Thesis for the degree of doctor of philosophy

Systems Biology of Recombinant Protein Production by Fungi

LIFANG LIU

Systems and Synthetic Biology

Department of Chemical and Biological Engineering

Chalmers University of Technology

Göteborg, Sweden 2014

Systems Biology of Recombinant Protein Production by Fungi

LIFANG LIU

PhD Thesis

Systems and Synthetic Biology

Department of Chemical and Biological Engineering

Chalmers University of Technology

SE-412 96 Göteborg

Sweden

Telephone +46 (0) 31-772 1000

Cover illustration

Schematic illustration of two major intracellular trafficking pathways

PREFACE

This dissertation serves as a partial fulfilment of the requirement to obtain the PhD degree at the Department of Chemical and Biological Engineering, Chalmers University of Technology, Sweden. The PhD research was performed in the Systems and Synthetic Biology group under the supervision of Professor Jens Nielsen. The work enabled systems biology approaches to investigate protein production mechanisms in two fungi. This work was funded by the European Research Council ERC project INSYSBIO (grant no. 247013), the Chalmers Foundation, and the Novo Nordisk Foundation.

Lifang Liu

May 2014

ABSTRACT

Systems biology has emerged as a highly potent tool for studying biological processes over the last decades. However, its application to complex metabolic processes such as protein secretion is still at the infant stage. *Saccharomyces cerevisiae* and *Aspergillus oryzae* are two important fungal cell factories which occupy significant proportions of recombinant protein productions, whereas various bottlenecks and undiscovered mechanisms limit their full potential as robust hosts.

In this thesis, systems biology approaches were applied to explore these two organisms in respect of protein production. By utilizing and engineering the yeast endogenous heme synthesis, we demonstrated the possibility for efficient production of complex proteins (*e.g.* multimer with a prosthetic group) by yeast. Applying inverse metabolic engineering, we identified many genomic variants that may contribute to improve protein secretion in yeast. Specifically, we examined the effect of a single point mutation on *VTA1* encoding a regulatory protein in the MVB pathway in endocytosis. Our result suggests that the *VTA1S196I* mutation might help to accelerate nutrient uptake via endocytosis, which subsequently enhanced protein synthesis and secretion. Oxygen is an important element associated with normal cellular metabolism as well as protein production. We studied how Rox1p, a heme-dependent transcription repressor of many hypoxia-induced genes, affect protein production in yeast, under aerobic conditions. By knocking out *ROX1*, we observed a 100% increase in the α-amylase production. Through genome wide transcriptome analysis we identified several Rox1p targets and based on this suggested their roles in improving protein productions. Lastly, applying comparative genomics study, we enriched the list of core protein components involved in the secretory machinery of *A. oryzae*. To verify the list, several high α-amylase producing strains were constructed. The transcriptional responses of these strains to α-amylase production were studied using microarray, through which several strategies including overexpressing the up-regulated cell wall proteins EglD and Cwp1 and knocking out the genes encoding extracellular proteins competing for the secretory pathway, were proposed.

LIST OF PUBLICATIONS

I.   Balanced globin protein expression and heme biosynthesis improve production of human hemoglobin in *Saccharomyces cerevisiae*.

Liu, L., Martinez, J. L., Liu, Z., Petranovic, D., Nielsen, J. Met Eng. 21(0): 9-16, 2014

II.   Systems Biology Analysis of Amylase Producing Yeast Strains.

Liu, Z*., **Liu, L***., Österlund, T* *et al*. Resubmitted after revision to AEM, 2014

III.   Improving heterologous protein secretion in aerobic conditions by activating hypoxia induced genes in *Saccharomyces cerevisiae*.

Liu, L., Liu, Z., Petranovic, D., Nielsen, J. Submitted to AEM, 2014

IV.   Genome-scale analysis of the high-efficient protein secretion system of *Aspergillus oryzae*.

Liu, L*., Feizi, A*., Österlund, T*., Carsten, Hjort., Nielsen, J. Submitted to BMC Systems Biology, 2014

Additional work during the PhD study but not included in this thesis

V.   Pharmaceutical protein production by yeast: towards production of human blood proteins by microbial fermentation.

Martinez, J. L.*, **Liu, L***, Petranovic, D, Nielsen, J. Curr Opin Biotechnol, 2012

VI.   Engineering oxygen sensing regulation results in an enhanced recombinant human hemoglobin production by *Saccharomyces cerevisiae.*

Martinez, J. L.*, **Liu, L***, Petranovic, D, Nielsen, J. Manuscript, 2014

*Equal contribution

CONTRIBUTION SUMMARY

I. DP and JN conceived the study. <u>LL</u> performed the experiments, analysed the data and wrote the manuscript. JM assisted experiments and writing. DP, JM and JN edited the manuscript.

II. JN conceived the study. LZ performed UV-mutagenesis and screening. TÖ performed genome sequencing and transcriptome analysis. <u>LL</u> verified mutant candidates and analysis the data. JN and DP supervised the study. <u>LL</u>, LZ and TÖ wrote the manuscript. JN and DP edited the manuscript.

III. <u>LL</u> conceived the study, performed the experiments, analysed the data and wrote the manuscript. LZ contributed the amylase overexpression plasmid. DP and JN supervised the project and edited the manuscript.

IV. <u>LL</u> performed the experiments and analysed the data. AF enriched the protein secretory component list. TÖ performed data analysis. CH supervised the strain construction. JN conceived the project and supervised the work. <u>LL</u>, AF, and TÖ wrote the paper. JN edited the manuscript.

V. JM and <u>LL</u> wrote the manuscript. DP and JN conceived the study and edited the manuscript.

VI. JM conceived the study, performed the experiments, analyzed the data and wrote the manuscript. <u>LL</u> co-performed the work. LL, DP and JN edited the manuscript.

TABLE OF CONTENTS

General introduction

1.1 Saccharomyces cerevisiae

The budding yeast *Saccharomyces cerevisiae* (will be referred to as yeast in this thesis except when otherwise specified) is one of the simplest forms of fungi belonging to the phylum of Ascomycota. Due to its extensive applications in preparation of bread and alcoholic beverages it is also known as Baker's and Brewer's yeast and therefore easily obtained the GRAS (Generally Regarded As Safe) status issued by the FDA (U.S. Food and Drug Administration). *S. cerevisiae* was the first approved genetically modified organism (GMO) to produce human food additives (calf chymosin for making cheese) (Walker, 1998) as well as the first GMO to be directly used in human food (beer) (Gopal and Hammond, 1992; Verstrepen *et al*., 2006). The applications of yeast have now been extended beyond food industries. As a facultative microbe, *S. cerevisiae* can grow both aerobically and anaerobically, and at the latter condition it produces ethanol as the dominant fermentation product. It has therefore been serving as the key biofuel cell factory to produce bioethanol at an industrial level. Another important application is the production of biopharmaceuticals (human insulin, hepatitis virus vaccines and human papilloma virus vaccines) (Martinez *et al*., 2012). In 2012, the sales of insulin and its analogs reached $18.92 billion, dominated by Novo Nordisk (40% market share) who uses *S. cerevisiae* as the leading producer (Huang *et al*., 2014). By 2009, around 18.5% of the industrially approved recombinant biotherapies are produced by *S. cerevisiae*, second to *Escherichia coli* which represents around 30% of the production (Ferrer-Miralles *et al*., 2009). Other applications, such as productions of biochemicals, biofuels (*e.g.* isopropanol, isoprenoids, flavonoids and terpenoids) (Chemler *et al*., 2006; Dugar and Stephanopoulos, 2011; Lee *et al*., 2008) and industrial enzymes (Hou *et al*., 2012c), are awaiting further developments.

Aside from being an important cell factory, *S. cerevisiae* has been serving as a model organism to study various cellular mechanisms in eukaryotes, for the facts that it i) can grow on defined medium which gives a good control of experiments and analysis; ii) has a short generation time, easy to manipulate and upscale; iii) has well-known genome and genetic engineering tools/techniques; iv) may share homologous proteins or mechanisms in other organisms that are of research or application interests. For example, studying Aβ (beta-amyloid peptide) trafficking in yeast may contribute to understand Alzheimer's disease in human (Treusch *et al*., 2011).

Being unicellular, *S. cerevisiae* was the first eukaryal organism whose genome was completely sequenced (Goffeau *et al*., 1996). Approximately 6000 genes on 16 chromosomes have been identified and the information is continuously updated in the yeast genome database (SGD). The yeast strains used in this thesis are all of the CEN.PK series (Entian and Kötter, 2007), which were designed to fit metabolic engineering applications in both

industrial and academic researches. The genome information of the wild type strain CEN.PK113-7D has been obtained recently (Nijkamp *et al*., 2012), which can be very helpful for our understanding and engineering the strain.

1.2 Aspergillus oryzae

Approximately half of the industrial enzymes are produced by filamentous fungi (Lubertozzi and Keasling, 2009; Machida *et al*., 2008). The koji mold *Aspergillus oryzae* has been widely used to produce amylase, glucoamylase, protease and lipase, *etc*. Actually, the first patent on microbial enzyme production was awarded to *A. oryzae* for producing Taka-diastase (U.S. Patent 525,823, in 1895) and the first example of producing commercial heterologous enzymes, a lipase for laundry detergent (Christensen *et al*., 1988) was also by *A. oryzae*. The fungus also has a long application history in food fermentations *e.g.* to make miso (soybean paste), sake (rice wine) and shoyou (soy sauce) in Asia, and therefore holds the GRAS status as well.

As other filamentous fungi, *A. oryzae* forms multinucleate conidia which further develop into tubular hyphae during germination. The hyphae are composed of cylindrical compartments separated by perforated walls called septa. Protein secretion has been indicated to happen mainly at hyphae tips as the organelles essential for protein secretion such as the endoplasmic reticulum (ER) and the Golgi apparatus were both found dominantly in apical regions (Kuratsu *et al*., 2007; Maruyama and Kitamoto, 2007). Furthermore, the ER in *A. oryzae* is often found in the cell periphery of subapical regions and near septa (Maruyama and Kitamoto, 2007), suggesting the occurrence of protein secretion at septa in addition to hyphae tips.

A. oryzae lacks a sexual life cycle, even though many genes are implicated in the fungal mating process and it contains an alpha mating-type gene in the genome (Galagan *et al*., 2005). It has been noticed that expression vectors cannot be stably retained over long time of production, and it is less stringent for the plasmid-less nuclei to tolerant severe growth defect conditions in the coenocytic fungal hyphae. Genetic modifications are therefore usually achieved through chromosomal integration, and here the dominant mechanism is no-homologous end joining (NHEJ) rather than the homologous recombination (HR) mechanism dominant in yeast. By NHEJ the expression cassette with selection marker is randomly integrated into the fungal chromosome in tandem copies (usually 2-10 copies). These characteristics have greatly prevented extensive genetic studies on *A. oryzae* using conventional approaches, and this explains why the biological knowledge on this commercially important fungus is still relatively limited. Sequencing of the wild type *A. oryzae* RIB40 genome was completed in 2005 (Machida *et al*., 2005). The genome consists of eight chromosomes with an entire genome size of 37.6 Mb (12,063 ORFs) (Machida *et al*., 2008) which is relatively bigger than that of several other *Aspergilli*, including *Aspergillus nidulans* (30.1Mb, 10,586 ORFs) (Galagan *et al*., 2005), *Aspergillus fumigatus* (29.4MB,

2

9,887 ORFs)(Nierman *et al*., 2005) and *Aspergills niger* (34.85 Mb, 14,086 ORFs) (Pel *et al*., 2007). The larger genome size is mainly reflected in the number of metabolic genes, including secretory hydrolases, transporters and primary and secondary metabolism genes, among which the expansion of secondary metabolism genes is most prominent (Machida *et al*., 2008). By constructing a Expressed Sequence Tag (EST) library, Vongsangnak *et al* enriched the numbers of the predicted genes in the *A. oryzae* genome and by using the improved annotated genome they constructed the first genome scale metabolic model of *A. oryzae* in 2008 (Vongsangnak *et al*., 2008).

II. Recombinant Protein Production

The total biopharmaceutical sales reached about $125 billion in 2012 (Huang *et al*., 2014) which account for approximately 40% of the total pharmaceutical sales nowadays (Martinez *et al*., 2012). With a predicted annual increase of 6.8% the market for industrial enzymes is expected to reach $8 billion in 2015 (Group, 2011). Both *S. cerevisiae* and filamentous fungi occupy significant proportions of recombinant protein productions (RPPs) (Figure 1), whereas various bottlenecks and undiscovered mechanisms limit their full potential.

In this thesis, we aimed to increase RPPs in the two fungal cell factories mentioned above, from which we expected to get novel insights into the mechanisms involved in protein production (especially secretion) which hopefully could provide guidance for further strain engineering.
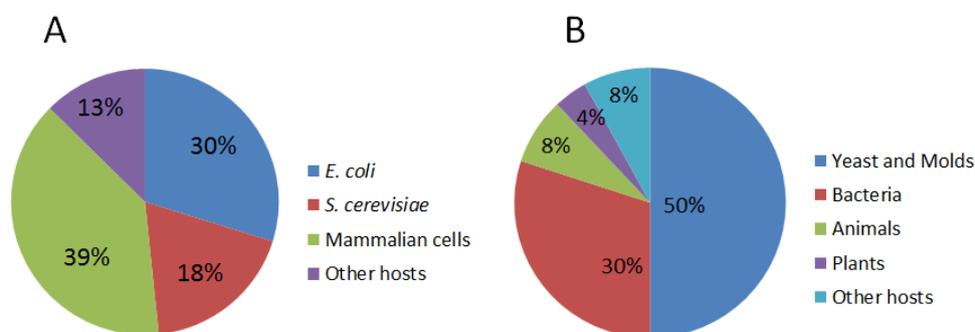
Figure 1. Recombinant protein productions in different host systems. (A) Biopharmaceuticals licensed by FDA and EMEA by 2009 (Ferrer-Miralles *et al*., 2009)(B) Industrial enzymes (Demain and Vaishnav, 2009).

2.1 Secretory protein production: α-amylase as example

α-amylase from *A. oryzae* was selected as the main model protein in this thesis to study protein production mechanisms, not only because it is an industrially important enzyme of great commercial value, but more importantly because it carries several characteristics (3 domains, 478 amino acids, 4 disulfide bonds, and 1 glycosylation site, *etc*) (Randez-Gil and Sanz, 1993) for studying protein secretion. Processes dealing with these post translational modifications (PTMs) are commonly considered as being bottlenecks for protein secretion.

2.1.1 Endocytosis

There are two pathways transferring macromolecules, especially proteins and lipids, in and out of the cells, namely exocytosis and endocytosis, respectively (Figure 2). Endocytosis occurs at the cell surface where cargos (proteins, lipids, and other molecules) from the extracellular milieu and plasma membrane (PM) are internalized and delivered through a set of endosomes to the vacuole for degradation (Segev, 2009). In addition to nutrient uptake, endocytosis also serves to regulate various processes initiated at the cell surface including

4

signalling from cell surface receptors, removing damaged proteins from the membrane and recycling of membrane proteins involved in exocytosis back to the secretory pathway (Lewis *et al*., 2000) (Sudha Kumari and Mayor, 2010). Cargo internalization from PM and sorting into the early endosome is mediated by clathrin and actin in *S. cerevisiae*. The process has been well described by Toret *et al* (Toret and Drubin, 2006). Subsequent cargo sorting to the late endosome is achieved through a sequential recruitment of three protein complexes called endosomal sorting complex required for transport (ESCRTs, I to III) to the late endosome membrane and the specific interactions with the associated proteins. The resulting inward invaginated form of late endosome is called multivesicular body (MVB) and therefore the process is called the MVB pathway. Sorting of cargo proteins into the MVB pathway is ubiquitin-dependent (Katzmann *et al*., 2001). At least 24 gene products have been implicated in the MVB sorting pathway (Saksena *et al*., 2007). These proteins are referred to as class E Vps (vacuolar protein sorting) proteins. Deletion of class E VPS genes in *S. cerevisiae* has resulted in mislocalization of MVB cargoes to the vacuole membrane surface (Raymond *et al*., 1992) and accumulation of endosomal cargoes in large aberrant structures next to the vacuole (Vida *et al*., 1993). Deletion of class E VPS genes in *A. oryzae* results in defective vacuole formation, impaired hyphal growth and conidiation (Tatsumi *et al*., 2006). The exocytic and endocytic pathways are connected by bi-directional transport between the Golgi and the endosomes (Conibear and Stevens, 1998). Exocytic v-SNARE Snc1p which mediates fusion of exocytic vesicles with the PM has been proven to be recycled by endocytosis (Lewis *et al*., 2000) implicating that endocytosis is a crucial process for protein production, in terms of supporting secretion (Shoji *et al*., 2008).

