Universidade do Minho
Escola de Engenharia

Tatiana Sofia da Quinta Vidal Aguiar

Understanding the biotechnological potential of *Ashbya gossypii*

Julho de 2013

Tatiana Sofia da Quinta Vidal Aguiar **Understanding the biotechnological potential of *Ashbya gossypii***

UMinho | 2013

Universidade do Minho
Escola de Engenharia

Tatiana Sofia da Quinta Vidal Aguiar

Understanding the biotechnological potential of *Ashbya gossypii*

Tese de Doutoramento em Bioengenharia

Trabalho realizado sob a orientação da
Doutora Lucília Domingues
e da
Professora Merja Penttilä

Julho de 2013

Autor

Tatiana Sofia da Quinta Vidal Aguiar

Título da tese

Understanding the biotechnological potential of *Ashbya gossypii*

Orientadoras

Doutora Lucília Domingues

Professora Merja Penttilä

Ano de conclusão 2013

Doutoramento em Bioengenharia

Universidade do Minho, 23 de Julho de 2013

_____

ACKNOWLEDGEMENTS

I begin by acknowledging the Portuguese government for having financially supported the work presented in this thesis through the Fundação para a Ciência e a Tecnologia (FCT) MIT-Portugal Program PhD grant SFRH/BD/39112/2007 and Project AshByofactory PTDC/EBB-EBI/101985/2008 co-funded by FEDER through COMPETE – Programa Operacional Factores de Competitividade (FCOMP-01-0124-FEDER-009701).

I also thank the Centre of Biological Engineering (CEB) of the University of Minho and the Technical Research Centre of Finland (VTT Biotechnology) for the working conditions provided.

My deepest gratitude goes to my supervisors, Lucília Domingues (CEB) and Merja Penttilä (VTT), for creating the conditions for this PhD project to go on. I am thankful for all the possibilities given to me to develop novel skills and for the vast knowledge generously shared with me. A special recognition is due to Lucília for the thoughtful guidance since the beginning of my scientific carrier and for the ceaseless support and encouragement throughout these years.

My warmest thanks go also to Marilyn Wiebe for coordinating my work at VTT and for all the prompt help, comments, expertise and friendship offered throughout this PhD work.

My most sincere thanks are due to my co-authors Cláudia Dinis (CEB), Frederico Magalhães (CEB), Carla Oliveira (CEB), Orquídea Ribeiro (CEB), Hannu Maaheimo (VTT), Mikko Arvas (VTT) and Anamari Heiskanen (Glycos Finland). I thank their contributions to the research work and writing of the manuscripts. Dominik Mojzita (VTT) and Mari Häkkinen (VTT) are acknowledged for their assistance with the microarray sample preparation and data acquisition in Chapter 1. I also thank Tiina Pakula for useful tips regarding the execution of the proteomic analyses in Chapter 1 and Aili Grundstöm for the technical assistance provided. Glycos Finland Ltd is acknowledged for the mass spectrometry analyses in Chapter 2 and Olli Aitio (Finnish Biological NMR Center, University of Helsinki) for helpful assistance with the interpretation of the spectroscopic data in Chapter 2.

All my colleagues at CEB are thanked for the excellent working atmosphere created and for their readiness to provide help. To Sofia Costa (thanks for the brainstormings), Diana

Mendes, Francisco Pereira, Daniel Machado, Daniel Gomes and Joana Cunha I address a special thank for their invaluable friendship and for the enriching life experiences we enjoyed, providing balance to life inside the lab. Thank you all for cheering up my days.

I am also grateful to everybody working at VTT Biotechnology that kindly helped me during my visits there. A special thank is due to Jaana Uusitalo for receiving me at her lab and to Merja A., Tarja L., Eila L., Michael B., Juha T., Yanming W., Anu T., Susanna K. Juha-Pekka P. and Toni P. for having contributed to make my visiting periods at VTT a personally enjoyable and fruitful experience.

To all my friends from Braga and Helsinki I am thankful for all the good moments shared. Thank you for making this period of my life so wonderful and memorable.

A special thank is addressed to Pedro for all the love and support throughout these years and to his family, for the kind help offered during the preparation of this manuscript.

