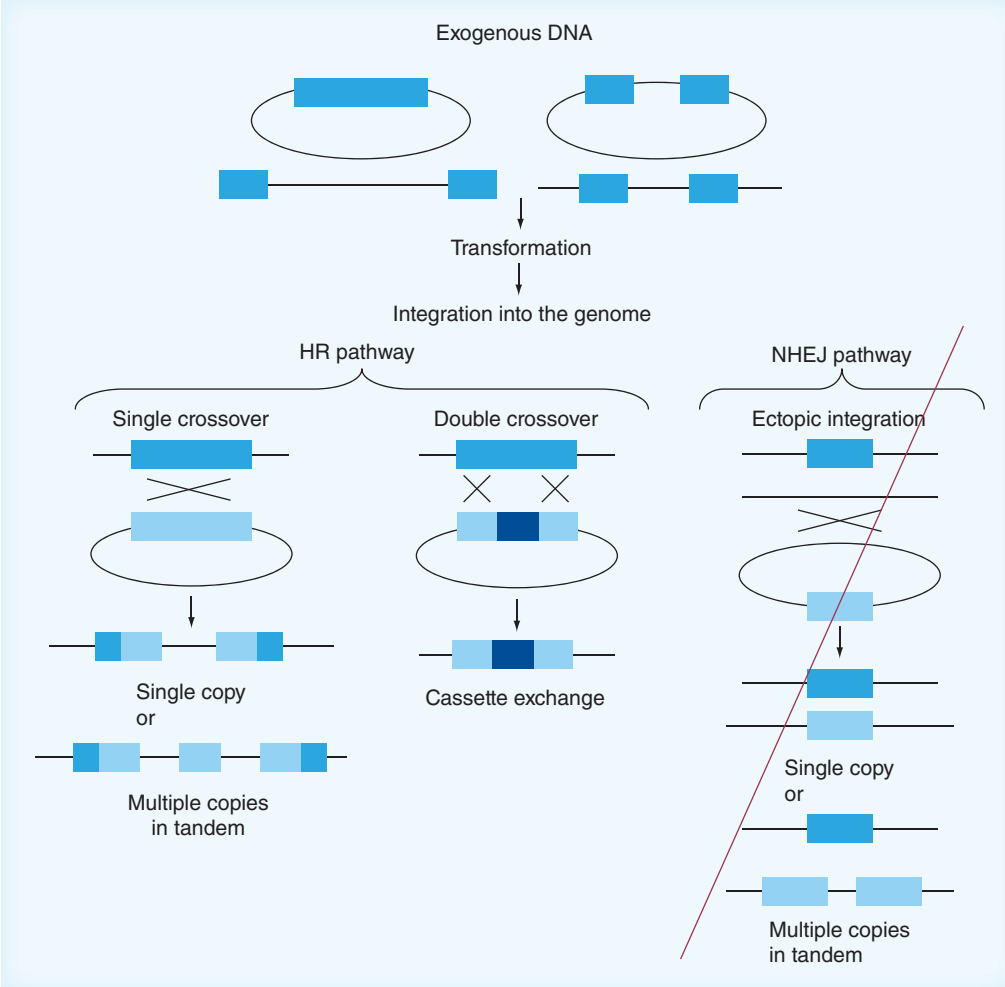

Figure 1. General integration events occurring during a gene deletion experiment. Via the HR pathway, gene deletion constructs are obtained. False-positive transformants are gathered by the action of the NHEJ pathway. Homologous regions are indicated as light-shaded boxes, and an inserted selection marker as a dark-shaded box. Recombination events are marked as crosses. HR: Homologous recombination; NHEJ: Nonhomologous end-joining.

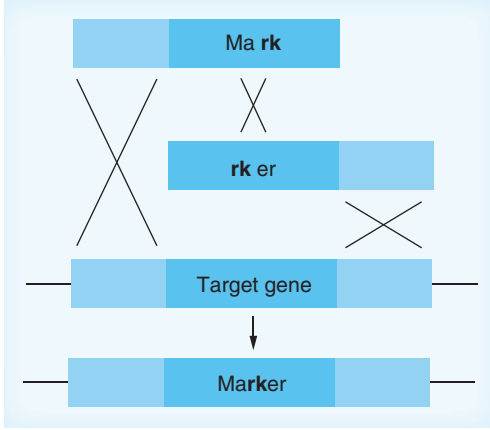

Figure 2. For the split-marker approach, a selection marker gene is split into two fragments. A homologous recombination (symbolized by a cross) event is necessary to activate the gene, which also favors the homologous recombination at the adjacent sites. Lighter shaded boxes are flanking regions of the targeted gene.

draft sequence was available before that point from Integrated Genomics (IL, USA), which allowed some earlier systems biology work.

Even before this, transcriptomic analyses were performed based on heterologous hybridization of *P. pastoris* samples to *S. cerevisiae* microarrays [71]. This work revealed a downregulation of the tricarboxylic acid (TCA) cycle and ribosomal biogenesis-related genes upon the shift to methanol-grown fed batch, which was obviously related to the lower specific growth rate. In addition, it formed the basis for direct application of transcriptomic data to strain improvement, allowing the identification of genes whose overexpression could enhance the secretion of heterologous proteins up to 2.5-fold [49]. Moreover, the microarray signal intensities formed a basis to predict expression levels and thus promoter strengths of *P. pastoris* genes, so that 24 new promoters with a wide range of expression strengths could be identified [19].

Based on the first available genome sequence, *P. pastoris*-specific DNA microarrays could be developed and applied to a detailed study of the UPR in *P. pastoris* [72], revealing a number of regulatory differences to *S. cerevisiae*. High similarity was observed for genes involved in protein translocation, folding and early glycosylation steps, whereas more distal steps of glycosylation and secretion, as well as lipid metabolism, were more diversely regulated when comparing the two yeasts. Based on this work, a detailed transcriptomic and proteomic study of cellular reactions to environmental stress factors was performed with different microorganisms, among them *P. pastoris* and *S. cerevisiae*. The effects of different media osmolarity and oxygen supply were compared at both transcriptome and proteome levels, revealing strong transcriptional regulation of core metabolic genes, correlating with proteomic data. High osmolarity led to the upregulation of genes related to the UPR and downregulation of genes involved in the TCA cycle and cell wall biogenesis [73]. Hypoxic conditions led to major regulatory shifts of central carbon metabolism both on the transcript and protein level, as well as changes in lipid metabolism, stress responses and protein folding and secretion [74]. *S. cerevisiae* seems to regulate these processes at a post-transcriptional or even post-translational level, as the transcriptome was much less perturbed by oxygen limitation [75]. The differential proteomes were analyzed at different osmolarities [73] and oxygen supply [74] by 2D SDS gel electrophoresis and fluorescence labeling, followed by spot picking from 2D gels and liquid chromatography–mass spectrometry-based identification of differentially regulated spots. The differential proteomes of a wild-type and a production strain at different temperatures highlighted that TCA cycle enzymes, oxidative stress-related proteins and chaperones are reduced at low temperature [76], which explains the higher biomass yield and specific productivity at low (20–25°C) compared with high (30°C) temperatures.

While differential proteomics reveals the regulation of cellular functions at the protein level, absolute proteomes are also of value, describing the equipment of a cell with proteins under given conditions. The response of *P. pastoris* to methanol induction was studied recently at the proteome level [77], describing a drastic change in the proteome from growth on glycerol to methanol. It is mainly proteins of the methanol dissimilation pathway that are highly induced, indicating that energy production dominates over biomass formation under these circumstances. Additionally, proteins related to UPR and ERAD were upregulated upon prolonged cultivation on methanol, which was attributed to the strong overproduction of recombinant hepatitis B surface antigen. UPR and ERAD proteins were also among the intracellular binding partners of a recombinant antibody Fab fragment in *P. pastoris*, as identified by a coimmunoprecipitation-based proteomic study [78]. This Fab fragment was targeted to secretion; however, cytosolic chaperones were also among the binding partners, leading to the

conclusion that post-transcriptional translocation may be the predominant pathway for transport to the ER, at least for these overproduced recombinant proteins. This matches with the fact that cytosolic chaperones were also among the proteins that helped to increase Fab fragment secretion upon co-overproduction [49].

Microbial secretomes are the extracellular subsets of their proteomes and provide valuable information on the lifestyle of an organism and its communication with the environment. For host platforms used for the production of secreted proteins, the secretome provides valuable information on the potential contamination of the product with host cell protein and the risk of product degradation by potentially present proteases. The secretomes of *P. pastoris* grown on glucose [6] and methanol [79] were identified. While a significant overlap was observed between cultures grown on glucose and on glycerol, culture supernatants of methanol-grown cells contain significantly more different proteins, and most of them are not predicted to be secreted. While the release of intracellular proteins from intact cells has been described [80], it is plausible that the more stressful growth conditions on methanol lead to increased cell lysis and a higher load of host cell protein in the culture supernatant. In order to reduce this load, the major secreted protein that was identified in both studies was characterized and deleted [69]. This protein, named Epx1, has no clear function and was deleted without deleterious effect, providing a production platform that secretes a reduced host cell protein load.

Metabolic modeling

Mathematical modeling of cellular metabolism has developed to become a key part of systems biology analysis and description of living cells. Its application in biotechnology leads mainly to the prediction of cell engineering targets in order to enhance product formation [81]. Genome-scale metabolic models of *P. pastoris* have been published in recent years [82–84] and have been applied to predict protein productivity upon changes of the substrate and oxygen supply. These *P. pastoris* metabolic models are a key requisite for the interpretation of metabolomics analysis, such as metabolic flux analysis [85,86] and multiparallel metabolite quantification [87,88]. A strong impact of modeling and measuring cellular metabolism on *P. pastoris* research can be anticipated, but it will depend on the further development of bioinformatics data interpretation tools.

The age of 'high throughput': increasing the quantities of analysis & decreasing the time of development

An obvious inherent characteristic of microbial cell factories, such as *P. pastoris*, is their small size, as every individual cell comprises the entire machinery that renders it a functional cell factory on its own. This enables the development of high-throughput methods. Scale-down of culture volume enables more experiments to be carried out in parallel. This leads from parallel shake-flask cultures to hundreds or thousands of parallel cultures in plate format, up to the analysis of billions of single cells (e.g., by flow cytometry).

As outlined below, high-throughput methods can be used to identify optimal or optimized production strains, but also to screen products with regards to specified characteristics, and to gain biological knowledge of the production strain.

High throughput for strain development & optimization

As it is often hardly predictable which microorganism is the best producer of the desired protein, systems are being developed for the parallel expression of proteins in various expression systems to gain rapid insight into which organism is the best starting point for further optimizations. Dortay *et al.* developed a platform for parallel protein expression *in vivo* and *in vitro*, which combines ligation-independent cloning with detection of the expressed proteins through fusions to iRFP [89]. For each protein in question, two PCR fragments are generated and inserted in parallel into ten expression vectors suitable for protein expression in microbial hosts, including *Escherichia coli*, *Kluyveromyces lactis*, *P. pastoris*, the protozoon *Leishmania tarentolae* and an *in vitro* transcription/translation system.

In other cases, the production host *P. pastoris* is defined from the beginning of the project, yet laborious strain development and screening is required to obtain a good producer. High-throughput methodologies are increasingly important in assisting the development of production strains. A wide collection of new vectors and new cloning systems have been developed for high-throughput recombinant gene expression in *P. pastoris* [90,91]. In particular, an expression system based on *in vivo* HR of various expression cassettes with a plasmid within the transformed yeast cell has been proposed [92]. Thereby, multiple constructs are easily produced in parallel. The usefulness of the

system is shown by the successful heterologous expression of seven membrane proteins.

The identification of high-producing cell lines is critical for heterologous protein production in general. Several cultivation techniques have been adopted for the high-throughput cultivation of *P. pastoris*. These include cultivation in 24-well plates [93] or even smaller sets [94]. However, a typical production process involves high cell density cultivations of the yeast with controlled pH and dissolved oxygen concentrations, which is not usually obtainable in microtiter plates. Miniaturized bioreactor systems have been described, ranging from 80 ml to as small as 100 µl [95]. As it turns out, neither solution fits for all examples, so the suggested procedure is a combination of deep well plate cultures with small-scale bioreactors [95].

At the end of every bioprocess, the product has to be purified and quantified. With the development of parallel fermentations of strains generated by high-throughput methods, the downstream processing also has to be developed in a high-throughput format, otherwise the cultures cannot be analyzed. Jiang *et al.* describe such a development for the analysis and purification of monoclonal antibodies, which have been produced by *P. pastoris* in a 96-well plate format [96]. The monoclonal antibody titer and quality obtained from the proposed method are comparable to those from conventional column chromatography. Hundreds of expression screening samples can be processed in a day. Not only accurately determined titers are obtained, but also milligram quantities of monclonal antibodies are generated for quality assessment, such as purity, folding, glycosylation or antigen binding affinity.

Screening of enhanced protein production strains employing microengraving brings parallel strain analysis to the next level [97]. A population of production strains with sufficient heterogeneity (e.g., induced by mutagenesis) is deposited into an array of sub-nanoliter wells with a density of approximately one cell per well. The wells are covered with a glass slide that is capable of capturing the product of interest. After a defined amount of time, the glass slide is removed and analyzed in a similar fashion to a microarray with regards to the amount of product. The best-performing cells can then be recovered, cultivated on a larger scale and analyzed or subjected to a further round of screening. With this technique, Panagiotou *et al.* were able to isolate a strain producing a Fc fragment with 265% improved product titer [97].

While parallel testing of host–protein combinations remains the method of choice for recombinant protein production, a better understanding of the mechanisms that promote or impede the production of a given protein in different hosts would help to reduce the required effort for production. Boettner *et al.* chose a high-throughput approach to search for parameters with a major impact on heterologous protein expression [98]. They analyzed 79 different human proteins expressed intracellularly in *P. pastoris*. Three analyzed sequence-based factors proved to have a statistically significant correlation with the expression level. These are the abundance of AT-rich regions in the coding sequence, the isoelectric point of the recombinant protein and the occurrence of a protein homolog in yeast. Interestingly, some often-discussed factors, such as codon usage or GC content, did not show any significant impact on protein yield.

Library approaches

A prerequisite for the construction of optimal production strains are appropriate tools for genetic engineering. Promoters that give rise to a defined and stable expression of the gene for the product or accessory genes for strain tuning are one such tool (see '*P. pastoris* as a host for protein production' section).

Typical genome-scale libraries include cDNA overexpression and gene-knockout libraries, which have been employed, for example, to screen for genes enhancing protein secretion or for pexophagy-related genes [26,99].

Engineering of protein glycosylation, as mentioned above, requires the correct successive localization of enzymes synthesizing and cleaving the modified glycosyl moieties in the ER and Golgi. In order to target the enzymes performing appropriate reactions similar to those in human cells, GlycoFi, Inc. (NH, USA), constructed combinatorial libraries, fusing various fungal targeting sequences with human enzymatic domains together with a high-throughput screening protocol for proper glycosylation [100].

Flow cytometry & cell surface display

Expression of proteins on the cell surface of yeasts has a wide range of applications in biotechnology [101]. For yeast surface display, the target protein is genetically fused with an anchor protein and retained on the cell wall instead of being released into the supernatant. This has a variety of advantages, as it connects a phenotype, such as

secreted protein amount or characteristics (e.g., ligand binding), with a genotype. Yeast surface display allows the screening of large numbers of genotypes for their phenotypes with different techniques. Flow cytometry is one of the most powerful methods in this context [102]. The analysis of surface-displayed proteins is used for strain optimization with regards to production capacity or directed evolution of proteins for increased affinity or thermal stability, screening of antibody libraries, epitope mapping or the direct use of the cells as whole-cell biocatalysts or vaccines [101].

Combining cell surface display with fluorescence-activated cell sorting allows the very-high-throughput screening of production hosts with the aim of isolating cells with increased secretion capacity for heterologous proteins. Besides isolating cells with improved properties, it is also possible to gain insight into which groups of genes are of particular relevance for the cellular characteristic in question. Therefore, high-throughput methods such as DNA microarrays or sequencing are applied. The method combining fluorescence-activated cell sorting and DNA microarray-assisted clone identification was termed genome-scale analysis of library sorting [26]. A *P. pastoris* cDNA library was coexpressed in a strain displaying the Fab fragment of a monoclonal antibody against HIV-1. Three genes were identified, increasing the relative expression level of the surface-displayed model protein by up to 145%, while the microarray-based monitoring of the enrichment of genes causing enhanced protein secretory capacity led to novel insights into the bottlenecks of protein secretion.

P. pastoris as host/vehicle for substance testing

Yeasts are the preferred vehicles for protein production and can be used for high-throughput screening regarding characteristics of the produced protein. Diehl *et al.* reported the use of *P. pastoris* to establish a novel bioassay for the identification of matrix metalloproteinase (MMP) inhibitors [103]. Malfunctions of human MMPs are connected to a wide range of diseases, including diabetes mellitus, arthritis, cardiovascular diseases and cancer. However, the development of specific inhibitors for this class of proteins has been shown to be difficult, because the MMP had to be purified from human cells, which was expensive, and a sensitive bioassay was missing. Overexpression and surface anchoring of human MMPs on *P. pastoris* enabled Diehl *et al.* to develop an assay that can be easily performed in a microtiter plate format. In short, the cells with the (auto)activated MMP on the surface are first incubated with the substance to be tested as the inhibitor and then with fluorescein-labeled gelatin. Activity of the MMP liberates fluorescein, which is easily and quantitatively detectable. This bioassay should hasten the progression of inhibitor development, as it enables quantitative comparison of MMP inhibitor capacity.

A similar approach has been chosen by Hu *et al.* to develop a screening method for human dipeptidyl peptidase-IV inhibitors [104]. Recombinant production of human dipeptidyl peptidase-IV in the establishment of *P. pastoris* allowed a microtiter plate-based assay to be established, independent from difficult protein isolation procedures.

P. pastoris as a model for human cell biology & disease

Peroxisome biosynthesis & degradation

Peroxisomes are eukaryotic organelles featuring densely packed enzymes coated with a single membrane. Their name derives from the common feature of producing or degrading peroxides and other reactive oxygen species, and they share the enzymes for acyl-CoA β-oxidation. Peroxisomal dysfunction in humans is responsible for severe diseases [105].

In methylotrophic yeasts, peroxisomes house the methanol metabolism enzymes, such as alcohol oxidase (producing formaldehyde and hydrogen peroxide) and catalase, which detoxify the peroxide. Dependent on substrate availability, methylotrophic yeasts can rapidly build up or degrade their peroxisomes. This controlled proliferation or degradation, and the high abundance in induced conditions, has rendered *P. pastoris* and *Hansenula polymorpha* valuable model systems for genetic and physiological studies of peroxisome biology [106].

Peroxisomes have been shown to proliferate by fission, which is regarded as the normal mode of proliferation, while the relevance of *de novo* biosynthesis in normally growing cells is in debate. *De novo* formation of peroxisomes is still not fully understood. Peroxisomal membranes (lipids) and membrane proteins are assembled in the ER in distinct subcomplexes, which bud from the ER as preperoxisomal vesicles and fuse to form functional peroxisomes [4]. Peroxisomal matrix proteins are imported into the organelle after their translation in the cytosol. Translocation across the peroxisomal

membrane requires peroxisomal targeting signals (PTS). Three PTS types have been characterized, which are quite conserved among eukaryotes: PTS1 and PTS2 are responsible for the translocation of luminal peroxisome proteins, while mPTS mediates translocation of peroxisomal membrane proteins [107].

Degradation of peroxisomes, also referred to as pexophagy, has recently been reviewed [105,108], describing mechanisms involving genes and their regulation. Micropexophagy is characterized by vacuolar engulfment of the degradation target, while macropexophagy is initiated by entrapment of degradation targets by a double membrane, which then fuses with the vacuole. Schroder *et al.* have used *P. pastoris* to identify pexophagy-related genes by random gene interruption mediated by restriction enzyme-mediated integration [99].

Using *P. pastoris* as a model system, peroxisome function, autophagy and fusion of autophagosome-like vesicles to the plasma membrane were identified as conserved mechanisms involved in the unconventional secretion of acyl-CoA binding protein (human ACBP and yeast Abp1), which plays an important role in the regulation of neural processes, steroidogenesis and insulin secretion in humans [109].

Protein folding & secretion: *P. pastoris* as a model for human cells

Besides being a good production host for recombinant proteins, *P. pastoris* is also used as a model system to study the secretory pathway, as its ER and Golgi apparatus structures resemble those of higher eukaryotes, in contrast to the commonly used yeast model *S. cerevisiae* [2]. In contrast to *S. cerevisiae,* where the Golgi is distributed throughout the entire cell, *P. pastoris* forms ordered Golgi stacks similar to the mammalian Golgi. These Golgi stacks are located next to discrete transitional ER sites, which makes it easy to study an entire ER–Golgi complex at once [110]. Figure 3 shows typical ER and Golgi localization in *P. pastoris*. Applying 3D tomography and 4D fluorescence videomicroscopy, coworkers from the Glick laboratory revealed the presence of single Golgi cisternae (three to four per stack), the formation of a *cis*-, medial or *trans*-Golgi network and showed that the individual cisternae of the Golgi mature as a means of anterograde intra-Golgi transport [111,112]. Further details on state-of-the-art knowledge of Golgi transport have been summarized previously [113]. Very recent results from the Glick laboratory showed a further phenomenon conserved in *P. pastoris* and mammalian cells, namely the interaction of two secretory proteins, Sec12 and Sec16, involved in COPII assembly at transitional ER sites [114]. Thus, *P. pastoris* serves as a model system for studying the mechanisms that create tubular smooth ER sites and Golgi stacks. Novel components of tubular smooth ER sites/ER-to-Golgi transport were identified that are not present in *S. cerevisiae* [115].

Moreover, it was proven that the transcriptional reactions to ER stress conditions (e.g., activation of the UPR) in *P. pastoris* are more similar to mammalian cells than those of *S. cerevisiae* [72]. Cells with an activated UPR exhibit a more reducing redox milieu in both compartments, and reduction of the cytosolic redox state is a response to increased oxidative protein folding in the ER, as demonstrated by targeting redox-sensitive GFP variants to the ER or the cytosol of *P. pastoris*, respectively [116].

As a wide variety of aging-related human diseases, including diabetes mellitus, atherosclerosis and neurodegenerative diseases such as Alzheimer's disease, amyotrophic lateral sclerosis and Parkinson's disease has been associated with protein misfolding or aggregation in the ER and the loss of redox control [117,118], *P. pastoris* may serve as a valuable model system to study the inter-relation between ER stress and cellular redox homeostasis, and contribute to better understanding of the development of aging-related diseases. In this respect, *P. pastoris* has been applied to uncover novel effects of the amyloid precursor protein of Alzheimer's disease on intracellular copper levels and copper efflux, and the biological function of clioquinol treatment, which both also affected expression of antioxidants [119]. Copper homeostasis and enhanced manganese concentrations were also revealed to promote prion formation [120].

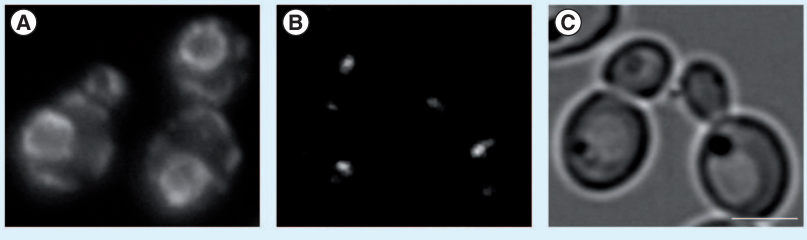

Figure 3. Subcellular localization of endoplasmic reticulum and late Golgi apparatus in *Pichia pastoris*. (A) Fluorescent image of *Pichia pastoris* stained with an anti-HDEL antibody (endoplasmic reticulum marker). **(B & C)** Fluorescent and bright field images of *P. pastoris* expressing Sec7-eGFP (late Golgi marker). Scale bar: 3 µm.

Conclusion

Biopharmaceutical drug candidates produced in *P. pastoris* are entering clinical studies and the market at an increasing pace, similar to the increasing market entry of process enzymes and research reagents. This development reflects the fact that *P. pastoris* is the most widely used yeast production system for heterologous proteins in research. It can be anticipated that successful glycoengineering will further boost this development. However, a number of bottlenecks, mainly in protein folding and secretion, still limit the productivity of this production platform. Public availability of the *P. pastoris* genome sequence has enabled the research community to study the yeast's physiology and genetics efficiently, and to develop a toolset for research and protein production, thus fuelling the development of improved strains. Beyond protein production, *P. pastoris* is gaining increasing interest as a model system for basic research, for (e.g., on eukaryotic protein folding and secretion).

Future perspective

Given the fact that *P. pastoris* is used ten-times more often for heterologous protein production than *S. cerevisiae* or any other yeast in the current published literature, it can be assumed that the biopharmaceutical application of this production system will rapidly increase, adopting a place among the top three production systems besides Chinese hamster ovary cells and *E. coli*. Advanced strain and process development will strongly increase productivity, and achievements in the engineering of glycosylation, folding and assembly of complex proteins will expand the range of target proteins to include those that are currently produced only in mammalian cells.

Rapid and cheap genome sequencing has 'democratized' research on nonconventional organisms. We expect that *P. pastoris*, among some other yeast species sharing more similarity with higher eukaryotes, will gain much more attention as model systems for human cell biology and disease.

Executive summary

Pichia pastoris *as a host for protein production*

- *Pichia pastoris* is widely established for heterologous protein production, and the number of approved biopharmaceuticals produced in this host is growing annually.
- In addition to established promoters from the methanol or the glycolytic pathway, new strong and/or regulated promoters are described.
- A method for efficient transformation as a prerequisite for many screening approaches is described, as well as targeted gene copy-number amplification.
- Protein folding and secretion are rate-limiting steps for efficient production of heterologous proteins. Methods to improve these steps are highlighted.
- Optimized cultivation media and fermentation processes are provided.

Novel methods in P. pastoris *research*

- With the availability of the genome sequence of *P. pastoris*, gene deletion has gained a great deal of interest. The split-marker system is highlighted as an efficient gene-knockout method.
- Transcriptomics and proteomics are employed for genome-scale investigation of *P. pastoris* physiology and as a source of information for strain engineering.
- High-throughput methods have been developed for analysis and screening. Their application to identify improved production strains, promoters, process steps and for drug testing are summarized.

P. pastoris *as a model for human cell biology & disease*

- *P. pastoris* is a valuable model to study peroxisome proliferation and degradation due to its high abundance on methanol media and its controllable synthesis and degradation.
- A number of protein folding and secretion processes in nonconventional yeasts are more similar to those of higher eukaryotes than in *Saccharomyces cerevisiae*, so *P. pastoris* is used as an efficient model to study these processes.