In contrast, the exocytic pathway carries proteins and lipids from the ER through Golgi apparatus to the PM, following which the secretion of recombinant proteins usually happens. The exocytic pathway is regarded as secretion in general in this thesis, and is discussed in details in below sections.
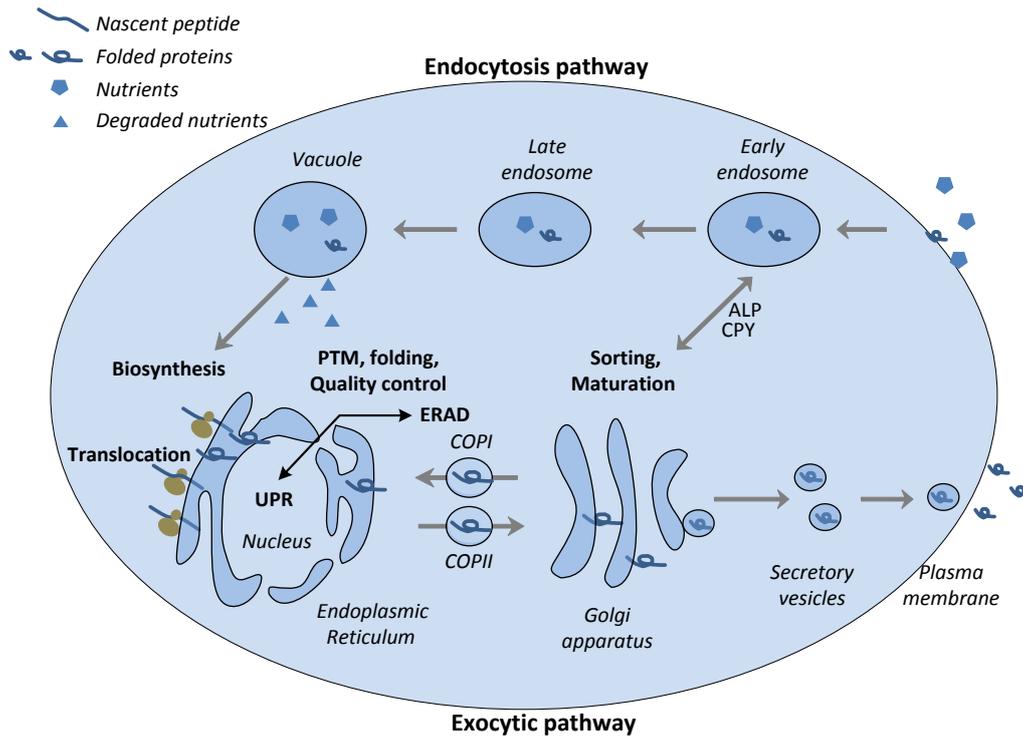
Figure 2. Two major intracellular trafficking pathways. The endocytic pathway internalizes cargos from the cell milieu or the periplasmic space through a series of endosomes and finally degrades the cargos in the vacuole. The exocytic pathway carries proteins and lipids from the ER through the Golgi to the PM. The steps include translocation, post-translational modifications, ER quality control, Golgi processing, sorting, vesicle trafficking and tethering *etc*. The endocytic and exocytic pathways are interconnected.

2.1.2 Transcription, translation and protein synthesis

According to the central dogma, protein production is initiated by transcription followed by translation of mRNA resulting in assembly of amino acids into polypeptides with the assistance from tRNA following the genetic codes on mRNA on the ER-bound ribosomes. Therefore, modulating transcription and translation efficiency is usually the first concern when designing expression systems for producing heterologous proteins, among which strategies such as i) selecting strong promoters (Liu *et al*., 2012), ii) increasing gene copy numbers (Verdoes *et al*., 1995), iii) overexpressing relevant transcription factors (Nabel and Baltimore, 1987), and iv) engineering the 5' untranslated regions to promote translation initiation (Koda *et al*., 2004) *etc* are included. It is worth mentioning that strong promoter and high copy numbers usually results in higher protein production (Verdoes *et al*., 1995), however, due to reasons such as saturation of the transcription factor (Kelly and Hynes, 1987), overloading to the secretory pathway, and reduced cell growth due to protein production induced stress, the volumetric yield may not always be satisfactory with this strategy. More studies have been therefore focused on the ER and the post ER secretory mechanisms.

2.1.3 ER processing

After translation, newly synthesized polypeptides targeting for secretion have to pass through several essential steps before getting matured and reaching their final destinations among which ER is the first checkpoint where a set of covalent modifications such as signal sequence removal, disulfide bond formation, N-glycosylation, GPI anchoring, sorting and degradation (Dobson, 2004; Idiris *et al*., 2010) happen.

Signal sequence removal

Leader sequences are very important for guiding protein along the secretory pathway. The *S. cerevisiae* α-mating factor prepro leader has been identified as a better leader for secreting α-amylase than the synthetic Yap3-TA57 leader (Liu *et al*., 2012) and therefore chosen to guide the amylase secretion in the yeast strains in this thesis. The pre-leader is responsible for directing the newly synthesized peptide through translocation into the ER and the pro-leader is supposed to increase the solubility of the recombinant protein (Kjeldsen *et al*., 1999) and the trafficking efficiency through the inter-organelle transport and vacuolar targeting (Rakestraw *et al*., 2009). The pre leader sequence is usually cleaved off after exiting ER, while the pro leader sequence is cleaved by the Kex2p proteinase in the Golgi apparatus in the late secretory pathway, and more often the cleavage is not complete (Kjeldsen, 2000). Secretion in *S. cerevisiae* usually results in hyper-glycosylation and therefore leader sequences are often mutated to reduce the amount of hyper-glycosylated proteins (Kjeldsen *et al*., 1998). Screening a library of mutant α-mating factor leader peptides has resulted in up to 16-fold enhanced secretion of a single-chain antibody fragment (scFv) (Rakestraw *et al*., 2009). Combining with strain engineering, a 180-fold improvement in the secretion of full-length, glycosylated functional human IgG1 was achieved (Rakestraw *et al*., 2009).

Protein folding

Being translocated into the ER through the Sec61p translocon either co- or post-translationally, the nascent proteins are either bound to the ER chaperone calnexin (encoded by *CNE1* in yeast) or the protein binding protein BiP (encoded by *KAR2* in yeast) for folding to their native configurations. It is very rare that the protein binds simultaneously to both calnexin and BiP and it seems the closer the N-glycan sites to the N-terminus of the protein, the higher the tendency to use calnexin folding system (Molinari and Helenius, 2000).

i)   **CNX cycle**: A preformed oligosaccharide (Glc3Man9GlcNAc2) is attached to the asparagine side chains on the Asparagine-X-Serine/ Threonine motif of the nascent peptide. The three glucose moieties are cleaved off by glucosidase I and II, and then one glucose residue is added back by UDP-glucose glycoprotein glucosyltransferase (UGGT), the resulted monoglucosylated (Glc3Man9GlcNAc2) protein can then be retained in the ER by interacting with calnexin which has lectin activity through the glucose residue until the protein is properly folded and further glycosylated.

Glucosidase II then removes the terminal glucose by which the calnexin is released for another round of inspection (Ruddock and Molinari, 2006).

ii) **BiP folding system**: similar to the CNX system, BiP binds transiently to the newly synthesized polypeptides to prevent them from misfolding until they are properly folded. Belonging to the Hsp70 (heat shock protein of 70 KDa) molecular chaperone family, BiP consists of an N-terminal ATPase domain and a C-terminal substrate binding domain that has affinity for hydrophobic patches (Flynn *et al*., 1991). A conformational change on Bip occurs when ATP is hydrolyzed to ADP through which the substrate (peptide) is released. Substrates can undergo cycles of BiP binding and release through ATP hydrolysis. During each round of dispatching, the substrate has the opportunity to undergo the folding pathway until it no longer presents any BiP binding motif (Gething, 1999). BiP's ATPase activity is regulated by DnaJ co-chaperones and other cofactors including Sil1p, Lhs1p, Scj1p, Jem1p, overexpression of which have shown increased secretion of recombinant human albumin, human transferrin and granulocyte-macrophage colony-stimulating factor in *S. cerevisiae* (Payne *et al*., 2008).

iii) **Disulfide bond formation**: Correct disulfide bond formation further increases the stability of the native protein conformation. Missing or mispairing produces severe misfolded proteins and induces stress. Illustrated in Figure 3, the process is driven by the flavine adenine dinucleotide (FAD) bound Ero1p. FAD is synthesized in the cytosol and enters the ER freely. Depletion of riboflavin in yeast results in a severe defect in oxidative folding. FAD bound Ero1p oxidizes protein disulfide isomerase (PDI) which then subsequently oxidizes the dithiols on the folding protein to form disulfide bonds. Electrons are then transferred from PDI to Ero1p and finally to molecular oxygen during which reactive oxygen species (ROS) may be generated (Bader *et al*., 1999). Ero1p possesses seven conserved cysteine residues, and PDI contains two thioredoxin like active sites, both of which are likely to be involved in the catalysis of electron transfer. There are four PDI homologs in yeast, which are Eug1p, Mpd1p, Mpd2p and Eps1p. In addition to PDI, these homologs may also perform disulfide isomerization or reduction with the assistance from reduced glutathione (GSH) resulting in the oxidative form of glutathione (GSSG) (Tu and Weissman, 2004). *In vitro* experiment suggested that Ero1p can discriminate between PDI and the homologs. Besides, there is an alternative folding enzyme parallel to Ero1p, namely Erv2p, Detailed mechanisms by Erv2p has not been clearly defined yet (Tu and Weissman, 2002).
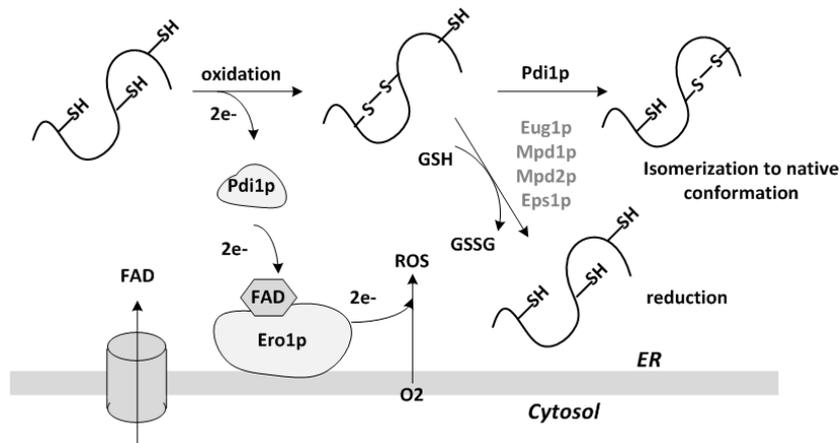
Figure 3. Disulfide bond formation catalyzed by PDI and Ero1p in yeast.

Unfolded protein response

Prolonged binding to either calnexin or the BiP complex causes repetitive rounds of oxidative protein folding attempts, resulting in aggregated ROS and stress in the ER. To cope with this stress and ensure survival of the cell, a signal is sent across the ER membrane into the nucleus and the cytoplasmic space inducing the so called unfolded protein response (UPR). Genes regulated by UPR involve in many functions associated with the ER, including protein translocation and glycosylation, lipid and inositol metabolism, protection against oxidative stress, or even later stages of secretion, including ER–Golgi transport, Golgi glycosylation, vacuolar targeting and exocytosis, in addition to degradation of the mis-folded proteins from the ER (Kimata *et al*., 2006; Saloheimo *et al*., 2003; Travers *et al*., 2000). Many genes involved in these processes or functions are up-regulated by UPR (Kimata *et al*., 2006). In contrast, some genes encode cell surface or extracellular proteins were found down-regulated by UPR in yeast. The authors suggested that this UPR-mediated gene repression may alleviate the load of client proteins targeted to the ER (Kimata *et al*., 2006). Actually, down-regulation of genes encoding secreted enzymes in response to secretion stress (RESS) is more frequently seen in filamentous fungi (Al-Sheikh et al., 2004; Pakula et al., 2003) than in yeast, as they produce much more secreted proteins. Some study indicated the RESS effect may be independent of UPR (Al-Sheikh et al., 2004). The folding stress causing UPR can be experimentally induced, *e.g.* by tunicamycin (blocks protein glycosylation) and dithiothreitol (DTT) (impairs disulfide bond formation).

The UPR utilizes a unique regulatory mechanism not known in any other signal transduction pathway. In yeast three proteins are required for the activation of the UPR, namely Ire1p, Hac1p and Rlg1p. Ire1p is a transmembrane kinase with the N-terminus in the ER lumen and the carboxyl terminus in the cytosol. The N-terminus of Ire1p senses the stress in ER and in the event of oxidative stress the protein gets oligomerized. This enables the middle kinase domain to activate the endonuclease in the tail domain, which cuts off the intron in the *HAC1* mRNA. Rlg1p ligates the two exons of the *HAC1* mRNA so that it can be translated into

Hac1p, which translocates to the nucleus and activates the up-regulation of UPR target genes (Kaufman, 1999; Patil and Walter, 2001; Sidrauski and Walter, 1997). The Hac1p binding region (UPRE) has been found in the ER-resident chaperones, *e.g. KAR2*, *PDI*, and *FKB2* (Mori *et al*., 1998). In the filamentous fungi *Trichoderma reesei*, *A. nidulans* and *A. niger*, *hacA* (homolog to yeast *HAC1*) undergoes an analogous splicing reaction of a shorter (20 nucleotides) intron during UPR. In addition, the *hacA* mRNA is truncated at the 5' flanking region during UPR induction (Saloheimo *et al*., 2003).

ER-associated degradation

The CNX and BiP systems ensure that only correctly folded, modified and assembled proteins travel further along the secretory pathway. If the UPR is insufficient to relieve the stress, misfolded or unmodified proteins are eventually eliminated by the ER-associated degradation (ERAD) system, through which they are relocated into the cytosol via the translocon (Sec61) or/and Der1p (Ye *et al*., 2004), and subsequently labelled by ubiquitin and finally recognized and degraded by the 26S proteasome (Meusser *et al*., 2005). The mark of the substrate by ubiquitin requires a Ub (ubiquitin) activating enzyme, a Ub conjugating enzyme and a Ub ligase (Doa10p) in addition to the ubiquitin itself (Meusser *et al*., 2005). In *S. cerevisiae*, Hrd1p-Hrd3p complex together with Sec61p mediate transferring of the targeted proteins for degradation (Gauss *et al*., 2006).

Up-regulation of essential components in ERAD pathway has been noticed in response to higher secretion of Fab fragment in *P. pastoris* (Gasser *et al*., 2007). Impairing *derA* a putative ERAD component in *A. niger* resulted in a 6-fold increase in the intracellular amount of glucoamylase–glucuronidase (a recombinant fusion protein), indicating a delay in protein degradation (Carvalho *et al*., 2011).

Engineering ER residential proteins

The effect of overexpressing the molecular chaperones (BiP, calnexin, and assisting chaperons), folding enzymes (Pdi1p, Ero1p), and the UPR regulator (Hac1p) have been reviewed by Idiris *et al* (Idiris *et al*., 2010). In general, the overexpression effects are protein or host specific meaning that in some cases, protein secretion can be significantly increased (BiP in (Harmsen *et al*., 1996) and the BiP chaperones in (Payne *et al*., 2008)), while in other cases, it had no or even negative effects (BiP (van der Heide *et al*., 2002)).

Glycosylation

Glycosylation appears to be important for pharmaceutical proteins as well as industrial enzymes as their activities or behaviours are much dependent on the modification patterns. Approximately 70% of the approved therapeutic proteins are glycoproteins.

The initial N-glycosylation is rather conserved which starts from the ER where a core N-glycan (2N-acetylglucosamines and 5 mannose residues) is transferred to the nascent

polypeptide on the asparagine residue of the consensus Asparagine-X-Serine/ Threonine motif, with X representing any amino acid but not Proline (Bause, 1983). The N-glycosylation is completed by the ER-resident oligosaccharyl transferases (OSTs). Another variety is the O-linked glycosylation which occurs at the hydroxyl groups of serine and threonine and is catalysed by O-mannosyltransferases (PMTs) (Hou *et al.*, 2012c). After ER quality control, the nascent protein is transferred to the Golgi where the glycan pattern is further processed. From this step, the glycosylation processing differs significantly between yeast and human being that the yeast glycan chain is less complex and rich of mannose. This confers a short half-life *in vivo* making yeast less suitable for expressing therapeutic glycoproteins (Wildt and Gerngross, 2005). Attempts to prevent hypermannosylation in yeast by inactivation mannosyltransferases Och1p and Mnn1p resulted in severe growth defect in *S. cerevisiae* (Zhou *et al.*, 2007). Currently, humanization of glycosylation pathway has been mainly performed in *P. pastoris* and in combination with other secretion and folding related engineering strategies, high level secretion of up to 1g/L full-length antibodies with correct human glycosylation was achieved (Potgieter *et al.*, 2009).

Glycosylation is another important factor determining folding state in addition to disulfide isomerization, and if the protein fails to achieve correct conformation, a specific 1, 2 α-mannosidase (mns1p) will remove the 1,2 α-mannose units and target the substrate to ERAD (Carvalho *et al.*, 2011). Htm1p, Yos9p and Mnl1p have also been shown to target the de-mannosylated proteins to ERAD (Clerc *et al.*, 2009).

2.1.4   Post ER processing

After successful folding and assembly in the ER, the secretory proteins will be recognized and exported by the coat protein complex II (COPII) to membrane-bound vesicles, which further travel the cargos to the Golgi. In the Golgi, the protein is further matured, including hyper-mannosylation and signal peptide removal *etc*. After Golgi processing, proteins can be sorted to different directions, such as retrograde to ER, transport to early and late endosome, to vacuole, to plasma membrane and to extracellular space, depending on their properties. Retrograde from the Golgi to ER is important for retrieval of the escaped ER resident proteins and return the unprocessed cargo proteins. The process is mediated by the COPI proteins. At each trafficking step, the membrane of the transport vesicles fuses to the target membrane and then delivers the cargo proteins through the assistance from a series of soluble N-ethylmaleimide-sensitive factor receptor (SNARE) complex proteins (Jahn and Scheller, 2006). Engineering of the SNARE proteins Sly1p and Sec1p has been shown to enhance α-amylase secretion. Insulin precursor secretion was improved by overexpressing Sec1p but not Sly1p (Hou *et al.*, 2012b). Over-expression of Sso1p and Sso2p increased amylase secretion in yeast as well (Ruohonen *et al.*, 1997).

2.2 Complex protein production: Human Hb as example

Cytoplasmic expression may lead to very high expression level of proteins as potential limitation for secretion pathways such as strict quality control and slow translocation between organelles are excluded (Mattanovich *et al*., 2012). One example was the very high expression of (22 g/L) rubber tree hydroxynitrile lyase from high cell density *P. pastoris* culture (Hasslacher *et al*., 1997). However, there are several obvious disadvantages such as efforts associated with breaking cells down and additional downstream processing, and the formation of insoluble aggregates where the protein is inactive (Binder *et al*., 1991), *etc*.

In comparison to protein production through the secretory pathway, we investigated the possibility of producing complex protein with human hemoglobin as example in *S. cerevisiae* due to i) human hemoglobin has potential medical application as blood substitute ii) yeast endogenous heme can be utilized to form complete proteins and therefore iii) it represents a good example for producing heme bound proteins *e.g.* catalases and cytochrome P450 enzymes that have various applications by yeast.