Finally, my most special thanks go to my parents, Natália and João, who have raised me with the greatest care and attention, and provided me all the love and support that only parents can.

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

Governo da República Portuguesa

COMPETE

QUADRO DE REFERÊNCIA ESTRATÉGICO NACIONAL

UNIÃO EUROPEIA
Fundo Europeu de Desenvolvimento Regional

Dedico esta tese
aos meus pais
e avô José

ABSTRACT

Understanding the biotechnological potential of *Ashbya gossypii*

Ashbya gossypii (syn. *Eremothecium gossypii*) is a filamentous Saccharomycete which has long been known in the scientific and industrial communities, first as a cotton pathogen and subsequently as a riboflavin overproducer. This fungus has the smallest free living eukaryotic genome known, which shares a high degree of gene homology and gene order conservation with that of *Saccharomyces cerevisiae*. It has haploid nuclei and is prone to genetic manipulation, allowing the use of simple PCR-based gene targeting strategies and free propagation of plasmids containing replicons from *S. cerevisiae*. Moreover, it grows well in a variety of defined, complex and waste-based media, and has a long history of safe use in the industrial production of riboflavin (vitamin B2). These unique features led to expanded interest in *A. gossypii* as a "minimal" host for the production of, yet unexploited, valuable compounds other than riboflavin, namely heterologous proteins. However, although two heterologous proteins have already been successfully secreted by *A. gossypii*, little is still known about the protein secretion ability of this fungus.

To further understand the biotechnological potential of *A. gossypii* as a cell factory organism, this thesis primarily focused on the characterization of the *A. gossypii* protein secretory pathway at the genomic, transcriptomic and proteomic levels. Based on experimental observations and on the data from the genomic and transcriptomic analyses, a hydrolytic enzyme, invertase, was deduced to be natively secreted by *A. gossypii* and molecularly characterized. To further address *A. gossypii* as a heterologous protein producer, the *β*-galactosidase from *Aspergillus niger* was expressed in this fungus under the regulation of different native and heterologous promoters. In addition, a new molecular tool for use in *A. gossypii* was developed to generate mutant strains free of exogenous selection markers, allowing the creation of improved *A. gossypii* strains suitable for industrial applications.

The results presented in this thesis demonstrate that the amount and variety of proteins natively secreted by *A. gossypii* to the culture medium is rather low, being more similar to that of yeast than to that of other filamentous fungi. Similarly, the *N*-glycosylation patterns produced by *A. gossypii* are generally more similar to those produced by yeast than to those produced by other filamentous fungi. However, extensive hyperglycosylation only occurs in certain culture conditions. Like other filamentous fungi, *A. gossypii* also seems to be able to trim its *N*-glycans.

A conventional unfolded protein response (UPR) was not activate in *A. gossypii* in response to heterologous protein secretion nor to dithiothreitol (DTT)-induced secretion stress, as generally observed in other fungi. However, the transcriptional responses of *A. gossypii* to DTT-induced stress indicate that alternative mechanisms exist in this fungus to cope with protein secretion stress.

The *A. gossypii* invertase was demonstrated to be encoded by the *AFR529W* (*AgSUC2*) gene, which is functionally complemented by the *S. cerevisiae SUC2* (*ScSUC2*) gene. The signal sequences of both AgSuc2p and ScSuc2p were able to direct the secretion of invertase into the culture medium in *A. gossypii*. Similarly to the invertases of other fungi, the expression of the *A. gossypii* invertase is regulated by the sugars present in the medium. These results expanded our knowledge about the *A. gossypii* native secretion capacities, being invertase the second hydrolytic enzyme natively secreted by this fungus to be experimental characterized.

The *β*-galactosidase from *A. niger* was successfully expressed in *A. gossypii* under the regulation of different promoters. The native *TEF* promoter revealed to be the best promoter for overexpressing heterologous *β*-galactosidase in *A. gossypii*, inducing 2-fold higher secreted activity than the *A. gossypii GPD* promoter and 7-fold higher than the *S. cerevisiae PGK1* and *ADH1* promoters. The levels of active *β*-galactosidase secreted by a *S. cerevisiae* laboratorial strain transformed with the same plasmids were up to 37 times lower than those obtained in *A. gossypii*. The secretion of active *β*-galactosidase by *A. gossypii* was approximately 1.5-fold higher in glycerol- than in glucose-containing medium. These results highlight the potential of *A. gossypii* as a heterologous protein producer and open new opportunities to further optimize its secretion capacities using this enzyme as a model. As its activity is easy to detect, the screening for improved secretion will be facilitated.