Conclusion

- Biopharmaceutical drug candidates produced in *P. pastoris* are entering clinical studies and the market at an increasing pace, reflecting the leading position of this production system among yeast hosts in research. Other products include process enzymes and research reagents.
- Successful glycoengineering will further boost this development.
- Genome sequencing of *P. pastoris* has enabled the research community to study the yeast's physiology and genetics efficiently, and to develop a toolset for research and protein production.
- Due to the specialized evolution of *S. cerevisiae,* it has become obvious that other yeasts, such as. *P. pastoris,* are important models for basic research (e.g., on eukaryotic protein folding and secretion).

Acknowledgements

The authors would like to thank the BOKU-VIBT Imaging Center for access to Leica fluorescence microscope devices.

Financial & competing interests disclosure

Research on Pichia pastoris *in our laboratory is supported by the Austrian Science Fund (FWF), the Austrian Research Promotion Agency, the European Science Foundation (ESF) and by the Federal Ministry of Economy, Family and Youth (BMWFJ), the Federal Ministry of Transport, Innovation and Technology (bmvit), the Styrian Business Promotion Agency (SFG), the Standortagentur Tirol and ZIT – Technology Agency of the City of Vienna through the COMET-Funding Program managed by the Austrian Research Promotion Agency (FFG). J Nocon is funded by the doctoral program 'BioToP – Biomolecular Technology of Proteins' (Austrian Science Fund, FWF Project W1224). Further support by Polymun Scientific GmbH, Biomin Research Center, Boehringer-Ingelheim RCV, Lonza AG, Biocrates Life Sciences AG, VTU Technology GmbH and Sandoz GmbH is acknowledged. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.*

No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as:
▪ of interest
▪▪ of considerable interest

1. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast* 22(4), 249–270 (2005).
2. Papanikou E, Glick BS. The yeast Golgi apparatus: insights and mysteries. *FEBS Lett.* 583(23), 3746–3751 (2009).
3. Corchero JL, Gasser B, Resina D *et al.* Unconventional microbial systems for the cost-efficient production of high-quality protein therapeutics. *Biotechnol Adv.* doi:10.1016/j.biotechadv.09.001 (2012) (Epub ahead of print).
4. Ma C, Agrawal G, Subramani S. Peroxisome assembly: matrix and membrane protein biogenesis. *J. Cell Biol.* 193(1), 7–16 (2011).
5. De Schutter K, Lin YC, Tiels P *et al.* Genome sequence of the recombinant protein production host *Pichia pastoris. Nat. Biotechnol.* 27(6), 561–566 (2009).

▪▪ **Genome sequence of a widely used commercial *Pichia pastoris* strain plus analysis of protein secretion and glycosylation pathways.**

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

▪▪ **Genome sequence of a *P. pastoris* wild-type strain plus analysis of substrate uptake and the secreted proteome.**

7. Kuberl A, Schneider J, Thallinger GG *et al.* High-quality genome sequence of *Pichia pastoris* CBS7435. *J. Biotechnol.* 154(4), 312–320 (2011).

▪ **Genome sequence of a *P. pastoris* wild-type strain with improved annotation.**

8. Cregg JM, Russell KA. Transformation. *Methods Mol. Biol.* 103, 27–39 (1998).
9. Sunga AJ, Tolstorukov I, Cregg JM. Posttransformational vector amplification in the yeast *Pichia pastoris. FEMS Yeast Res.* 8(6), 870–876 (2008).
10. Gil DF, Garcia-Fernandez R, Alonso-del-Rivero M *et al.* Recombinant expression of ShPI-1A, a non-specific BPTI-Kunitz-type inhibitor, and its protection effect on proteolytic degradation of recombinant human miniproinsulin expressed in *Pichia pastoris. FEMS Yeast Res.* 11(7), 575–586 (2011).
11. Marx H, Mecklenbräuker A, Gasser B, Sauer M, Mattanovich D. Directed gene copy number amplification in *Pichia pastoris* by vector integration into the ribosomal DNA locus. *FEMS Yeast Res.* 9(8), 1260–1270 (2009).
12. Cereghino J, Cregg J. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris. FEMS Microbiol. Rev.* 24(1), 45–66 (2000).
13. Mochizuki S, Hamato N, Hirose M *et al.* Expression and characterization of recombinant human antithrombin III in *Pichia pastoris. Protein Expr. Purif.* 23(1), 55–65 (2001).
14. Shen S, Sulter G, Jeffries TW, Cregg JM. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris. Gene* 216(1), 93–102 (1998).
15. Tschopp JF, Brust PF, Cregg JM, Stillman CA, Gingeras TR. Expression of the *lacZ* gene from two methanol-regulated promoters in *Pichia pastoris. Nucleic Acids Res.* 15(9), 3859–3876 (1987).
16. Menendez J, Valdes I, Cabrera N. The *ICL1* gene of *Pichia pastoris*, transcriptional regulation and use of its promoter. *Yeast* 20(13), 1097–1108 (2003).
17. Ahn J, Hong J, Park M *et al.* Phosphate-responsive promoter of a *Pichia pastoris* sodium phosphate symporter. *Appl. Environ. Microbiol.* 75(11), 3528–3534 (2009).
18. Hartner F, Ruth C, Langenegger D *et al.* Promoter library designed for fine-tuned gene expression in *Pichia pastoris. Nucleic Acids Res.* 36(12), e76 (2008).
19. Stadlmayr G, Mecklenbrauker A, Rothmuller M *et al.* Identification and characterisation of novel *Pichia pastoris* promoters for heterologous protein production. *J. Biotechnol.* 150(4), 519–529 (2010).
20. Waterham HR, Digan ME, Koutz PJ, Lair SV, Cregg JM. Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. *Gene* 186(1), 37–44 (1997).
21. Baumann K, Maurer M, Dragosits M, Cos O, Ferrer P, Mattanovich D. Hypoxic fed-batch cultivation of *Pichia pastoris* increases specific and volumetric productivity of recombinant proteins. *Biotechnol. Bioeng.* 100(1), 177–183 (2008).
22. Maurer M, Kühleitner M, Gasser B, Mattanovich D. Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with *Pichia pastoris. Microb. Cell Fact.* 5, 37 (2006).
23. Qin X, Qian J, Yao G, Zhuang Y, Zhang S, Chu J. GAP promoter library for fine-tuning of gene expression in *Pichia pastoris. Appl. Environ. Microbiol.* 77(11), 3600–3608 (2011).
24. Ahn J, Hong J, Lee H *et al.* Translation elongation factor 1-alpha gene from *Pichia pastoris*: molecular cloning, sequence, and use of its promoter. *Appl. Microbiol. Biotechnol.* 74(3), 601–608 (2007).
25. Gasser B, Dragosits M, Mattanovich D. Engineering of biotin-prototrophy in *Pichia pastoris* for robust production processes. *Metab. Eng.* 12(6), 573–580 (2010).
26. Stadlmayr G, Benakovitsch K, Gasser B, Mattanovich D, Sauer M. Genome-scale analysis of library sorting (GALibSo): Isolation of secretion enhancing factors for recombinant protein production in *Pichia pastoris. Biotechnol. Bioeng.* 105(3), 543–555 (2010).

▪ **Combination of cDNA library sorting with genome-scale analysis of sort pools based on DNA microarrays.**

27. Cos O, Ramón R, Montesinos J, Valero F. Operational strategies, monitoring and

control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: a review. *Microb. Cell Fact.* 5, 17 (2006).

28. Zhu J, Zhang MQ. SCPD: a promoter database of the yeast *Saccharomyces cerevisiae*. *Bioinformatics* 15(7–8), 607–611 (1999).
29. Damasceno LM, Huang CJ, Batt CA. Protein secretion in *Pichia pastoris* and advances in protein production. *Appl. Microbiol. Biotechnol.* 93(1), 31–39 (2012).
30. Inan M, Aryasomayajula D, Sinha J, Meagher M. Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. *Biotechnol. Bioeng.* 93(4), 771–778 (2006).
31. 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.* 85(4), 367–375 (2004).
32. Daly R, Hearn MT. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J. Mol. Recognit.* 18(2), 119–138 (2005).
33. Anelli T, Sitia R. Protein quality control in the early secretory pathway. *EMBO J.* 27(2), 315–327 (2008).

▪▪ **Protein quality control plays a huge role in the production of secreted recombinant proteins. Basic research on *Saccharomyces cerevisiae* and mammalian cells is an indispensable source of knowledge.**

34. Braakman I, Bulleid NJ. Protein folding and modification in the mammalian endoplasmic reticulum. *Annu. Rev. Biochem.* 80, 71–99 (2011).
35. Brodsky JL, Skach WR. Protein folding and quality control in the endoplasmic reticulum: recent lessons from yeast and mammalian cell systems. *Curr. Opin. Cell Biol.* 23(4), 464–475 (2011).
36. Margittai E, Sitia R. Oxidative protein folding in the secretory pathway and redox signaling across compartments and cells. *Traffic* 12(1), 1–8 (2011).

▪ **The importance of the mechanisms of disulfide bond formation for recombinant protein production is receiving increasing attention. Basic mechanisms in different eukaryotes are described here.**

37. Hoseki J, Ushioda R, Nagata K. Mechanism and components of endoplasmic reticulum-associated degradation. *J. Biochem.* 147(1), 19–25 (2010).
38. Kohno K. Stress-sensing mechanisms in the unfolded protein response: similarities and differences between yeast and mammals. *J. Biochem.* 147(1), 27–33 (2010).
39. Malhotra JD, Kaufman RJ. The endoplasmic reticulum and the unfolded protein response. *Semin. Cell Dev. Biol.* 18(6), 716–731 (2007).
40. Vembar SS, Brodsky JL. One step at a time: endoplasmic reticulum-associated degradation. *Nat. Rev. Mol. Cell Biol.* 9(12), 944–957 (2008).
41. Idiris A, Tohda H, Kumagai H, Takegawa K. Engineering of protein secretion in yeast: strategies and impact on protein production. *Appl. Microbiol. Biotechnol.* 86(2), 403–417 (2010).
42. Damasceno LM, Anderson KA, Ritter G, Cregg JM, Old LJ, Batt CA. Co-overexpression of chaperones for enhanced secretion of a single-chain antibody fragment in *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* 74(2), 381–389 (2007).
43. Gasser B, Maurer M, Gach J, Kunert R, Mattanovich D. Engineering of *Pichia pastoris* for improved production of antibody fragments. *Biotechnol. Bioeng.* 94(2), 353–361 (2006).
44. Vad R, Nafstad E, Dahl L, Gabrielsen O. Engineering of a *Pichia pastoris* expression system for secretion of high amounts of intact human parathyroid hormone. *J. Biotechnol.* 116(3), 251–260 (2005).
45. Huo X, Liu Y, Wang X *et al.* Coexpression of human protein disulfide isomerase (hPDI) enhances secretion of bovine follicle-stimulating hormone (bFSH) in *Pichia pastoris*. *Protein Expr. Purif.* 54(2), 234–239 (2007).
46. Tsai CW, Duggan PF, Shimp RL Jr, Miller LH, Narum DL. Overproduction of *Pichia pastoris* or plasmodium falciparum protein disulfide isomerase affects expression, folding and *O*-linked glycosylation of a malaria vaccine candidate expressed in *P. pastoris*. *J. Biotechnol.* 121(4), 458–470 (2006).
47. Guerfal M, Ryckaert S, Jacobs PP *et al.* The *HAC1* gene from *Pichia pastoris*: characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins. *Microb. Cell Fact.* 9, 49 (2010).
48. Resina D, Maurer M, Cos O *et al.* Engineering of bottlenecks in *Rhizopus oryzae* lipase production in *Pichia pastoris* using the nitrogen source-regulated FLD1 promoter. *N. Biotechnol.* 25(6), 396–403 (2009).
49. 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.* 73(20), 6499–6507 (2007).
50. Marx H, Sauer M, Resina D *et al.* Cloning, disruption and protein secretory phenotype of the GAS1 homologue of *Pichia pastoris*. *FEMS Microbiol. Lett.* 264(1), 40–47 (2006).
51. Liu SH, Chou WI, Lin SC, Sheu CC, Chang MD. Molecular genetic manipulation of *Pichia pastoris* SEC4 governs cell growth and glucoamylase secretion. *Biochem. Biophys. Res. Commun.* 336(4), 1172–1180 (2005).
52. Whyteside G, Alcocer MJ, Kumita JR *et al.* Native-state stability determines the extent of degradation relative to secretion of protein variants from *Pichia pastoris*. *PLoS One* 6(7), e22692 (2011).
53. Dalken B, Jabulowsky RA, Oberoi P, Benhar I, Wels WS. Maltose-binding protein enhances secretion of recombinant human granzyme B accompanied by *in vivo* processing of a precursor MBP fusion protein. *PLoS One* 5(12), e14404 (2010).
54. Pfeffer M, Maurer M, Kollensperger G, Hann S, Graf AB, Mattanovich D. Modeling and measuring intracellular fluxes of secreted recombinant protein in *Pichia pastoris* with a novel 34S labeling procedure. *Microb. Cell Fact.* 10, 47 (2011).
55. Love KR, Politano TJ, Panagiotou V, Jiang B, Stadheim TA, Love JC. Systematic single-cell analysis of *Pichia pastoris* reveals secretory capacity limits productivity. *PLoS One* 7(6), e37915 (2012).
56. Zhang W, Zhao HL, Xue C *et al.* Enhanced secretion of heterologous proteins in *Pichia pastoris* following overexpression of *Saccharomyces cerevisiae* chaperone proteins. *Biotechnol. Prog.* 22(4), 1090–1095 (2006).
57. Hamilton S, Gerngross T. Glycosylation engineering in yeast: the advent of fully humanized yeast. *Curr. Opin. Biotechnol.* 18(5), 387–392 (2007).

▪▪ **The efforts to create a *P. pastoris* platform for the production of fully humanized glycoproteins may enable the replacement of mammalian cells with microorganisms as production hosts.**

58. De Pourcq K, De Schutter K, Callewaert N. Engineering of glycosylation in yeast and other fungi: current state and perspectives. *Appl. Microbiol. Biotechnol.* 87(5), 1617–1631 (2010).
59. Bollok M, Resina D, Valero F, Ferrer P. Recent patents on the *Pichia pastoris* expression system: expanding the toolbox for recombinant protein production. *Recent Pat. Biotechnol.* 3(3), 192–201 (2009).
60. Li P, Anumanthan A, Gao XG *et al.* Expression of recombinant proteins in *Pichia pastoris*. *Appl. Biochem. Biotechnol.* 142(2), 105–124 (2007).

61. Mattanovich D, Branduardi P, Dato L, Gasser B, Sauer M, Porro D. Recombinant protein production in yeasts. *Methods Mol. Biol.* 824, 329–358 (2012).
62. Barnett JA, Payne RW, Yarrow D. *Yeasts: Characteristics and Identification (3rd Edition).* Cambridge University Press, UK (2000).
63. Ninomiya Y, Suzuki K, Ishii C, Inoue H. Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. *Proc. Natl Acad. Sci. USA* 101(33), 12248–12253 (2004).
64. Näätsaari L, Mistlberger B, Ruth C, Hajek T, Hartner FS, Glieder A. Deletion of the *Pichia pastoris* KU70 homologue facilitates platform strain generation for gene expression and synthetic biology. *PLoS One* 7(6), e39720 (2012).
65. Critchlow SE, Jackson SP. DNA end-joining: from yeast to man. *Trends Biochem. Sci.* 23(10), 394–398 (1998).
66. Gu Y, Jin S, Gao Y, Weaver DT, Alt FW. Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding activity, and inability to support V(D)J recombination. *Proc. Natl Acad. Sci. USA* 94(15), 8076–8081 (1997).
67. Meyer V, Arentshorst M, El-Ghezal A *et al.* Highly efficient gene targeting in the *Aspergillus niger kusA* mutant. *J. Biotechnol.* 128(4), 770–775 (2007).
68. Fairhead C, Llorente B, Denis F, Soler M, Dujon B. New vectors for combinatorial deletions in yeast chromosomes and for gap-repair cloning using 'split-marker' recombination. *Yeast* 12(14), 1439–1457 (1996).
69. Heiss S, Maurer M, Hahn R, Mattanovich D, Gasser B. Identification and deletion of the major secreted protein of *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* doi:10.1007/s00253-012-4260-4264 (2012) (Epub ahead of print).
70. Graf A, Dragosits M, Gasser B, Mattanovich D. Yeast systems biotechnology for the production of heterologous proteins. *FEMS Yeast Res.* 9(3), 335–348 (2009).
71. Sauer M, Branduardi P, Gasser B *et al.* Differential gene expression in recombinant *Pichia pastoris* analysed by heterologous DNA microarray hybridisation. *Microb. Cell Fact.* 3(1), 17 (2004).
72. Graf A, Gasser B, Dragosits M *et al.* Novel insights into the unfolded protein response using *Pichia pastoris* specific DNA microarrays. *BMC Genomics* 9, 390 (2008).
73. Dragosits M, Stadlmann J, Graf A *et al.* The response to unfolded protein is involved in osmotolerance of *Pichia pastoris*. *BMC Genomics* 11, 207 (2010).
74. Baumann K, Carnicer M, Dragosits M *et al.* A multi-level study of recombinant *Pichia pastoris* in different oxygen conditions. *BMC Syst. Biol.* 4, 141 (2010).
75. Baumann K, Dato L, Graf AB *et al.* The impact of oxygen on the transcriptome of recombinant *S. cerevisiae* and *P. pastoris*: a comparative analysis. *BMC Genomics* 12(1), 218 (2011).
76. Dragosits M, Stadlmann J, Albiol J *et al.* The effect of temperature on the proteome of recombinant *Pichia pastoris*. *J. Proteome Res.* 8(3), 1380–1392 (2009).
77. Vanz AL, Lunsdorf H, Adnan A *et al.* Physiological response of *Pichia pastoris* GS115 to methanol-induced high level production of the hepatitis B surface antigen: catabolic adaptation, stress responses, and autophagic processes. *Microb. Cell Fact.* 11(1), 103 (2012).

▪ **Detailed analysis of the radical reorganization of *P. pastoris* cells upon the shift to methanol media.**

78. Pfeffer M, Maurer M, Stadlmann J *et al.* Intracellular interactome of secreted antibody Fab fragment in *Pichia pastoris* reveals its routes of secretion and degradation. *Appl. Microbiol. Biotechnol.* 93(6), 2503–2512 (2012).
79. Huang CJ, Damasceno LM, Anderson KA, Zhang S, Old LJ, Batt CA. A proteomic analysis of the *Pichia pastoris* secretome in methanol-induced cultures. *Appl. Microbiol. Biotechnol.* 90(1), 235–247 (2011).
80. Miura N, Kirino A, Endo S *et al.* Tracing putative trafficking of the glycolytic enzyme enolase via SNARE-driven unconventional secretion. *Eukaryot. Cell* 11(8), 1075–1082 (2012).
81. Blazeck J, Alper H. Systems metabolic engineering: genome-scale models and beyond. *Biotechnol. J.* 5(7), 647–659 (2010).
82. Sohn SB, Graf AB, Kim TY *et al.* Genome-scale metabolic model of methylotrophic yeast *Pichia pastoris* and its use for *in silico* analysis of heterologous protein production. *Biotechnol. J.* 5(7), 705–715 (2010).
83. Chung BK, Selvarasu S, Andrea C *et al.* Genome-scale metabolic reconstruction and *in silico* analysis of methylotrophic yeast *Pichia pastoris* for strain improvement. *Microb. Cell Fact.* 9, 50 (2010).
84. Caspeta L, Shoaie S, Agren R, Nookaew I, Nielsen J. Genome-scale metabolic reconstructions of *Pichia stipitis* and *Pichia pastoris* and *in silico* evaluation of their potentials. *BMC Syst. Biol.* 6, 24 (2012).
85. Heyland J, Fu J, Blank LM, Schmid A. Quantitative physiology of *Pichia pastoris* during glucose-limited high-cell density fed-batch cultivation for recombinant protein production. *Biotechnol.Bioeng.* 107(2), 357–368 (2010).
86. Jorda J, Jouhten P, Camara E, Maaheimo H, Albiol J, Ferrer P. Metabolic flux profiling of recombinant protein secreting *Pichia pastoris* growing on glucose:methanol mixtures. *Microb. Cell Fact.* 11, 57 (2012).
87. Carnicer M, Pierich AT, Dam JV *et al.* Quantitative metabolomics analysis of amino acid metabolism in recombinant *Pichia pastoris* under different oxygen availability conditions. *Microb. Cell Fact.* 11(1), 83 (2012).
88. Neubauer S, Haberhauer-Troyer C, Klavins K *et al.* U(13) C cell extract of *Pichia pastoris*: a powerful tool for evaluation of sample preparation in metabolomics. *J. Sep. Sci.* 35(22), 3091–3105 (2012).
89. Dortay H, Akula UM, Westphal C, Sittig M, Mueller-Roeber B. High-throughput protein expression using a combination of ligation-independent cloning (LIC) and infrared fluorescent protein (IFP) detection. *PLoS One* 6(4), e18900 (2011).
90. Rao B, Zhong X, Wang Y, Wu Q, Jiang Z, Ma L. Efficient vectors for expression cloning of large numbers of PCR fragments in *P. pastoris*. *Yeast* 27(5), 285–292 (2010).
91. Sasagawa T, Matsui M, Kobayashi Y *et al.* High-throughput recombinant gene expression systems in *Pichia pastoris* using newly developed plasmid vectors. *Plasmid* 65(1), 65–69 (2011).
92. Mizutani K, Yoshioka S, Mizutani Y, Iwata S, Mikami B. High-throughput construction of expression system using yeast *Pichia pastoris*, and its application to membrane proteins. *Protein Expr. Purif.* 77(1), 1–8 (2011).
93. Boettner M, Prinz B, Holz C, Stahl U, Lang C. High-throughput screening for expression of heterologous proteins in the yeast *Pichia pastoris*. *J. Biotechnol.* 99(1), 51–62 (2002).
94. 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.* 5(2), 179–189 (2004).
95. Barnard GC, Kull AR, Sharkey NS *et al.* High-throughput screening and selection of yeast cell lines expressing monoclonal antibodies. *J. Ind. Microbiol. Biotechnol.* 37(9), 961–971 (2010).
96. Jiang Y, Li F, Button M *et al.* A high-throughput purification of monoclonal antibodies from glycoengineered *Pichia*

pastoris. *Protein Expr. Purif.* 74(1), 9–15 (2010).

97. Panagiotou V, Love KR, Jiang B, Nett J, Stadheim T, Love JC. Generation and screening of *Pichia pastoris* strains with enhanced protein production by use of microengraving. *Appl. Environ. Microbiol.* 77(9), 3154–3156 (2011).
98. Boettner M, Steffens C, von Mering C, Bork P, Stahl U, Lang C. Sequence-based factors influencing the expression of heterologous genes in the yeast *Pichia pastoris*: a comparative view on 79 human genes. *J. Biotechnol.* 130(1), 1–10 (2007).
99. Schroder LA, Glick BS, Dunn WA. Identification of pexophagy genes by restriction enzyme-mediated integration. *Methods Mol. Biol.* 389, 203–218 (2007).
100. Nett JH, Stadheim TA, Li H *et al.* A combinatorial genetic library approach to target heterologous glycosylation enzymes to the endoplasmic reticulum or the Golgi apparatus of *Pichia pastoris*. *Yeast* 28(3), 237–252 (2011).
101. Tanaka T, Yamada R, Ogino C, Kondo A. Recent developments in yeast cell surface display toward extended applications in biotechnology. *Appl. Microbiol. Biotechnol.* 95(3), 577–591 (2012).
102. Mattanovich D, Borth N. Applications of cell sorting in biotechnology. *Microb. Cell Fact.* 5, 12 (2006).
103. Diehl B, Hoffmann TM, Mueller NC, Burkhart JL, Kazmaier U, Schmitt MJ. Novel yeast bioassay for high-throughput screening of matrix metalloproteinase inhibitors. *Appl. Environ. Microbiol.* 77(24), 8573–8577 (2011).
104. Hu CX, Huang H, Zhang L *et al.* A new screening method based on yeast-expressed human dipeptidyl peptidase IV and discovery of novel inhibitors. *Biotechnol. Lett.* 31(7), 979–984 (2009).
105. Till A, Lakhani R, Burnett SF, Subramani S. Pexophagy: the selective degradation of peroxisomes. *Int. J. Cell Biol.* 2012, 512721 (2012).
106. Gould SJ, McCollum D, Spong AP, Heyman JA, Subramani S. Development of the yeast *Pichia pastoris* as a model organism for a genetic and molecular analysis of peroxisome assembly. *Yeast* 8(8), 613–628 (1992).
107. van der Klei I, Veenhuis M. Protein targeting to yeast peroxisomes. *Methods Mol. Biol.* 390, 373–391 (2007).
108. Dunn WA Jr, Cregg JM, Kiel JA *et al.* Pexophagy: the selective autophagy of peroxisomes. *Autophagy* 1(2), 75–83 (2005).
109. Manjithaya R, Anjard C, Loomis WF, Subramani S. Unconventional secretion of *Pichia pastoris* Acb1 is dependent on GRASP protein, peroxisomal functions, and autophagosome formation. *J. Cell Biol.* 188(4), 537–546 (2010).
110. Rossanese O, Soderholm J, Bevis B *et al.* Golgi structure correlates with transitional endoplasmic reticulum organization in *Pichia pastoris* and *Saccharomyces cerevisiae*. *J. Cell Biol.* 145(1), 69–81 (1999).
111. Mogelsvang S, Gomez-Ospina N, Soderholm J, Glick BS, Staehelin LA. Tomographic evidence for continuous turnover of Golgi cisternae in *Pichia pastoris*. *Mol. Biol. Cell* 14(6), 2277–2291 (2003).
112. Losev E, Reinke CA, Jellen J, Strongin DE, Bevis BJ, Glick BS. Golgi maturation visualized in living yeast. *Nature* 441(7096), 1002–1006 (2006).

▪▪ **Live cell imaging highlights that Golgi cisternae mature from early to late Golgi states, thus transporting secretory proteins.**

113. Glick BS, Nakano A. Membrane traffic within the Golgi apparatus. *Annu. Rev. Cell Dev. Biol.* 25, 113–132 (2009).
114. Montegna EA, Bhave M, Liu Y, Bhattacharyya D, Glick BS. Sec12 binds to Sec16 at transitional ER sites. *PLoS One* 7(2), e31156 (2012).
115. Esaki M, Liu Y, Glick B. The budding yeast *Pichia pastoris* has a novel Sec23p homolog. *FEBS Lett.* 580(22), 5215–5221 (2006).
116. Delic M, Rebnegger C, Wanka F *et al.* Oxidative protein folding and unfolded protein response elicit differing redox regulation in endoplasmic reticulum and cytosol of yeast. *Free Radic. Biol. Med.* 52(9), 2000–2012 (2012).