III.   Systems Biology Platforms

The complexity of the cellular machineries, in terms of composition and regulation, has always been a challenge for understanding the biological problems of a living organism (Papini *et al.*, 2010). During the last decades, with extensive developments in molecular biology, especially genome sequencing and high throughput experimental technologies, it became possible to collect comprehensive data and get holistic insights into the target system, based on which the concept of systems biology was thus emerged (Kitano, 2002). The consensus definition over many interpretations of systems biology is the global study of biological processes at cell, organism or even community level in order to illuminate the linkage between molecules and large-scale physiology (Barrett *et al.*, 2006; Mustacchi *et al.*, 2006).

Depending on the objective of a study, one can choose from two types of systems biology approaches, namely top-down and bottom-up respectively. The top-down systems biology is a data-driven process in a sense that biological information is extracted from large data sets including different omics data obtained from high throughput analysis. In contrast, the bottom-up systems biology is based on detailed or specific knowledge, which can be translated into mathematical formulation (model) to simulate the biological system or to verify an existing model. Even though bottom up modeling is usually only able to describe a subset of the complete biological system due to lack of quantitative information (components and interactions) it has been used for long before the top-down method was developed with the availability of large scale data sets. Today, the two approaches are often applied jointly to generate hypothesis for improving strain behaviors (Nielsen and Jewett, 2008) including cell factories used to produce recombinant proteins.

3.1 Bottom-up vs. protein production

There have been many mathematical models constructed by the means of bottom up approaches for simulating the processes involved in protein production. Dong *et al*. studied how the location and spacing of the rare aminoacyl-tRNA (which was termed slow codons) can dramatically affect the protein production rate (Dong *et al*., 2007). A series of studies focused on how glycosylation (site occupancy, oligosaccharide modifications) and co-/ post-translational processes are interacted and affect glycoprotein's secretion (Shelikoff *et al*., 1996; Umaña and Bailey, 1997). Raden *et al*. constructed a kinetics model describing the functions of the key enzymes Ire1p and Kar2p during the UPR (Raden *et al*., 2005). More studies tried to correlate strain growth with heterologous protein production (Sidorenko *et al*., 2008; Tholudur *et al*., 1999), in which metabolic flux analysis was usually taken into consideration in addition to the production of related parameters. Feizi *et al* recently constructed the first genome-scale model for the yeast secretory pathway, where 163 secretory components are divided into 16 subsystems based on their PTM features. The

metabolic requirements associated with processing of each specific protein in the secretory machinery can be estimated using the model (Feizi *et al*., 2013).

3.2 Top-down vs. protein production

Entering the genomics and functional genomics era, the application of mathematics models has gone beyond the hypothesis driven scope, where it can also be applied to analyze global data in addition to serve as a scaffold to extract information (Nielsen and Jewett, 2008). The global data is usually achieved through high-throughput experimental techniques involved in different omics studies.

Genomics studies, being advanced with the development of high-speed DNA sequencing technologies, enable identification and analysis of an entire cellular network on the fundamental genome level. For strain improvement, it is often applied together with random or evolved mutagenesis followed by intelligent screening processes and assisted with other omics analysis, such as whole genome transcriptome profiling (DNA microarray, RNA seqencing), proteome profiling (2D-gel electrophoresis combined with mass spectrography), metabolome and fluxome profiling (isotopically labeling substrates in combination with GC/MS, LC/MS and NMR technologies) *etc* (Lee *et al*., 2005). The metabolome and fluxome only gives valuable information if rates of individual reaction steps can be assessed. Due to the complexity with the protein secretory pathway, transcriptomics and proteomics are more frequently applied to understand the cellular responses to protein productions (Vijayendran and Flaschel, 2010).

Owing to the well-annotated genome, transcriptome analysis is often applied to study RPPs in yeast. By studying the transcriptional responses to DTT induction, Travers *et al* revealed an intimate coordination between the UPR and the ERAD responses. They found that efficient ERAD requires intact UPR and UPR induction increases ERAD capacity. Furthermore, loss of ERAD leads to constitutive UPR induction, and simultaneous loss of ERAD and UPR greatly decreases cell viability (Travers *et al*., 2000). Overexpressing *HSF1* that is essential for the heat shock response (HSR) resulted in significant improvement on recombinant insulin precursor and amylase productions. Transcriptome analysis indicated that the ER stress was relieved through *HSF1* overexpression, mainly through up-regulation of protein folding related genes and repression of transcription and translation in yeast (Hou *et al*., 2012a). It has been shown that amylase production was favored when cultured anaerobically. Through reporter metabolite analysis based on the transcriptome data, the authors proposed fumarate as the electron acceptor for protein folding under anaerobic conditions in yeast (Liu *et al*., 2013).

Research on *A. oryzae* has been mainly focused on genomics, comparative genomics, proteomics and secretomics, as there are still many un-annotated (<10%) ORFs in the genome. By comparison with similar and annotated species and performing secretome or

proteome analysis, more analogous genes or proteins may be identified. Genome analysis has been demonstrated in many papers trying to elucidate the high protein secretion capacity of *Aspergilli* (Kobayashi *et al*., 2007; Machida, 2002; Machida *et al*., 2008; Pel *et al*., 2007). Putative genes are mainly predicted based on their homologies to known genes in the public databases. Specifically, Geysens *et al* studied the processes of protein folding, secretion stress and glycosylation on the gene levels through genomics comparison among four different fungi, including *A. niger*, *A. nidulans*, *S. cerevisae*, and *P. pastoris* (Geysens *et al*., 2009). Considering that most cellular metabolic activities are directly or indirectly mediated by proteins, proteome profiling allows a systematic understanding of the molecular events occur in an organism under various physiological states. Even though, due to technique and cost related issues, not all proteins have been identified yet. The state of proteome profiling in *Aspergilli* has been reviewed by Kim *et al* (Kim *et al*., 2007; Kim *et al*., 2008).

Microarrays for *Aspergilli* have not yet been designed by commercial companies and therefore the research has been mainly driven by university groups that have designed arrays (Andersen and Nielsen, 2009). Sims *et al*. mapped 20 known genes of *A. niger* involved in the secretory pathway to identify homologous sequences in *A. nidulans*. The putative exons of these sequences were examined by microarray, through which hypothetical orthologs of the ER chaperones were found to be up-regulated. The transcriptional level of these genes were similar to that of the *A. niger* orthologs (Sims *et al*., 2004). Through comparative genomics and transcriptomics studies Vongsangnak *et al* found that the MAL locus responsible for maltose utilization in *A. oryzae* does not exist in *A. niger*. They report that *A. niger* requires a different regulatory system that involves the AmyR regulator for glucoamylase induction to utilize maltose (Vongsangnak *et al*., 2009).

As a substitute to microarray, high-throughput RNA-sequencing (RNA-Seq) technology has become a powerful and cost-efficient tool for transcriptome analysis. It is able to i) accurately quantify gene expression levels at low background; ii) detect a larger dynamic range of gene expression levels than microarray which lacks sensitivity for very low or very high expressions; iii) reveal precisely the boundaries of untranslated regions and sequence variations such as alternative splicing (AS) and gene fusion (Wang *et al*., 2010). Wang *et al* applied RNA-seq to study the *A. oryzae* transcriptome, through which they identified numerous novel transcripts, newly identified exons and untranslated regions. By finding differentially expressed genes involved in protein secretion system of *A. oryzae* in solid and liquid cultures, they gained further insight into the protein secretion pathway (Wang *et al*., 2010).

IV.   Case studies

4.1 Paper I. Recombinant hemoglobin production in yeast

Hemoglobin is a tetramer metalloprotein essential for oxygen transport in human red blood cells. The protein is comprised of two α and two β subunits with each subunit wrapping a prosthetic heme molecule. Heme is comprised of a protoporphyrin IX and a ferrous iron which has high affinity for binding oxygen (Belcher *et al*., 2010; Benesch and Benesch, 1963). The structure of protoporphyrin IX is common among divergent species such as human and yeast, and therefore one can utilize the endogenous yeast heme to produce recombinant human hemoglobin (Kumar, 1995). Expression of recombinant hemoglobin has been successfully demonstrated in different organisms including S. *cerevisiae* and *E. coli* (Behringer *et al*., 1989; Coghlan *et al*., 1992; Dieryck *et al*., 1997; Hoffman *et al*., 1990).

There are basically two problems associated with producing recombinant hemoglobin in yeast. One is the insufficiency of the heme cofactor. Incorporation of the heme molecule has been suggested to happen co-translationally, and could therefore facilitate the formation of the proper tertiary structure on the ribosome (Komar *et al*., 1997). In addition to heme supply, balancing the expression ratio of the α and β globins is also crucial as insufficiency of either peptide will prevent timely assembling and result in degradation. Eight steps are involved in the heme biosynthesis pathway (Figure 4) with *HEM2*, *HEM3* and *HEM12* been reported rate-limiting in different experiments (Hoffman *et al*., 2003; Keng, 1992).

In Paper I, we applied a metabolic engineering strategy to improve endogenous heme level by overexpressing the rate-limiting enzymes in the heme biosynthesis pathway. The expressions of the α and β globin peptides were also balanced in combination with the elevated heme synthesis.

Expression of the heme genes (*HEM2*, *HEM3* and *HEM12*) and the globin genes (*HBA* and *HBB*) were controlled by promoters of varied strengths on separate vectors harboring different autotrophic markers (Figure 5). The strains transformed with heme vectors were first evaluated to find the best heme producing vectors. The strains overexpressing globin vectors were also evaluated, however, likely due to insufficient heme, we did not detect globin bands on SDS-PAGE. In the third step, we combined the best heme overproducing vectors with the globin constructs and identified the best combinations for high level human hemoglobin production.
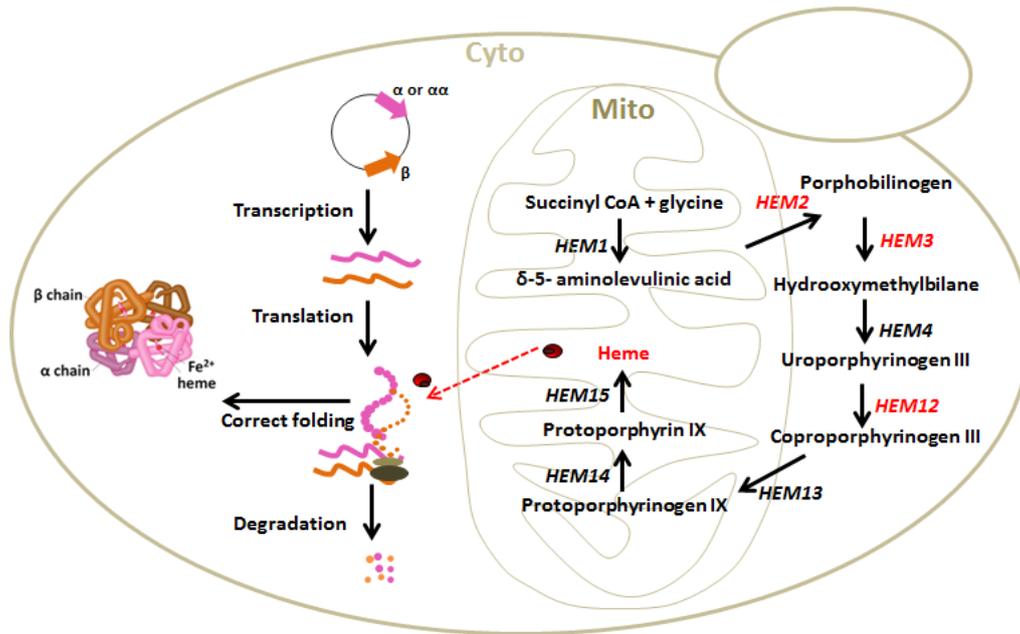
Figure 4. Strategy for engineering hemoglobin production in yeast. Yeast genes *HEM2*, *HEM3* and *HEM12* (highlighted in red) in the heme biosynthesis pathway were overexpressed for improved heme levels in the cytosol. Human α and β globins were overproduced using plasmid expression. Correct folding and assembling resulted in corrected folded hemoglobin, whereas failure in assembling resulted in degradation.
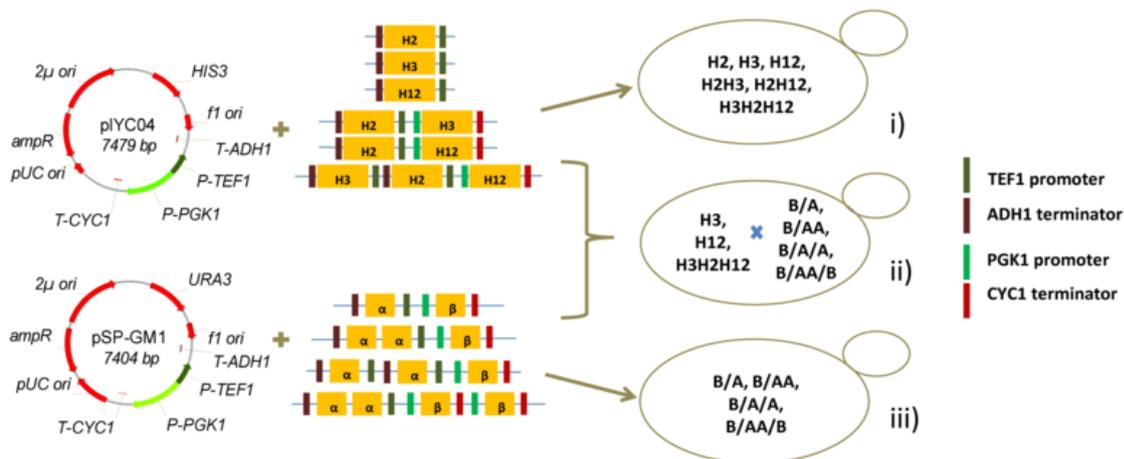
Figure 5. Strain construction. CEN.PK 113-11C as the expression host being transformed with i) heme overexpression plasmids: gene *HEM2*, *HEM3* and *HEM12* were inserted into the expression cassette *TEF1-ADH1*, either individually or in different combinations in the cassettes *TEF1-ADH1* and *PGK1-CYC1* on pIYC04. ii) globin overexpression plasmids: *HBA*, *HBB*, and *HBAA* (two *HBA* connected by a glycine linker) were inserted into pSP-GM1 in the cassettes of *TEF1-ADH1* and *PGK1-CYC1* in different combinations. iii) heme and globin overexpression plasmids: The top three best heme/coproporphyrin producing plasmids H3, H12 and H3H2H12 were transformed in different combinations together with the globin overexpression plasmids B/A, B/AA, B/A/A and B/AA/B.

17
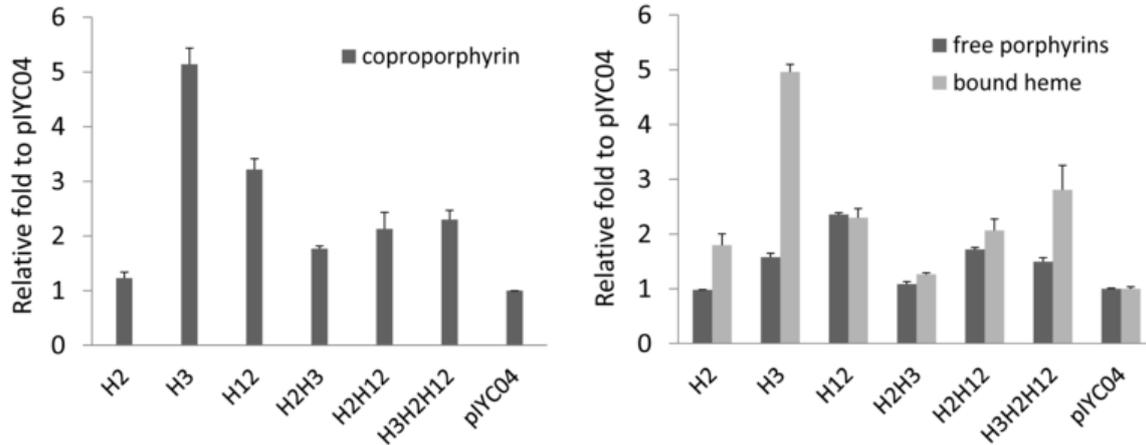
Figure 6. Porphyrin/heme levels in heme over-expression strains.

Overexpressing the heme biosynthesis genes successfully enhanced endogenous heme levels. *HEM3* overexpression alone (strain H3) displayed the best effects in terms of the levels of intermediate coproporphyrin, free porphyrins and bound heme in cells compared to the other combinations tested (Figure 6). The *HEM12* (strain H12) and *HEM3*, *HEM2* and *HEM12* triple overexpression plasmids (strain H3H2H12) produced lesser but still substantial amounts of heme and were therefore chosen for evaluation of hemoglobin production together with the *HEM3* plasmid.
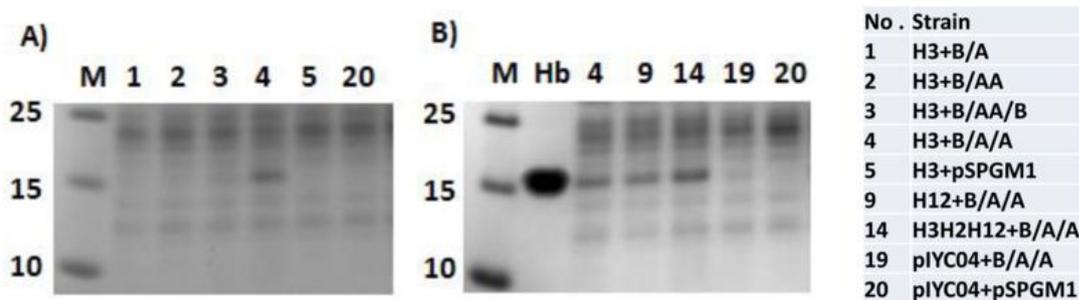
Figure 7. Selection of the best hemoglobin producing strains on SDS-PAGE.