The Cre-*lox*P recombination system, which has been widely used in other organisms, was successfully adapted for use in *A. gossypii*, allowing the removal and reuse of selection marker genes in targeted engineering of this fungus. The set of disruption cassettes and plasmids constructed greatly expand the possibilities for genetically engineer *A. gossypii*, being these suitable for use in both laboratorial and industrial strains, as they do not required any predetermined genetic background. In the future, targeted improvement of *A. gossypii* strains for industrial applications will benefit from this molecular tool.

RESUMO

Compreensão do potencial biotecnológico de *Ashbya gossypii*

O *Ashbya gossypii* (sin. *Eremothecium gossypii*) é um Saccharomycete filamentoso há muito conhecido nas comunidades científica e industrial, primeiro como patogénio do algodão e subsequentemente como super-produtor de riboflavina. Este fungo tem o mais pequeno genoma eucariótico não parasitário conhecido, o qual partilha um elevado grau de homologia e conservação de ordem génica com o genoma da *Saccharomyces cerevisiae*. Ele tem núcleos haplóides e é de fácil manipulação genética, permitindo o uso de estratégias de manipulação genética direcionada simples e a livre propagação de plasmídeos contendo sequências de replicação de *S. cerevisiae*. Além do mais, cresce bem numa variedade de meios definidos, complexos e baseados em resíduos, e tem um longo historial de utilização segura na produção industrial de riboflavina (vitamina B2). Estas características únicas levaram a um interesse alargado em *A. gossypii* como hospedeiro "mínimo" para a produção de outros compostos de interesse, ainda inexplorados, para além da riboflavina, nomeadamente proteínas heterólogas. No entanto, apesar de duas proteínas heterólogas já terem sido secretadas por *A. gossypii* com sucesso, pouco ainda se conhece acerca da capacidade de secreção de proteínas deste fungo.

Para melhor compreender o potencial biotecnológico de *A. gossypii* como fábrica celular, esta tese focou-se primariamente na caracterização da via de secreção de proteínas do *A. gossypii* aos níveis genómico, transcritómico e proteómico. Com base em observações experimentais e nos dados das análises genómicas e transcritómicas, uma enzima hidrolítica, invertase, foi prevista ser nativamente secretada por *A. gossypii* e molecularmente caracterizada. Para melhor avaliar o *A. gossypii* como produtor de proteínas heterólogas, a *β*-galactosidase de *Aspergillus niger* foi expressa neste fungo sob a regulação de diferentes promotores nativos e heterólogos. Adicionalmente, uma nova ferramenta molecular para uso em *A. gossypii* foi desenvolvida para gerar estirpes mutantes livres de marcadores de selecção exógenos, permitindo a criação de estirpes de *A. gossypii* melhoradas para aplicações industriais.

Os resultados apresentados nesta tese demonstram que a quantidade e variedade de proteínas nativamente secretadas por *A. gossypii* para o meio de cultura é relativamente baixa, sendo mais comparável à das leveduras do que à de outros fungos filamentosos. De modo semelhante, os padrões de *N*-glicosilação produzidos por *A. gossypii* são genericamente mais semelhantes aos produzidos pelas leveduras do que aos produzidos por outros fungos filamentosos. No entanto, hiperglicosilação extensa só ocorre em determinadas condições de

cultura. Tal como outros fungos filamentosos, o *A. gossypii* parece ter também a capacidade de produzir *N*-glicanos truncados.

Uma *unfolded protein response* (UPR) convencional não foi ativada em *A. gossypii* em resposta à secreção de proteínas heterólogas nem ao *stress* de secreção induzido por ditiotreitol (DTT), como geralmente observado noutros fungos. No entanto, as respostas transcricionais do *A. gossypii* ao stress induzido por DTT indicam que existem mecanismos alternativos neste fungo para lidar com o *stress* de secreção proteica.