▪▪ **The unfolded protein response and unstressed oxidative protein folding follow different strategies, either to reduce or oxidize endoplasmic reticulum-resident proteins. Both processes are accompanied by a reduction of the cytosol.**

117. Yoshida H. ER stress and diseases. *FEBS J.* 274(3), 630–658 (2007).
118. Naidoo N. ER and aging-protein folding and the ER stress response. *Ageing Res. Rev.* 8(3), 150–159 (2009).
119. Treiber C, Simons A, Strauss M *et al.* Clioquinol mediates copper uptake and counteracts copper efflux activities of the amyloid precursor protein of Alzheimer's disease. *J. Biol. Chem.* 279(50), 51958–51964 (2004).
120. Treiber C, Simons A, Multhaup G. Effect of copper and manganese on the *de novo* generation of protease-resistant prion protein in yeast cells. *Biochemistry* 45(21), 6674–6680 (2006).
121. Shekhar C. *Pichia* power: India's biotech industry puts unconventional yeast to work. *Chem. Biol.* 15(3), 201–202 (2008).
122. Tolner B, Smith L, Hillyer T *et al.* From laboratory to Phase I/II cancer trials with recombinant biotherapeutics. *Eur. J. Cancer* 43(17), 2515–2522 (2007).
123. Anders RF, Adda CG, Foley M, Norton RS. Recombinant protein vaccines against the asexual blood stages of *Plasmodium falciparum*. *Hum. Vaccin.* 6(1), 39–53 (2010).
124. Meyer HP, Brass J, Jungo C, Klein J, Wenger J, Mommers R. An emerging star for therapeutic and catalytic protein. *Bioprocess Int.* 6(S6), 10–21 (2008).
125. Ghosalkar A, Sahai V, Srivastava A. Secretory expression of interferon-alpha 2b in recombinant *Pichia pastoris* using three different secretion signals. *Protein Expr. Purif.* 60(2), 103–109 (2008).
126. Schaefer JV, Pluckthun A. Engineering aggregation resistance in IgG by two independent mechanisms: lessons from comparison of *Pichia pastoris* and mammalian cell expression. *J. Mol. Biol.* 417(4), 309–335 (2012).
127. Guo JP, Ma Y. High-level expression, purification and characterization of recombinant *Aspergillus oryzae* alkaline protease in *Pichia pastoris*. *Protein Expr. Purif.* 58(2), 301–308 (2008).
128. Tanaka H, Okuno T, Moriyama S, Muguruma M, Ohta K. Acidophilic xylanase from *Aureobasidium pullulans*: efficient expression and secretion in *Pichia pastoris* and mutational analysis. *J. Biosci. Bioeng.* 98(5), 338–343 (2004).
129. Kottmeier K, Ostermann K, Bley T, Rodel G. Hydrophobin signal sequence mediates efficient secretion of recombinant proteins in *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* 91(1), 133–141 (2011).
130. Eiden-Plach A, Zagorc T, Heintel T, Carius Y, Breinig F, Schmitt MJ. Viral preprotoxin signal sequence allows efficient secretion of green fluorescent protein by *Candida glabrata*, *Pichia pastoris*, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe*. *Appl. Environ. Microbiol.* 70(2), 961–966 (2004).
131. Khasa YP, Conrad S, Sengul M, Plautz S, Meagher MM, Inan M. Isolation of *Pichia pastoris PIR* genes and their utilization for cell surface display and recombinant protein secretion. *Yeast* 28(3), 213–226 (2011).
132. Jungo C, Schenk J, Pasquier M, Marison IW, von Stockar U. A quantitative analysis of the benefits of mixed feeds of sorbitol and

methanol for the production of recombinant avidin with *Pichia pastoris*. *J. Biotechnol.* 131(1), 57–66 (2007).

133. Resina D, Cos O, Ferrer P, Valero F. Developing high cell density fed-batch cultivation strategies for heterologous protein production in *Pichia pastoris* using the nitrogen source-regulated FLD1 promoter. *Biotechnol. Bioeng.* 91(6), 760–767 (2005).

134. Carnicer M, Baumann K, Toplitz I *et al.* Macromolecular and elemental composition analysis and extracellular metabolite balances of *Pichia pastoris* growing at different oxygen levels. *Microb. Cell Fact.* 8(1), 65 (2009).

135. Pichia *Protocols (Methods in Molecular Biology)*. Cregg JM (Ed.). Humana Press, NY, USA (2007).

136. Delic M, Mattanovich D, Gasser B. Monitoring intracellular redox conditions in the endoplasmic reticulum of living yeasts. *FEMS Microbiol. Lett.* 306(1), 61–66 (2010).

137. Zhang P, Zhang W, Zhou X, Bai P, Cregg JM, Zhang Y. Catabolite repression of Aox in *Pichia pastoris* is dependent on hexose transporter PpHxt1 and pexophagy. *Appl. Environ. Microbiol.* 76(18), 6108–6118 (2010).

Prielhofer et al. BMC Genomics (2015) 16:167
DOI 10.1186/s12864-015-1393-8

BMC Genomics

RESEARCH ARTICLE Open Access

Pichia pastoris regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level

Roland Prielhofer[1,2], Stephanie P Cartwright[3], Alexandra B Graf[2,4], Minoska Valli[1,2], Roslyn M Bill[3*], Diethard Mattanovich[1,2*] and Brigitte Gasser[1,2]

Abstract

Background: The methylotrophic, Crabtree-negative yeast *Pichia pastoris* is widely used as a heterologous protein production host. Strong inducible promoters derived from methanol utilization genes or constitutive glycolytic promoters are typically used to drive gene expression. Notably, genes involved in methanol utilization are not only repressed by the presence of glucose, but also by glycerol. This unusual regulatory behavior prompted us to study the regulation of carbon substrate utilization in different bioprocess conditions on a genome wide scale.

Results: We performed microarray analysis on the total mRNA population as well as mRNA that had been fractionated according to ribosome occupancy. Translationally quiescent mRNAs were defined as being associated with single ribosomes (monosomes) and highly-translated mRNAs with multiple ribosomes (polysomes). We found that despite their lower growth rates, global translation was most active in methanol-grown *P. pastoris* cells, followed by excess glycerol- or glucose-grown cells. Transcript-specific translational responses were found to be minimal, while extensive transcriptional regulation was observed for cells grown on different carbon sources. Due to their respiratory metabolism, cells grown in excess glucose or glycerol had very similar expression profiles. Genes subject to glucose repression were mainly involved in the metabolism of alternative carbon sources including the control of glycerol uptake and metabolism. Peroxisomal and methanol utilization genes were confirmed to be subject to carbon substrate repression in excess glucose or glycerol, but were found to be strongly de-repressed in limiting glucose-conditions (as are often applied in fed batch cultivations) in addition to induction by methanol.

Conclusions: *P. pastoris* cells grown in excess glycerol or glucose have similar transcript profiles in contrast to *S. cerevisiae* cells, in which the transcriptional response to these carbon sources is very different. The main response to different growth conditions in *P. pastoris* is transcriptional; translational regulation was not transcript-specific. The high proportion of mRNAs associated with polysomes in methanol-grown cells is a major finding of this study; it reveals that high productivity during methanol induction is directly linked to the growth condition and not only to promoter strength.

Keywords: *Pichia pastoris*, Methylotrophic yeast, Crabtree-negative yeast, Polysome profiling, Microarray analysis, Transcriptome, Glucose repression, Carbon substrate repression, Methanol induction

* Correspondence: R.M.Bill@aston.ac.uk; diethard.mattanovich@boku.ac.at
[3]School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, UK
[1]Department of Biotechnology, BOKU University of Natural Resources and Life Sciences Vienna, Muthgasse 18, 1190 Vienna, Austria
Full list of author information is available at the end of the article

Background

Pichia pastoris (syn. *Komagataella* sp.) is a methylotrophic yeast that is widely used for the production of heterologous proteins and metabolites; it is also used as a model organism for the study of peroxisome biosynthesis and degradation, as well as for the analysis of protein secretion (see [1], and references therein). Its ability to use methanol as a carbon and energy source, its non-fermentative utilization of glucose and its efficient growth on glycerol are key metabolic features that make it attractive for bioprocess development.

Recently, Liang et al. [2] comprehensively annotated the *P. pastoris* transcriptome and identified novel untranslated regions (UTR), alternative splicing sites (AS), internal ribosome entry sites (IRES), upstream ATGs (uATGs) and upstream ORFs (uORFs). Transcriptional profiling of a recombinant strain harboring *Rhizomucor miehei* lipase (RML) under the control of the methanol-driven P_{AOX1} promoter revealed that cells grown on methanol induce genes involved in protein production and energy metabolism more than cells grown on glycerol. Methanol utilization takes place in peroxisomes; genes such as the alcohol oxidases (*AOX1, AOX2*), formaldehyde dehydrogenase (*FLD*), dihydroxyacetone synthase (*DAS1, DAS2*) and peroxisomal genes (e. g. *PEX1*) were all found to be induced on methanol.

The specific growth rate of a culture, which was kept constant in the study by Liang et al. [2], is also known to play a fundamental role in gene regulation and consequently in protein production. High growth rates were previously suggested to be beneficial for protein production in *P. pastoris* due to the up-regulation of genes related to gene expression and translation, while catabolic processes (e.g. autophagy, transport to the peroxisome and mitochondrial degradation, many of them under the control of *TOR* signalling), were shown to correlate negatively with increasing growth rate [3].

Less is known about the specific regulation of carbon substrate utilization, with the notable exception of *Saccharomyces cerevisiae*. Most studies in *S. cerevisiae* have been performed on glucose-grown cells under respiro-fermentative or fermentative growth conditions [4] or on non-fermentable carbon-sources such as glycerol or galactose. The shift from glucose to glycerol leads to extensive transcriptomic remodelling [5], a global translational down-regulation [6] and reduced growth rates. In contrast, the Crabtree-negative yeast, *P. pastoris*, maintains its respiratory metabolism even under conditions of excess glucose (such as that used in batch cultivations) and exhibits similar growth rates and substrate uptake kinetics when grown on either glucose or glycerol [7]. Shifts from glycerol to methanol, which is metabolized even more slowly with lower maximal specific growth rates, are often used in bioprocesses that employ *P. pastoris*.

Transcriptional regulators involved in glucose repression have been identified and studied in the methylotrophic yeasts *P. pastoris* and *Hansenula polymorpha*, and in the lactose-utilizing yeast *Kluyveromyces lactis* [8-13]. Glucose repression of methanol utilization genes is established as a feature of methylotrophic yeasts such as *Candida boidinii*, *H. polymorpha*, *Pichia methanolica*, and *P. pastoris* [14], but the degree of repression/de-repression by different carbon sources is species-dependent. For example, different modes of regulation have been described for key enzymes of methanol metabolism pathways such as alcohol oxidase, dihydroxyacetone synthase and formaldehyde dehydrogenase (summarized in [14,15]). Understanding the molecular mechanisms underpinning the unique carbon substrate utilization properties of *P. pastoris* is now required in order to more fully understand this valuable host organism.

The regulation of gene expression is often analyzed at the level of transcription, although it is well established that altered transcript levels are not necessarily reflected by the corresponding protein levels [16]. For example, the protein level of more than 70% of *S. cerevisiae* protein-coding genes is transcriptionally regulated, but this drops to only about 50% in *E. coli* [17] and is even lower in humans [18]. In order to obtain a more complete view of the regulation of gene expression in *P. pastoris*, we analyzed both transcriptional and translational responses of cells grown in glucose-, glycerol- or methanol-containing media. Microarray analysis was done on the total mRNA pool as well as on mRNAs that had been fractionated based upon ribosome occupancy. We adapted published methods for polysome profiling [6,19]: translationally quiescent mRNAs were defined as being associated with single ribosomes (monosomes); actively-translated mRNAs with multiple ribosomes (polysomes) [20]. The hybridization of a microarray with these mRNA fractions as well as the total mRNA population provided insight into how efficiently individual mRNA translation and global transcriptional responses are affected by carbon source utilization.

Results and discussion

P. pastoris strain X-33 was cultivated in shake flasks under four different bioprocess conditions (Table 1): excess glycerol or glucose (batch culture conditions; these cells were harvested during exponential growth); limiting glucose (using slow glucose-releasing silica disks or feed beads in fed-batch mode, [21,22]); and periodic methanol addition (methanol induction conditions). Cells grown in excess glucose or glycerol or those grown in methanol had growth rates close to μ_{max}: 0.23 h^{-1} for the former and 0.1 h^{-1} for the latter conditions. Cells in limiting glucose conditions grew at $\mu = 0.015\ h^{-1}$.

For polysome fractionation, cells were treated with cycloheximide, harvested and quickly chilled for sample preparation. Isolates were used for polysome profiling to

Table 1 *Pichia pastoris* cultivations in buffered synthetic media supplemented with different carbon substrates

Condition	ID	Start-OD_{600}	Cultivation substrate	Cultivation time [h]	Harvest-OD_{600}	μ [h^{-1}]	Bioprocess Step	Replicates
Excess glucose	D	0.1	2% glucose	23.3	10.0 (1.0)	0.23 (0.004)	Glucose batch	3
Excess glycerol	G	0.1	2% glycerol	23.3	10.5 (1.3)	0.23 (0.001)	Glycerol batch	3
Methanol feed	M	1.5	0.5 and 0.6% methanol	24.5	8.6 (1.4)	0.10 (0.008)	Methanol shot/feed	3
Limiting glucose	X	1.5	0.25% glucose and feed beads	16.8	11.4 (0.6)	0.010 - 0.022	Glucose fed batch	3

Cultures with different biomass densities were fed with appropriate amounts of carbon substrate in order that the cells could be harvested at a similar OD_{600} [mean (sd)]. Growth rates (μ) [mean (sd)] were recorded; the values were highly reproducible and reflect growth of typical bioprocess phases, as shown.

obtain the profile data and to collect mono- and polysome fraction samples for mRNA extraction. mRNA was isolated from the fractionated and unfractionated isolates for microarray analysis; for each condition three biological replicates were analyzed.

The excess glucose condition, which is often used as a control for studies in *S. cerevisae*, was used as a control in our experiments.

Global transcript profiles are very similar for excess glucose or glycerol grown *P. pastoris* cells, while extensive transcriptional regulation is observed for cells grown on methanol or limiting glucose concentrations

Differentially expressed genes were identified from fold changes between total RNA samples (i.e. those from unfractionated isolates). Samples from the excess glucose condition were the control for all these experiments (cut-off criteria ±50% fold change and adjusted p-values < 0.05; [23]). Transcriptional fold changes for all genes are listed in Additional file 1: Table S1. The data in Figure 1 show that cells cultured in excess glycerol (G) or glucose (D) have a very similar transcriptome with just 265 genes differentially regulated; in contrast 817 genes are differentially regulated in methanol-grown cells (M) and 2,822 are differentially regulated in glucose-limited cells (X) (Figure 1A). The corresponding Gene Ontology (GO) terms are listed in Additional file 2. A high correlation between the two excess carbon source condition transcriptomes (G and D) was also observed by principal component analysis (PCA), which showed a good correlation of the biological replicates of each condition (Figure 2). The methanol-grown and glucose-limited cells were also found to share many differentially-regulated genes and hence seem to be more similar to each other than to the two excess conditions (Figure 1B).

Further analysis (Figure 1B, C) revealed that only a small sub-set of genes are differently expressed in response to glycerol as carbon source (10% of the 148 up-regulated and 15% of the 114 down-regulated genes), while most of the regulated genes are shared either with both (56%) or at least one (approx. 30%) of the two other conditions (methanol induction or limiting glucose). We defined genes that are differentially regulated in excess glycerol conditions plus at least one other condition (either methanol induction or limiting glucose) to be subject to "glucose repression". Genes that are differentially regulated in

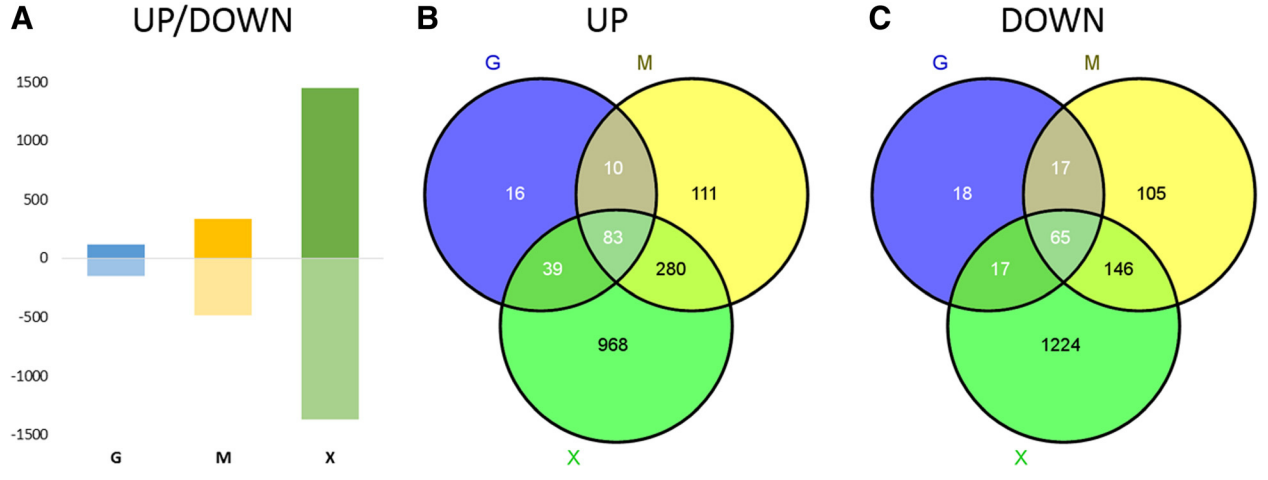

Figure 1 Differentially expressed genes. The bar chart **(A)** shows the number of differentially expressed genes in excess glycerol (G), methanol (M) and limiting glucose (X) compared to the excess glucose condition. Venn diagrams illustrate the number of up-regulated **(B)** and down-regulated genes **(C)** in the conditions and intersections. Significantly-regulated genes were identified from total RNA fold changes compared to the excess glucose condition (cutoff ±50% fold change and adjusted p-values < 0.05; [23]).

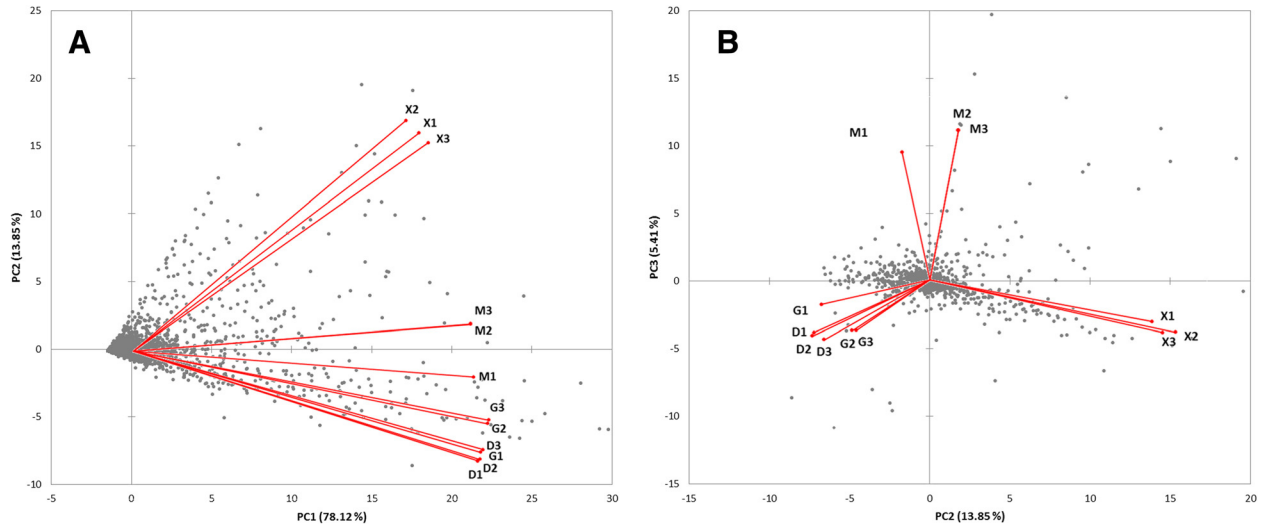

Figure 2 Principal component analysis (PCA) bi-plots of microarray intensities from the green channel. Red vectors indicate variable (condition) correlation of all analyzed replicates and the grey data points indicate observations (genes). Replicate correlation fits very well already before data normalization. The components one and two **(A)** and two and three **(B)** are compared, which explain 78, 14 and 5% of the total variation, respectively. Similar PCA biplots are obtained from microarray intensities of the red channel.

response to methanol induction or limiting glucose conditions, but are not differentially regulated between the two excess conditions were defined as being subject to "carbon substrate repression".

Polysome-mRNA association is lowest in glucose-limited cells and highest in methanol-grown cells

Isolates of cells subject to the different growth conditions in Table 1 were analyzed by polysome profiling, which characterizes the translational status of a cell according to the distribution of ribosomes across the mRNA pool. Profile curves showing the proportion of ribosomes that appear as individual sub-units (40S and 60S), monosomes or polysomes (where two or more ribosomes are associated with a given mRNA transcript) are shown in Figure 3. The ratios of the polysome to monosome peak areas (P:M ratios) in the profiles (Figure 3A) are presented in Figure 3B: mRNAs that are associated with polysomes are more highly-translated than mRNAs associated with monosomes [20]. The P:M ratio is therefore established as a relative measure of translational activity at a cellular level [24,25]. In our experiments, triplicate cultures gave reproducible values for each of the different growth conditions.

Due to their similar transcript profiles, the two fastest growing conditions (excess glycerol and excess glucose, $\mu \sim 0.23\ h^{-1}$) were anticipated to have similar P:M ratios. However, the excess glycerol condition had a higher P:M ratio (Figure 3) suggesting higher translational activity compared to cells grown under conditions of excess glucose. The P:M ratio was highest in cells grown on methanol, although the specific growth rate was significantly lower ($\mu \sim 0.10\ h^{-1}$) compared to the excess glycerol and excess glucose conditions. The condition with the lowest specific growth rate (limiting glucose, $\mu \sim 0.015\ h^{-1}$) had the lowest P:M ratio.

The transcription of translation-related genes in *P. pastoris* was previously shown to be tightly connected to growth rate in glucose-limited chemostat cultivations [3]. We found that this was also true when we analyzed the total RNA of unfractionated, slow-growing cells cultivated under limiting glucose conditions ($\mu \sim 0.015\ h^{-1}$). Under these conditions, most ribosomal and translation-related genes were found to be expressed at a lower level (Additional file 1: Table S2). Strikingly, we found that those genes were equally expressed in slow-growing methanol fed cells ($\mu \sim 0.1\ h^{-1}$) compared to excess glucose and glycerol ($\mu \sim 0.23\ h^{-1}$), suggesting that the whole translation machinery is up-regulated despite the slow growth rate on methanol. The methanol induction-, excess glucose- and excess glycerol- conditions operated near μ_{max} for their respective condition, which means that they possess a similar μ/μ_{max} ratio. Hence, the expression of growth-associated genes might respond to the ratio of μ/μ_{max}, rather than an absolute value of the specific growth rate (μ).

Despite the general transcriptional down-regulation of translation-related genes in *P. pastoris* cells grown in limiting glucose, the transcription of certain genes is induced

Certain genes required for ribosome biogenesis and its regulation, RNA processing and translationally silent messenger ribonucleoprotein complexes (mRNPs) were highly expressed in *P. pastoris* cells grown in limiting glucose, as

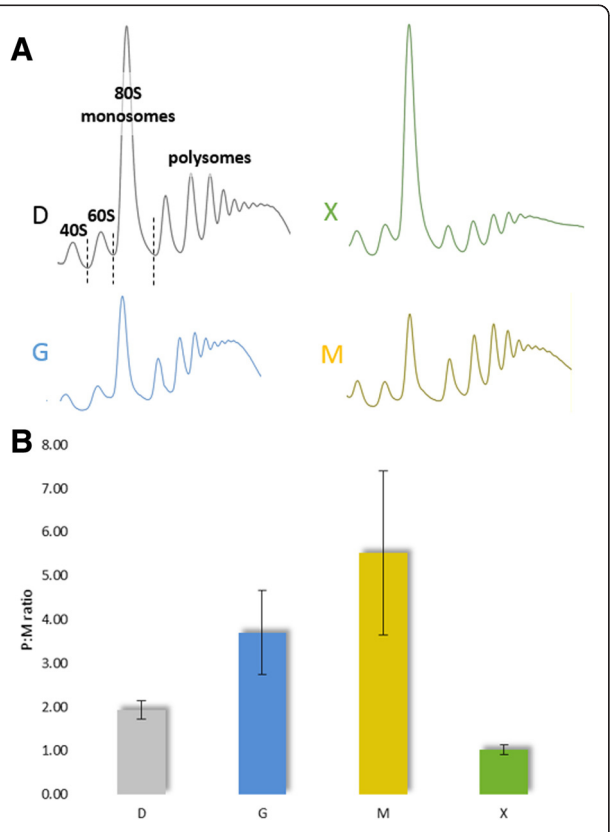

Figure 3 Polysome profiles and P:M ratios for *P. pastoris* grown in different conditions. (A) Representative polysome profiles and **(B)** a bar chart presenting P:M ratios (with sd) of the four different cultivation conditions (excess glucose, D; excess glycerol, G; limiting glucose, X; methanol, M). Corresponding peaks (40S, 60S, 80S/monosomes and polysomes) are indicated in the first (D) polysome profile. P:M ratios were calculated from areas beneath the profile curve using ImageJ.

determined by the analysis of total mRNA (Additional file 1: Table S2): *RPS22A* (protein component of the small (40S) ribosomal subunit, homologous to mammalian ribosomal protein S15A and bacterial S8, also up-regulated in methanol-fed cells); genes linked to ribosome association, interaction or biogenesis (*TMA108, DOT6, GDE1, TMA64, PAS_FragB_0030, YMR295C, MTC1, YOR019W, MTG1*); negative regulation of RNA polymerase III transcription and TOR signaling (*KNS1*); RRPE (ribosomal RNA processing element)-binding and glucose-induced transition from quiescence to growth (*STB3*); rRNA biogenesis (*DOT6*) and mitochondrial ribosome recycling (*RRF1*). Poly(A)-binding protein is also translation-associated, and the two genes are differently expressed (PAS_chr1-4_0283 is up- and PAB1 is down-regulated) in *P. pastoris* cells grown in limiting glucose. The gene encoding the translational activator *GIS2* that was also up-regulated in limiting glucose, plays an important role as activator of mRNAs with internal ribosome entry sites [26]. It binds to a specific subset of mRNAs, associates with polysomes and localizes to RNA processing bodies (P bodies) and to stress granules. The role of cap-independent translation in physiological adaptation to stress in *S. cerevisiae* has been reported previously [27]. P bodies are used to store translationally silent mRNPs [28], and glucose-limited *P. pastoris* cells were found to differentially express related genes. *DHH1* (the gene product of which functions in decapping and translational repression) was up-regulated, but *PAT1* and *EDC3*, with a similar function, were down-regulated in glucose-limited cells. Hence, although limiting glucose decreases global translation, certain transcripts may be translated as a part of specific stress responses.