HEM3 overexpression vector was first co-transformed with all the globin overexpression plasmids (Figure 7A). By co-overexpression with H3, the globin construct B/A/A showed the highest protein expression, whereas there was nearly no band of the right size (around 16 KDa) when H3 was expressed in combination with other globin constructs. Based on this result, B/A/A was co-overexpressed with the other two heme overexpression plasmids H12 and H3H2H12 and clear bands were also observed (Figure 7B). Expression of B/A/A with the empty plasmid pIYC04 did not show clear expression, further supported the hypothesis that endogenous heme levels are insufficient to achieve high heterologous globin production (Smith *et al*., 2011).
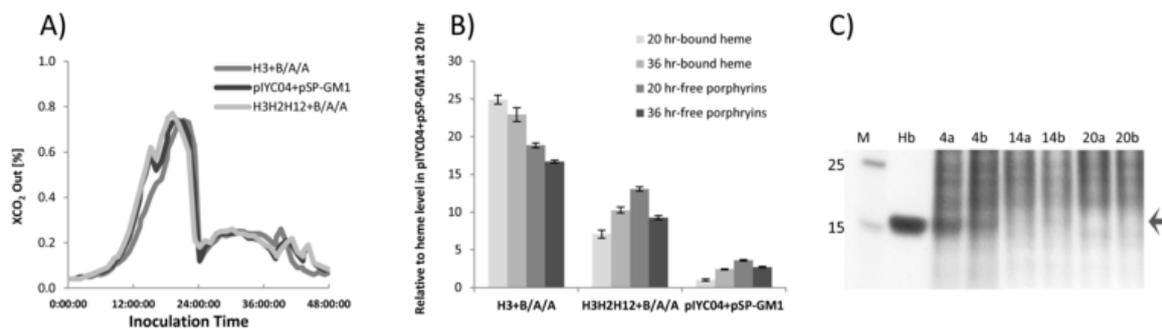
Figure 8. Heme and hemoglobin production during batch fermentations. (A) Offgas profile: showing that glucose phase finished at around 20 hrs and ethanol phase started at around 36 hrs. (B) Heme and porphyrin levels in strains after inoculation for 20 and 36hrs respectively. (C) SDS-PAGE analysis of the globin patterns produced at 20 and 36 hrs respectively. 4, 14, 20 represent strains H3+B/A/A, H3H2H12+B/A/A and pIYC04+pSP-GM1 respectively. a and b represent 20 and 36 hrs after inoculation.

The physiological properties of the recombinant hemoglobin strains were characterized in batch fermentations (Figure 8). The cell growth reached the end of the glucose phase and the beginning of the ethanol phase at around 20 and 36 hours after inoculation respectively (Figure 8A). Interestingly, we found that the heme level in H3+B/A/A significantly decreased before the ethanol phase started (Figure 8B) indicating a negative correlation between heme/porphyrin synthesis and respiration. In contrast, in H3H2H12+B/A/A and pIYC04+pSP-GM1, the heme level kept improving after the glucose phase was finished with a decrease in the free porphyrin level suggesting that porphryin synthesized in the glucose phase were gradually converted into heme. The correlation between the heme levels and the hemoglobin production was also measured by SDS-PAGE (Figure 8C). Clearly, the level of hemoglobin produced in the glucose phase is higher than in the beginning of the ethanol phase (about 1.3 fold) for H3+B/A/A, confirming that efficient heterologous hemoglobin production requires a low rate of respiration.

At the end of cultivation, strain H3+B/A/A had the most active hemoglobin accounting for 4.09% of the whole cell soluble protein followed by H3H2H12+B/A/A with 3.82%. 2.8% of active protein was detected in strain pIYC04+pSP-GM1 which has neither heme nor globin overexpressed. The activity measured in this strain might come from the endogenous flavohemoglobin (Zhao *et al*., 1996).

4.2 Paper II. Inverse metabolic engineering

The concept of inverse metabolic engineering was introduced by Bailey (Bailey *et al*., 1996) to identify novel targets for metabolic engineering through characterization of mutagenesis libraries or strains exposed to adaptive laboratory evolution. In paper II this concept was used for improving protein secretion. By combining UV-random mutagenesis and selection for

growth on starch, mutant strains able to produce up to 5 fold increased levels of heterologous amylase compared with the reference strain were obtained (Figure 9).
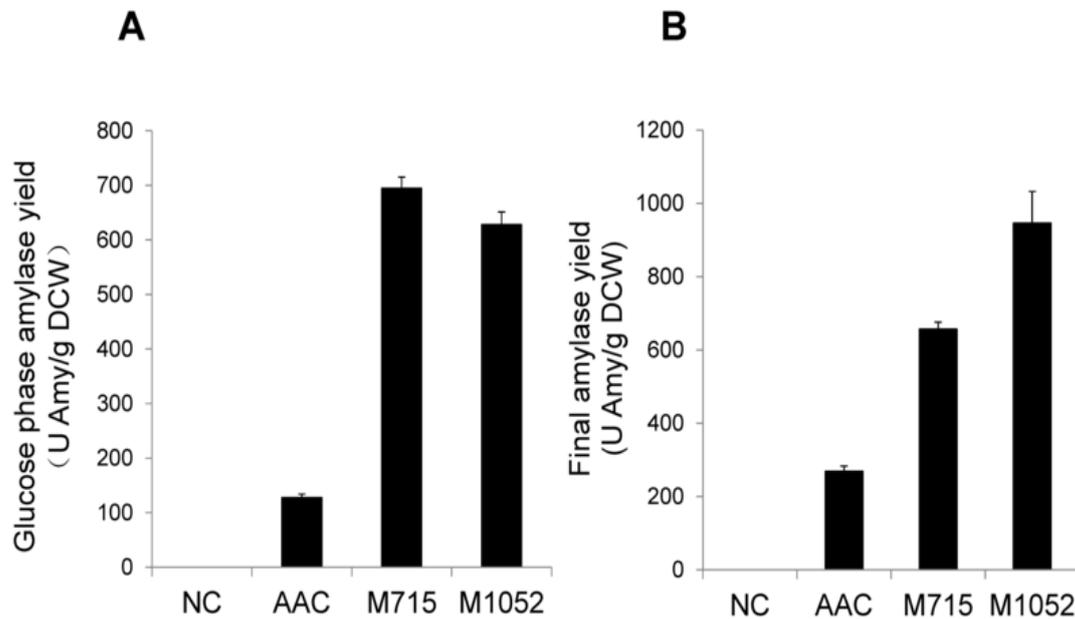
Figure 9. Protein yield during batch fermentations. (A) Amylase yield on cell mass during the exponential phase. (B) Final Amylase yield. NC: negative control, produces no amylase; AAC: amylase producing strain subjected to UV mutagenesis; M715 and M1052: two mutant strains derived from AAC after UV mutagenesis

Through whole genome sequencing, we identified genomic variations that could be associated with higher amylase secretion. Several single point mutations were evaluated by introducing the mutations to the starting strain and checking the performance of the resulting strains with respect to amylase production. Interestingly, we successfully identified a single point mutation in *VTA1*, coding a regulatory protein interacting with several Vps proteins in the MVB sorting pathway in endocytosis. By applying this modification alone, the amylase secretion could be improved by 35%, and the cell growth was increased by 21% in shake flask cultivations (Figure 10).
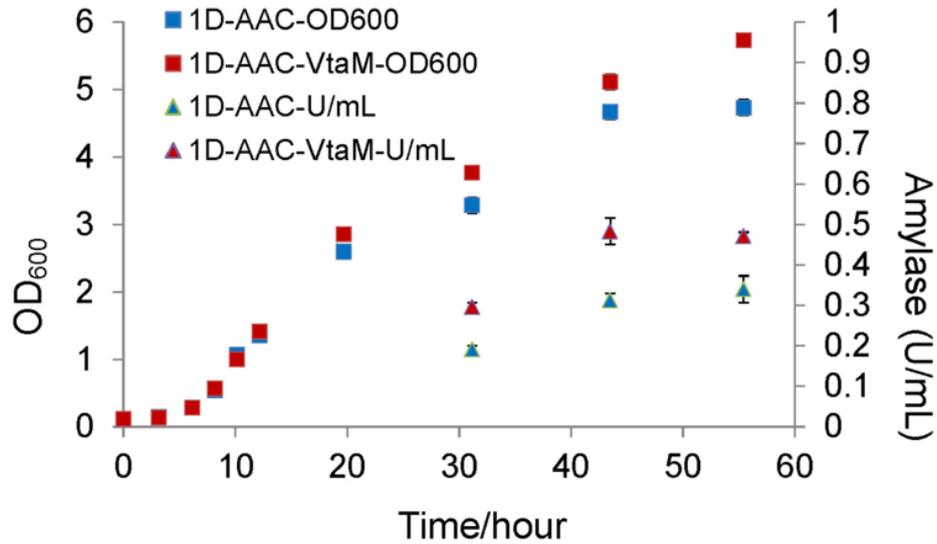
Figure 10. Effect of *VTA1S196I* mutation on α-amylase production in shake flask fermentation. 1D-AAC: AAC strain with *URA3* knocked out to facilitate application of point mutation; 1D-AAC-VtaM: 1D-AAC with *VTAS196I* point mutation.

To understand what effect the *VTA1S196I* mutation may have on endocytosis, we added different concentrations of bovine serum albumin (BSA) into the Delft medium, which lacks peptone and nutrient rich yeast nitrogen source compared to the YPD medium, and compared strain growth and amylase production. Cell growths (Figure 11A) in the mutated strain were slightly better in all three media than in the control strain. The amylase productions (Figure 11B) were slightly higher in the Delft medium and the Delft + 1g/L BSA medium until around 55 hours. After that point, the main nutrients (glucose and added BSA) supporting cell growth as well as amylase production may have been depleted and the yeast cells may start to take up the produced amylase through endocytosis. With more BSA (2g/L) presented, amylase uptake was reduced, as indicated by the higher amylase amount (0.8 U/mL) in 1D-AAC-VtaM than 1D-AAC (0.4 U/mL) after 60 hours. Naturally *S. cerevisiae* produces very low amounts of extracellular proteins, including proteases (Mattanovich *et al*., 2012). The basal expression of protease (s) should be of similar level in 1D-AAC and 1D-AAC-VtaM as the only difference between these two strains is the *VTA1S196I* mutation. If the presence of more BSA (2g/L) prevented protease degradation on amylase, there should be more amylase left in the Delft+2g/L BSA medium than in the Delft+1g/L BSA medium for both strains. However, we only see this phenomenon in strain 1D-AAC-VtaM (0.3U/L in Delft+1g/L BSA, and 0.8U/L in Delft+2g/L BSA). The amylase activity remained in 1D-AAC was similar (0.4U/L) in both Delft+1g/L BSA and Delft+2g/L BSA. Based on this analysis, we are confident that the reduction on amylase activity was mainly resulted from endocytosis but not proteolysis.

After confirming the effect of *VTA1S196I* to promote endocytosis, we further proposed that the structural stabilization of this protein may have been changed due to mutation from a

polar uncharged serine to a non-polar hydrophobic isoleucine on the linker region where a series of neutral or polar amino acids (Asp192, His193, Gln194, Thr195, Ser196 and Asp197) reside. A more favorable configuration for Vta1p's interaction with the client proteins, especially Vps4p, may have been created through which the late endosome sorting process is also accelerated. A faster recycling of the nutrients through subsequent degradation in the vacuole is hence expected for amino acid biosynthesis (for protein) as well as for cell growth.
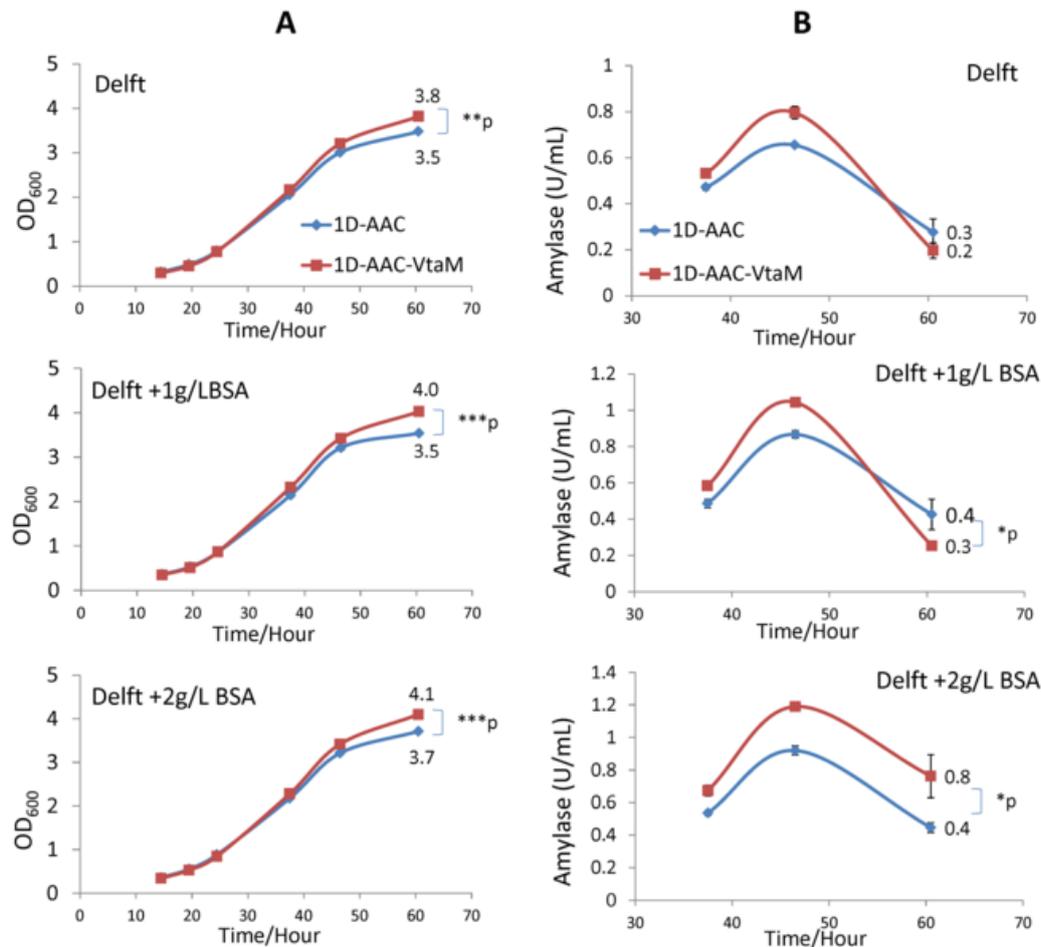
Figure 11. Effect of *VTA1S196I* mutation on amylase production in Delft medium supplemented with BSA. (A) Growth profiles. (B) Amylase productions.

We also performed transcriptome analysis (Figure 12) as a complement to the genomic variant identification analysis in order to get a global view of the transcriptional changes associated with the improved amylase production caused by UV mutagenesis.
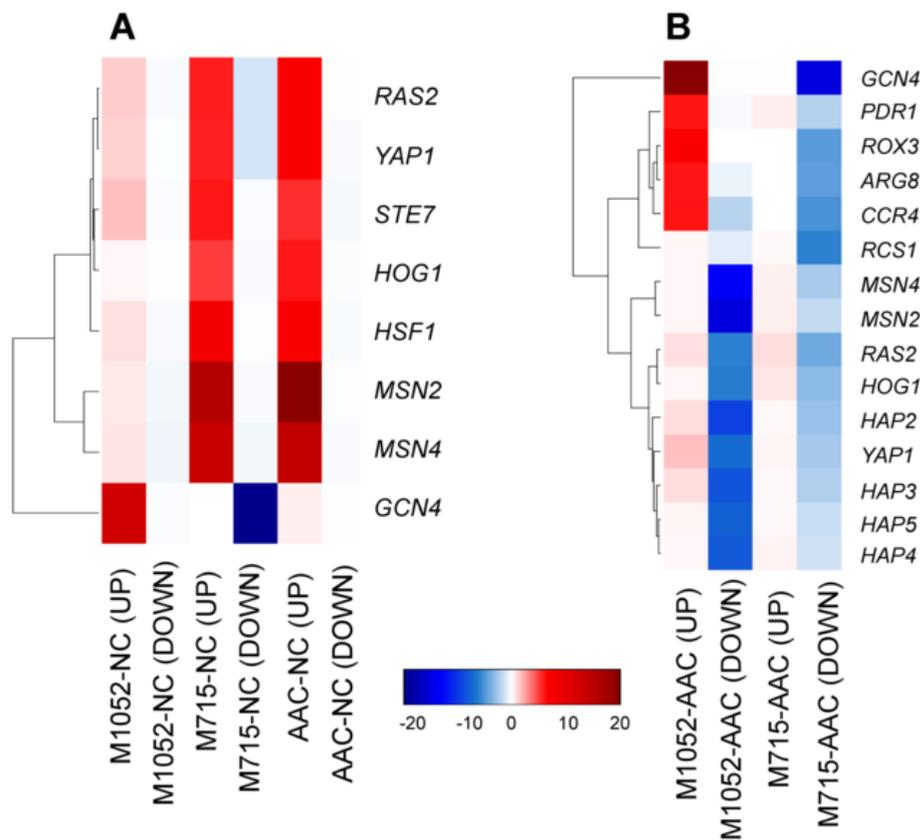
Figure 12. Reporter TF analysis. (A) amylase producing strains AAC, M1052 and M715 compared to the non-producing strain NC. (B) mutation strains M1052 and M715 compared to the mother strain AAC.

Most genes related to stress responses, such as genes regulated by oxidative stress (transcriptionally regulated by Yap1p), osmotic stress (regulated by Hog1p), and general stress (regulated by Msn2p and Msn4p) were up-regulated (Figure 12A) when comparing all amylase producing strains AAC, M715 and M1052 against the non-producing strain NC, indicating that increased protein production resulted in the induction of cellular stress responses. In contrast, the same classes of genes were down-regulated (Figure 12B) when comparing mutant strains M715 and M1052 with the non-mutated amylase producing strain AAC, suggesting that the mutations existing in the mutant strains contributed to a reduction of cellular stress induced by RPP.

Interestingly, genes regulated by Gcn4p, which regulates amino acid biosynthesis during amino acid starvation, were significantly up-regulated in M1052 but down-regulated in M715 when they are compared to both NC and AAC. A high recombinant protein producing strain has an increasing demand for amino acids. The up-regulation of amino acid biosynthesis genes in M1052 may be due to a response to certain amino acid starvation. Down-regulation of amino acid biosynthesis has been noticed in a *HAC1* deletion strains producing recombinant amylase and insulin precursor where the UPR-ensured efficient protein folding was affected (Tyo *et al*., 2012). One possibility could be that the increased amylase

production in the glucose phase caused perturbation in the UPR and thus generating similar responses as observed in a *HAC1* deletion strain.

4.3 Paper III. Oxygen sensing vs. protein production in *S. cerevisiae*

Oxygen is an important element associated with normal cellular metabolism as well as protein production in a role to assist oxidative protein folding. However, several studies have reported that anaerobic conditions seem to be more favorable in terms of producing recombinant proteins. For example, low oxygen availability has been proven to promote recombinant α-amylase (Liu *et al.*, 2013) and glucoamylase (Cha *et al.*, 1997) production in *S. cerevisiae* and on human Fab fragment production in *P. pastoris* (Baumann *et al.*, 2011; Baumann *et al.*, 2007; Cha *et al.*, 1997).

Rox1p is a heme-dependent transcription factor that represses the transcription of many hypoxia-induced genes under aerobic conditions (Kwast *et al.*, 1997; Kwast *et al.*, 2002; Ter Linde and Steensma, 2002). About one-third of the anaerobically induced genes contain a putative Rox1p binding site in their promoter regions (Kwast *et al.*, 2002). We hence studied the effect of knocking out *ROX1* on the production of recombinant proteins in *S. cerevisiae*.
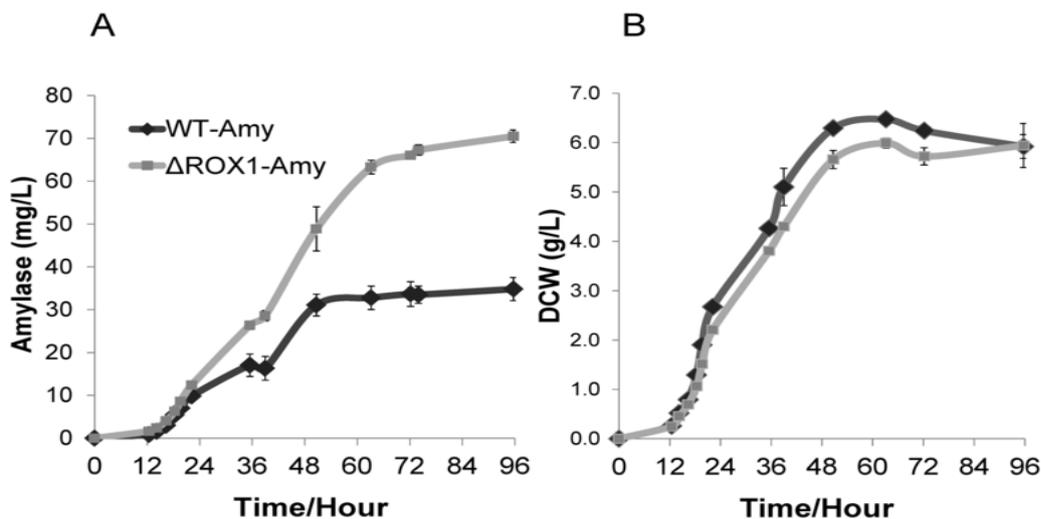
Figure 13. Aerobic batch fermentations of α-amylase producing strains. A) α-amylase volumetric yields. B) Biomass accumulations. WT-Amy: control strain producing amylase; ΔROX1-Amy: *rox1Δ* strain producing amylase.

Table 1. Physiological parameters of α-amylase producing strains.

Strain	μ_{max}[a]	Y_{SE}[b]	Y_{SG}[c]	Y_{SX}[e]	Y_{Samy}[f]	Y_{Xamy}[g]	Y_{Xamy}[h]	Q_{amy}[i]
WT-Amy	0.33±0.002	0.37±0.008	0.05±0.003	0.122±0.001	0.461±0.05	3.79±0.45	5.89±0.18	1.23±0.14
ΔROX1-Amy	0.31±0.004	0.35±0.005	0.08±0.003	0.122±0.002	0.746±0.03	7.14±0.34	10.2±0.9	2.23±0.12

[a]μ_{max}, specific growth rate on glucose (h^{-1}); [b]Y_{SE}, ethanol yield on glucose (g/g); [c]Y_{SG}, glycerol yield on glucose (g/g); [e]Y_{SX}, biomass yield on glucose (g/g); [f]Y_{Samy}, α-amylase yield on glucose (mg/g); [g]Y_{Xamy}, α-amylase yield on biomass during glucose phase (mg/g); [h]Y_{Xamy}, α-amylase yield on biomass during ethanol/acetate phase (mg/g); [i]Q_{amy}, α-amylase specific production rate (mg/g/h)

Intriguingly, we found a 100% increase in the α-amylase yield as well as productivity without obvious reduction on cell growth (Figure 13 and Table 1). Varied levels of improvements were also observed for the productions of insulin precursor (IP) and endogenous invertase. Through transcriptomic analysis, we noticed *ANB1*, coding the anaerobically induced translational elongation factor, was significantly up-regulated in the ΔROX1-Amy strain compared to the WT-Amy strain. Since translation is the first step in protein production, we performed overexpression and deletion experiments to evaluate the effect of Anb1p on protein production.
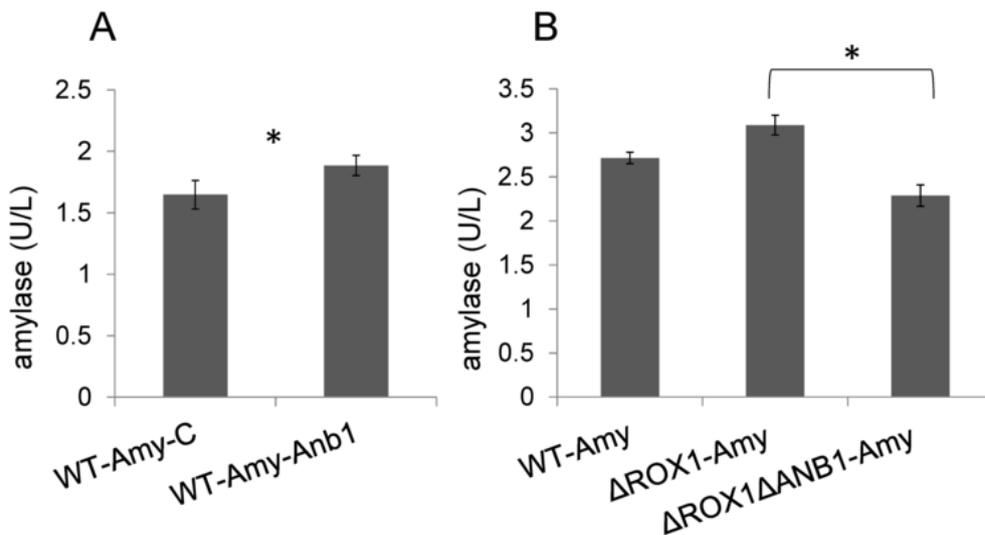
Figure 14. Effects of Anb1 on α-amylase production. (A). α-amylase production in the Anb1p (WT-Amy-Anb1) co-overexpression strain compared to the control strain (WT-Amy-C). (B). α-amylase production in the WT strain (WT-Amy), the *ROX1* deletion strain (ΔROX1-Amy) and the *ROX1* and *ANB1* double deletion strain (ΔROX1ΔANB1-Amy).

Shown in Figure 14, over-expression of *ANB1* resulted in a slight improvement in amylase production (12.7% at 72h in shake flask) but the deletion of *ANB1* reduced amylase production significantly (39% reduction on ΔROX1ΔANB1-Amy compared to ΔROX1-

Amy). These results indicated that the increased expression of *ANB1* may have an important role in amylase production, but it is not likely to be the only important regulator.

Furthermore, we noticed the GO terms (Figure 15A) related to lipid/ergosterol biosynthesis, structural constituent of cell wall and heme biosynthesis were significantly up-regulated. Accordingly, reporter TF (Figure 15B) analysis shows genes regulated by Upc2p (lipid/sterol biosynthesis, anaerobic cell wall mannoproteins) were significantly up-regulated. After disrupting the cell wall, we found the intracellular amylase contents in the two strains were of similar level, although slightly higher in the ΔROX1-Amy strain, suggesting the cell wall mannoproteins contributed to amylase secretion.
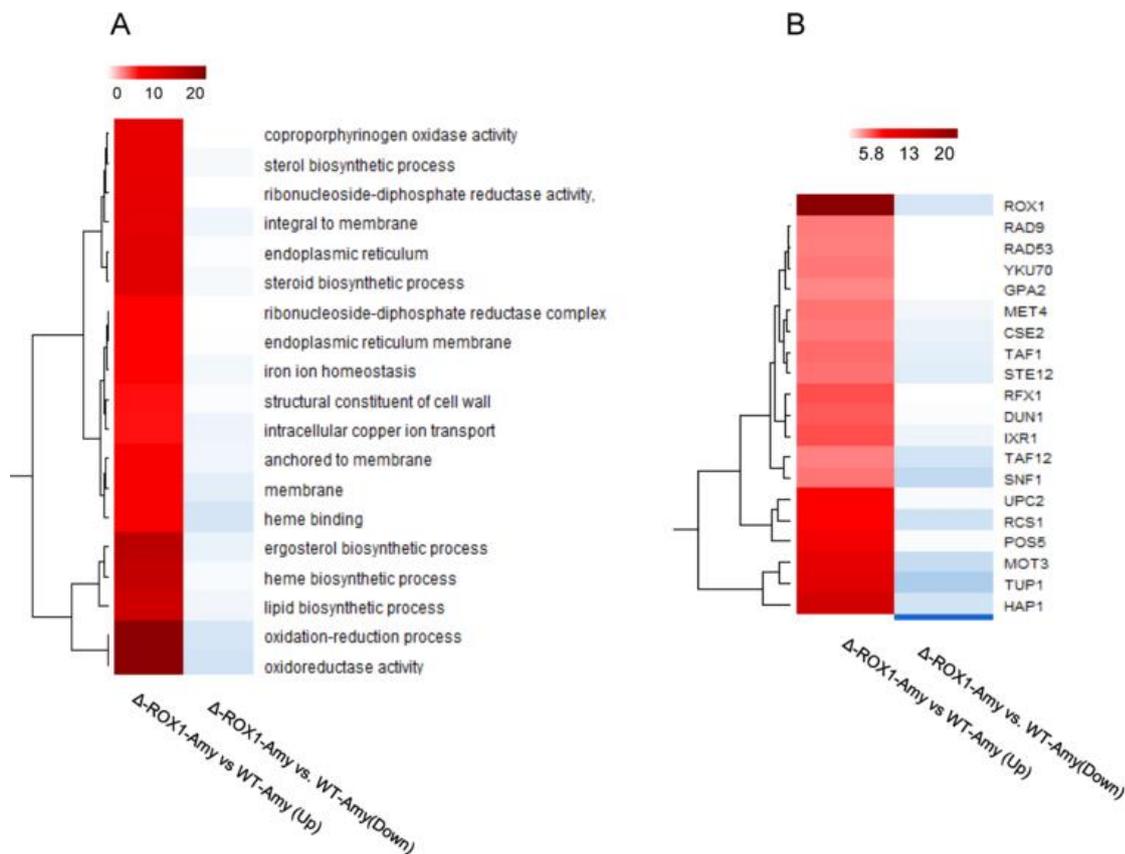
Figure 15. Reporter GO term and Reporter TF analysis. (A). Reporter GO term. Reporter p-value<10^{-6}. (B). Reporter TF. Reporter p-value<10^{-3}.

Furthermore, we extracted lipid classes and measured their concentrations in the two cell types. There was an overall enhancement on the contents of all the lipid classes measured, including phospholipids, neutral lipids steryl esters (SE) and triacylglycerols (TAG) in the ΔROX1-Amy cells compared to the WT-Amy cells. Besides, the fatty acids contents were also varied between the two strains. We therefore speculated that the composition of lipids might have a role in the composition and function of membranes that play essential roles during protein secretion, such as the ER membrane, membranes of Golgi and the trafficking vesicles, as well as the cell plasma membrane.

Specifically, since phospholipids account for main compositions of cell PM, we checked cell morphology by flow cytometry as well as microscope.
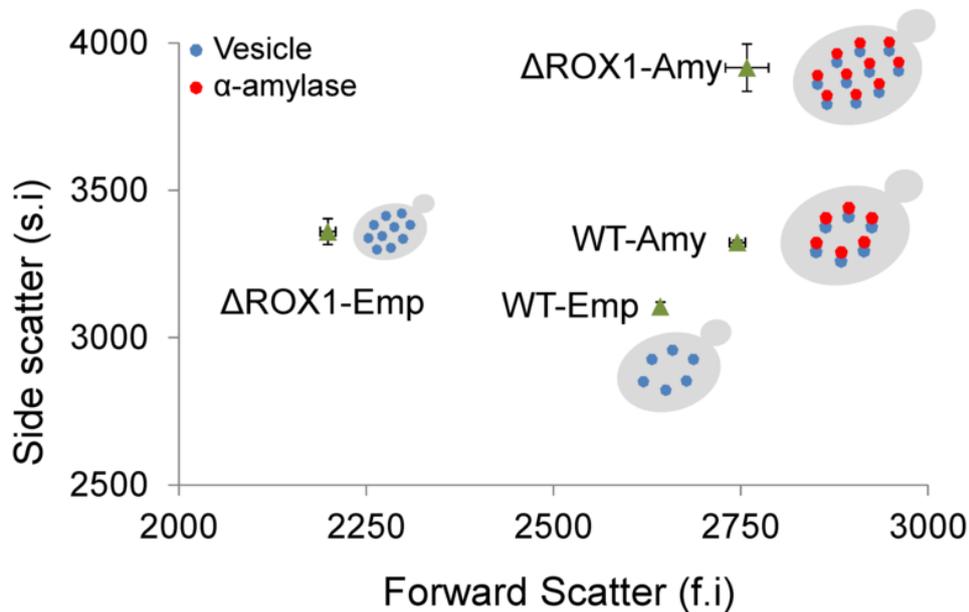
Figure 16. Flow cytometry estimation of cell size and intracellular cell densities. WT-Amy: reference strain expressing amylase; ΔROX1-Amy: *rox1Δ* strain expressing amylase; WT-Emp: reference strain containing empty vector; ΔROX1-Emp: *rox1Δ* strain containing empty vector. We illustrate schematically our speculations on cell sizes and the amount of α-amylase and vesicles aside.

Indicated in Figure 16 by the forward scatter intensity (f.i), after deleting *ROX1*, cell size reduced significantly (ΔROX1-Emp, f.i: 2199< WT-Emp, f.i: 2643, $p<10^{-6}$), however, when amylase was overexpressed, the cells increased "back" in size with ΔROX1-Amy (f.i: 2758) reaching even similar size as WT-Amy (f.i: 2746). Any internal and surface irregularities, including cytoplasmic granules, vesicles, and other organelles and membrane roughness will typically contribute to side scatter signals (51). Indicated by relative side scatter intensity (s.i), cell density increased significantly in the *ROX1* deleting cells (ΔROX1-Emp, s.i: 3360 > WT-Emp, s.i: 3105, $p<10^{-2}$) as well as in the amylase overproducing cells (WT-Amy, s.i: 3322 > WT-Emp, $p<10^{-3}$). The strain with both *ROX1* deletion and amylase overexpression (ΔROX1-Amy, s.i: 3916) exhibited the highest cell granularity. The increased lipid contents in the cells may very likely contribute to the increased cell density, in the form of lipid particles (SE and TAG) and in the enrichment of secretory vesicles. It is possible to envisage that the efficiency of protein secretion is also increased due to presence of more secretory vesicles in the ΔROX1-Amy strain.

In conclusion, we suggest in this study that the increased secretion efficiency in the *ROX1* deletion strain is due to a synergistic effect caused by derepression of Rox1p targets, including (i) Anb1p (translation and protein synthesis), (ii) lipid composition (membranes of compartments and vesicles involved in secretion), as well as (iii) mannoproteins on the cell

wall, even though it is not possible to conclude exactly to what extent the change in the membranes (due to changes in the lipid metabolism), including the vesicular transport, or the change in the cell wall composition (mannoproteins) contribute the most to the observed phenotype of increased protein secretion in the *ROX1* deletion strain.

4.4 Paper IV. Secretory machinery reconstruction in *A. oryzae*

Even though the high capacity of the secretion machinery of filamentous fungi has been widely exploited for the production of homologous and heterologous proteins, systemic analysis of its secretion system is lacking, most likely due to the poorly annotated proteome. In paper IV, using comparative genomics approaches, the list of the functional components involved in the secretory machinery of *A. oryzae* was enriched to cover 389 proteins. The resources include i) a previously reported secretory model of *S. cerevisiae*, ii) functional components reported in other closely related fungal species such as *A. nidulans* and *A. niger*, and iii) formerly reported *A. oryzae* secretory components.

In order to evaluate the component list, and trace the secretory machinery response, we constructed two novel *A. oryzae* α-amylase over-producing strains A16 and CF32. Together with an earlier reported high α-amylase producer CF1.1 (Pedersen *et al*., 1999; Spohr *et al*., 1998) their performances on maltose medium were compared with the reference strain A1560 (Christensen *et al*., 1988) which produces basal level of amylase.
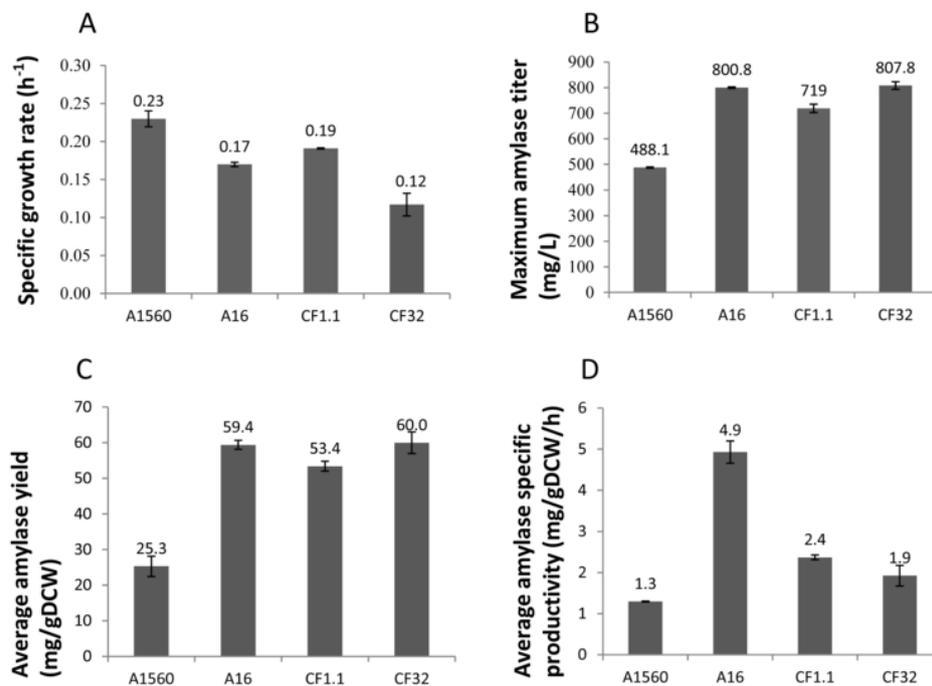
Figure 17. Physiological parameters of the α-amylase producing strains in batch cultivations. (A) Specific growth rate on maltose (h^{-1}), (B) Maxium α-amylase titer (mg/L) (C) Average α-amylase yield (mg/g DCW), (D) Average specific α-amylase productivity (mg/g DCW/h).