Growth conditions have a minimal influence on transcript-specific translational regulation

We next examined the fractionated mRNAs by microarray analysis. We normalized the abundance of each transcript in the polysome fraction to that of the total RNA, which we termed the "translational state". In order to confirm the integrity of the RNA fractions, microarray signal intensities of the monosome, polysome and total RNA samples from the limiting glucose condition were compared as previously described [29]. The $\log_{10}$ intensity values of total RNA correlated with $\log_{10}$ of the sums of intensities in the monosome- and polysome-bound mRNA with a correlation coefficient of $R^2 = 0.963$ (see Additional file 3). Translational states of individual transcripts for the excess glycerol, limiting glucose and methanol induction conditions were normalized to the excess glucose condition in order to identify transcripts with changed translational states (shown in Figure 4 and Additional file 4). This identified an increased or decreased abundance of transcripts that are actively translated in the polysome fraction. Translational states of individual genes ranged from 0.08-fold (in limiting glucose conditions) to 3.05-fold (in methanol). No transcripts were totally excluded from the polysome fractions, which is in agreement with a study published by Arava et al. [30].

Only 16 transcripts had different translational states (8 increased and 8 decreased) in response to excess glycerol compared to the excess glucose condition, while more differences were found for the glucose-limited and methanol-grown cells. In excess glycerol-grown cells, *RPL2A, TEF2, RPS4B, ENO1, FBA1-1, RPL5, RPL11B* and *TDH3* had decreased translational states compared to cells grown in excess glucose. These genes are annotated with GO terms "biosynthetic/metabolic process" and "translation". Both, the transcript level and the translational state was found to be decreased for transcripts of the glycolytic fructose 1,6-bisphosphate aldolase (*FBA1-1*), glyceraldehyde-3-phosphate dehydrogenase (*TDH3*) and phosphopyruvate hydratase (*ENO1*) in excess glycerol.

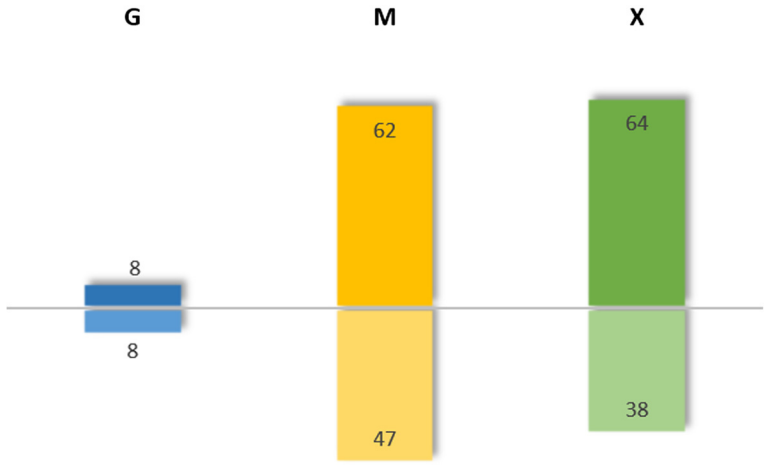

Figure 4 Translationally-enriched and depleted genes. Bar chart representing the number of translationally enriched and depleted genes in excess glycerol, limiting glucose and methanol conditions related to the excess glucose condition (cutoff ±50% change of the translational state and adjusted p-values < 0.05).

This suggests that specific translational down-regulation reinforces the transcriptional down-regulation of these genes in response to excess glycerol.

In methanol-grown cells, genes required for methanol utilization (MUT), were strongly up-regulated at the transcriptional level, but had a decreased translational state compared to excess glucose. Hence translational regulation appears to counteract the strong transcriptional up-regulation of most of these genes. Such "post-transcriptional buffering" has also been observed in two *Saccharomyces* species [31]. Neither significantly enriched GO terms nor other patterns could be found in the other gene groups with altered translational states.

Translational states are linked to ORF length and transcript abundance

We analyzed the translation states of individual transcripts compared to total mRNA for all growth conditions. Enriched gene groups were initially identified (Table 2); closer inspection revealed that the groups had closely correlated open reading frame (ORF) lengths, which has been reported previously for other organisms [32-34]. Liang et al. [2] identified *P. pastoris* gene ORFs, uORFs, UTRs and introns by sequencing, and found ORF lengths from 141 to 14853 bp, with an average of 1444 ± sd = 1032 bp (median of 1203 bp). We used this information to define three gene groups according to ORF length (Table 3): long and short genes, comprising the upper and lower quartile of all genes, and the remaining 50% of medium-length genes. Translation efficiency is also known to be affected by codon usage, so we included synonymous codon usage order (SCUO), which was obtained from the CodonO platform [35]; higher values indicate more codon bias, meaning less random codon use in a gene's coding region. The three gene groups significantly differ in transcript level, translational states, codon usage bias (SCUO) and 5´UTR frequency: Short genes are highly transcribed (as measured by transcript abundance) and translated (high translational states), rarely possess a 5´UTR and have an enhanced codon usage bias (Table 3).

Statistical tests (Fishers exact test, chi square test and regression analysis) were used to verify these relationships.

Table 2 Translational regulation of functional gene groups for *P. pastoris* cells grown in excess glucose conditions

Functional group	Genes in group	Significantly regulated genes	Average translational $\log_2$ ratio of significantly regulated genes	Average ORF length of significantly regulated genes [bp]
Secretion: chaperones	79	31	0.225	885
Antioxidant	21	7	0.160	476
Transport(er)	60	22	0.137	1669
Pexophagy	23	9	−0.082	2302
Autophagy	69	25	−0.117	1690
Vacuole	105	48	−0.151	1781
Mitochondria	110	23	−0.165	1541
TCA	20	10	−0.339	1544
Secretion: glycosylation	46	28	−0.344	1884

Average translational states and ORF length of functional gene groups for *P. pastoris* cells grown in excess glucose. Translational trends were similar in the other conditions.

Table 3 *P. pastoris* gene statistics of long, medium and short genes

	Long	Medium	Short	All
Number of genes	1262	2538	1265	5065
ORF length [bp]	>1807	770-1807	<770	141-14853
Mean ORF length [bp]	2786	1235	524	1444
Median ORF length [bp]	2412	1206	540	1203
Mean expression intensity	5081	7141	12092	7864
Median expression intensity	2600	2591	3416	2736
Mean SCUO	0.078	0.105	0.198	0.123
Median SCUO	0.069	0.093	0.165	0.096
Genes with 5′UTR	628	257	29	914
Genes with 5′UTR [%]	50%	10%	2%	18%
5′UTR length mean	238	253	320	245
Mean translational state	−0.22	−0.01	0.18	−0.02

Based on the information published by Liang et al. [2], all *P. pastoris* genes were split into 3 groups comprising the 25% longest (>1807 bp), the 25% shortest (<770 bp) and the remaining (50%, <1807 and >770 bp) medium length genes. Gene groups are not exactly the same size because they were split by length cut-off (some genes possess equal ORF lengths). 5′UTR information was also taken from Liang et al. [2]. Expression intensities were obtained from our total RNA microarray data which were normalized as described in the Methods section. Synonymous codon usage order (SCUO) was obtained from the CodonO platform [35].

ORF length was shown to have a negative correlation with transcript abundance (gene expression intensity by microarray) and codon usage bias, so short genes are more highly transcribed than longer ones (regression analysis, p-value < $1.5e^{-11}$) and more codon biased (non-linear regression, p-value < $2.2e^{-16}$). The correlation of ORF length with translational states and 5′UTR length was found to be significantly positive (p-value < $2.2e^{-16}$ for both). Hence, short genes are more-highly translated and rarely have a 5′UTR, while longer genes are less-highly translated and often possess a 5′UTR (Figure 5).

Transcriptional regulation responding to different carbon sources correlates with expression of corresponding transcription factors

As mentioned above, excess glucose was used as a calibrator to calculate the transcriptional regulation in the other conditions (see Additional file 1: Table S1 for respective values for all genes). Concerning global transcriptional control systems, we could identify *P. pastoris* gene expression responding to glucose repression, carbon catabolite repression elicited by excess glucose and glycerol, as well as control by methanol availability. Limiting glucose triggers extensive transcriptional responses due to carbon limitation and low growth rate, which correlate well with the regulation patterns described by Rebnegger et al. [3] recently. Corresponding to the important role of glycogen metabolism in slow growing conditions [36], we found genes encoding glycogen synthase (*GSY2*), phosphoglucomutase (*PGM2*) and other glycogen metabolism genes (*UGP1, NTH1, ATH1, GLG1, GLC3, GLC7*) up-regulated in limiting glucose.

Glucose repression signalling is mainly mediated through the central kinase Snf1, which controls the expression of important transcription factors such as Mig1, Sip4, Rds2, Cat8 and Adr1 [37], thereby playing an important role in the utilization of non-fermentable carbon sources in *S. cerevisiae* [38]. We found the transcripts of many genes involved in catabolite (de)repression to be induced in limiting glucose, especially *CAT8-2*, which is about 39-fold up-regulated compared to excess glucose (and about 7-fold up-regulated on methanol). In addition, almost all genes that are reported to be controlled by *CAT8* in *S. cerevisiae* [39] are also up-regulated.

Interestingly, 2 homologs of Mig1 are found in the *P. pastoris* genome, one of which is about 9-fold up-regulated in response to methanol and limiting glucose (*MIG1-1*), while the second one is down-regulated on all other tested carbon sources compared to glucose (*MIG1-2*);

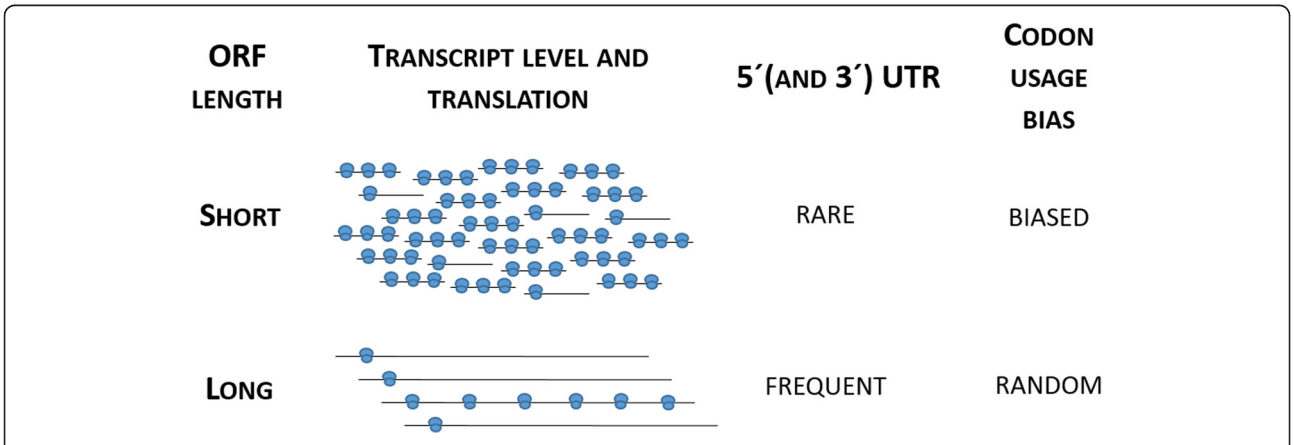

Figure 5 Schematic illustration of relations between transcript level, translation, UTR frequency and codon usage bias in *P. pastoris* genes. In contrast to genes with long coding sequences, shorter genes are more highly expressed, more efficiently translated, possess UTR's less frequently and are more codon biased than longer genes.

it is possible that it acts as a carbon catabolite or glucose repressor similar to *CRE1* in *Trichoderma reesei* [40] or *CREA* in *Aspergillus nidulans* [41].

The homologue of *S. cerevisiae Activator of Ferrous Transport*, *AFT1,* was found to have induced expression levels in excess glycerol, methanol and limiting glucose conditions and has been reported to play a role in the regulation of carbon repressed genes in *P. pastoris* recently [42]. The transcription factors PAS_chr4_0324, CTH1, PAS_chr1-1_0422, PAS_chr3_1209, PAS_chr1-1_0122 were related to excess conditions.

Among the most strongly-induced genes in methanol and limiting glucose conditions, several transcription factors are present (Table 4). Of these, the Zn(II)2Cys6 zinc cluster protein PAS_chr3_0836, which has an 80-fold higher transcript level on methanol and 120-fold higher transcript level under limiting glucose compared to excess glucose, has significant sequence homology to *H. polymorpha MPP1* [43]. Mpp1 was suggested to be the master regulator of methanol-responsive genes in *H. polymorpha* [43,44]. Since PAS_chr3_0836 is also located in a similar chromosomal arrangement (next to *DAS1/2*; PAS_chr3_0832 and PAS_chr3_0834) to *H. polymorpha*, we propose that it is the *P. pastoris* homologue of *HpMPP1*. Pp*MXR1* encoding a transcription factor that is necessary for the activation of many genes in response to methanol [8] is induced in all three conditions compared to excess glucose. We suggest that *PpMXR1*, similar to its *S. cerevisiae* homolog *ADR1,* is needed for the activation (de-repression) of genes for alternative carbon sources including the MUT genes that are repressed in the presence of excess glucose and glycerol, but that Mpp1 is the transcriptional activator of peroxisomal import and matrix proteins required for methanol utilization in *P. pastoris*. This awaits experimental verification in future.

Other previously-characterized transcription factors acting on methanol metabolism, *ROP* (repressor of phosphoenolpyruvate carboxykinase; PAS_chr3_0554, [10]) and *TRM1* (positive regulation of methanol, PAS_chr4_0203) are induced only on methanol, but not on limiting glucose, confirming their specific involvement in methanol metabolism (reviewed by [15]).

Glucose and carbon catabolite repression regulate the expression of genes involved in glycolysis, gluconeogenesis and the metabolism of alternative carbon sources

The expression of genes related to carbon source uptake and initial metabolism is strongly regulated at the level of transcription. The respective transcriptional control of genes such as glucose sensors and transporters (low- and high-affinity), hexokinase, and glycerol- and methanol utilization are shown in Table 5 and Figure 6.

We found glycolytic *P. pastoris* genes involved in upper and lower glycolysis to have lower expression levels in all three conditions compared to excess glucose. Glycolytic genes are known to be weakly regulated at the level of transcription in *S. cerevisiae* [45], but transcriptional regulation has been previously described for Crabtree-negative yeasts such as *P. pastoris* and *K. lactis*, and was assumed to coincide with their limited glucose uptake [46,47]. As expected, the genes encoding the key gluconeogenic enzymes fructose-1,6-bisphosphatase (*FBP1*) and PEP carboxykinase (*PCK1*) are less expressed in excess glucose (compared to the other conditions). The transition between those two pathways is associated with *Gid2/*Rmd5-dependent ubiquitin-proteasome linked elimination of the key enzyme fructose-1,6-bisphosphatase [48]. *Vid24/*GID4, encodes a previously-identified key regulator of *GID2/*Rmd5 that is strongly up-regulated in methanol fed cells. A hypothetical gene (PAS_chr1-1_0399), also strongly induced on methanol and limiting glucose, could encode the homolog of Rmd5: it contains a C3HC4 RING finger domain.

In *S. cerevisiae,* fermentative glucose- or catabolite-repressed growth is described for cells grown on excess glucose [49]. Upon glucose depletion or in the presence of non-fermentable carbon sources, such as glycerol or ethanol, extensive reprogramming of gene expression allows *S. cerevisiae* to take up alternative carbon sources and enhances activity of the glyoxylate cycle, the tricarboxylic acid (TCA) cycle and gluconeogenesis [5]. We found several *P. pastoris* genes encoding enzymes involved in the metabolism of alternative carbon sources to be less expressed during growth on glycerol, methanol and/or limiting glucose (Additional file 1: Table S3 and Figure 6). Among them, the non-annotated ORFs PAS_chr4_0338, PAS_chr4_0339 and PAS_chr4_0341 could be identified to be homologs of *LRA1*, *2* and *4*. The encoded enzymes are part of the alternative pathway of L-rhamnose catabolism present in *Pichia* (*Scheffersomyces*) *stipitis* [50] and most probably allow *P. pastoris* to utilize rhamnose as sole carbon source [51]. Interestingly, *PpLRA2* and *4* flank an uncharacterized fungal-specific Zn2/Cys6 transcription factor (PAS_chr4_0340), which is up-regulated in response to methanol and limiting glucose (Table 4). Increased transcript levels in comparison to excess glucose can also be seen for many TCA cycle genes, isocitrate lyase (*ICL1*) involved in the glyoxylate cycle (Figure 6) and genes involved in channeling alternative carbon sources into the TCA cycle (e.g. the cytosolic aldehyde dehydrogenase isoforms *ALD4-2* and PAS_chr4_0470). Interestingly, some genes encoding proteins present as isoenzymes such as *ACO1/2*, *IPD1/2* and *ACS1/2* are oppositely regulated in all the de-repressed conditions.

Respiration is repressed in excess glucose conditions during fermentative growth in *S. cerevisiae* [5,52,53], thus respiration-associated functions such as oxidative phosphorylation, mitochondrial electron transport and ATP generation are induced upon glucose depletion. Unlike

Table 4 Transcriptional regulation of transcriptional regulators

Short name Pp	Description	G-D logFC	G-D adjPV	M-D logFC	M-D adjPV	X-D logFC	X-D adjPV
PAS_chr4_0340	Fungal specific transcription factor domain; Zn2/Cys6 DNA-binding domain	0.35	*	**0.72**	***	**1.50**	***
CAT8-2	Zinc cluster transcriptional activator; necessary for derepression of a variety of genes under non-fermentative growth conditions in *S. cerevisiae*	−0.07		**2.72**	***	**5.27**	***
YAP1	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance	0.27		**1.13**	***	**1.64**	***
PAS_chr1-4_0516	Putative transcription factor	**0.94**		**7.81**	***	**7.86**	***
MPP1	Fungal Zn2/Cys6 DNA-binding domain; homolog to *Hansenula polymorpha* transcription factor involved in peroxisome biogenesis/degradation	**0.90**	***	**6.34**	***	**6.99**	***
AFT1	Transcription factor, possibly involved in carbohydrate metabolism	**2.17**	***	**3.68**	***	**5.16**	***
YPR022C-3	Putative transcription factor	**1.57**	***	**2.33**	***	**4.55**	***
PAS_chr3_0348	Helix-loop-helix DNA-binding domain	0.06		0.29		**3.69**	***
ADR1/MXR1	Carbon source-responsive zinc-finger transcription factor, required for transcription of the glucose-repressed gene *ADH2*, of peroxisomal protein genes, and of genes required for ethanol, glycerol, and fatty acid utilization	**1.34**	***	**1.61**	***	**2.16**	***
RSF2/ROP	Zinc-finger protein; involved in transcriptional control of both nuclear and mitochondrial genes in *S. cerevisiae*	−0.10		**1.85**	***	−0.24	
PpTRM1	$Zn(II)_2Cys_6$-type transcription factor involved in the positive regulation of methanol utilization genes in P. pastoris and C. boidinii	-0.14		0.74	***	0.34	*
SNF1	AMP-activated serine/threonine protein kinase; found in a complex containing Snf4p and members of the Sip1p/Sip2p/Gal83p family; required for transcription of glucose-repressed genes, thermotolerance, sporulation, and peroxisome biogenesis in *S. cerevisiae*	0.39	**	**0.61**	**	**1.42**	***
SNF2	Catalytic subunit of the SWI/SNF chromatin remodeling complex involved in transcriptional regulation; contains DNA-stimulated ATPase activity	0.13		0.40	**	−0.37	**
SNF4	Activating gamma subunit of the AMP-activated Snf1p kinase complex	0.19		0.35		**0.76**	***
MIG1-1	Transcription factor involved in glucose repression in *S. cerevisiae*; regulated by the Snf1p kinase and the Glc7p phosphatase;	0.57	*	**1.09**	**	**3.09**	***
MIG1-2	Transcription factor involved in glucose repression in *S. cerevisiae*; regulated by the Snf1p kinase and the Glc7p phosphatase;	***−0.76***	**	***−1.23***	***	−0.56	***
SIP2	One of three beta subunits of the Snf1 kinase complex in *S. cerevisae*	0.00		−0.14		**0.65**	***
RDS2	Transcription factor involved in regulating gluconeogenesis and glyoxylate cycle genes; member of the zinc cluster family of proteins; confers resistance to ketoconazole in *S. cerevisiae*	−0.07		0.20		**0.83**	***
PAS_chr1-3_0274	Fungal specific transcription factor; Zn2/Cys6 DNA-binding domain	0.11		0.29		**0.90**	***
PAS_chr4_0324	Fungal specific transcription factor; Zn2/Cys6 DNA-binding domain	***−3.07***	***	***−2.99***	***	***−3.47***	***
CTH1	Member of the CCCH zinc finger family	***−2.54***	***	***−2.81***	***	***−2.92***	***
PAS_chr1-1_0422	Myb/SANT-like DNA-binding domain	−0.13		−0.57		***−2.56***	***
PAS_chr3_1209	Helix-loop-helix DNA-binding domain	0.16		−0.21		***−2.56***	***
PAS_chr1-1_0122	Helix-loop-helix DNA-binding domain	***−0.93***		−0.57		***−2.33***	***

Log_2 fold changes and adjusted P-values (* adjPV < 0.1; ** adjPV < 0.05; *** adjPV < 0.01) are shown (see Additional file 1: Table S1 for detailed data). Up-regulated genes are in bold letters, down-regulated genes in bold and italics.

S. cerevisiae, Crabtree-negative yeasts are dependent on respiratory processes even in excess glucose. Consequently, the expression of mitochondrial genes is not induced in the presence of non-fermentable carbon-sources in *P. pastoris* (Additional file 1: Table S4). However, several subunits of respiratory complex I [54], which is not present in *S. cerevisiae*, appear to be de-repressed.

Methanol utilization and peroxisomal genes are subject to carbon substrate repression

Unexpectedly, the transcript levels of most genes involved in methanol utilization (MUT) are not only highly induced in methanol-grown cells but also in glucose-limited cells (Table 6). The transcript level of *AOX1* is almost equally high in both conditions. This observation correlates well

Table 5 Transcriptional regulation of sugar transporters and sensors

Short name Pp	Description	G-D logFC	G-D adjPV	M-D logFC	M-D adjPV	X-D logFC	X-D adjPV
Pp*HXT1*	*P. pastoris* major low affinity glucose transporter (major facilitator superfamily)	***−1.31***		***−3.34***	***	***−0.82***	*
ITR2	Myo-inositol transporter	−0.40	*	***−0.88***	*	***−0.62***	***
PAS_c034_0021	Major facilitator superfamily, related to *STL1*	***−0.59***	**	0.10		−0.55	***
PAS_chr2-1_0006	Major facilitator superfamily, Quinate permease (Quinate transporter) - similar to *S. stipitis*	−0.06		***−0.80***		−0.01	
YBR241C	Putative transporter, member of the sugar porter family	0.12		−0.16		0.26	
Pp*HXT2*	*P. pastoris* putative low affinity glucose transporter of the major facilitator superfamily	−0.10		−0.10		−0.09	
STL1-1	Glycerol proton symporter of the plasma membrane, subject to glucose-induced inactivation in *S. cerevisiae*	0.08		−0.11		**1.23**	***
STL1-2	Glycerol proton symporter of the plasma membrane, subject to glucose-induced inactivation in *S. cerevisiae*	−0.27		0.40		**2.08**	***
SNF3	*P. pastoris* plasma membrane glucose sensor Gss1, regulates glucose transport	0.16		0.44		**1.60**	***
PAS_chr3_1076	Glycerol proton symporter of the plasma membrane, related to *RGT2*	0.37		**0.65**	**	**0.62**	**
PAS_chr3_1099	Glycerol proton symporter of the plasma membrane, related to *STL1* or *RGS2*	0.34		**0.80**	**	**1.33**	***
MAL31	Maltose permease, high-affinity maltose transporter (alpha-glucoside transporter)	0.09		**0.81**	***	**0.68**	***
GTH1	*P. pastoris* major high affinity glucose transporter; similar to *K. lactis HGT1*	0.17		**1.09**	***	**6.14**	***
Pp*HGT1*	*P. pastoris* high affinity glucose transporter - similar to *K. lactis HGT1*	**0.59**		**0.86**	**	**4.91**	***
PAS_chr4_0828	Myo-inositol transporter with strong similarity to the major myo-inositol transporter Itr1p, member of the sugar transporter superfamily	**2.35**	***	**3.65**	***	**7.30**	***
HXK1	Hexokinase isoenzyme 1; a cytosolic protein that catalyzes phosphorylation of glucose during glucose metabolism; expression in *S. cerevisiae* is highest during growth on non-glucose carbon sources	0.30		−0.21		**1.69**	***
HXK2	Hexokinase isoenzyme 2; catalyzes phosphorylation of glucose in the cytosol; predominant hexokinase during growth on glucose in *S. cerevisiae*	−0.12		0.18		0.03	
GLK1	Glucokinase; catalyzes the phosphorylation of glucose at C6; expression regulated by non-fermentable carbon sources in *S. cerevisiae*	***−0.99***	**	***−2.58***	***	−0.34	

Log_2 fold changes and adjusted P-values (* adjPV < 0.1; ** adjPV < 0.05; *** adjPV < 0.01) are shown (see Additional file 1: Table S1 for detailed data). Up-regulated genes are in bold letters, down-regulated genes in bold and italics.

with pre-induction expression from the *AOX1* promoter in the glycerol-fed batch prior to methanol addition [55-57], and high Aox1 protein levels in glucose-limited chemostats [58,59]. Repression of *AOX1* expression was previously determined in *P. pastoris* grown on glucose, glycerol, ethanol and acetate [60], with glycerol repression being specific for *P. pastoris AOX1/2*, but not for alcohol oxidase genes in related yeasts such as *H. polymorpha* or *C. boidinii* [14].