Figure 17 shows the physiological parameters for the α-amylase producing strains in batch fermentation in maltose medium. The three recombinant strains grew slower than the reference strain A1560, with CF32 grew the slowest, followed by A16 and then CF1.1. CF32 and A16 had the highest average yields and final titers. In consideration of the growth rate, A16 appeared to be the best strain as it exhibited the highest productivity.

We studied the transcriptional responses of these strains to α-amylase production using microarray. After normalization and statistical analysis we found 1212, 653 and 1709 genes to be differentially (adj. p-value < 0.05) expressed when comparing A16, CF1.1 and CF32 to A1560. Figure 18 presents the global transcriptional responses. Many genes with GO-term annotations related to protein secretion were significantly up-regulated in all three comparisons, including protein N-linked glycosylation, ER translocation and folding, signal peptide processing, ER to Golgi and the retrograde Golgi to ER vesicle trafficking *etc*. The down-regulated GO-terms are mainly related to amino acid biosynthesis, which could be due to the feedback inhibition from overloaded proteins in the ER.
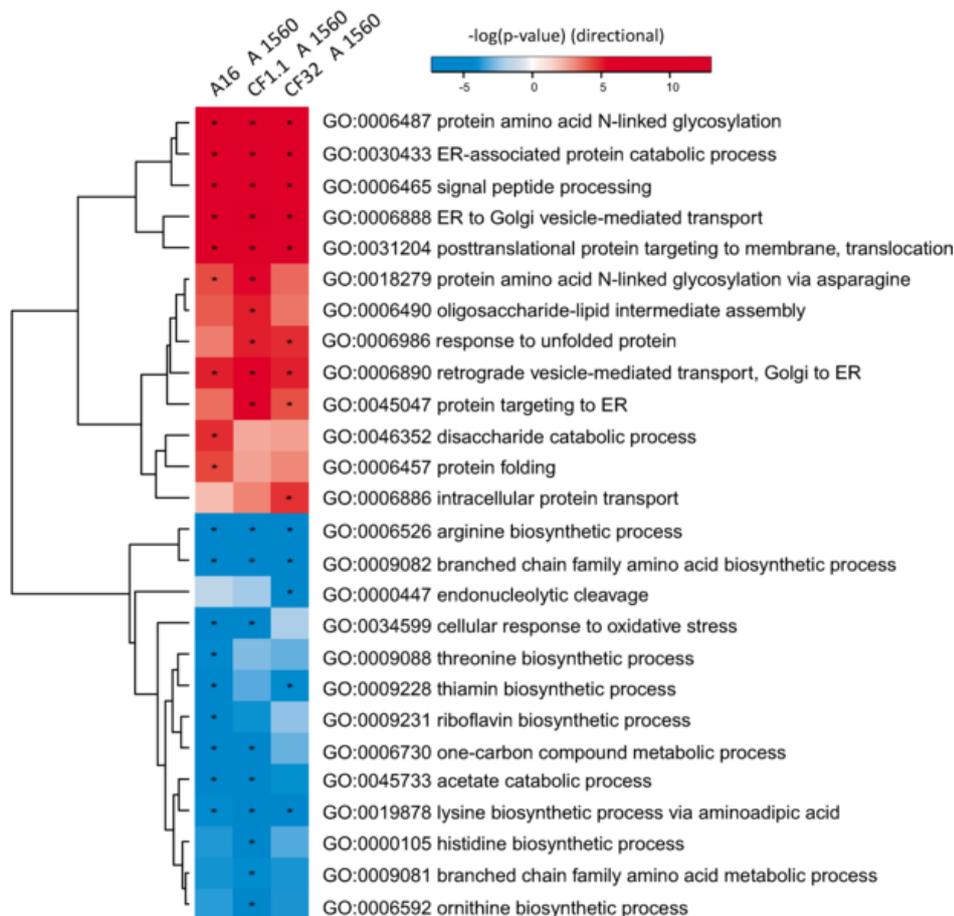
Figure 18. Reporter GO-terms for the three α-amylase over-producing strains compared to the wild type A1560. Red color indicates that the genes belonging to the GO-term are up-regulated and blue color indicates down-regulation. The intensity corresponds to significance. GO-terms with reporter p-values smaller than 10^{-4} are indicated by asterisks.

To get a detailed mechanistic picture of the protein secretory response at the molecular level we mapped the gene transcriptional profiles to the reconstructed *A. oryzae* secretory machinery. With the most complete secretory component list, we are able to monitor the transcriptional response of the secretory machinery to the ultimate extend. 51 out of the 369 components were significantly changed (adj p <0.05) in all three comparisons. Their expression profiles are described in Figure 19 based on their classification in the defined subsystems.

Clearly, in comparison with the reference strain A1560, the subsystems responsible for the α-amylase PTMs, especially the components involved in ER processing (translocation, glycosylation, Dolichol pathway, Folding, UPR, ERAD, GPI biosynthesis, and trafficking between ER and Golgi) were significantly up-regulated, particularly in the CF32 strain which had the slowest growth compared to the other strains. There are several novel secretory components identified through blast search with the yeast secretory components, including homologs to yeast *ERD2* (AO090102000650), *YOS9* (AO090023000334), *JEM1* (AO090020000010), *MNS1* (AO090003000057) and *GAB1* (AO090023000750), among which we found the up-regulation of the *ERD2* homolog is very interesting. Erd2p mediates retrieval of the ER residential proteins such as Bip1p, Pdi1p, Ero1p, and Fkb2p, from the Golgi through binding of the *ERD2* receptor to the C-terminal peptide sequence HEDL on these soluble ER proteins (Semenza *et al.*, 1990). In yeast, this process has been suggested as non-essential as it shares functional redundancy with Ire1p for maintaining normal levels of ER residential proteins (Beh and RoSE, 1995). We found the transcriptional level of *IRE1* homolog in *A. oryzae* was not changed in all the three overproducing strains, which may indicate that *A. oryzae* has a different mechanism to retrieve ER residential proteins than yeast, and the *A. oryzae* homolog of *ERD2* seems to play a more important role than the *A. oryzae* homolog of yeast *IRE1*. Additionally, the transcript level of the *HAC1* homolog was also not changed, while the folding chaperones including homologs of *CNE1*, *KAR2*, *PDI1*, *MPD1*, *FKP2* and the *KAR2*'s co-chaperones *SCJ1* and *JEM1* were all significantly up-regulated. This may indicate that the *IRE1*-mediated *HAC1* splicing is not the only mechanism for activating UPR in *A. oryzae*.
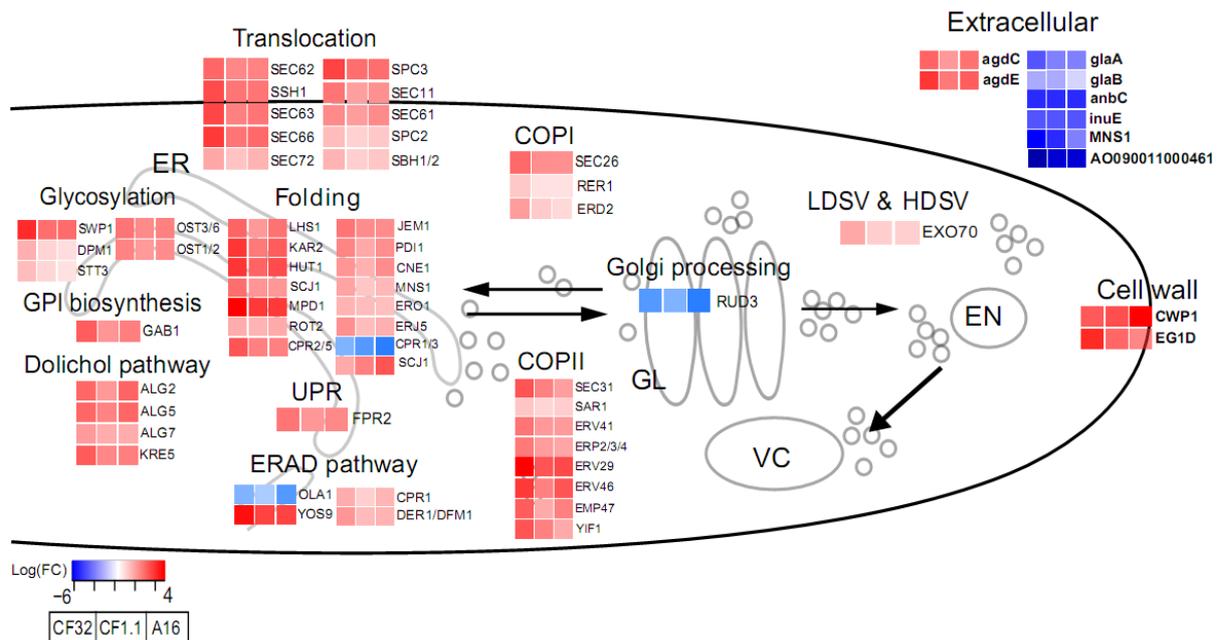
Figure 19. Transcriptional profiles of the significantly changed genes in all three strains compared to the wild type. Gene expressions are described based on log fold change compared to A1560, with red color indicates up-regulation and blue color indicates down-regulation. Genes identified from the secretome study are highlighted in bolds.

The transcriptional response of the secretome to a specific heterologous protein is important in a sense that if the target protein overloads the secretory machinery, the cell probably needs to change its secretome profile to adapt to the processing capacity of the secretory machinery. To examine this idea we extracted the putative *A. oryzae* secretome from the Fungal Secretome Database (FSD) and compared their expression profiles in response to amylase overproduction in the three strains compared to the wild type A1560.

Reporter GO-Slim enrichment analysis (Figure 20) roughly clustered the putative secretome into two groups, with genes encoding secretory proteins localized to the ER, the Golgi apparatus, the cytoplasm and the membranes distinctively up-regulated in all three comparisons. The genes encoding proteins secreted to the extracellular region, the vacuole and the cell wall were either significantly up- or down- regulated, which shows that the endogenously secreted proteins are responding to the over-expression of α-amylase.
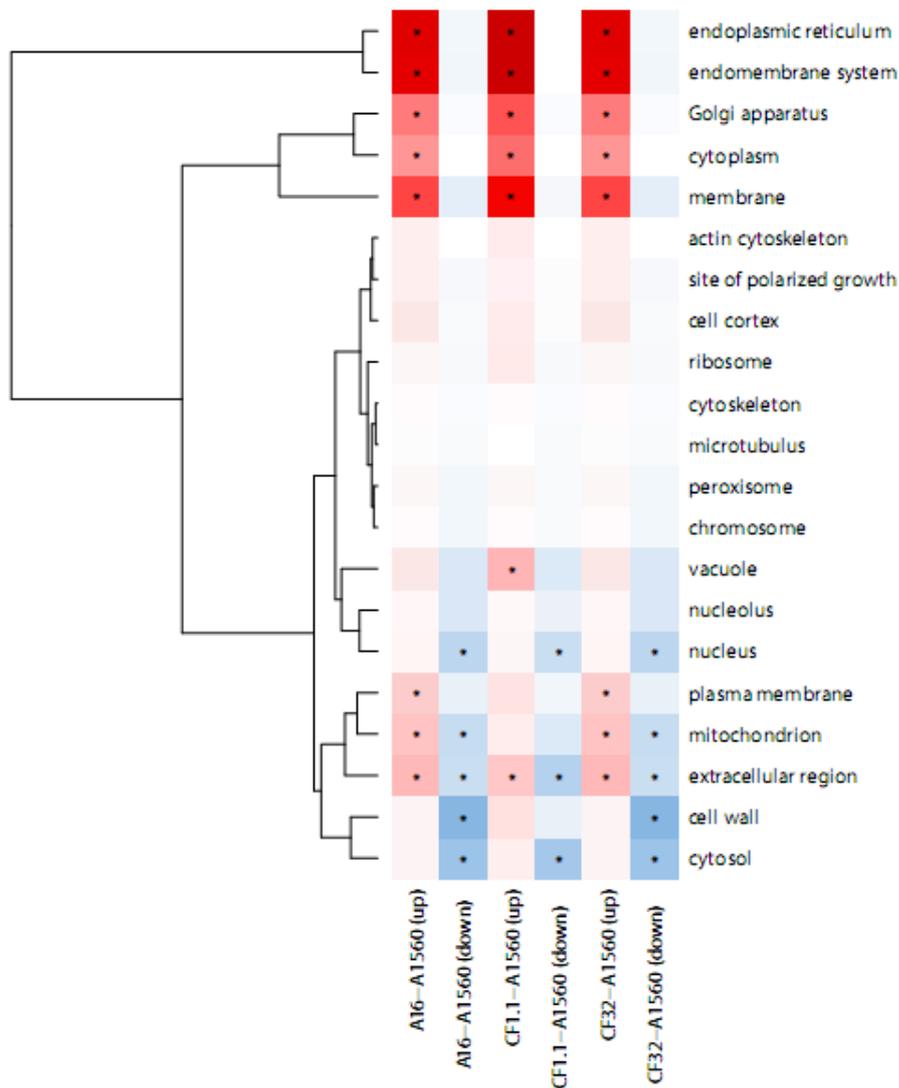
Figure 20. Reporter GO-Slim analysis for the secretome genes. Red color indicates that the genes belonging to the GO-Slim term are up-regulated and blue color indicates down-regulation. GO-Slim terms with reporter p-values smaller than 0.05 are indicated by asterisks.

Interestingly, we found two genes predicted to reside in the fungal cell wall namely AO090011000119 (homolog of *A. niger cwpA*) and AO090701000717 (homolog of *A. nidulans eglD*) were significantly up-regulated in all three strains. *eglD* encodes a putative endoglucanase discovered in the conidial cell wall of *A. nidulans* carrying an expansin like domain (Bouzarelou *et al*., 2008). Expansins exhibit wall loosening activity and are involved in plant cell expansion and other developmental events. The expansins are highly conserved among plants and fungi (Cosgrove, 2000). EglD in *A. nidulans* has been indicated to be involved in fungal cell wall remodeling during germination (Bouzarelou *et al*., 2008). The significant up-regulation of the *eglD A. oryzae* homolog AO090701000717 in the α-amylase overproduction strains might be the consequence of α-amylase overloading to the secretory pathway. Remodeling of the cell wall through *eglD* up-regulation may possibly help to loose cell wall structure and facilitate α-amylase secretion. There is no direct explanation for the

32

up-regulation of the *cwpA*, however since the surface properties of fungi are primarily determined by the presence of cell wall mannoproteins (De Groot *et al*., 2005), the expression of *cwpA* may also be altered in response to *eglD* up-regulation to facilitate the remodeling of the fungal cell wall and ultimately the protein secretion.

In conclusion, by providing a far more complete secretory component list of *A. oryzae* we were able to monitor the whole secretory pathway at molecular level in response to α-amylase overproductions. The defined *A. oryzae* component list offers a better platform to trace the secretory machinery responses on genomics (gene variation), transcriptomics, and proteomics levels, based on which we discussed several interesting mechanisms in response to α-amylase overproduction. Additionally, this study generated a list of targets for genetic manipulation. For example, overexpressing the up-regulated *eglD* and *cwpA* encoding cell wall proteins, and knocking out the genes encoding extracellular proteins competing for the secretory pathway may help to increase protein secretion in this industrially important fungus.

V. Summary and Perspectives

Current strategies for engineering protein secretion mainly involve i) increasing protein expression, ii) engineering protein folding and quality control in ER, iii) engineering post-translational glycosylation, iv) engineering protein trafficking pathway, v) minimizing post-secretory proteolytic degradation (Idiris *et al.*, 2010).

In this thesis, by applying systems biology approaches we extended the strategy list for improving heterologous protein productions in both fungi. Specifically, in paper II, we demonstrated that enhancing endocytosis through gene modification on the MVB pathway may play a role in improving protein secretion. One distinct advantage associated with this *VTA1S196S* mediated endocytosis enhancement is that not only the volumetric yield of α-amylase but also the cell growth was promoted. With adequate nutrients, especially in the industrially preferred fed-batch fermentations, we could expect a continuously high production of the protein of interest. In paper III, we demonstrated a novel strategy to improve protein production under aerobic cultivation by knocking out *ROX1*. This approach offers an ideal strategy to overcome the disadvantages associated with anaerobic fermentations *e.g.* supplementation of anaerobic growth factors and slow growth for improving heterologous protein productions in *S. cerevisiae*. In paper IV, we studied the transcriptome of high amylase producing *A. oryzae* strains, through which we proposed several strategies to improve protein production in this industrially important organism, including overexpressing the up-regulated cell wall proteins EglD and CwpA and knocking out the genes encoding extracellular proteins competing for the secretory pathway.

Aside from the demonstrated successes for improving protein productions, there have also been many interesting phenomena that are worth discussing and pointing to further investigations.