Although it was assumed that some MUT genes might also be regulated by catabolite de-repression [15], the extent of this regulatory pathway has not been shown experimentally in *P. pastoris*. Early observations reported that the mRNA levels of *AOX1* upon de-repression was only 1-2% of the methanol-induced mRNA levels [61], while *FLD* expression was assumed not to be under glucose repression control [62]. On the contrary we see a high level of de-repression in cells grown on limiting glucose (Table 6). This contradiction might be explained by the fact that in our set up, the cells are actively growing, while previous experiments employed glucose-exhausted stationary-phase cells for studies of de-repression. Upon (constant)

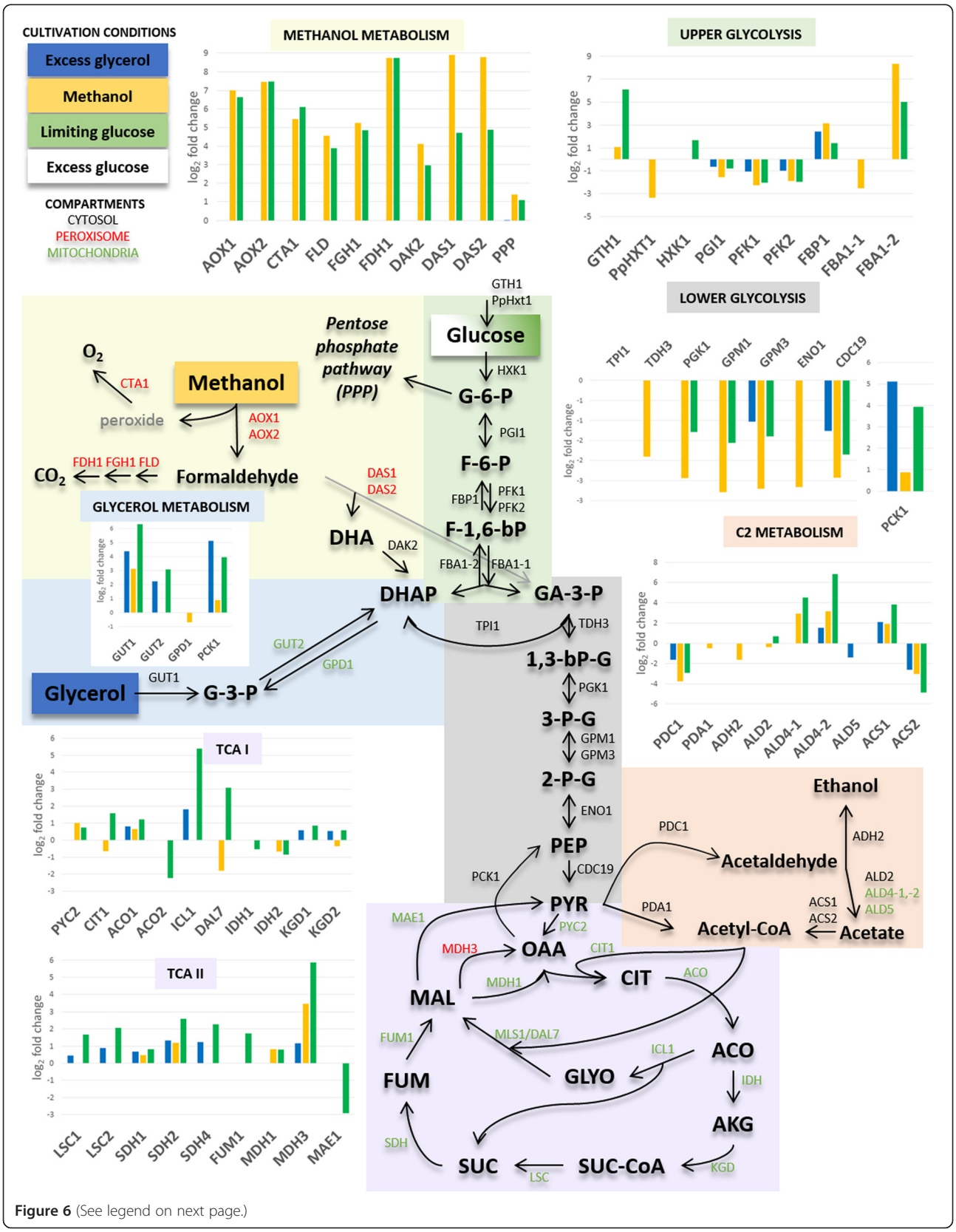

Figure 6 (See legend on next page.)

(See figure on previous page.)
Figure 6 Central carbon metabolism pathways in *Pichia pastoris*. Transcriptional $\log_2$ fold changes of genes significantly regulated in excess glycerol, methanol and limiting glucose compared to excess glucose are presented in bar charts (cutoff ±50% fold change and adjusted p-values < 0.05; [23]). According to cellular localization, peroxisomal, cytosolic and mitochondrial enzymes are colored in red, black and green, respectively. **Metabolites:** G-6-P: glucose 6-phosphate; F-1,6-P: fructose 1,6-phosphate; DHA(P): dihydroxy acetone (phosphate); G-3-P: glycerol 3-phosphate; GA-3-P: glyceraldehyde 3-phopshate; 1,3-bPG: 1,3-bisphosphoglycerate; 3-PG: 3-phosphoglycerate; 2-PG: 2-phosphoglycerate; PEP: phosphoenolpyruvate; PYR: pyruvate; OAA: oxaloacetate; CIT: citrate; ICIT: isocitrate; AKG: alpha-keto glutarate; SUC: succinate; SUC-CoA: succinyl-Coenzyme A; FUM: fumerate; MAL: malate; GLYO: glyoxylate; **Enzymes:** AOX1/2: alcohol oxidase; CTA1: catalase A; FLD: bifunctional alcohol dehydrogenase and formaldehyde dehydrogenase; FGH1: S-formylglutathione hydrolase; FDH1: formate dehydrogenase; DAK2: dihydroxyacetone kinase; DAS1/2: dihydroxyacetone synthase; GUT1: glycerol kinase; GUT2: glycerol-3-phosphate dehydrogenase; GPD1: glycerol-3-phosphate dehydrogenase; PCK1: phosphoenolpyruvate carboxykinase; GTH1: high-affinity glucose transporter; HXT1: low-affinity glucose transporter; HXK1: hexokinase; PGI1: phosphoglucose isomerase; PFK1/2: phosphofructokinase; FBP1: fructose-1,6-bisphosphatase; FBA1-1/1-2: fructose 1,6-bisphosphate aldolase; TPI1: triose phosphate isomerase; TDH3: glyceraldehyde-3-phosphate dehydrogenase; PGK1: 3-phosphoglycerate kinase; GPM1/3: phosphoglycerate mutase; ENO1: enolase I, phosphopyruvate hydratase; CDC19: pyruvate kinase; PDC1 pyruvate decarboxylase; PDA1: E1 alpha subunit of the pyruvate dehydrogenase (PDH) complex; ALD2: cytoplasmic aldehyde dehydrogenase; ALD4-1/4-2/5: mitochondrial aldehyde dehydrogenase; ACS1/2: acetyl-coA synthetase; PYC2: pyruvate carboxylase; CIT1: citrate synthase; ACO1/2: aconitase; ICL1: isocitrate lyase; DAL7: malate synthase; IDH1/2: isocitrate dehydrogenase; KGD1: alpha-ketoglutarate dehydrogenase complex; KGD2: dihydrolipoyl transsuccinylase; LSC1: succinyl-CoA ligase; SDH1/2/4: succinate dehydrogenase; FUM1: fumarase; MDH1: mitochondrial malate dehydrogenase; MDH3: malate dehydrogenase; MAE1: mitochondrial malic enzyme.

methanol addition e.g. in fed batch or chemostat, MUT gene transcript levels are on average 55-fold higher compared to glucose-limited growth conditions (unpublished data). However, our data highlight that different degrees of carbon catabolite repression are acting on individual MUT genes; for example *DAS1/2* are less de-repressed than *AOX1/2*. This strongly points towards – yet unidentified – transcriptional regulators being involved in induction/repression of the individual MUT genes in addition to the global methanol regulator *PpMXR1* (summarized by [15]). Induction of peroxisomal protein synthesis was observed in *S. cerevisiae* grown on glycerol as sole carbon source [5], which appears to be different from the situation in *P. pastoris*. In the present study, up-regulation of peroxisomal gene transcript levels occurs in glucose-limited and methanol-grown cells but not in excess glycerol (Table 6), which may also be associated with the specific repression exerted by glycerol on MUT gene expression; it might be speculated that the zinc cluster protein Cat8-2 (Table 4) is the responsible transcription factor for this.

Peroxisomal processes such as methanol utilization and beta-oxidation are associated with the formation of H_2O_2, requiring the action of antioxidants. *YAP1*, the oxidative stress response transcription factor, and many of its target genes [63] were found to be significantly up-regulated in methanol-grown cells and/or more pronounced in limiting glucose. While it was previously shown that Yap1 is required for ROS detoxification and sufficient growth on methanol [64], the strong up-regulation of *YAP1* in glucose-limited conditions was unexpected. Interestingly, starvation is linked to the expression of genes encoding oxidative stress functions in bacteria and yeast [65,66]. The protective effect of antioxidants is proposed to have a beneficial effect in cells with nutrient limitation.

The expression of fatty acid β-oxidation genes is up-regulated in *P. pastoris* cells responding to limiting glucose

Peroxisomal protein expression and fatty acid oxidation were previously reported to be regulated by Snf1 kinase through Adr1 action [67,68]. At least three other transcription factors act in concert with Adr1 in *S. cerevisiae* [68], but two of them – Oaf1 and Pip2 – cannot be found in *P. pastoris*. Instead, the putative fungal specific transcription factor PAS_chr1-3_0274 (Zn2/Cys6 domain) represents a homolog to FarA/B, the transcriptional activators of fatty acid utilization in *Aspergillus spp*., and *C. albicans* and *Y. lipolytica* Ctf1 [69]. The elevated transcript levels of PAS_chr1-3_0274 in limiting glucose are reflected by the strong induction of fatty acid utilization genes (e.g. all genes involved in beta-oxidation *FAA2*, *FOX2*, *POT1*, *POX1*, *ECI1*, *SPS19*, *PXA1* and *PXA2* have on average 100-fold higher transcript levels in limiting glucose, while only having approximately 2-fold higher transcript levels on methanol or glycerol in comparison to excess glucose). A similar regulation pattern was also observed for the non-annotated genes PAS_chr2-1_0249, PAS_FragB_0022, PAS_chr2-2_0403 and PAS_chr1-1_0108, indicating a possible involvement in beta-oxidation. Indeed, PAS_FragB_0022, PAS_chr2-1_0249 and PAS_chr1-1_0108 contain predicted PTS1 targeting signals [70], the latter having strong sequence homology to the peroxisome-targeted non-specific lipid transfer protein Pox18 present in *Candida tropicalis* and *Candida maltosa* [71,72]. Additionally, many genes connected to synthesis and degradation of triacylglycerol (TAG; metabolic pathway based on [73]) are regulated mainly in response to limiting glucose, which probably leads to the accumulation of free fatty acids which can then be degraded by beta-oxidation. Genes encoding fatty acid synthases (*FAS1*, *FAS2*) needed for *de novo* fatty acid

Table 6 Transcriptional regulation of genes involved in methanol metabolism and peroxisome formation

Short name Pp	Description	G-D logFC	G-D adjPV	M-D logFC	M-D adjPV	X-D logFC	X-D adjPV
AOX1	Alcohol oxidase (*Pichia pastoris*)	0.28		**7.00**	***	**6.64**	***
AOX2	Alcohol oxidase (*Pichia pastoris*)	0.39		**7.44**	***	**7.48**	***
CTA1	Catalase A, breaks downhydrogen peroxide in the peroxisomal matrix	**1.48**	*	**5.45**	***	**6.11**	***
DAK2	Dihydroxyacetone kinase, required for detoxification of dihydroxyacetone (DHA)	−0.18		**4.13**	***	**2.97**	***
DAS1	Dihydroxyacetone synthase variant 1	0.21		**8.91**	***	**4.72**	***
DAS2	Dihydroxyacetone synthase variant 2	0.10		**8.78**	***	**4.89**	***
FDH1	NAD(+)-dependent formate dehydrogenase, protect cells from formate	0.44		**8.74**	***	**8.75**	***
FGH1	S-formylglutathione hydrolase; involved in the detoxification of formaldehyde	**0.65**		**5.25**	***	**4.86**	***
FLD	glutathione-dependent formaldehyde dehydrogenase	0.34		**4.56**	***	**3.89**	***
PEX1	AAA-peroxin	0.50	***	**2.56**	***	**2.75**	***
PEX10	Peroxisomal membrane E3 ubiquitin ligase	0.33		**3.64**	***	**4.19**	***
PEX11	Peroxisomal membrane protein	**1.01**	***	**5.40**	***	**5.57**	***
PEX12	C3HC4-type RING-finger peroxin and E3 ubiquitin ligase	0.36	**	**2.50**	***	**3.75**	***
PEX13	Integral peroxisomal membrane protein	0.55	*	**4.39**	***	**3.90**	***
PEX14	Peroxisomal membrane peroxin	0.23		**3.14**	***	**3.90**	***
PEX17	Peroxisomal membrane peroxin	−0.26		**2.26**	***	**2.96**	***
PEX19	Chaperone and import receptor for newly-synthesized class I PMPs	−0.07		**0.75**	***	**2.10**	***
PEX2	RING-finger peroxin and E3 ubiquitin ligase	**0.75**	***	**3.48**	***	**3.63**	***
PEX20	Peroxin 20	**0.74**	***	**1.03**	***	**3.97**	***
PEX22	Putative peroxisomal membrane protein	0.11		0.55	*	**0.85**	***
PEX25	Peripheral peroxisomal membrane peroxin	−0.19		**1.09**	***	**3.29**	***
PEX28	Peroxisomal integral membrane peroxin	0.04		0.23		**1.55**	***
PEX29	Peroxisomal integral membrane peroxin	−0.24		−0.16		0.48	***
PEX3	Peroxisomal membrane protein (PMP)	0.37	**	**2.27**	***	**1.30**	***
PEX30	Peroxisomal integral membrane protein	0.10		0.09		0.47	***
PEX31	Peroxisomal integral membrane protein	0.36		**0.93**	*	**2.29**	***
PEX4	Peroxisomal ubiquitin conjugating enzyme	**0.76**	***	**2.03**	***	**4.45**	***
PEX5	Peroxisomal membrane signal receptor	0.29		**4.63**	***	**4.87**	***
PEX6	AAA-peroxin	**0.82**	***	**3.53**	***	**2.62**	***
PEX7	Peroxisomal signal receptor	−0.22		0.30		**1.98**	***
PEX8	Intraperoxisomal organizer of the peroxisomal import machinery	0.42	**	**2.93**	***	**3.52**	***
PEX26	Peroxisomal membrane protein	**0.94**	***	**3.16**	***	**4.63**	***
PEX11C	Ortholog of PEX11	0.36		**3.45**	***	**1.53**	***

Log_2 fold changes and adjusted P-values (* adjPV < 0.1; ** adjPV < 0.05; *** adjPV < 0.01) are shown (see Additional file 1: Table S1 for detailed data). Up-regulated genes are in bold letters, down-regulated genes in bold and italics.

biosynthesis are only up-regulated in methanol-grown cells, while all sterol biosynthesis genes with the exception of *ERG10*, which encodes the first step of the pathway (acetyl-CoA C-acetyltransferase), are down-regulated in limiting glucose. Potential interaction partners which are also strongly induced in glucose-limited and methanol-grown cells could be the putative transcription factor *SUT2* (PAS_chr1-4_0516*)* and *MPP1*, which was previously described to regulate peroxisomal matrix proteins and peroxins in *Hansenula polymorpha* [43].

Conclusions

Our current knowledge of translational regulation comes from studies on *S. cerevisiae* cells [74-77], where stress conditions have been found to induce a global translational down-regulation that is mediated by translation initiation factors (eIFs). The specific regulation of defined mRNAs is dependent on regulatory UTR- binding protein complexes and miRNAs [78]. A significant finding emerging from this work is that the response of *P. pastoris* to different carbon sources (glycerol, glucose and

methanol) is regulated mainly at the transcriptional level. Furthermore, we found translational regulation to be global rather than transcript-specific in the analyzed conditions.

Strikingly, cells grown on excess glycerol or glucose have a very similar transcriptome in contrast to the situation in *S. cerevisiae*, which undergoes extensive changes when shifting between those two catabolites [5,6]. We have also identified genes that are subject to glucose repression in *P. pastoris*. Global gene regulation patterns in glucose-limited cells differ strongly from cells grown in excess glycerol, which is a de-repressing carbon source. While this may be partly associated with the reduced growth rate of glucose-limited cells, transcriptional de-repression of genes of the methanol utilization pathway, peroxisome biogenesis and fatty acid β-oxidation is specific to glucose-limited growth (apart from methanol induction). The transcription factor(s) responsible for this regulatory function remain(s) to be identified.

Finally, we have shown that translational regulation is global rather than transcript-specific for *P. pastoris* cells in different growth conditions. Cells growing on methanol exhibited the highest P:M ratio – which might also account for the superior protein production capacities observed in this condition. Despite the lower growth rate, transcription of genes encoding ribosomal constituents and parts of the translational machinery is not affected on methanol, indicating an increased global translation which is also reflected in the degree of polysome-associated mRNAs in the polysome profiles. The high abundance of methanol utilization enzymes [14] in combination with peroxisome proliferation [79] increases the burden on the translation machinery in methanol-grown cells. Indeed, *P. pastoris* has increased cellular protein content during methylotrophic growth (Buchetics, Russmayer et al. manuscript in preparation).

Methods

Yeast strain and growth conditions

Pichia pastoris wildtype (X-33, *HIS4*$^+$, Mut$^+$, Invitrogen) was used for this study. In liquid culture, cells were cultivated in shake flasks at 25°C on a rotary shaker at 180 rpm. YP media without carbon source (20 g L^{-1} peptone and 10 g L^{-1} yeast extract) and synthetic media (buffered M2 minimal media, pH set to 6.0, see Delic et al. [80]) with carbon source were used for pre- and main cultures, respectively. Four different cultivation strategies (Table 1) were applied for the analysis of distinct growth phases: carbon excess (starting with 2% glycerol or glucose), methanol induction (repeated batch) or glucose-limitation (12 mm glucose feed beads, Kuhner, CH).

Cultivations with excess glycerol and glucose were inoculated to an OD of 0.1 and started with 2% carbon source, while methanol fed and glucose-limited cultivations were started with an OD of 1.5 and 0.5% or 0.25% carbon source, respectively. For the cultivation on methanol, another pulse of 0.6% methanol was given after 16 hours, about 8 hours before harvesting the culture. Limiting glucose was applied by using glucose feed beads, which are polymer particles releasing glucose at a non-linear rate of $1.63 \cdot t^{0.74}$ mg per disc. In order to generate a growth rate of about 0.015 h^{-1}, 9 feed beads were added to 40 mL culture. The cells were harvested after 16 hours, at which time point the beads liberate 5.32 mg glucose per hour. Growth rate is calculated considering the average biomass concentration (3.3 g/L DCW), the average glucose feed rate (5.32 mg/h) and the low substrate yield coefficient $Y_{X/S}$ (0.37 g/g) at low growth rates (see [3]). Assuming that any of the three variables would deviate up to 35%, the growth rate would still be within the range of 0.010 – 0.022 h^{-1}. All cultivations were performed in triplicates and harvested at an OD of about 10 (Table 1).

Polysome isolation and analysis

The method for polysome isolation and analysis for *P. pastoris* was adapted from previously published methods [6,19]. RNA is prone to degradation, so working with pre-cooled and RNase-free materials is required. Polysomes were fixed by the addition of 0.1 mg cycloheximide (fresh solution of 10 mg/mL DEPC water) per mL main culture (at an $OD_{600} \sim 10$, synthetic M2 media). The cultures were incubated for another 15 minutes on the shaker and then rapidly chilled by pouring into a 50 mL falcon tube containing 10 mL frozen DEPC-treated water and by using an ice water bath. Then the cells were recovered by 2 centrifugation steps (5300 × g, 4°C, 5 minutes) and a washing step with 10 mL cold lysis buffer (10 mM Tris–HCl pH 7.5, 0.1 M NaCl, 30 mM $MgCl_2$, 50 μg/mL cycloheximide, 200 μg/mL heparin, 1% DEPC) in between. Resuspended cells (500 μL cold lysis buffer, or more if too dense) were mixed with about 1 mL baked acid washed glass beads in ribolyzer/breaking tubes and applied in a Fast Prep (pre-cooled to −80°C, Thermo Fisher Scientific, UK) for 3 minutes at 50 RPM. The lysate was transferred into fresh RNase-free tubes, cleared by centrifugation (13 K RPM, 4°C, 15 min) and analyzed using a Nanodrop spectrophotometer (Thermo Fisher Scientific, UK).

Sucrose gradients were prepared by stacking and freezing (−80°C) of each 2 mL 50%, 40%, 30%, 20% and 10% sucrose (in sucrose gradient buffer: 50 mM NH_4Cl, 50 mM Tris-OAc pH7, 12 mM $MgCl_2$) in ultracentrifuge tubes. Gradients (stored at −80°C, thawed o/n at 4°C) were carefully loaded with polysome isolate corresponding to 150 μg RNA and centrifuged at 38 K RPM and 4°C for 2 hours in a SW40 Beckman rotor. The gradient station (Biocomp, CAN) was cleaned with ethanol (70%) and DEPC-treated water prior to gradient analysis, then blanked with water and used at a speed of 0.34 mm/s. The profile was

recorded and fractions were collected. ImageJ was used to calculate P:M ratios from the profiles, which is a measure of cellular translational activity.

RNA isolation

Monosome and polysome fractions (each about 5 mL) were separated according to the live polysome profile and collected in ice-cold tubes containing 15 mL 6 M guanidine hydrochloride (resulting in ~4 M final concentration), mixed with 2.5 volumes ice-cold 100% ethanol and precipitated o/n at −20°C. Tubes were centrifuged at 3400 × g and 4°C for one hour, supernatant was removed entirely (apply short spin for residual liquid) and pellets were carefully air-dried for 5 minutes (this step can be repeated to pool material from 2 or more gradients). In order to isolate total RNA, polysome isolate corresponding to 150 μg RNA was directly mixed with guanidine hydrochloride and processed as described above. RNA was purified from the pellets using RNeasy mini kit (Qiagen, DE). Therefore, 100 μL DEPC-treated water was used for resuspension, mixed with 350 μL buffer RLT and further processed according to the manufacturer's protocol. In the last step, 70 μL RNAse-free water was used to elute the RNA and the sample quality was checked by Nanodrop spectrophotometer and bioanalyzer analysis or gel electrophoresis.

Microarray & data analysis

In-house *P. pastoris* DNA microarrays (Agilent platform, AMAD-ID: 034821, design and general processing as described by [23]) were used. cRNA synthesis, hybridization and scanning were done according to the Agilent protocol for 2-color expression arrays. Each sample was hybridized against an RNA reference pool sample in dye swap. The microarray data were not background normalized. Within the arrays, loess-normalization was done for the color-effect. Quantile normalization was done between the arrays, the limma package (R-project) was used to calculate fold-changes, and p-value correction was done for multiple testing using the false discovery rate controlling method of [81]. Raw microarray data are provided in Additional file 5. Venn diagrams were created using the web-based tool Venny [82] and gene ontology (GO) term enrichment analysis was conducted with GO term finder and Saccharomyces Genome Database (SGD) annotations.

Principal component analysis was performed with the Excel plug-in XLSTAT.

Synonymous codon usage order (SCUO) analysis was performed online using the CondonO platform [35].

The statistical analysis was done in R using the standard functions fisher.test, chisq.test, and lm for the regression [83]. The implementation of the Fisher test obtains the p-values directly if a 2 by 2 table is present [84], otherwise a network implementation based on FEXACT was used [85]. For the group comparisons a test on normality was performed (Shapiro-Wilk-test) and Wilcoxon-Rank tests were performed since normality was not given.

Additional files

Additional file 1: Transcriptional regulation of all *P. pastoris* genes in excess glycerol, limiting glucose and methanol fed conditions compared to excess glucose condition is listed in Table S1. Separate lists for genes related to translation/ribosomes/RNA processing **(Table S2),** de-repression in excess glycerol + limiting glucose and/or methanol **(Table S3),** mitochondrial genes **(Table S4).** Log_2 fold changes and adjusted P-values (numerical values are shown in Table_S1; asterisks indicate the significance level in Table_S2-S4: * adjPV < 0.1; ** adjPV < 0.05; *** adjPV < 0.01) are shown.

Additional file 2: Enriched GO terms in differentially expressed genes in *P. pastoris* cells grown in excess glycerol, limiting glucose and methanol fed cells compared to excess glucose. Result details are provided: FDR (false discovery rate), corrected p-value and false positives.

Additional file 3: Correlation of the log_{10} mean intensity of total RNA and the log_{10} of the sum of intensities in monosome and polysome RNA.

Additional file 4: Significant translationally enriched and depleted *P. pastoris* transcripts in excess glycerol, limiting glucose and methanol fed cells compared to excess glucose.

Additional file 5: Raw microarray data of all spot replicates on the array. Fold changes of all sample replicates are shown from the green and red channel in relation to the reference pool sample.

Abbreviations

G: Excess glycerol condition; D: Excess glucose condition; X: Limiting glucose condition; M: Methanol condition; μ: Specific growth rate; GO term: Gene ontology term; ORF: Open reading frame; UTR: Untranslated region; bp: Base pairs; TCA cycle: Tricarboxylic acid cycle; Pp: *Pichia pastoris*; Hp: *Hansenula polymorpha*; Sc: *Saccharomyces cervisiae*; MUT genes: Methanol utilization genes; P bodies: Processing bodies; sd: Standard deviation; adjPV: Adjusted p-value; logFC: Logarithmic (base 2) fold change.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

RP performed the experimental work, data analysis, contributed to the study design and drafted the manuscript. SC assisted with all experimental work during polysome profiling. AG performed bioinfomatic data analysis and investigated the statistical significance of the results. MV helped with *P. pastoris* gene annotation. BG and DM contributed to the study design and data analysis, coordinated the project and contributed to drafting the manuscript. RB planned and supervised the experimental work, contributed to the study design and drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This work has been supported by the Federal Ministry of Science, Research and Economy (bmwfw), the Federal Ministry of Traffic, Innovation and Technology (bmvit), the Styrian Business Promotion Agency SFG, the Standortagentur Tirol and ZIT - Technology Agency of the City of Vienna through the COMET-Funding Program managed by the Austrian Research Promotion Agency FFG. RP thanks Erasmus for support with an internship grant. RMB and SPC thank BBSRC for funding through a targeted priority studentship to SPC.