1. How do physiological parameters affect protein production during batch fermentations?

Two physiological parameters, namely the specific growth rate and the glycerol yield have been noticed to be especially sensitive to recombinant protein productions. Except for the increased growth in the endocytosis enhanced cells (Paper II), strain growth was always slower in the better recombinant protein producing strains compared with the reference strains in this thesis. One reasonable speculation could be that the recombinant proteins compete for both nutrients (*e.g.* carbon for amino acid synthesis) and cellular machineries (*e.g.* transcription and translation) with the native proteins that catalyse reactions important for normal cellular metabolism leading to impaired cell growth and reduced biomass formation. Considering this, enhanced endocytosis offers an effective way to fasten nutrient uptake and recycling, enabling delivery of more resources for both requirements.

Glycerol yield was usually higher in the better RPP strains than in the reference strain. Amino acid biosynthesis generates a large amounts of cytosolic NADH (Rigoulet *et al*., 2004) which need to be neutralized to reduce cellular stress. Production of one mole of glycerol from glucose consumes one mole of NADH offering an effective means to balance the redox. During aerobic cultivations, oxygen consumption rate for protein production is usually high (Tyo *et al*., 2012). If the oxygen supply is insufficient during oxidative protein folding, the cell may have to produce glycerol to oxidize cytosolic NADH. Under anaerobic conditions, glycerol production is naturally high (van Dijken and Scheffers, 1986), which may provide better redox hemostasis environment for the cells to cope with overproduced proteins.

2. How does heme level influence respiration? Is there a competition between respiration and hemoglobin for heme molecules?

Heme is involved in many cellular processes, including oxygen sensing, forming hemoproteins such as catalases and P450 enzymes, and assisting electron transfer during oxidative protein folding and respiration *etc*. The extraction by oxalic acid mainly functions for cytosolic porphyrin/heme (either free or protein-bound). We observed a negative correlation between porphyrins/heme/hemoglobin levels and respiration in the best hemoglobin producing strain in Paper I, indicating that respiration is competing for heme molecules with hemoglobin. This assumption was further supported by overexpressing the heme and globin constructs in a *HAP1* deletion strain (Paper VI, not included in this thesis). In this strain both heme and hemoglobin levels were significantly enhanced, however, in contrast, the *HAP1* deletion strain was not able to respire. Since Hap1p regulates many cytochrome synthesis genes (*e.g.* *CYC1* and *CYT1*) by directly binding to their upstream regulatory regions (Schneider and Guarente, 1991), deleting *HAP1* results in respiration defect as the strain is not able to synthesis these essential enzymes in the electron transport chain. Consequently, the heme pool was shifted to synthesis of hemoglobin and other hemoproteins.

It could be an opposite scenario in the *ROX1* deletion strain though, as Rox1p functions as a repressor but not activator and by knocking it out genes repressed under aerobic conditions are activated, among which *HEM13*, *HEM14, COX5b* and *CYC7* are included. The heme level is thus expected to increase and the respiration may be even intensified. Actually, many genes on the complex IV of the mitochondria respiration chain were up-regulated in the ΔROX1-Amy strain. In this sense, the ΔROX1-Amy strain not only benefits from anaerobic metabolism (as discussed in paper III for enhancing protein production), but also from an intensified aerobic metabolism (in terms of maintaining decent metabolism and growth).

3. Endogenous invertase as an indicator to study protein secretion.

Different proteins require different mechanisms to cope with their productions. For example, protein has many disulfide bonds may need more efficient oxidative folding in ER. In the

ΔROX1-Amy strain, genes essential for UPR (*KAR2, HAC1 and IRE1*) and oxidative folding (*ERO1*, *PDI1*, and *ERV1*) were not differentially expressed compared to WT-Amy, indicating that more efficient folding in ER was not the main reason for the significant improvement on α-amylase production in the ΔROX1-Amy strain (even though there was slight up-regulation with *EUG1* and *ERV2* which are homologs of *PDI1* and *ERV1* respectively, that may help amylase folding). Since invertase does not contain any disulfide bonds (thus does not need an effective oxidative folding), while its production was also improved in the ΔROX1-Amy strain, we could push our speculation further to the post-ER processes (as discussed in the paper), which is in good connection with previous studies demonstrating the importance of engineering post-ER vesicle trafficking for enhancing RPP in yeast (Hou *et al.*, 2012b; Ruohonen *et al.*, 1997).

4. Opening for future projects

As we applied systems biology approaches to study strain behaviors in paper II, III and IV, large scale datasets are involved to understand strain behaviors and identify engineering targets. This thesis work hence generated many opportunities for further improving strain performance. For example, besides the targets mentioned in paper IV for *A. oryzae*, there are more SNVs and INDELs on the coding and upstream regulatory regions of many others genes identified through genome sequencing in Paper II. In combination with literature review and the transcriptome data, more targets could be proposed for inverse metabolic engineering.

In addition to oxidative protein folding, ROS may be generated from mitochondria respiration, sterol and unsaturated fatty acid biosynthesis and many more processes. In paper III, we used Dihydrorhodamine 123 to stain the ROS containing cells, and found that the ROS level in the *ROX1* deletion strain is much higher than in the reference cells. Interestingly, with amylase produced, the ROS level seems to be reduced (ROS: ΔROX1> ΔROX1-Amy>WT-Amy (Data not shown)). As mentioned, amino acid biosynthesis generates large amount of NADH. Is the reduced ROS in amylase producing strains due to NADH neutralization? It may be interesting to check how redox is balanced in the ΔROX1-Amy strain and how this may contribute to improve protein production. It may also be interesting to check how the ΔROX1-Amy strain behaves under anaerobic conditions.

In conclusion, this thesis work has generated a series of vector and host platforms, provided deeper understanding of the protein production mechanisms, and proposed promising targets for future engineering.

VI. Acknowledgement

nights or during weekends. Yiming Zhang, Xin Chen, Yongjin Zhou, Jiufu Qin, Guodong Liu, Zongjie Dai, Ruifei Wang, Shuobo Shi, Lei Shi, Boyang Ji, Zheng Wang, Cheng Zhang, Yating Hu, Haichuan Ji, Natapol Pornputtapong, Kwanjeera Wanichthanarak, Kuk-ki Hong, Luis Caspeta-Guadarrama, Kanokarn Kocharin, Juan Valle, Clara Navarrete, Bouke de Jong, Klaas Buijs, Stefan Tippmann, Anastasia Krivoruchko, Payam Ghiaci, Nina Johansson, Adil Mardinoglu and all of you in sysbio group, thank you so much for being around and always be supportive.

A special acknowledgement goes to my master supervisor Professor Uffe Mortensen at DTU and my former boss Dr. Wenping Wu and Ms Lan Tang at Novozymes. I am very grateful for their encouragements and advices when I decided to pursue this study.

I owe my deepest gratitude to my parents, who have been giving me forever and unconditional love. I hope this PhD thesis can be a tiny reward to make them happy and proud. Finally, to my husband, thank you for everything and I hope we will never be apart anymore.

References

Al-Sheikh, H., Watson, A. J., Lacey, G. A., Punt, P. J., MacKenzie, D. A., Jeenes, D. J., Pakula, T., Penttila, M., Alcocer, M. J., Archer, D. B., 2004. Endoplasmic reticulum stress leads to the selective transcriptional downregulation of the glucoamylase gene in *Aspergillus niger*. Mol Microbiol. 53, 1731-42.

Andersen, M. R., Nielsen, J., 2009. Current status of systems biology in *Aspergilli*. Fungal Genetics and Biology. 46, S180-S190.

Bader, M., Muse, W., Ballou, D. P., Gassner, C., Bardwell, J. C., 1999. Oxidative protein folding is driven by the electron transport system. Cell. 98, 217-227.

Bailey, J. E., Sburlati, A., Hatzimanikatis, V., Lee, K., Renner, W. A., Tsai, P. S., 1996. Inverse metabolic engineering: a strategy for directed genetic engineering of useful phenotypes. Biotechnology and Bioengineering. 52, 109-121.

Barrett, C. L., Kim, T. Y., Kim, H. U., Palsson, B. Ø., Lee, S. Y., 2006. Systems biology as a foundation for genome-scale synthetic biology. Current Opinion in Biotechnology. 17, 488-492.

Baumann, K., Adelantado, N., Lang, C., Mattanovich, D., Ferrer, P., 2011. Protein trafficking, ergosterol biosynthesis and membrane physics impact recombinant protein secretion in Pichia pastoris. Microb Cell Fact. 10, 1-15.

Baumann, K., Maurer, M., Dragosits, M., Cos, O., Ferrer, P., Mattanovich, D., 2007. Hypoxic fed batch cultivation of *Pichia pastoris* increases specific and volumetric productivity of recombinant proteins. Biotechnology and Bioengineering. 100, 177-183.

Bause, E., 1983. Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. Biochem. J. 209, 331-336.

Beh, C. T., RoSE, M. D., 1995. Two redundant systems maintain levels of resident proteins within the yeast endoplasmic reticulum. Proceedings of the National Academy of Sciences. 92, 9820-9823.

Behringer, R. R., Ryan, T. M., Reilly, M. P., Asakura, T., Palmiter, R. D., Brinster, R. L., Townes, T. M., 1989. Synthesis of functional human hemoglobin in transgenic mice. Science. 245, 971.

Belcher, J. D., Beckman, J. D., Balla, G., Balla, J., Vercellotti, G., 2010. Heme degradation and vascular injury. Antioxid Redox Signal. 12, 233-48.

Benesch, R., Benesch, R. E., 1963. Some Relations between Structure and Function in Hemoglobin. J Mol Biol. 6, 498-505.

Binder, M., Schanz, M., Hartig, A., 1991. Vector-mediated overexpression of catalase A in the yeast *Saccharomyces cerevisiae* induces inclusion body formation. European journal of cell biology. 54, 305-312.

Bouzarelou, D., Billini, M., Roumelioti, K., Sophianopoulou, V., 2008. EglD, a putative endoglucanase, with an expansin like domain is localized in the conidial cell wall of *Aspergillus nidulans*. Fungal Genetics and Biology. 45, 839-850.

Carvalho, N. D. S. P., Arentshorst, M., Kooistra, R., Stam, H., Sagt, C. M., van den Hondel, C. A., Ram, A. F. J., 2011. Effects of a defective ERAD pathway on growth and heterologous protein production in *Aspergillus niger*. Appl Microbiol Biot. 1-17.

Cha, H. J., Choi, S. S., Yoo, Y. J., Bentley, W. E., 1997. Enhancement of production of cloned glucoamylase under conditions of low aeration from recombinant yeast using a *SUC2* promoter. Process Biochemistry. 32, 679-684.

Chemler, J. A., Yan, Y., Koffas, M. A., 2006. Biosynthesis of isoprenoids, polyunsaturated fatty acids and flavonoids in *Saccharomyces cerevisiae*. Microb Cell Fact. 5, 20.

Christensen, T., Woeldike, H., Boel, E., Mortensen, S. B., Hjortshoej, K., Thim, L., Hansen, M. T., 1988. High level expression of recombinant genes in *Aspergillus oryzae*. Nature Biotechnology. 6, 1419-1422.

Clerc, S., Hirsch, C., Oggier, D. M., Deprez, P., Jakob, C., Sommer, T., Aebi, M., 2009. Htm1 protein generates the N-glycan signal for glycoprotein degradation in the endoplasmic reticulum. The Journal of Cell Biology. 184, 159.

Coghlan, D., Jones, G., Denton, K. A., Wilson, M. T., Chan, B., Harris, R., Woodrow, J. R., Ogden, J. E., 1992. Structural and functional characterisation of recombinant human haemoglobin A expressed in *Saccharomyces cerevisiae*. Eur J Biochem. 207, 931-6.

Conibear, E., Stevens, T. H., 1998. Multiple sorting pathways between the late Golgi and the vacuole in yeast. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research. 1404, 211-230.

Cosgrove, D. J., 2000. Loosening of plant cell walls by expansins. Nature. 407, 321-326.

De Groot, P. W., Ram, A. F., Klis, F. M., 2005. Features and functions of covalently linked proteins in fungal cell walls. Fungal Genetics and Biology. 42, 657-675.

Demain, A. L., Vaishnav, P., 2009. Production of recombinant proteins by microbes and higher organisms. Biotechnology Advances. 27, 297-306.

Dieryck, W., Pagnier, J., Poyart, C., Marden, M. C., Gruber, V., Bournat, P., Baudino, S., Mérot, B.,

1997. Human haemoglobin from transgenic tobacco. Nature. 386, 29.

Dobson, C. M., Principles of protein folding, misfolding and aggregation. Seminars in Cell & Developmental Biology, Vol. 15. Elsevier, 2004, pp. 3-16.

Dong, J., Schmittmann, B., Zia, R. K., 2007. Towards a model for protein production rates. Journal of Statistical Physics. 128, 21-34.

Dugar, D., Stephanopoulos, G., 2011. Relative potential of biosynthetic pathways for biofuels and bio-based products. Nature Biotechnology. 29, 1074-1078.

Entian, K.-D., Kötter, P., 2007. 25 Yeast Genetic Strain and Plasmid Collections. In: Ian, S., Michael, J. R. S., Eds.), Methods in microbiology. vol. Volume 36. Academic Press, pp. 629-666.

Feizi, A., Österlund, T., Petranovic, D., Bordel, S., Nielsen, J., 2013. Genome-scale modeling of the protein secretory machinery in yeast. PloS one. 8, e63284.

Ferrer-Miralles, N., Domingo-Espń, J., Corchero, J. L., Vázquez, E., Villaverde, A., 2009. Microbial factories for recombinant pharmaceuticals. Microb Cell Fact. 8, 17.

Flynn, G. C., Pohl, J., Flocco, M. T., Rothman, J. E., 1991. Peptide-binding specificity of the molecular chaperone BiP. Nature. 353, 726-730.

Galagan, J. E., Calvo, S. E., Cuomo, C., Ma, L.-J., Wortman, J. R., Batzoglou, S., Lee, S.-I., Baştürkmen, M., Spevak, C. C., Clutterbuck, J., 2005. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature. 438, 1105-1115.

Gasser, B., Maurer, M., Rautio, J., Sauer, M., Bhattacharyya, A., Saloheimo, M., Penttilä, M., Mattanovich, D., 2007. Monitoring of transcriptional regulation in *Pichia pastoris* under protein production conditions. BMC Genomics. 8, 179.

Gauss, R., Jarosch, E., Sommer, T., Hirsch, C., 2006. A complex of Yos9p and the HRD ligase integrates endoplasmic reticulum quality control into the degradation machinery. Nature cell biology. 8, 849-854.

Gething, M.-J., Role and regulation of the ER chaperone BiP. Seminars in Cell & Developmental Biology, Vol. 10. Elsevier, 1999, pp. 465-472.

Geysens, S., Whyteside, G., Archer, D. B., 2009. Genomics of protein folding in the endoplasmic reticulum, secretion stress and glycosylation in the *Aspergilli*. Fungal Genetics and Biology. 46, S121-S140.

Goffeau, A., Barrell, B., Bussey, H., Davis, R., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J., Jacq, C., Johnston, M., 1996. Life with 6000 genes. Science. 274, 546-567.

Gopal, C., Hammond, J., Use of genetically-modified yeasts for beer production. Eur Brew Conv Proc Congr, Vol. 5, 1992, pp. 297-306.

Group, T. F., 2011. World Enzymes, Industry Study with Forecasts for 2015 & 2020.

Harmsen, M., Bruyne, M., Raue, H., Maat, J., 1996. Overexpression of binding protein and disruption of the PMR1 gene synergistically stimulate secretion of bovine prochymosin but not plant thaumatin in yeast. Appl Microbiol Biot. 46, 365-370.

Hasslacher, M., Schall, M., Hayn, M., Bona, R., Rumbold, K., Lückl, J., Griengl, H., Kohlwein, S. D., Schwab, H., 1997. High-Level Intracellular Expression of Hydroxynitrile Lyase from the Tropical Rubber TreeHevea brasiliensisin Microbial Hosts. Protein expression and purification. 11, 61-71.

Hoffman, M., Gora, M., Rytka, J., 2003. Identification of rate-limiting steps in yeast heme biosynthesis. Biochem Biophys Res Commun. 310, 1247-53.

Hoffman, S. J., Looker, D. L., Roehrich, J. M., Cozart, P. E., Durfee, S. L., Tedesco, J. L., Stetler, G. L., 1990. Expression of fully functional tetrameric human hemoglobin in *Escherichia coli*. Proc Natl Acad Sci U S A. 87, 8521-5.

Hou, J., Österlund, T., Liu, Z., Petranovic, D., Nielsen, J., 2012a. Heat shock response improves heterologous protein secretion in *Saccharomyces cerevisiae*. Appl Microbiol Biot. 1-10.

Hou, J., Tyo, K., Liu, Z., Petranovic, D., Nielsen, J., 2012b. Engineering of vesicle trafficking improves heterologous protein secretion in *Saccharomyces cerevisiae*. Metab Eng. 14, 120-7.

Hou, J., Tyo, K. E. J., Liu, Z., Petranovic, D., Nielsen, J., 2012c. Metabolic Engineering of Recombinant Protein Secretion by *Saccharomyces cerevisiae*. FEMS yeast research. n/a-n/a.

Huang, M., Bao, J., Nielsen, J., 2014. Biopharmaceutical protein production by *Saccharomyces cerevisiae*: current state and future prospects.

Idiris, A., Tohda, H., Kumagai, H., Takegawa, K., 2010. Engineering of protein secretion in yeast: strategies and impact on protein production. Appl Microbiol Biot. 86, 403-417.

Jahn, R., Scheller, R. H., 2006. SNAREs—engines for membrane fusion. Nature Reviews Molecular Cell Biology. 7, 631-643.

Katzmann, D. J., Babst, M., Emr, S. D., 2001. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. Cell. 106, 145-155.

Kaufman, R. J., 1999. Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls. Genes & development. 13, 1211-1233.

Kelly, J. M., Hynes, M., 1987. Multiple copies of the amdS gene of *Aspergillus nidulans* cause titration of trans-acting regulatory proteins. Current genetics. 12, 21-31.

Keng, T., 1992. *HAP1* and *ROX1* form a regulatory pathway in the repression of HEM13 transcription in *Saccharomyces cerevisiae*. Molecular and cellular biology. 12, 2616.

Kim, Y., Nandakumar, M. P., Marten, M. R., 2007. Proteomics of filamentous fungi. TRENDS in Biotechnology. 25, 395-400.