Author details

[1]Department of Biotechnology, BOKU University of Natural Resources and Life Sciences Vienna, Muthgasse 18, 1190 Vienna, Austria. [2]Austrian Centre of Industrial Biotechnology (ACIB), Muthgasse 11, 1190 Vienna, Austria. [3]School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, UK. [4]School of Bioengineering, University of Applied Sciences FH Campus Wien, Vienna, Austria.

Received: 14 October 2014 Accepted: 24 February 2015
Published online: 11 March 2015

References

1. Gasser B, Prielhofer R, Marx H, Maurer M, Nocon J, Steiger M, et al. Pichia pastoris: protein production host and model organism for biomedical research. Future Microbiol. 2013;8:191–208.
2. Liang S, Wang B, Pan L, Ye Y, He M, Han S, et al. Comprehensive structural annotation of Pichia pastoris transcriptome and the response to various carbon sources using deep paired-end RNA sequencing. BMC Genomics. 2014;9(4):511–25.
3. Rebnegger C, Graf AB, Valli M, Steiger MG, Gasser B, Maurer M, et al. In Pichia pastoris, growth rate regulates protein synthesis and secretion, mating and stress response. Biotechnol J. 2014;9(4):511–25.
4. Dashko S, Zhou N, Compagno C, Piskur J. Why, when, and how did yeast evolve alcoholic fermentation? FEMS Yeast Res. 2014;14(6):826–32.
5. Roberts GG, Hudson AP. Transcriptome profiling of Saccharomyces cerevisiae during a transition from fermentative to glycerol-based respiratory growth reveals extensive metabolic and structural remodeling. Mol Genet Genomics. 2006;276(2):170–86.
6. Kuhn KM, DeRisi JL, Brown PO, Sarnow P. Global and specific translational regulation in the genomic response of Saccharomyces cerevisiae to a rapid transfer from a fermentable to a nonfermentable carbon source. Mol Cell Biol. 2001;21(3):916–27.
7. Mattanovich D, Graf A, Stadlmann J, Dragosits M, Redl A, Maurer M, et al. Genome, secretome and glucose transport highlight unique features of the protein production host Pichia pastoris. Microb Cell Fact. 2009;8:29.
8. Lin-Cereghino GP, Godfrey L, de la Cruz BJ, Johnson S, Khuongsathiene S, Tolstorukov I, et al. Mxr1p, a key regulator of the methanol utilization pathway and peroxisomal genes in Pichia pastoris. Mol Cell Biol. 2006;26(3):883–97.
9. Polupanov AS, Nazarko VY, Sibirny AA. Gss1 protein of the methylotrophic yeast Pichia pastoris is involved in glucose sensing, pexophagy and catabolite repression. Int J Biochem Cell Biol. 2012;44(11):1906–18.
10. Kumar NV, Rangarajan PN. The zinc finger proteins Mxr1p and repressor of phosphoenolpyruvate carboxykinase (ROP) have the same DNA binding specificity but regulate methanol metabolism antagonistically in Pichia pastoris. J Biol Chem. 2012;287(41):34465–73.
11. Leao-Helder AN, Krikken AM, Lunenborg MG, Kiel JA, Veenhuis M, van der Klei IJ. Hansenula polymorpha Tup1p is important for peroxisome degradation. FEMS Yeast Res. 2004;4(8):789–94.
12. Stasyk OG, van Zutphen T, Ah Kang H, Stasyk OV, Veenhuis M, Sibirny AA. The role of Hansenula polymorpha MIG1 homologues in catabolite repression and pexophagy. FEMS Yeast Res. 2007;7(7):1103–13.
13. Ozimek P, Lahtchev K, Kiel JA, Veenhuis M, van der Klei IJ. Hansenula polymorpha Swi1p and Snf2p are essential for methanol utilisation. FEMS Yeast Res. 2004;4(7):673–82.
14. Hartner FS, Glieder A. Regulation of methanol utilisation pathway genes in yeasts. Microb Cell Fact. 2006;5:39.
15. Vogl T, Glieder A. Regulation of Pichia pastoris promoters and its consequences for protein production. N Biotechnol. 2013;30(4):385–404.
16. Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat Rev Genet. 2012;13(4):227–32.
17. Lu P, Vogel C, Wang R, Yao X, Marcotte EM. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nat Biotechnol. 2007;25(1):117–24.
18. Vogel C, Abreu RS, Ko D, Le SY, Shapiro BA, Burns SC, et al. Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. Mol Syst Biol. 2010;6:400.
19. Bonander N, Darby RA, Grgic L, Bora N, Wen J, Brogna S, et al. Altering the ribosomal subunit ratio in yeast maximizes recombinant protein yield. Microb Cell Fact. 2009;8(1):10.
20. Beilharz TH, Preiss T. Translational profiling: the genome-wide measure of the nascent proteome. Brief Funct Genomic Proteomic. 2004;3(2):103–11.
21. Jeude M, Dittrich B, Niederschulte H, Anderlei T, Knocke C, Klee D, et al. Fed-batch mode in shake flasks by slow-release technique. Biotechnol Bioeng. 2006;95(3):433–45.
22. Prielhofer R, Maurer M, Klein J, Wenger J, Kiziak C, Gasser B, et al. Induction without methanol: novel regulated promoters enable high-level expression in Pichia pastoris. Microb Cell Fact. 2013;12(1):5.
23. Graf A, Gasser B, Dragosits M, Sauer M, Leparc G, Tuechler T, et al. Novel insights into the unfolded protein response using Pichia pastoris specific DNA microarrays. BMC Genomics. 2008;9(1):390.
24. Hartwell LH, McLaughlin CS. Temperature-sensitive mutants of yeast exhibiting a rapid inhibition of protein synthesis. J Bacteriol. 1968;96(5):1664–71.
25. Hirschmann WD, Westendorf H, Mayer A, Cannarozzi G, Cramer P, Jansen RP. Scp160p is required for translational efficiency of codon-optimized mRNAs in yeast. Nucleic Acids Res. 2014;42(6):4043–55.
26. Sammons MA, Samir P, Link AJ. Saccharomyces cerevisiae Gis2 interacts with the translation machinery and is orthogonal to myotonic dystrophy type 2 protein ZNF9. Biochem Biophys Res Commun. 2011;406(1):13–9.
27. Gilbert WV, Zhou K, Butler TK, Doudna JA. Cap-independent translation is required for starvation-induced differentiation in yeast. Science. 2007;317(5842):1224–7.
28. Parker R, Sheth U. P bodies and the control of mRNA translation and degradation. Mol Cell. 2007;25(5):635–46.
29. Pradet-Balade B, Boulme F, Beug H, Mullner EW, Garcia-Sanz JA. Translation control: bridging the gap between genomics and proteomics? Trends Biochem Sci. 2001;26(4):225–9.
30. Arava Y, Boas FE, Brown PO, Herschlag D. Dissecting eukaryotic translation and its control by ribosome density mapping. Nucleic Acids Res. 2005;33(8):2421–32.
31. McManus CJ, May GE, Spealman P, Shteyman A. Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. Genome Res. 2014;24(3):422–30.
32. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science. 2009;324(5924):218–23.
33. Arava Y, Wang Y, Storey JD, Liu CL, Brown PO, Herschlag D. Genome-wide analysis of mRNA translation profiles in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A. 2003;100(7):3889–94.
34. Marín A, Gallardo M, Kato Y, Shirahige K, Gutiérrez G, Ohta K, et al. Relationship between G + C content, ORF-length and mRNA concentration in Saccharomyces cerevisiae. Yeast. 2003;20(8):703–11.
35. Angellotti MC, Bhuiyan SB, Chen G, Wan XF. CodonO: codon usage bias analysis within and across genomes. Nucleic Acids Res. 2007;35(Web Server issue):W132–136.
36. Boender LG, van Maris AJ, de Hulster EA, Almering MJ, van der Klei IJ, Veenhuis M, et al. Cellular responses of Saccharomyces cerevisiae at near-zero growth rates: transcriptome analysis of anaerobic retentostat cultures. FEMS Yeast Res. 2011;11(8):603–20.
37. Kim JH, Roy A, Jouandot 2nd D, Cho KH. The glucose signaling network in yeast. Biochim Biophys Acta. 2013;1830(11):5204–10.
38. Turcotte B, Liang XB, Robert F, Soontorngun N. Transcriptional regulation of nonfermentable carbon utilization in budding yeast. FEMS Yeast Res. 2010;10(1):2–13.
39. Tachibana C, Yoo JY, Tagne JB, Kacherovsky N, Lee TI, Young ET. Combined global localization analysis and transcriptome data identify genes that are directly coregulated by Adr1 and Cat8. Mol Cell Biol. 2005;25(6):2138–46.
40. Portnoy T, Margeot A, Linke R, Atanasova L, Fekete E, Sandor E, et al. The CRE1 carbon catabolite repressor of the fungus Trichoderma reesei: a master regulator of carbon assimilation. BMC Genomics. 2011;12:269.
41. Dowzer CE, Kelly JM. Analysis of the creA gene, a regulator of carbon catabolite repression in Aspergillus nidulans. Mol Cell Biol. 1991;11(11):5701–9.
42. Ruth C, Buchetics M, Vidimce V, Kotz D, Naschberger S, Mattanovich D, et al. Pichia pastoris Aft1 - a novel transcription factor, enhancing recombinant protein secretion. Microb Cell Fact. 2014;13(1):120.
43. Leao-Helder AN, Krikken AM, van der Klei IJ, Kiel JA, Veenhuis M. Transcriptional down-regulation of peroxisome numbers affects selective peroxisome degradation in Hansenula polymorpha. J Biol Chem. 2003;278(42):40749–56.
44. van Zutphen T, Baerends RJ, Susanna KA, de Jong A, Kuipers OP, Veenhuis M, et al. Adaptation of Hansenula polymorpha to methanol: a transcriptome analysis. BMC Genomics. 2010;11:1.
45. Daran-Lapujade P, Rossell S, van Gulik WM, Luttik MA, de Groot MJ, Slijper M, et al. The fluxes through glycolytic enzymes in Saccharomyces cerevisiae are predominantly regulated at posttranscriptional levels. Proc Natl Acad Sci U S A. 2007;104(40):15753–8.
46. Baumann K, Dato L, Graf AB, Frascotti G, Dragosits M, Porro D, et al. The impact of oxygen on the transcriptome of recombinant S. cerevisiae and P. pastoris - a comparative analysis. BMC Genomics. 2011;12(1):218.

47. Lopez ML, Redruello B, Valdes E, Moreno F, Heinisch JJ, Rodicio R. Isocitrate lyase of the yeast Kluyveromyces lactis is subject to glucose repression but not to catabolite inactivation. Curr Genet. 2004;44(6):305–16.
48. Santt O, Pfirrmann T, Braun B, Juretschke J, Kimmig P, Scheel H, et al. The yeast GID complex, a novel ubiquitin ligase (E3) involved in the regulation of carbohydrate metabolism. Mol Biol Cell. 2008;19(8):3323–33.
49. Gancedo JM. Yeast carbon catabolite repression. Microbiol Mol Biol Rev. 1998;62(2):334–61.
50. Watanabe S, Saimura M, Makino K. Eukaryotic and bacterial gene clusters related to an alternative pathway of nonphosphorylated L-rhamnose metabolism. J Biol Chem. 2008;283(29):20372–20382.
51. Kurtzman CP. Description of Komagataella phaffii sp. nov. and the transfer of Pichia pseudopastoris to the methylotrophic yeast genus Komagataella. Int J Syst Evol Microbiol. 2005;55(Pt 2):973–976.
52. Duenas-Sanchez R, Gutierrez G, Rincon AM, Codon AC, Benitez T. Transcriptional regulation of fermentative and respiratory metabolism in Saccharomyces cerevisiae industrial bakers' strains. FEMS Yeast Res. 2012;12(6):625–36.
53. Kresnowati MT, van Winden WA, Almering MJ, ten Pierick A, Ras C, Knijnenburg TA, et al. When transcriptome meets metabolome: fast cellular responses of yeast to sudden relief of glucose limitation. Mol Syst Biol. 2006;2:49.
54. Bridges HR, Fearnley IM, Hirst J. The subunit composition of mitochondrial NADH:ubiquinone oxidoreductase (complex I) from Pichia pastoris. Mol Cell Proteomics. 2010;9(10):2318–26.
55. Khatri NK, Gocke D, Trentmann O, Neubauer P, Hoffmann F. Single-chain antibody fragment production in Pichia pastoris: Benefits of prolonged pre-induction glycerol feeding. Biotechnol J. 2011;6(4):452–62.
56. Sauer M, Branduardi P, Gasser B, Valli M, Maurer M, Porro D, et al. Differential gene expression in recombinant Pichia pastoris analysed by heterologous DNA microarray hybridisation. Microb Cell Fact. 2004;3(1):17.
57. Bawa Z, Routledge SJ, Jamshad M, Clare M, Sarkar D, Dickerson I, et al. Functional recombinant protein is present in the pre-induction phases of Pichia pastoris cultures when grown in bioreactors, but not shake-flasks. Microb Cell Fact. 2014;13(1):127.
58. Dragosits M, Stadlmann J, Graf A, Gasser B, Maurer M, Sauer M, et al. The response to unfolded protein is involved in osmotolerance of Pichia pastoris. BMC Genomics. 2010;11:207.
59. Dragosits M, Stadlmann J, Albiol J, Baumann K, Maurer M, Gasser B, et al. The effect of temperature on the proteome of recombinant Pichia pastoris. J Proteome Res. 2009;8(3):1380–92.
60. Inan M, Meagher MM. Non-repressing carbon sources for alcohol oxidase (AOX1) promoter of Pichia pastoris. J Biosci Bioeng. 2001;92(6):585–9.
61. Tschopp JF, Brust PF, Cregg JM, Stillman CA, Gingeras TR. Expression of the lacZ gene from two methanol-regulated promoters in Pichia pastoris. Nucleic Acids Res. 1987;15(9):3859–76.
62. Shen S, Sulter G, Jeffries TW, Cregg JM. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast Pichia pastoris. Gene. 1998;216(1):93–102.
63. Delic M, Graf AB, Koellensperger G, Troyer C, Hann S, Mattanovich D, et al. Overexpression of the transcription factor Yap1 modifies intracellular redox conditions and enhances recombinant protein secretion. Microbial Cell. 2014;1:376–386.
64. Yano T, Yurimoto H, Sakai Y. Activation of the oxidative stress regulator PpYap1 through conserved cysteine residues during methanol metabolism in the yeast Pichia pastoris. Biosci Biotechnol Biochem. 2009;73(6):1404–11.
65. McDougald D, Gong L, Srinivasan S, Hild E, Thompson L, Takayama K, et al. Defences against oxidative stress during starvation in bacteria. Antonie Van Leeuwenhoek. 2002;81(1–4):3–13.
66. Petti AA, Crutchfield CA, Rabinowitz JD, Botstein D. Survival of starving yeast is correlated with oxidative stress response and nonrespiratory mitochondrial function. Proc Natl Acad Sci U S A. 2011;108(45):E1089–1098.
67. Ratnakumar S, Young ET. Snf1 dependence of peroxisomal gene expression is mediated by Adr1. J Biol Chem. 2010;285(14):10703–14.
68. Young ET, Dombek KM, Tachibana C, Ideker T. Multiple pathways are co-regulated by the protein kinase Snf1 and the transcription factors Adr1 and Cat8. J Biol Chem. 2003;278(28):26146–58.
69. Ramirez MA, Lorenz MC. The transcription factor homolog CTF1 regulates {beta}-oxidation in Candida albicans. Eukaryot Cell. 2009;8(10):1604–14.
70. Neuberger G, Maurer-Stroh S, Eisenhaber B, Hartig A, Eisenhaber F. Prediction of peroxisomal targeting signal 1 containing proteins from amino acid sequence. J Mol Biol. 2003;328(3):581–92.
71. Tan H, Okazaki K, Kubota I, Kamiryo T, Utiyama H. A novel peroxisomal nonspecific lipid-transfer protein from Candida tropicalis. Gene structure, purification and possible role in beta-oxidation. Eur J Biochem. 1990;190(1):107–12.
72. Hwang CW, Yano K, Takagi M. Sequences of two tandem genes regulated by carbon sources, one being essential for n-alkane assimilation in Candida maltosa. Gene. 1991;106(1):61–9.
73. Kohlwein SD, Veenhuis M, van der Klei IJ. Lipid droplets and peroxisomes: key players in cellular lipid homeostasis or a matter of fat–store 'em up or burn 'em down. Genetics. 2013;193(1):1–50.
74. Lackner DH, Schmidt MW, Wu S, Wolf DA, Bähler J. Regulation of transcriptome, translation, and proteome in response to environmental stress in fission yeast. Genome Biol. 2012;13(4):R25.
75. Gerashchenko MV, Lobanov AV, Gladyshev VN. Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. Proc Natl Acad Sci U S A. 2012;109(43):17394–9.
76. Halbeisen RE, Gerber AP. Stress-dependent coordination of transcriptome and translatome in yeast. PLoS Biol. 2009;7(5):e1000105.
77. Payne T, Hanfrey C, Bishop AL, Michael AJ, Avery SV, Archer DB. Transcript-specific translational regulation in the unfolded protein response of Saccharomyces cerevisiae. FEBS Lett. 2008;582(4):503–9.
78. Gebauer F, Hentze MW. Molecular mechanisms of translational control. Nat Rev Mol Cell Biol. 2004;5(10):827–35.
79. Wriessnegger T, Gübitz G, Leitner E, Ingolic E, Cregg J, de la Cruz B, et al. Lipid composition of peroxisomes from the yeast Pichia pastoris grown on different carbon sources. Biochim Biophys Acta. 2007;1771(4):455–61.
80. Delic M, Mattanovich D, Gasser B. Monitoring intracellular redox conditions in the endoplasmic reticulum of living yeasts. FEMS Microbiol Lett. 2010;306(1):61–6.
81. Benjamini Y, Heller R, Yekutieli D. Selective inference in complex research. Philos Trans A Math Phys Eng Sci. 2009;367(1906):4255–71.
82. Oliveros JC. VENNY. An Interactive Tool for Comparing lists with Venn Diagrams, BioinfoGP, CNB-CSIC. 2007 http://bioinfogp.cnb.csic.es/tools/venny.
83. Chambers JM, Hastie TJ. Linear Models. In: Statistical Models in S. Wadsworth & Brooks/Cole. 1992.
84. Fisher RA. The Design of Experiments. Edinbourgh: Oliver and Boyd; 1935.
85. Mehta, CR & Patel NR. ALGORITHM 643: FEXACT: a FORTRAN subroutine for Fisher's exact test on unordered r×c contingency tables. ACM Transactions on Mathematical Software (TOMS). 1986;12:154–161.

P_{GTH1} promoter variants reveal key regulatory elements and enable improved protein production in *Pichia pastoris*

Roland Prielhofer[1], Diethard Mattanovich[1], Brigitte Gasser[1]

[1]Department of Biotechnology, BOKU University of Natural Resources and Life Sciences Vienna, Muthgasse 18, 1190 Vienna, Austria

Abstract

Protein production in *Pichia pastoris* often applies methanol-induced gene promoters such as P_{AOX1} to drive the expression of the target gene. The use of methanol has major drawbacks, so there is a demand for alternative promoters with good induction properties such as the P_{GTH1} promoter which we reported recently. In order to further increase its potential, we investigated its regulation in more detail by screening of promoter variants harboring deletions and mutations. Thereby we could identify the main regulatory region and important transcription factor binding sites of P_{GTH1}. Concluding from that, the transcription factors Mxr1, Rgt1, Cat8-1, Cat8-2 and Mig1 likely contribute to the regulation of the promoter. We also created a P_{GTH1} variant with greatly enhanced induction compared to the wild type promoter.

Keywords: *Pichia pastoris*; yeast, recombinant protein production, carbon source, promoter engineering, transcription factor binding sites

Background

The methylotrophic yeast *Pichia pastoris* (syn. *Komagataella* sp.) is a well-established protein production host. Numerous strain engineering approaches for *P. pastoris* improved the productivity for various recombinant proteins (Gasser *et al.*, 2013, Ahmad *et al.*, 2014, Byrne, 2015) and effort was also dedicated to establish novel promoters for production purposes (Vogl *et al.*, 2014, Ahn *et al.*, 2009, Prielhofer *et al.*, 2013, Mellitzer *et al.*, 2014). Promoters are key features for the expression of a particular gene: transcription of RNA of a downstream (3´) ORF is driven by the upstream (5´) promoter sequence. RNA polymerase II (RNAPII) is responsible for transcription of mRNA in eukaryotes. RNAPII promoters consist of a core promoter and several cis-acting DNA elements: proximal promoter, enhancers, silencers and boundary/insulator elements (Marsman & Horsfield, 2012, Narlikar & Ovcharenko, 2009, Phillips-Cremins & Corces, 2013). Yeast core promoters are typically located close (-75/+50 bp) to the main transcription initiation site, frequently contain improper TATA boxes (up to 2 bases difference to the TATA consensus sequence) and lack promoter elements such as Inr and DPE which are typically found in other organisms (Lubliner *et al.*, 2013). Transcriptional regulation responds to different conditions and is conducted through cis-acting elements and corresponding regulatory proteins (transcription factors (TFs)).

For biotechnological applications, strong promoters which allow either constitutive or regulated/inducible gene expression are commonly used (Mattanovich *et al.*, 2012). Production processes utilizing *P. pastoris* favorably apply carbon source dependent promoters such as the methanol-inducible alcohol oxidase promoter P_{AOX1} (Tschopp *et al.*, 1987). Thereby, the growth phase can be separated from the potentially burdening

protein production phase. We recently reported another set of promoters, which is also controlled by the carbon source, but does not require methanol for induction: These promoters share the feature of repression by excess glycerol and induction by limiting glucose (Prielhofer et al., 2013). The promoter of the high affinity glucose transporter Gth1 (P_{GTH1} =P_{G1}), which is the strongest of these promoters, is fully induced below 0.05 g/L glucose and repressed at higher glucose concentrations.

Glucose uptake characteristics are dependent on the presence of high and low affinity glucose transporters (van Urk *et al.*, 1989). The Crabtree positive yeast *Saccharomyces cerevisiae* features a huge capacity of glucose uptake permitting fermentative glucose metabolism even in the presence of oxygen. On the contrary, Crabtree-negative yeasts including *P. pastoris* are characterized by a lower glucose uptake rate and a respiratory metabolism of glucose. Correspondingly, in *S. cerevisiae* seventeen hexose transport (HXT) genes are present (*HXT1-17*) that are expressed depending on the glucose concentration (Rolland et al., 2002). Only two *HXT* homologs (PpHxt1 and PpHxt2) are found in *P. pastoris* (Zhang *et al.*, 2010). Thereof, PpHxt1 was identified to be the major low-affinity transporter in *P. pastoris* (Zhang et al., 2010, Mattanovich et al., 2009). High affinity glucose transport is mediated by PpHgt1 (PAS_chr3_0023, the homolog of *Kluyveromyces lactis* Hgt1) and Gth1 (PAS_chr1-3_0011, the gene controlled by P_{GTH1}) (Mattanovich et al., 2009, Prielhofer et al., 2013). Due to their energy-dependent high-affinity glucose transport systems, respiratory yeasts are able to take up glucose at much lower extracellular concentrations than *S. cerevisiae* (K_M of high-affinity transporters in µM range in *Pp* vs. mM in *Sc*) (Boles & Hollenberg, 1997). Transcription of both Pp*HGT1* and *GTH1* was strongly induced in glucose limiting conditions compared to glucose

surplus, while the gene encoding the low-affinity glucose transporter Pp*HXT1* showed higher transcript levels at higher glucose concentrations (Prielhofer *et al.*, 2015). Furthermore, the fundamental difference in glucose uptake behavior between different yeasts is also reflected in evolved functions of transcriptional regulators. Exemplarily, *Pp*Mxr1, the homolog of *Sc*Adr1, has specialized for methanol metabolism (Lin-Cereghino *et al.*, 2006), while the homolog of the activator of ferrous transport (Aft1) is involved in the regulation of carbohydrate-responsive genes, but does not regulate the iron-regulon in *P. pastoris* (Ruth *et al.*, 2014).

P. pastoris promoter studies and random mutagenesis of P_{AOX1} and of the promoter of glyceraldehyde-3-phosphate dehydrogenase P_{GAP} resulted in libraries of promoter variants possessing different activities, altered induction behavior compared to the wild-type promoter and in the identification of several important transcription factor binding sites (TFBS) of P_{AOX1} (Hartner *et al.*, 2008, Qin *et al.*, 2011). The interesting induction behavior of P_{GTH1} prompted us to analyze its key promoter features. Rather than to generate libraries by random mutagenesis, we constructed length variants and variants lacking certain TFBS to study their impact on induction and repression. Based on these promoter variants, we were able to identify key regulatory elements and to engineering stronger P_{GTH1} variants with unchanged regulation properties.

Results and Discussion

In order to identify the relevant regulatory region of P_{GTH1}, we cloned 8 shortened P_{GTH1} variants starting from the alternative 5´-positions -858, -663, -492, -371, -328, -283, -211 and -66 to position -1 upstream of the ATG (see **Figure 1** and **Additional file 1**, numbering based on the annotation of the *GTH1* gene locus PAS_chr1-3_0011 in the *P. pastoris* strain GS115). These shortened promoter variants were screened for eGFP expression in deep well plates to test for the repression- (glycerol) and induction properties (glucose feed beads) in comparison to the original 965 bp version of P_{GTH1} (**Figure 2**). No difference in eGFP signal was found for all length variants in the repressing condition, showing that promoter repression was not restricted in any of the shortened variants. After 48 hours of induction, the expression capacity remained fully functional for the promoter variants down to a length of 328 bp. The 283 bp- variant was only about two thirds as strong as the original P_{GTH1} promoter. The two shortest length variants (211 and 66 bp) appeared to be almost nonfunctional. These results indicate that the region between position -400 and -200 contains important regulatory features.