Kim, Y., Nandakumar, M. P., Marten, M. R., 2008. The state of proteome profiling in the fungal genus *Aspergillus*. Brief Funct Genomic Proteomic. 7, 87-94.

Kimata, Y., Ishiwata‐Kimata, Y., Yamada, S., Kohno, K., 2006. Yeast unfolded protein response pathway regulates expression of genes for anti‐oxidative stress and for cell surface proteins. Genes to Cells. 11, 59-69.

Kitano, H., 2002. Systems biology: a brief overview. Science. 295, 1662-1664.

Kjeldsen, T., 2000. Yeast secretory expression of insulin precursors. Appl Microbiol Biot. 54, 277-286.

Kjeldsen, T., Frost Pettersson, A., Hach, M., 1999. The role of leaders in intracellular transport and secretion of the insulin precursor in the yeast *Saccharomyces cerevisiae*. Journal of Biotechnology. 75, 195-208.

Kjeldsen, T., Hach, M., Balschmidt, P., Havelund, S., Pettersson, A. F., Markussen, J., 1998. Prepro-Leaders Lacking N-Linked Glycosylation for Secretory Expression in the Yeast *Saccharomyces cerevisiae*. Protein expression and purification. 14, 309-316.

Kobayashi, T., Abe, K., Asai, K., Gomi, K., Juvvadi, P. R., Kato, M., Kitamoto, K., Takeuchi, M., Machida, M., 2007. Genomics of *Aspergillus oryzae*. Biosci Biotech Bioch. 71, 646.

Koda, A., Minetoki, T., Ozeki, K., Hirotsune, M., 2004. Translation efficiency mediated by the 5'untranslated region greatly affects protein production in *Aspergillus oryzae*. Appl Microbiol Biot. 66, 291-296.

Komar, A. A., Kommer, A., Krasheninnikov, I. A., Spirin, A. S., 1997. Cotranslational folding of globin. J Biol Chem. 272, 10646-51.

Kumar, R., 1995. Recombinant Hemoglobins as Blood Substitutes - a Biotechnology Perspective. P Soc Exp Biol Med. 208, 150-158.

Kuratsu, M., Taura, A., Shoji, J.-y., Kikuchi, S., Arioka, M., Kitamoto, K., 2007. Systematic analysis of SNARE localization in the filamentous fungus *Aspergillus oryzae*. Fungal Genetics and Biology. 44, 1310-1323.

Kwast, K. E., Burke, P. V., Brown, K., Poyton, R. O., 1997. *REO1* and *ROX1* are alleles of the same gene which encodes a transcriptional repressor of hypoxic genes in *Saccharomyces cerevisiae*. Current genetics. 32, 377-383.

Kwast, K. E., Lai, L. C., Menda, N., James III, D. T., Aref, S., Burke, P. V., 2002. Genomic analyses of anaerobically induced genes in *Saccharomyces cerevisiae*: functional roles of Rox1 and other factors in mediating the anoxic response. Journal of Bacteriology. 184, 250.

Lee, S. K., Chou, H., Ham, T. S., Lee, T. S., Keasling, J. D., 2008. Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. Curr Opin Biotechnol. 19, 556-63.

Lee, S. Y., Lee, D.-Y., Kim, T. Y., 2005. Systems biotechnology for strain improvement. TRENDS in Biotechnology. 23, 349-358.

Lewis, M. J., Nichols, B. J., Prescianotto-Baschong, C., Riezman, H., Pelham, H. R., 2000. Specific retrieval of the exocytic SNARE Snc1p from early yeast endosomes. Molecular biology of the cell. 11, 23-38.

Liu, Z., Österlund, T., Hou, J., Petranovic, D., Nielsen, J., 2013. Anaerobic α-amylase production and secretion with fumarate as the final electron acceptor in yeast. Applied and Environmental Microbiology.

Liu, Z., Tyo, K. E. J., Mart ńez, J. L., Petranovic, D., Nielsen, J., 2012. Different expression systems for production of recombinant proteins in *Saccharomyces cerevisiae*. Biotechnology and Bioengineering.

Lubertozzi, D., Keasling, J. D., 2009. Developing *Aspergillus* as a host for heterologous expression. Biotechnology Advances. 27, 53-75.

Machida, M., 2002. Progress of *Aspergillus oryzae* genomics. Advances in applied microbiology. 51, 81-106.

Machida, M., Asai, K., Sano, M., Tanaka, T., Kumagai, T., Terai, G., Kusumoto, K., Arima, T., Akita, O., Kashiwagi, Y., 2005. Genome sequencing and analysis of *Aspergillus oryzae*. Nature. 438, 1157-1161.

Machida, M., Yamada, O., Gomi, K., 2008. Genomics of *Aspergillus oryzae*: learning from the history of Koji mold and exploration of its future. DNA research. 15, 173-183.

Martinez, J. L., Liu, L., Petranovic, D., Nielsen, J., 2012. Pharmaceutical protein production by yeast: towards production of human blood proteins by microbial fermentation. Curr Opin Biotechnol.

Maruyama, J. i., Kitamoto, K., 2007. Differential distribution of the endoplasmic reticulum network in filamentous fungi. Fems Microbiol Lett. 272, 1-7.

Mattanovich, D., Branduardi, P., Dato, L., Gasser, B., Sauer, M., Porro, D., 2012. Recombinant protein production in yeasts. Methods in molecular biology (Clifton, NJ). 824**,** 329.

Meusser, B., Hirsch, C., Jarosch, E., Sommer, T., 2005. ERAD: the long road to destruction. Nature cell biology. 7**,** 766-772.

Molinari, M., Helenius, A., 2000. Chaperone selection during glycoprotein translocation into the endoplasmic reticulum. Science. 288**,** 331-333.

Mori, K., Ogawa, N., Kawahara, T., Yanagi, H., Yura, T., 1998. Palindrome with Spacer of One Nucleotide Is Characteristic of the cis-Acting Unfolded Protein Response Element in *Saccharomyces cerevisiae*. Journal of Biological Chemistry. 273**,** 9912-9920.

Mustacchi, R., Hohmann, S., Nielsen, J., 2006. Yeast systems biology to unravel the network of life. Yeast. 23**,** 227-238.

Nabel, G., Baltimore, D., 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature. 326**,** 711-713.

Nielsen, J., Jewett, M. C., 2008. Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. FEMS yeast research. 8**,** 122-131.

Nierman, W. C., Pain, A., Anderson, M. J., Wortman, J. R., Kim, H. S., Arroyo, J., Berriman, M., Abe, K., Archer, D. B., Bermejo, C., 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature. 438**,** 1151-1156.

Nijkamp, J. F., van den Broek, M., Datema, E., de Kok, S., Bosman, L., Luttik, M. A., Daran-Lapujade, P., Vongsangnak, W., Nielsen, J., Heijne, W. H., 2012. De novo sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN. PK113-7D, a model for modern industrial biotechnology. Microb Cell Fact. 11**,** 36.

Pakula, T. M., Laxell, M., Huuskonen, A., Uusitalo, J., Saloheimo, M., Penttila, M., 2003. The effects of drugs inhibiting protein secretion in the filamentous fungus *Trichoderma reesei*. Evidence for down-regulation of genes that encode secreted proteins in the stressed cells. J Biol Chem. 278**,** 45011-20.

Papini, M., Salazar, M., Nielsen, J., 2010. Systems Biology of Industrial Microorganisms

Biosystems Engineering I. In: Wittmann, C., Krull, R., Eds.). vol. 120. Springer Berlin / Heidelberg, pp. 51-99.

Patil, C., Walter, P., 2001. Intracellular signaling from the endoplasmic reticulum to the nucleus: the unfolded protein response in yeast and mammals. Current opinion in cell biology. 13**,** 349-355.

Payne, T., Finnis, C., Evans, L., Mead, D., Avery, S., Archer, D., Sleep, D., 2008. Modulation of chaperone gene expression in mutagenized *Saccharomyces cerevisiae* strains developed for recombinant human albumin production results in increased production of multiple heterologous proteins. Applied and Environmental Microbiology. 74**,** 7759-7766.

Pedersen, H., Carlsen, M., Nielsen, J., 1999. Identification of enzymes and quantification of metabolic fluxes in the wild type and in a recombinant *Aspergillus oryzae* strain. Applied and Environmental Microbiology. 65**,** 11-19.

Pel, H. J., De Winde, J. H., Archer, D. B., Dyer, P. S., Hofmann, G., Schaap, P. J., Turner, G., de Vries, R. P., Albang, R., Albermann, K., 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nature Biotechnology. 25**,** 221-231.

Potgieter, T. I., Cukan, M., Drummond, J. E., Houston-Cummings, N. R., Jiang, Y., Li, F., Lynaugh, H., Mallem, M., McKelvey, T. W., Mitchell, T., 2009. Production of monoclonal antibodies by glycoengineered *Pichia pastoris*. Journal of Biotechnology. 139**,** 318-325.

Raden, D., Hildebrandt, S., Xu, P., Bell, E., Doyle, F., Robinson, A., 2005. Analysis of cellular response to protein overexpression. IEE Proceedings-Systems Biology. 152**,** 285-289.

Rakestraw, J. A., Sazinsky, S. L., Piatesi, A., Antipov, E., Wittrup, K. D., 2009. Directed evolution of a secretory leader for the improved expression of heterologous proteins and full length antibodies in *Saccharomyces cerevisiae*. Biotechnology and Bioengineering. 103**,** 1192-1201.

Randez-Gil, F., Sanz, P., 1993. Expression of *Aspergillus oryzae* α-amylase gene in *Saccharomyces cerevisiae*. Fems Microbiol Lett. 112**,** 119-124.

Raymond, C. K., Howald-Stevenson, I., Vater, C., Stevens, T., 1992. Morphological classification of the yeast vacuolar protein sorting mutants: evidence for a prevacuolar compartment in class E vps mutants. Molecular biology of the cell. 3**,** 1389-1402.

Rigoulet, M., Aguilaniu, H., Avéret, N., Bunoust, O., Camougrand, N., Grandier-Vazeille, X., Larsson, C., Pahlman, I.-L., Manon, S., Gustafsson, L., 2004. Organization and regulation of the cytosolic NADH metabolism in the yeast *Saccharomyces cerevisiae*. Molecular and cellular biochemistry. 256**,** 73-81.

Ruddock, L. W., Molinari, M., 2006. N-glycan processing in ER quality control. Journal of cell science. 119**,** 4373.

Ruohonen, L., Toikkanen, J., Outola, M., Soderlund, H., Keranen, S., 1997. Enhancement of protein secretion in *Saccharomyces cerevisiae* by overproduction of Sso protein, a late-acting

component of the secretory machinery. Yeast. 13, 337-351.

Saksena, S., Sun, J., Chu, T., Emr, S. D., 2007. ESCRTing proteins in the endocytic pathway. Trends in biochemical sciences. 32, 561-573.

Saloheimo, M., Valkonen, M., Penttilä, M., 2003. Activation mechanisms of the *HACI*-mediated unfolded protein response in filamentous fungi. Molecular microbiology. 47, 1149-1161.

Schneider, J. C., Guarente, L., 1991. Regulation of the yeast *CYT1* gene encoding cytochrome c1 by *HAP1* and *HAP2/3/4*. Molecular and cellular biology. 11, 4934-4942.

Segev, N., 2009. Trafficking Inside Cells: Pathways, Mechanisms, and Regulation. Springer.

Semenza, J. C., Hardwick, K. G., Dean, N., Pelham, H. R., 1990. *ERD2*, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. Cell. 61, 1349-1357.

Shelikoff, M., Sinskey, A., Stephanopoulos, G., 1996. A modeling framework for the study of protein glycosylation. Biotechnology and Bioengineering. 50, 73-90.

Shoji, J., Arioka, M., Kitamoto, K., 2008. Dissecting cellular components of the secretory pathway in filamentous fungi: insights into their application for protein production. Biotechnology letters. 30, 7-14.

Sidorenko, Y., Antoniukas, L., Schulze-Horsel, J., Kremling, A., Reichl, U., 2008. Mathematical Model of Growth and Heterologous Hantavirus Protein Production of the Recombinant Yeast *Saccharomyces cerevisiae*. Engineering in Life Sciences. 8, 399-414.

Sidrauski, C., Walter, P., 1997. The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. Cell. 90, 1031-1039.

Sims, A. H., Gent, M. E., Robson, G. D., Dunn-Coleman, N. S., Oliver, S. G., 2004. Combining transcriptome data with genomic and cDNA sequence alignments to make confident functional assignments for *Aspergillus nidulans* genes. Mycological research. 108, 853-857.

Smith, B. J., Gutierrez, P., Guerrero, E., Brewer, C. J., Henderson, D. P., 2011. Development of a method to produce hemoglobin in a bioreactor culture of *Escherichia coli* BL21(DE3) transformed with a plasmid containing Plesiomonas shigelloides heme transport genes and modified human hemoglobin genes. Appl Environ Microbiol. 77, 6703-5.

Spohr, A., Carlsen, M., Nielsen, J., Villadsen, J., 1998. α-Amylase production in recombinant *Aspergillus oryzae* during fed-batch and continuous cultivations. J Ferment Bioeng. 86, 49-56.

Sudha Kumari, S. M., Mayor, S., 2010. Endocytosis unplugged: multiple ways to enter the cell. Cell research. 20, 256-275.

Tatsumi, A., Kikuma, T., Arioka, M., Kitamoto, K., 2006. *Aovps24*, a homologue of *VPS24*, is required for vacuolar formation which could maintain proper growth and development in the filamentous fungus *Aspergillus oryzae*. Biochem Bioph Res Co. 347, 970-978.

Ter Linde, J. J. M., Steensma, H. Y., 2002. A microarray‐assisted screen for potential Hap1 and Rox1 target genes in *Saccharomyces cerevisiae*. Yeast. 19, 825-840.

Tholudur, A., Ramirez, W. F., McMillan, J. D., 1999. Mathematical modeling and optimization of cellulase protein production using Trichoderma reesei RL-P37. Biotechnology and Bioengineering. 66, 1-16.

Toret, C. P., Drubin, D. G., 2006. The budding yeast endocytic pathway. Journal of cell science. 119, 4585-4587.

Travers, K. J., Patil, C. K., Wodicka, L., Lockhart, D. J., Weissman, J. S., Walter, P., 2000. Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. Cell. 101, 249-258.

Treusch, S., Hamamichi, S., Goodman, J. L., Matlack, K. E., Chung, C. Y., Baru, V., Shulman, J. M., Parrado, A., Bevis, B. J., Valastyan, J. S., 2011. Functional links between Aβ toxicity, endocytic trafficking, and Alzheimer's disease risk factors in yeast. Science. 334, 1241-1245.

Tu, B. P., Weissman, J. S., 2002. The FAD-and O (2)-dependent reaction cycle of Ero1-mediated oxidative protein folding in the endoplasmic reticulum. Molecular cell. 10, 983.

Tu, B. P., Weissman, J. S., 2004. Oxidative protein folding in eukaryotes mechanisms and consequences. The Journal of Cell Biology. 164, 341-346.

Tyo, K. E. J., Liu, Z., Petranovic, D., Nielsen, J., 2012. Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress. BMC biology. 10, 16.

Umaña, P., Bailey, J. E., 1997. A mathematical model of N-linked glycoform biosynthesis. Biotechnology and Bioengineering. 55, 890-908.

van der Heide, M., Hollenberg, C., van der Klei, I., Veenhuis, M., 2002. Overproduction of BiP negatively affects the secretion of *Aspergillus niger* glucose oxidase by the yeast *Hansenula polymorpha*. Appl Microbiol Biot. 58, 487-494.

van Dijken, J. P., Scheffers, W. A., 1986. Redox balances in the metabolism of sugars by yeasts. Fems Microbiol Lett. 32, 199-224.

Verdoes, J. C., Punt, P., Hondel, C., 1995. Molecular genetic strain improvement for the overproduction of fungal proteins by

filamentous fungi. Appl Microbiol Biot. 43, 195-205.

Verstrepen, K. J., Chambers, P. J., Pretorius, I. S., 2006. The development of superior yeast strains for the food and beverage industries: challenges, opportunities and potential benefits. Yeasts in Food and Beverages. Springer, pp. 399-444.

Vida, T. A., Huyer, G., Emr, S. D., 1993. Yeast vacuolar proenzymes are sorted in the late Golgi complex and transported to the vacuole via a prevacuolar endosome-like compartment. The Journal of Cell Biology. 121, 1245-1256.

Vijayendran, C., Flaschel, E., 2010. Impact of profiling technologies in the understanding of recombinant protein production. Biosystems Engineering II. Springer, pp. 45-70.

Vongsangnak, W., Olsen, P., Hansen, K., Krogsgaard, S., Nielsen, J., 2008. Improved annotation through genome-scale metabolic modeling of *Aspergillus oryzae*. BMC Genomics. 9, 245.

Vongsangnak, W., Salazar, M., Hansen, K., Nielsen, J., 2009. Genome-wide analysis of maltose utilization and regulation in *aspergilli*. Microbiology. 155, 3893-3902.

Walker, G. M., 1998. Yeast physiology and biotechnology. John Wiley & Sons.

Wang, B., Guo, G., Wang, C., Lin, Y., Wang, X., Zhao, M., Guo, Y., He, M., Zhang, Y., Pan, L., 2010. Survey of the transcriptome of *Aspergillus oryzae* via massively parallel mRNA sequencing. Nucleic Acids Res. 38, 5075-5087.

Wildt, S., Gerngross, T. U., 2005. The humanization of N-glycosylation pathways in yeast. Nat Rev Micro. 3, 119-128.

Ye, Y., Shibata, Y., Yun, C., Ron, D., Rapoport, T. A., 2004. A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. Nature. 429, 841-847.

Zhao, X. J., Raitt, D., P, V. B., Clewell, A. S., Kwast, K. E., Poyton, R. O., 1996. Function and expression of flavohemoglobin in *Saccharomyces cerevisiae*. Evidence for a role in the oxidative stress response. J Biol Chem. 271, 25131-8.

Zhou, J., Zhang, H., Liu, X., Wang, P. G., Qi, Q., 2007. Influence of N-glycosylation on *Saccharomyces cerevisiae* morphology: a golgi glycosylation mutant shows cell division defects. Current microbiology. 55, 198-204.