A high density of predicted carbon source-related TFBS marks the main regulatory region of P_{GTH1}

Next, the P_{GTH1} promoter sequence (1000 bp upstream of the gene PAS_chr1-3_0011) was searched for matrix families belonging to the matrix groups ´fungi´ and ´general core promoter elements´ using the MatInspector from Genomatix. 111 putative TFBS belonging to 46 different matrix families were found (**Figure 1** and **Additional file 2**). The most common matrix families in the analyzed sequence were monomeric Gal4-class

motifs (F$MGCM, 12 binding sites), homeodomain-containing transcriptional regulators (F$HOMD, 6 binding sites), fungal basic leucine zipper family (F$BZIP, 5 binding sites) and yeast GC-Box Proteins (F$YMIG, 5 binding sites). As anticipated from the results obtained with the length variants, we noticed a very high TFBS binding site density between position -400 to -200 with about two thirds of the mentioned TFBS (most common matrix families) occurring there (18 out of 28). Regarding general core promoter elements, no yeast- or fungi-related motifs were identified by the MatInspector, but a TATA box can be found starting at position -26.

A prominent motif was identified at position -390 to -375, which we termed TAT15 due to its sequence 5´-TA$(T)_{15}$-3´. Such poly(A:T) tracts in promoter regions are known to negatively affect nucleosome binding and to stimulate TF binding at nearby sites in yeast (Weingarten-Gabbay & Segal, 2014).

During the analysis of the genomic region of P_{GTH1}, we realized that its gene *GTH1* (PAS_chr1-3_0011 in *P. pastoris* strain GS115) has a different start annotation in the *P. pastoris* strains CBS7435 (P7435_Chr1-0007) and DSMZ70382 (PIPA00372). In contrast to GS115, the coding sequence is annotated to start 36 bp further downstream in the other two strains. We assume that the annotation of *GTH1* in strain GS115 is correct.

The carbon source-related transcription factors Mxr1, Rgt1, Cat8-1, Cat8-2 and Mig1 were revealed to contribute to the regulatory properties of P_{GTH1}

Transcription factor binding sites with predicted glucose- or carbon source dependency were selected for further analyses (see **Figure 3**, **Table 1**, and **Additional file 4**). P_{GTH1} variants with deletions of the respective regions were generated using overlap-extension

PCR. **Figure 3 (A)** shows all selected TFBS and **(B)** indicates all TFBS which are (partially) affected by the deletion (listed in **Table 1**). For some deletions (e.g. Δ9 and Δ10), some nucleotides of the respective TFBS were left untouched in order to keep closely neighboring TFBS functional and to separately examine their effect.

All TFBS deletion and TAT15 mutation variants were screened for eGFP expression in repressing (glycerol) and inducing conditions (glucose feed bead) (**Figure 4**). It is important to consider that individual TF/TFBS are usually not sufficient to fulfill a promoter´s regulation. TFBS deletions also imply that the promoter sequence can be affected by the newly formed adjoined sequence, by altered distances between TFBS or by changes of higher order properties (chromatin organization). The same TFBS at different positions of the promoter can have different functions, also because of other adjacent TFBS. At closely neightbouring TFBS, TFs might either act synergystically or restrict binding of other TFs due to steric hindrance.

TFBS of four different carbon source-related TF families were deleted in the P_{GTH1} promoter variants (see **Table 1** and **Additional file 4**): Yeast metabolic regulator (F$ADR; matrixes: F$ADR1.01), Monomeric Gal4-class motifs (F$MGCM; matrixes: F$RGT1.01, F$RGT1.02), Carbon source-responsive elements (F$CSRE, matrixes: F$CSRE.01, F$SIP4.01) and Yeast GC-Box Proteins (F$YMIG; matrixes: F$MIG1.01 and F$MIG1.02). The corresponding transcription factors in *S. cerevisiae* are Adr1, Rgt1, Sip4/Cat8 and Mig1, respectively.

Carbon source dependent promoters are controlled by glucose repression and/or induction by carbohydrates or other non-sugar carbon sources (a comprehensive review on carbon source dependent promoters in yeasts is provided by (Weinhandl *et al.*, 2014)).

Glucose repression is mainly conducted by the Snf1 protein kinase complex, the transcriptional repressor Mig1 and protein phosphatase 1. Downstream factors regulate e.g. respiratory genes (Hap4), gluconeogenesis genes (Cat8, Sip4) and glucose transporters (Rgt1) in *S. cerevisiae.*

P. pastoris has two Mig1 homologs, called Mig1-1 and Mig1-2, the second of which possibly acts as carbon catabolite repressor (Prielhofer et al., 2015). When glucose is available, Mig1 acts as repressor, while Rgt1 acts as transcriptional activator. To fulfill repressor function Mig1 gets dephosphorylated and imported into the nucleus where it recruits the corepressors Ssn6 and Tup1. In limiting glucose conditions, Rgt1 gets dephosphorylated and acts as transcriptional repressor (Kim *et al.*, 2013). Recently, Roy and colleagues reported that Rgt1 function is controlled by its phosphorylation state (Rgt1 has four phosphorylation sites) and that promoter induction does not require its dissociation, as typically seen for transcriptional repressors (Roy *et al.*, 2014).

The carbon source-responsive zinc-finger transcription factor Adr1 is required for transcriptional activation of the glucose-repressible alcohol dehydrogenase (*ADH2*) gene in *S. cerevisae* (Denis & Young, 1983). The Adr1 homolog in *P. pastoris* is Mxr1 (PAS_chr4_0487), the key regulator of methanol metabolism (Lin-Cereghino et al., 2006), and it was reported to be a positive acting transcription factor being essential for strong P_{AOX} induction on methanol (Hartner et al., 2008). The reported TFBS core motif 5’ CYCC 3’ for Mxr1 (Kranthi *et al.*, 2010) matches with both F$ADR1.01 sites found in the P_{GTH1} promoter sequence.

The carbon source response element (CSRE) is bound by the transcriptional activators Sip4 and Cat8 and functions to induce the expression of gluconeogenesis genes in *S. cerevisiae*

(Hiesinger *et al.*, 2001). Two *P. pastoris* homologs of *Sc*Cat8 can be found: Cat8-1 (PAS_chr2-1_0757) and Cat8-2 (PAS_chr4_0540), both also being the best blastp hits for *Sc*Sip4. Cat8-2 is weakly similar to *Sc*Cat8, and it potentially plays an important role during glucose derepression (Prielhofer et al., 2015).

Deletion variants of the P_{GTH1} promoter reveal TFBS responsible for its repression and induction

Out of the 5 deletion variants residing upstream (5´) of the main regulatory region of P_{GTH1} identified before (see dashed box in **Figure 3**), the variants P_{GTH1}-Δ1, -Δ2 and -Δ4 appear to have a beneficial effect on promoter strength while the deletion variants P_{GTH1}-Δ3 and Δ5 had no effect on GFP expression compared to the original P_{GTH1} promoter. These results suggest that 5´-shortening of the promoter might be beneficial for the engineering of P_{GTH1}. TFBS deletions within the main regulatory region of P_{GTH1} (P_{GTH1}-Δ6 to -Δ12, see **Figure 3**) had different impacts on eGFP expression, but none showed increased induction without losing the repression properties. Therefore, we assume that the main regulatory region of P_{GTH1} needs to be maintained in engineered P_{GTH1} promoter variants in order to retain its tight regulation. Without this region, much lower induction in limiting glucose is observed (P_{GTH1}-328 and P_{GTH1}-283, **Figure 2**).

Mig1 binding sites were deleted in P_{GTH1}-Δ3, -Δ4, -Δ10 and -Δ11 (F$MIG1.02 in Δ3, F$MIG1.01 in Δ4, Δ10 and Δ11), whereat P_{GTH1}-Δ10 and P_{GTH1}-Δ11 also include $ADR1.01 and F$RGT1.02 deletions, respectively. Slightly tighter repression was found for Δ3, while Δ4 had unchanged repression but enhanced eGFP levels after induction. Liberated repression seen for Δ10 and weaker promoter induction of Δ10 and Δ11 could also be

connected to F$RGT1 binding sites in this region (F$RGT1.01 and F$RGT1.02 deleted in Δ9 and Δ11). Also, Mig1 could play a bifunctional role in P_{GTH1} regulation: two *MIG1* genes are found in *P. pastoris* (*MIG1-1, MIG1-2*) and they were shown to be regulated contrariwise upon glucose availability (Prielhofer et al., 2015).

The deletion of F$ADR1.01 increased eGFP levels in the variant P_{GTH1}-Δ1, although Mxr1 (positive regulator of methanol metabolism in *Pp,* homolog of *ScADR1*) binding site deletion would be expected to rather weaken the promoter. Combined deletion of F$ADR1.01 with F$MIG1.01 in P_{GTH1}-Δ10 liberated promoter repression on glycerol and weakened its induction, which is a conclusive response for Mig1 TFBS deletion.

In the main regulatory region, the binding site F$RGT1.02 was deleted in the variants P_{GTH1}-Δ6 (two sites), -Δ7, -Δ8, -Δ11 and -Δ12, and F$RGT1.01 was deleted in Δ9. The variant harboring the deletion of the paired F$RGT1.02 site (Δ6, binding sites on opposite strands with a shift of 7 bp) showed a slightly liberated repression and reduced induction. The variants Δ7 and Δ8 contain very close F$RGT1.02 sites, whereat the first lies on the negative- and the second on the positive strand; also Δ8 contains the deletion of an F$SIP4.01 site. The first (Δ7) showed a slightly liberated repression and increased induction, while the second (Δ8) was much weaker induced (but had unchanged promoter repression). This indicates a strong role for the transcriptional activator(s) Cat8-1 and/or Cat8-2 (strongest homologs for *Sc*Sip4) for P_{GTH1} induction. The variant Δ9 was created to delete closely located F$RGT1.01 and F$CSRE.01 TFBS (binding sites on opposite strands) and the drastic loss of repression indicates a strong role of these TFBS to tightly control P_{GTH1}, most likely through binding of Rgt1, Cat8-1 and/or Cat8-2. The deletion of F$RGT1.02 in the variant P_{GTH1}-Δ12 did not have an effect on eGFP expression

performance. Interestingly, *CAT8-2* transcription is strongly upregulated in limiting glucose compared to glucose surplus , while *RGT1* and *CAT8-1* were not transcriptionally regulated in the tested conditions (Prielhofer et al., 2015).

P_{GTH1} promoter strength is dependent on the poly(A:T) tract TAT15

The TAT15 motif is located about 80 bp upstream (5´, position -390 to -374) of the main regulatory region of P_{GTH1}. Repeated sequencing of the 5´-region of GTH1 in *P. pastoris* GS115 or CBS7435 resulted in the detection of 15 ± 1 Ts in the TAT15 motif. To elucidate its impact on promoter performance, the TAT15 motif was selected for deletion (P_{GTH1}-ΔTAT15) or elongation (to T16, T18 and T20; P_{GTH1}-T16, P_{GTH1}-T18, P_{GTH1}-T20). Primers (see primers #37-42 in **Additional file 3**) were initially designed to obtain T18, T20 and T22, but variants with different lengths (T16, T20 and T18, respectively) were obtained and used. Deletion of the TAT15 motif indeed resulted in lower GFP signals, whereas its prolongation increased the expression strength of P_{GTH1}. This indicates that the use of a prolonged TAT15 motif would be beneficial for P_{GTH1} engineering.

Partial sequence duplications of P_{GTH1}´s main regulatory region significantly improve its expression strength

Two duplication variants (P_{GTH1}-D1240 and P_{GTH1}-D1427, the numbers state the lengths of the respective promoter variants) of the P_{GTH1} promoter were generated by PCR amplification of two sequence fragments (-472 to -188 and -472 to -1, see **Figure 3**) and insertion using the restriction sites PstI and BglII (positions 509-514 and 525-530). The

duplication sections start upstream of TFBS deleted in P_{GTH1}-Δ5 and end after the main regulatory region of P_{GTH1} for the first variant (P_{GTH1}-D1240), while the second duplication (P_{GTH1}-D1427) reaches until the 3´-end of the P_{GTH1} promoter. These variants were screened for eGFP expression in the same way as described for the TFBS deletion and TAT15 mutation variants. Both duplication variants showed more tight repression in excess glycerol and stronger induction upon limiting glucose (**Figure 6**).

We also tested the post-transformational stability of the duplication variant clone P_{GTH1}-D1240 #3 by performing three consecutive batch cultivations without selection pressure, which is equal to about 20 generations. We could verify that eGFP expression was stable over the whole cultivation time (data not shown). In comparison, a typical *P. pastoris* bioreactor process starts with OD_{600}=1 (~0.2-0.4 g/L YDM) in the batch phase and ends with ~100 g/L YDM after the fed batch phase and thereby takes about 10 generations.

Verification of P_{GTH1} promoter variant performance in fed batch bioreactor cultivation

In order to verify the performance of the generated promoter variants in bioprocess conditions, some variants were selected for fed batch cultivation based on their altered eGFP expression performance: P_{GTH1}-Δ2 (deletions of Rgt1 and CSRE TFBS) was the most enhanced variant upstream of the main regulatory region, and P_{GTH1}-T16 and P_{GTH1}-D1240 showed higher eGFP expression levels in limiting glucose without losing promoter repression in the glycerol condition. A bioreactor cultivation, which was started with a glycerol batch phase followed by a space-time yield optimized fed batch (Prielhofer et al., 2013, Maurer *et al.*, 2006) for one clone each and compared to the control strain P_{GTH1} #8 for eGFP expression (see **Figure 7** and **Table 2**). The gene copy number (GCN) of these

three clones were analyzed using quantitative Real-time PCR and resulted in one copy of the expression cassettte for all of them (data not shown). The strong improvement of the duplication variant P_{GTH1}-D1240 could be verified in bioreactor conditions, the clone P_{GTH1}-D1240 #3 showed a 50% increase in GFP fluorescence at the fed batch end, but the signal was already increased at the batch end. Other than in the screening, the clone P_{GTH1}-Δ2 #3 had a slightly increased signal at the batch end, and about 10% weakened signal at the fed batch end. The TAT15 mutation variant clone P_{GTH1}-T16 #3 showed the strongest signal at the batch end, and fell behind the duplication variant at the fed batch end, reaching about 20% improvement over the control P_{GTH1} #8, similar to the screening result. The different induction behavior of the clones in the batch phase is explained by derepression due to decreasing glycerol concentration throughout the batch phase (see **Figure 7A**). Overall, the fed batch cultivations could largely confirm the results obtained in small scale screening.

Achievements and Conclusions

Gene promoters with carbon source-dependent regulation are favorable for bioprocess applications because the production phase can be separated from growth. Potential promoter-based protein production improvement can be accomplished by finding the optimal growth conditions (e. g. (Maurer et al., 2006)) or by directly manipulating the promoter sequence (e. g. (Hartner et al., 2008)).

We constructed several P_{GTH1} promoter variants with shortened length, TFBS deletions, TAT15 motif mutations and fragment duplications. Thereby, we could identify the main regulatory region of P_{GTH1}, including its important TFBS. The analysis of TFBS deletions

indicates that the transcription factors Rgt1 and Cat8-1 and/or Cat8-2 play an essential role for P_{GTH1} repression and induction: two motifs consisting of F$RGT1 and F$CSRE binding at the same position on the opposite strands were deleted. Deletion of the first part (P_{GTH1}-Δ8, position -293 to -285; RGT1: (+)-310 to -299, CSRE: (-) -299 to-285) caused weakened promoter induction, while deletion of the second part (P_{GTH1}-Δ9, position -275 to -261; RGT1: (-) -275 to -259, CSRE: (+) -276 to -260) lead to decreased promoter repression. Thereby, we could clearly identify a regulatory motif which is essential and characteristic for P_{GTH1} regulation.

The role of the transcriptional regulators Mig1 (F$MIG1) and Mxr1 (F$ADR1) might be more important in other conditions such as excess glucose or methanol induction. Other transcription factors which bind in or close to that region might also contribute to P_{GTH1}´s regulation.

The poly(A:T) tracts are known to play a role in promoter sequences, and the TAT15 motif in P_{GTH1}, which is located upstream (position -390 to -375) of the main regulatory region, could be shown to be essential for P_{GTH1} promoter strength. Elongation of this motif to T16, T18 and T20 had a positive effect on promoter performance.

Deletion variants of P_{GTH1} revealed that 5´-shortening might be beneficial for promoter engineering as well. TFBS for Mxr1, Mig1, Rgt1 and Cat8 deleted upstream of the main regulatory region of P_{GTH1} improved eGFP expression, although this effect was not seen for the 5´-shortened promoter variants.

Two variants with partial sequence duplications reached greatly enhanced expression capacities compared to the wild type P_{GTH1}.

Thereby, we could assign distinct features of P_{GTH1} good expression performance, which is a solid basis for promoter engineering: 5´-shortening, TAT15 motif elongation and fragment duplication. P_{GTH1} variant performance in small scale screening could successfully be verified in fed batch cultivations.

Materials and Methods

Strains and cultivation

Escherichia coli DH10B (Invitrogen) was used for subcloning. It was routinely cultivated in petri dishes or shake flasks using media supplemented with 25 µg mL^{-1} Zeocin for selection. The wild type *P. pastoris* strain CBS2612 was used for protein production in this work.

The main culture for screenings was either done with buffered M2 minimal-, YP- or BM media and glucose feed beads (12 mm, Kuhner, CH) which provided the carbon source (Prielhofer et al., 2013).

YP media contained 20 g L^{-1} soy peptone and 10 g L^{-1} yeast extract, which can be supplemented with 12.6 g glycerol or 20 g glucose to obtain YPG and YPD, respectively. For cultivation on plates, 20 g L^{-1} agar-agar was added to the liquid medium. BM contained 13.4 g L^{-1} yeast nitrogen base (Cat.No. 291940, Becton Dickinson, FR) with ammonium sulfate, 0.4 mg L^{-1} biotin and 100 mM potassium phosphate buffer pH 6.0. Buffered M2 minimal media (see Delic et al. (2010)) was also set to pH 6.0.

Promoter cloning and transformation into *P. pastoris*

Cloning and transformation was done as described previously (Prielhofer et al., 2013). The P_{GTH1} promoter sequence (Prielhofer et al., 2013) was PCR- amplified (see primers #1-2 in **Additional file 3**) from *P. pastoris* genomic DNA and ligated into the pPuzzle vector (Stadlmayr *et al.*, 2010) upstream of the reporter gene eGFP using the ApaI (5`-GGGCCC-3`) and the SbfI (5`-CCTGCAGG-3`) restriction sites. Due to the use of multiple cloning sites (*SbfI*), ten bases were inserted between the promoter sequence and the eGFP start codon: 5`-*CCTGCAGG*CC-3`.

Genome integration of the expression plasmid was targeted to the 3´-flanking region of the *AOX1* gene of *P. pastoris* to avoid positional effects on reporter gene expression. Plasmids were linearized within the genome integration region prior to electroporation (2 kV, 4 ms, GenePulser, BioRad) into electrocompetent *P. pastoris*. In order to generate clones with low copy integration, low amounts of DNA were used for the transformation (< 1 µg DNA). Multicopy clones were excluded from the screening data, as they can easily be identified by their strongly enhanced eGFP fluorescence (Prielhofer et al., 2013). The GCN was analysed with Real-time PCR and resulted in one copy of the expression for P_{GTH1} #8, P_{GTH1}-Δ2 #3, P_{GTH1}-T16 #3 and P_{GTH1}-D1240 #3.

P. pastoris cells were first selected and cultivated on YPD agar and then inoculated in liquid YPG medium as pre-culture for screenings and fermentations. Antibiotic selection by Zeocin was applied on plates and in pre-culture at a concentration of 25 µg mL^{-1}.

Promoter sequence analysis

The P_{GTH1} promoter sequence (1000 bp upstream of the gene PAS_chr1-3_0011 according to the annotation in the *P. pastoris* strain GS115) was analysed for putative TFBS using MatInspector ((Cartharius *et al.*, 2005); release professional 8.1, September 2013, from Genomatix (Genomatix software suite v3.2)). The search was based on the MatInspector library Matrix Family Library Version 9.2 (October 2014) carried out with standard search parameters (matrix groups ´fungi´ and ´general core promoter elements´, core similarity 0.75, matrix similarity optimized).

Cloning of P_{GTH1} promoter variants

Shortened promoter versions were cloned by applying the same strategy as for the original 965 bp version of P_{GTH1} (Prielhofer et al., 2013) by using different forward primers (and the same backward primer). Shortened P_{GTH1} versions of 858, 662, 491, 370, 328, 283, 211, 127 and 66 bp (see **Figure 1** for schematic representation and **Additional file 3** for primer sequences #2-10) were PCR amplified and cloned upstream of eGFP .

TFBS deletions and TAT15 motif mutations (see **Figure 3**) of P_{GTH1} were cloned by applying overlap-extension PCR (see **Additional file 3** for primer sequences #11-42) from *P. pastoris* genomic DNA. For each variant two fragments (up- and downstream of the deletion) were PCR amplified which were then joined in second PCR step.

Fragment duplications (see **Figure 5**) within P_{GTH1} were cloned using the restriction sites PstI and BglII, both found in the P_{GTH1} sequence at the position 509-514 and 525-530. Therefore, a forward primer containing PstI and two different backward primers

containing BglII (primers #43-45, see **Additional file 3**) were used to amplify and insert fragments from -472 to –188 and -472 to -1 to generate the variants P_{GTH1}-D1240 and P_{GTH1}-D1427.

The P_{GTH1} promoter sequence and all P_{GTH1} variants are provided in **Additional file 2** in fasta format.

Screening, fed batch cultivation and eGFP expression analysis

Expression screenings were done in 24- deep well plate screenings at 25°C and with shaking at 280 rpm with 2 mL culture per well. Glucose feed beads (6mm, Kuhner, CH) were used to generate glucose-limiting growth conditions. Cells were analysed for eGFP expression during repression (YP + 1% glycerol) and induction (YP + 1 feed bead). Screenings were repeated to verify the reproducibility of the results.

For the screening of P_{GTH1} length variants, two clones each were cultivated in triplicates. Samples were taken at the end of the pre-culture and after 24 and 48 hours of the main culture (a second feed bead was added after 24 hours). For the P_{GTH1} deletion variants, -TAT15 mutants and –duplication variants, clones were pool cultivated (mixed culture of 5 to 9 clones) in 3 wells.

Fed batch bioreactor cultivations were performed as described before (Prielhofer et al., 2013). The batch phase of approximately 25 h was followed by a fed batch phase (glucose fed batch media) with a space-time yield optimized feeding rate according to Maurer et al. (Maurer et al., 2006). Samples were taken during the batch and fed batch phase, and analysed for eGFP levels.

Expression of eGFP in screenings was analysed by flow cytometry. Specific eGFP fluorescence referred to in this study is the fluorescence intensity related to the cell volume for each data point as described by Hohenblum et al. (Hohenblum *et al.*, 2003). For all graphs showing specific eGFP fluorescence, the geometric mean of the whole population was used. Please note that the specific eGFP fluorescence of two different screenings cannot be compared.

For bioreactor samples, a plate reader (Infinite 200, Tecan, CH) was used to determine eGFP fluorescence. Samples were diluted to an OD_{600} of 5 and fluorescence intensity was related to the bioreactor culture volume.

Tables

Table 1: Positions and TFBS deletions of P_{GTH1} TFBS deletion variants

P_{GTH1}-Δ	Position	TFBS Deletions[a] (TF Matrices)
1	**-785 to -777**	**F$ADR1.01**
2	**-628 to -612**	F$PHD1.03, **F$RGT1.02, F$CSRE.01**
3	**-586 to -568**	F$REB1.02, **F$MIG1.02**, F$MSN2.01, F$YAP1.02, F$TOS8.01
4	**-553 to -535**	**F$MIG1.01**, F$RAP1.06, F$AFT2.01
5	**-442 to -426**	**F$RGT1.02**, F$GZF3.01, F$PHD1.01
6	**-337 to -316**	F$ASG1.01, **F$RGT1.02, F$RGT1.02**, F$RDR1.01, F$GATA.01
7	**-310 to -299**	F$STE12.01, F$GAT1.01, **F$RGT1.02**, O$DMTE.01, F$OAF1.01
8	**-293 to -285**	F$OAF1.01, **F$RGT1.02**, F$GAL4.01, **F$SIP4.01**, F$RDR1.01, F$LAC9.01
9	**-275 to -261**	F$LEU3.02, **F$CSRE.01, F$RGT1.01**, F$TEA1.01
10	**-258 to -242**	F$REB1.02, F$MCM1.02, **F$MIG1.01**, F$ADR1.01
11	**-239 to -221**	**F$RGT1.02, F$MIG1.01**, F$TEA1.01, F$PPR1.01, F$PDRE.01, F$PPR1.01, F$PDRE.01
12	**-220 to -209**	F$HAP1.01, F$QA1F.01, **F$RGT1.02**, F$HAP1.01

[a] Targeted and affected TFBS in P_{GTH1} TFBS deletion variants (P_{GTH1}-Δ1 to Δ12) are listed. Targeted carbon source-related TFBS are shown in bold. Detailed information for all TFBS and for the deleted TFBS is provided in Additional file 1 and 4, respectively.

Table 2: Fed batch cultivation of P_{GTH1} and P_{GTH1} variants expressing eGFP

	Batch End				Fed Batch End			
Clone	**t[a] [h]**	**YDM[b] [g/L]**	**relative eGFP fluorescence**	**%[c]**	**t[a] [h]**	**YDM[b] [g/L]**	**relative eGFP fluorescence**	**%[c]**
P_{GTH1} #8	-5.3	9.8	44 +/- 1	**100**	19.5	118.6	**2005** +/- 36	**100**
P_{GTH1}-Δ2 #3	-4.6	11.0	51 +/- 1	**116**	19.5	110.6	**1819** +/- 43	**91**
P_{GTH1}-T16 #3	-3.0	14.2	70 +/- 1	**160**	19.5	113.1	**2383** +/- 24	**119**
P_{GTH1}-D1240 #3	-3.0	14.9	62 +/- 1	**141**	19.5	113.3	**2948** +/- 33	**147**

[a] The time points were set to 0 at the batch end. [b] The biomass concentrations (YDM) in the batch and fed batch were as expected.[c] A clone expressing eGFP under control of P_{GTH1} (#8) was compared to clones expressing under control of a P_{GTH1} deletion (P_{GTH1}-Δ2), a TAT15 mutation (P_{GTH1}-T16), and a duplication (P_{GTH1}-D1240) variant.

Figures

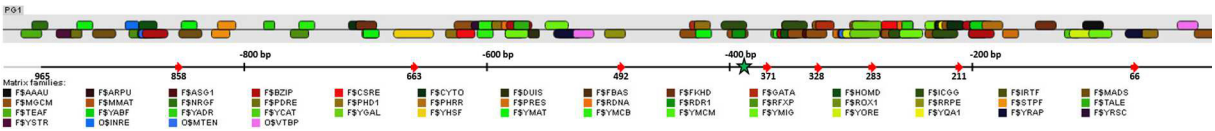

Figure 1: P_{GTH1} promoter sequence analysis for TFBS using MatInspector

The P_{GTH1} sequence is provided in **Additional file 3** (1000 bp upstream of the gene PAS_chr1-3_0011 according to the annotation in the *P. pastoris* strain GS115). Matrix families belonging to the matrix groups ´fungi´ and ´general core promoter elements´ are shown (detailed matrix match table is provided in **Additional file 1**). The green asterisk indicates the position of the prominent TAT15 motif (position -390 to -374). P_{GTH1} was initially amplified and cloned from position -965 to -1 (length of 965 bp). Alternative 5´-starts of the shortened P_{GTH1} promoter variants are labelled with red arrows and the length of the corresponding variant.

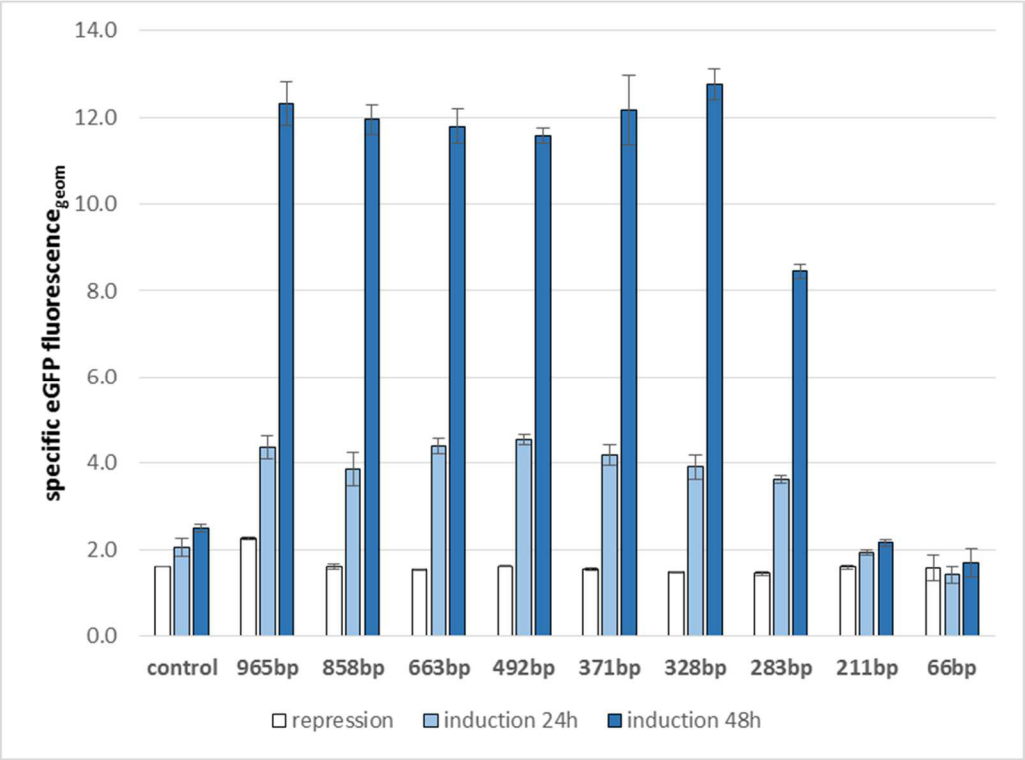

Figure 2: Screening data of the shortened P_{GTH1} promoter variants

The geometric mean of the population´s specific eGFP fluorescence (fluorescence related to cell volume) is shown for clones expressing eGFP under control of P_{GTH1} (clone #8, verified GCN of 1) or a shortened P_{GTH1} variant (each 2 clones cultivated in triplicates, selected in pre-screenings) in repressing and inducing growth conditions. Wild type *P. pastoris* cells were used as negative control. Samples were taken during the repressing pre-culture and after 24 and 48 hours induction with feed beads.

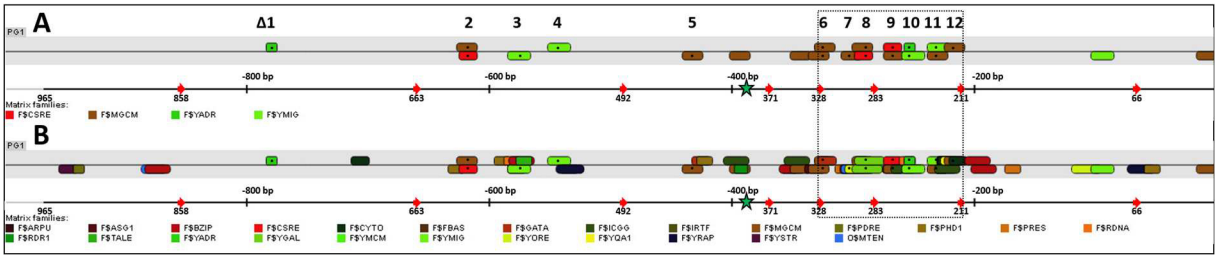

Figure 3: P_{GTH1} promoter sequence analysis for carbon source-related TFBS using MatInspector and selected TFBS for deletion

Black dots and corresponding numbers indicate TFBS which were selected for deletion (**listed in Table 1** and **Additional file 4**). Associated matrix families are shown in (A), and (B) illustrates all TFBS which might be affected by the deletions (matrix match detail information is given in **Additional file 1**). The black dashed box indicates the main regulatory region of P_{GTH1} which was identified by the screening of shortened P_{GTH1} variants. The green asterisk indicates the position of the prominent TAT15 motif which was also selected for deletion and for mutation.

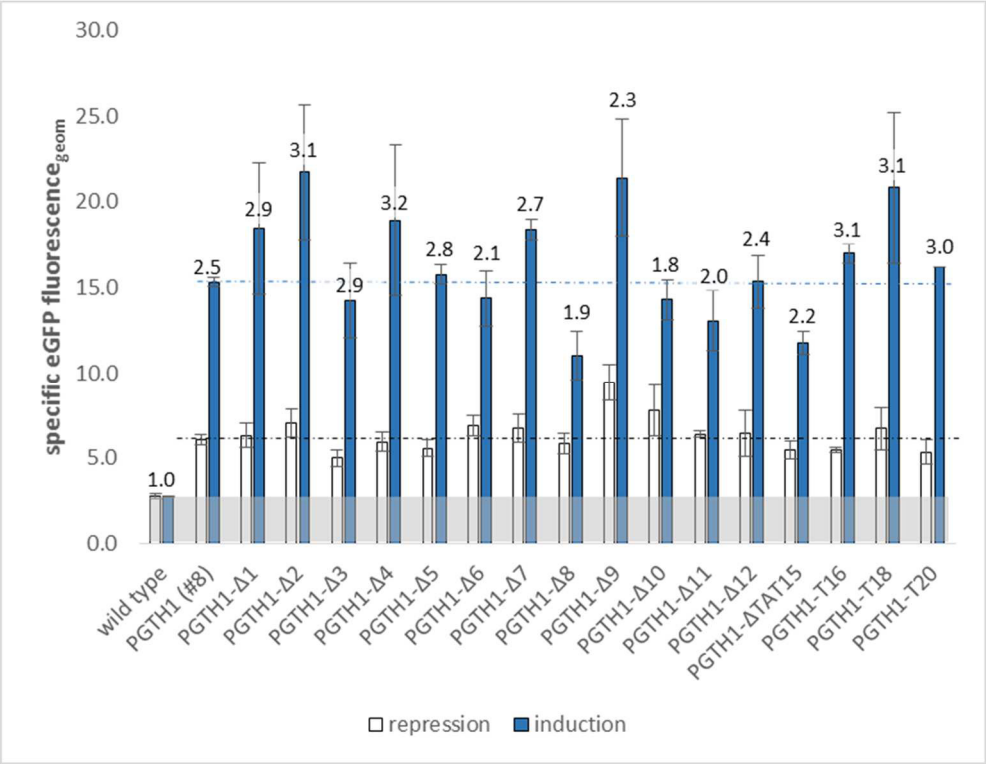

Figure 4: Screening data of the P_{GTH1} deletion and -TAT15 mutation variants

The geometric mean of the population´s specific eGFP fluorescence (fluorescence related to cell volume) is shown for clones expressing eGFP under the control of P_{GTH1} (clone #8, verified GCN of 1) or a P_{GTH1} variant (up to 9 clones were pool cultivated in 3 wells) in repressing and inducing growth conditions. Wild type *P. pastoris* cells were used as negative control.

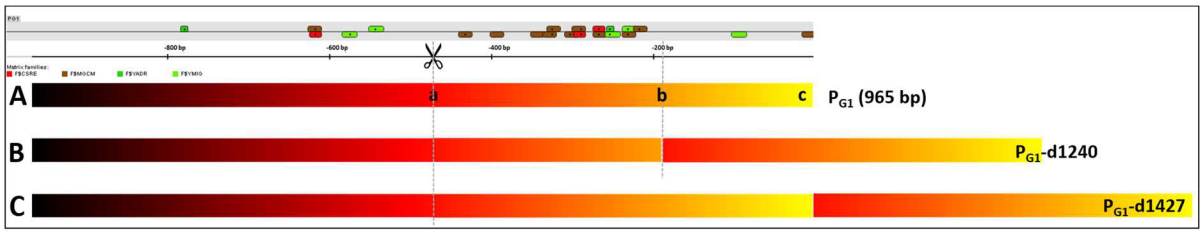

Figure 5: Schematic illustration of P_{GTH1} promoter duplication variants

The first cloned P_{GTH1} sequence (A) is 965 bp long (amplified from position 36 to 1000 of the sequence shown here). PstI and BglII restriction sites (located at positions 509-514 and 525-530, indicated with scissors) were used to generate P_{GTH1} variants P_{GTH1}-D1240 and P_{GTH1}-D1426 with duplicate fragments. Sequence fragments corresponding to a-b and a-c of sequence (A) were amplified with primers containing appropriate restriction sites and ligated into the site, thereby generating duplication variants with a length of 1240 bp (B) and 1427 bp (C).

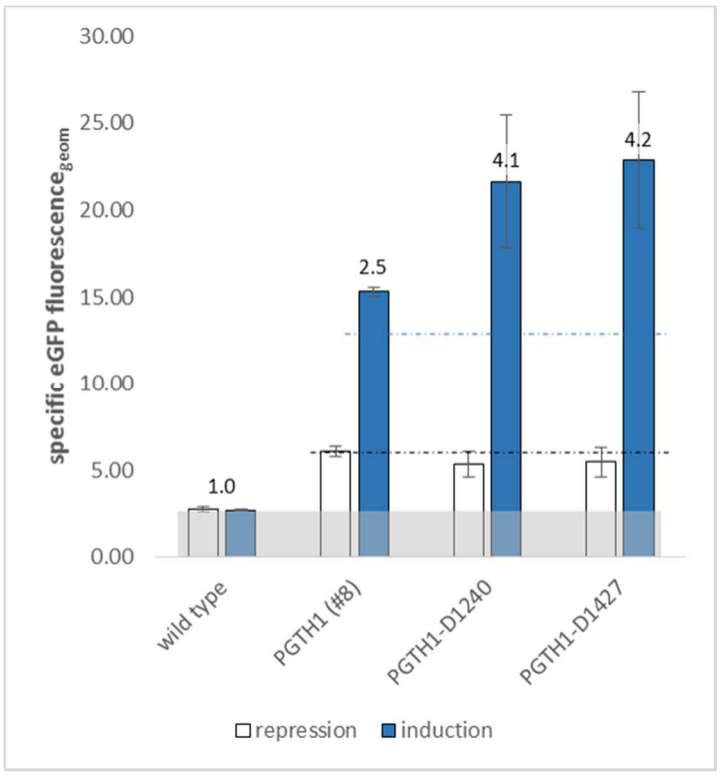

Figure 6: Screening data of the P_{GTH1} duplication variants

The geometric mean of the population´s specific eGFP fluorescence (fluorescence related to cell volume) is shown for clones expressing eGFP under the control of P_{GTH1} (clone #8, verified GCN of 1) or a P_{GTH1} variant (up to 9 clones were pool cultivated in 3 wells, selected in pre-screenings) in repressing and inducing growth conditions. Wild type *P. pastoris* cells were used as negative control.

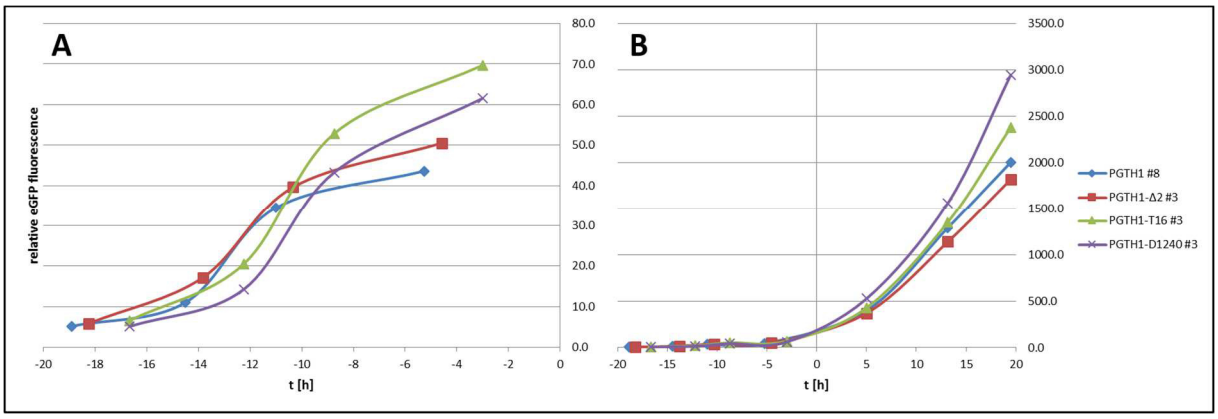

Figure 7: Fed batch cultivation of P_{GTH1} and P_{GTH1} variants expressing eGFP

Relative eGFP fluorescence was measured from bioreactor samples (diluted to similar biomass densities) using a plate reader and is shown over the feed time (batch end set to 0) in batch (A) and fed batch cultivation (B). A clone expressing eGFP under control of P_{GTH1} (#8) was compared to clones expressing under control of a P_{GTH1} deletion (P_{GTH1}-Δ2), a TAT15 mutation (P_{GTH1}-T16), and a duplication (P_{GTH1}-D1240) variant.

Additional Files

Additional file 1: TFBS identified in the P_{GTH1} promoter sequence using MatInspector. The color code in the left row matches the color code used in Figure 1 and 3. Targeted carbon source-related TFBS of the P_{GTH1} deletion variants are shown in bold and deleted positions are given in the right column.

Additional file 2: Fasta sequence files of the P_{GTH1} promoter and the generated P_{GTH1} variants. The Δ in file names was replaced by "_".

Additional file 3: Primer sequences

Additional file 4: Affected TFBS of the P_{GTH1} promoter sequence in the deletion mutants P_{GTH1}-Δ1 to Δ12. Sequence analysis was done using MatInspector from Genomatix. Glucose- and carbon- related TFBS which were selected for deletion are shown in bold and the corresponding ID (1-12) and deleted positions are stated in column 1 and 2. Detailed information about all TFBS of the P_{GTH1} sequence are given in **Additional file 2**.

Abbreviations

CSRE: carbon source response element, F$: fungi specific TF matrix, GCN: gene copy number, GOI: gene of interest, *Pp*: Pichia pastoris, *Sc*: Saccharomyces cerevisiae, TF: transcription factor(s), TFBS: transcription factor binding site(s), YDM: yeast dry mass

Authors' contributions

RP performed the experimental work, drafted the manuscript and contributed to study design. BG and DM contributed to data analysis, drafting the manuscript and to study design. All authors read and approved the final manuscript.

Acknowledgements

This project was financially supported by LONZA (Visp, CH).

References

Ahmad, M., M. Hirz, H. Pichler & H. Schwab, (2014) Protein expression in Pichia pastoris: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol* **98**: 5301-5317.

Ahn, J., J. Hong, M. Park, H. Lee, E. Lee, C. Kim, J. Lee, E.S. Choi & J.K. Jung, (2009) Phosphate-responsive promoter of a Pichia pastoris sodium phosphate symporter. *Appl Environ Microbiol* **75**: 3528-3534.

Boles, E. & C.P. Hollenberg, (1997) The molecular genetics of hexose transport in yeasts. In: FEMS Microbiol Rev. Netherlands, pp. 85-111.

Byrne, B., (2015) Pichia pastoris as an expression host for membrane protein structural biology. *Curr Opin Struct Biol* **32C**: 9-17.

Cartharius, K., K. Frech, K. Grote, B. Klocke, M. Haltmeier, A. Klingenhoff, M. Frisch, M. Bayerlein & T. Werner, (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* **21**: 2933-2942.

Delic, M., D. Mattanovich & B. Gasser, (2010) Monitoring intracellular redox conditions in the endoplasmic reticulum of living yeasts. *FEMS Microbiol Lett* **306**: 61-66.

Denis, C.L. & E.T. Young, (1983) Isolation and characterization of the positive regulatory gene ADR1 from Saccharomyces cerevisiae. *Mol Cell Biol* **3**: 360-370.

Gasser, B., R. Prielhofer, H. Marx, M. Maurer, J. Nocon, M. Steiger, V. Puxbaum, M. Sauer & D. Mattanovich, (2013) Pichia pastoris: protein production host and model organism for biomedical research. *Future Microbiol* **8**: 191-208.

Hartner, F.S., C. Ruth, D. Langenegger, S.N. Johnson, P. Hyka, G.P. Lin-Cereghino, J. Lin-Cereghino, K. Kovar, J.M. Cregg & A. Glieder, (2008) Promoter library designed for fine-tuned gene expression in Pichia pastoris. *Nucleic Acids Res* **36**: e76.

Hiesinger, M., S. Roth, E. Meissner & H.J. Schuller, (2001) Contribution of Cat8 and Sip4 to the transcriptional activation of yeast gluconeogenic genes by carbon source-responsive elements. *Curr Genet* **39**: 68-76.

Hohenblum, H., N. Borth & D. Mattanovich, (2003) Assessing viability and cell-associated product of recombinant protein producing Pichia pastoris with flow cytometry. *J Biotechnol* **102**: 281-290.

Kim, J.H., A. Roy, D. Jouandot, 2nd & K.H. Cho, (2013) The glucose signaling network in yeast. *Biochim Biophys Acta* **1830**: 5204-5210.

Kranthi, B.V., H.R. Kumar & P.N. Rangarajan, (2010) Identification of Mxr1p-binding sites in the promoters of genes encoding dihydroxyacetone synthase and peroxin 8 of the methylotrophic yeast Pichia pastoris. *Yeast* **27**: 705-711.

Lin-Cereghino, G.P., L. Godfrey, B.J. de la Cruz, S. Johnson, S. Khuongsathiene, I. Tolstorukov, M. Yan, J. Lin-Cereghino, M. Veenhuis, S. Subramani & J.M. Cregg, (2006) Mxr1p, a key regulator of the methanol utilization pathway and peroxisomal genes in Pichia pastoris. *Mol Cell Biol* **26**: 883-897.

Lubliner, S., L. Keren & E. Segal, (2013) Sequence features of yeast and human core promoters that are predictive of maximal promoter activity. *Nucleic Acids Res* **41**: 5569-5581.

Marsman, J. & J.A. Horsfield, (2012) Long distance relationships: enhancer-promoter communication and dynamic gene transcription. *Biochim Biophys Acta* **1819**: 1217-1227.

Mattanovich, D., P. Branduardi, L. Dato, B. Gasser, M. Sauer & D. Porro, (2012) Recombinant protein production in yeasts. *Methods Mol Biol* **824**: 329-358.

Mattanovich, D., A. Graf, J. Stadlmann, M. Dragosits, A. Redl, M. Maurer, M. Kleinheinz, M. Sauer, F. Altmann & B. Gasser, (2009) Genome, secretome and glucose transport highlight unique features of the protein production host Pichia pastoris. *Microb Cell Fact* **8**: 29.

Maurer, M., M. Kühleitner, B. Gasser & D. Mattanovich, (2006) Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with Pichia pastoris. *Microb Cell Fact* **5**: 37.

Mellitzer, A., C. Ruth, C. Gustafsson, M. Welch, R. Birner-Grunberger, R. Weis, T. Purkarthofer & A. Glieder, (2014) Synergistic modular promoter and gene optimization to push cellulase secretion by Pichia pastoris beyond existing benchmarks. *J Biotechnol* **191**: 187-195.

Narlikar, L. & I. Ovcharenko, (2009) Identifying regulatory elements in eukaryotic genomes. In: Brief Funct Genomic Proteomic. England, pp. 215-230.

Phillips-Cremins, J.E. & V.G. Corces, (2013) Chromatin insulators: linking genome organization to cellular function. *Mol Cell* **50**: 461-474.

Prielhofer, R., S.P. Cartwright, A.B. Graf, M. Valli, R.M. Bill, D. Mattanovich & B. Gasser, (2015) Pichia pastoris regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. *BMC Genomics* **16**: 1393.

Prielhofer, R., M. Maurer, J. Klein, J. Wenger, C. Kiziak, B. Gasser & D. Mattanovich, (2013) Induction without methanol: novel regulated promoters enable high-level expression in Pichia pastoris. *Microb Cell Fact* **12**: 5.

Qin, X., J. Qian, G. Yao, Y. Zhuang, S. Zhang & J. Chu, (2011) GAP promoter library for fine-tuning of gene expression in Pichia pastoris. *Appl Environ Microbiol* **77**: 3600-3608.

Rolland, F., J. Winderickx & J.M. Thevelein, (2002) Glucose-sensing and -signalling mechanisms in yeast. *FEMS Yeast Res* **2**: 183-201.

Roy, A., D. Jouandot, 2nd, K.H. Cho & J.H. Kim, (2014) Understanding the mechanism of glucose-induced relief of Rgt1-mediated repression in yeast. *FEBS Open Bio* **4**: 105-111.

Ruth, C., M. Buchetics, V. Vidimce, D. Kotz, S. Naschberger, D. Mattanovich, H. Pichler & B. Gasser, (2014) Pichia pastoris Aft1 - a novel transcription factor, enhancing recombinant protein secretion. *Microb Cell Fact* **13**: 120.

Stadlmayr, G., A. Mecklenbrauker, M. Rothmuller, M. Maurer, M. Sauer, D. Mattanovich & B. Gasser, (2010) Identification and characterisation of novel Pichia pastoris promoters for heterologous protein production. *J Biotechnol* **150**: 519-529.

Tschopp, J.F., P.F. Brust, J.M. Cregg, C.A. Stillman & T.R. Gingeras, (1987) Expression of the lacZ gene from two methanol-regulated promoters in Pichia pastoris. *Nucleic Acids Res* **15**: 3859-3876.

van Urk, H., E. Postma, W. Scheffers & J. van Dijken, (1989) Glucose transport in crabtree-positive and crabtree-negative yeasts. *J Gen Microbiol* **135**: 2399-2406.

Vogl, T., C. Ruth, J. Pitzer, T. Kickenweiz & A. Glieder, (2014) Synthetic core promoters for Pichia pastoris. *ACS Synth Biol* **3**: 188-191.

Weingarten-Gabbay, S. & E. Segal, (2014) The grammar of transcriptional regulation. *Hum Genet* **133**: 701-711.

Weinhandl, K., M. Winkler, A. Glieder & A. Camattari, (2014) Carbon source dependent promoters in yeasts. *Microb Cell Fact* **13**: 5.

Zhang, P., W. Zhang, X. Zhou, P. Bai, J.M. Cregg & Y. Zhang, (2010) Catabolite repression of Aox in Pichia pastoris is dependent on hexose transporter PpHxt1 and pexophagy. *Appl Environ Microbiol* **76**: 6108-6118.

Curriculum vitae

Mag. Roland Prielhofer

Born 4th April 1984 in Wels, Austria; Austrian Citizen, Residence in Vienna, Austria.
Email: roland.prielhofer@boku.ac.at

EDUCATION

2013	Short-term stay at Aston University in Birmingham, UK. Cooperation project on translational profiling in *Pichia pastoris* with the group of Prof. Roslyn Bill
2010 – present	PhD thesis at the Department of Biotechnology, University of Natural Resources and Life Sciences, in the group of Prof. Diethard Mattanovich Field of research: Strain engineering of *P. pastoris*
Jan. 2010	Studies finished with the Degree: Mag. rer. nat.
2009	Diploma thesis in the group of Univ.-Doz. Dr. Monika Bradl at the Center for Brain Research/Medical University Vienna, Division: Neuroimmunology, Head: Univ.-Prof. Hans Lassmann; Title: Dscam expression in rat microglia
2004 – 2010	Studies of Molecular Biology at the University of Vienna Emphasis: Immunology, Neuroscience and Bioinformatics
2003 – 2004	Civil service, Red Cross in Wels, Austria
1998 – 2003	Higher Technical Education Institute for chemical engineering in Wels, Austria

PUBLICATIONS

Prielhofer, R., M. Maurer, J. Klein, J. Wenger, C. Kiziak, B. Gasser & D. Mattanovich, (2013) Induction without methanol: novel regulated promoters enable high-level expression in *Pichia pastoris*. Microb Cell Fact 12: 5.

Gasser, B., **R. Prielhofer**, H. Marx, M. Maurer, J. Nocon, M. Steiger, V. Puxbaum, M. Sauer & D. Mattanovich, (2013) *Pichia pastoris*: protein production host and model organism for biomedical research. Future Microbiol 8: 191-208.

Prielhofer, R., S.P. Cartwright, A.B. Graf, M. Valli, R.M. Bill, D. Mattanovich & B. Gasser, (2015) *Pichia pastoris* regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. BMC Genomics 16: 167.

PATENT APPLICATIONS

Mattanovich D, Gasser B, Maurer M, **Prielhofer R**, Klein J, Wenger J (2012). Regulatable Promoter. US Patent 20140242636.

Mattanovich D, Gasser B, **Prielhofer R** (2013). Constitutive promoter. US Patent US20140274761.

CONFERENCE CONTRIBUTIONS

Prielhofer R, Maurer M, Klein J, Wenger J, Kiziak C, Gasser B, Mattanovich D **(2013):** Induction without methanol: Novel regulated promoters enable high- level expression in *Pichia pastoris*. RPP7, MAR 6-8, 2013, Laupheim, Germany **- Oral presentation**

Prielhofer R, Gasser B, Maurer M, Mattanovich D (**2012**): Identification and Characterization of novel promoters in *Pichia pastoris*. *Pichia* 2012, FEB 29 - MAR 3, 2012, Alpbach, Austria **- Poster**