A invertase de *A. gossypii* foi demonstrada ser codificada pelo gene *AFR529W* (*AgSUC2*), que é funcionalmente complementado pelo gene *SUC2* de *S. cerevisiae* (*ScSUC2*). As sequências sinal da AgSuc2p e ScSuc2p conseguiram direcionar em *A. gossypii* a secreção de invertase para o meio de cultura. Similarmente às invertases de outros fungos, a expressão da invertase de *A. gossypii* é regulada pelos açúcares presentes no meio. Estes resultados expandiram o nosso conhecimento acerca das capacidades de secreção nativas do *A. gossypii*, uma vez que a invertase é apenas a segunda enzima hidrolítica nativamente secretada por este fungo a ser experimentalmente caracterizada.

A *β*-galactosidase de *A. niger* foi espressa com sucesso em *A. gossypii* sob a regulação e diferentes promotores. O promoter nativo *TEF* revelou-se o melhor promotor para sobre-expressar *β*-galactosidase heteróloga em *A. gossypii*, induzindo 2 vezes mais atividade secretada do que o promotor *GPD* de *A. gossypii* e 7 vezes mais do que os promotores *PGK1* e *ADH1* de *S. cerevisiae*. Os níveis de *β*-galactosidase ativa secretada por uma estirpe laboratorial de *S. cerevisiae* transformada com os mesmos plasmídeos foram até 37 vezes mais baixos do que os obtidos em *A. gossypii*. A secreção de *β*-galactosidase ativa por *A. gossypii* foi cerca de 1.5 vezes mais alta em meio contendo glicerol em vez de glucose. Estes resultados evidenciam o potencial do *A. gossypii* como produtor de proteínas heterólogas e abrem novas oportunidades para optimizar as suas capacidades de secreção usando esta enzima como modelo. Como ela é de fácil deteção, o rastreio de melhorias na secreção será facilitado.

O sistema de recombinação Cre-*lox*P, que tem sido amplamente usado em diferentes organismos, foi adaptado com sucesso para utilização em *A. gossypii*, permitindo a remoção e reuso de marcadores de selecção em engenharia genética direcionada neste fungo. O conjunto de cassetes de deleção e plasmídeos construídos expandem as possibilidades para manipular geneticamente o *A. gossypii*, sendo adequados quer para uso em estirpes laboratoriais, quer em estirpes industriais, pois não requerem a existência de nenhum *background* genético predeterminado. No futuro, o melhoramento direcionado de estirpes de *A. gossypii* para aplicações industriais irá beneficiar desta ferramenta molecular.

LIST OF PUBLICATIONS

This thesis is based on the following original articles:

Aguiar TQ, Ribeiro O, Arvas M, Wiebe MG, Penttilä M, Domingues L. Genome-wide study of protein secretion and secretion stress in *Ashbya gossypii*. *Submitted.* [Chapter 2]

Aguiar TQ, Maaheimo H, Heiskanen A, Wiebe MG, Penttilä M, Domingues L Characterization of the *Ashbya gossypii* secreted *N*-glycome and genomic insights into its *N*-glycosylation pathway. *Submitted*. [Chapter 3]

Aguiar TQ, Dinis C, Magalhães F, Oliveira C, Wiebe MG, Penttilä M, Domingues L. Molecular and functional characterization of the *Ashbya gossypii* invertase. *Submitted*. [Chapter 4]

Magalhães F, Aguiar TQ, Oliveira C, Domingues L. High-level expression of *Aspergillus niger β*-galactosidase in *Ashbya gossypii*. *Submitted.* [Chapter 5]

Dinis C, Aguiar TQ, Domingues L. Cre-*lox*P based system for removal and reuse of selection markers in *Ashbya gossypii* targeted engineering. *Submitted*. [Chapter 6]

CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

Figure 2.3 – Overall variations in the *A. gossypii* transcriptome after addition of DTT. The data refers to the outcome of the LIMMA analysis for differentially expressed genes within 30 min, 1 h and 4 h of exposure to DTT, when compared to the time immediately before DTT addition (fold change > log2(1.5) and p-value < 0.01). The Venn diagrams indicate the number of genes that had increased (A) and reduced (B) expression after 30 min, 1 h and 4 h of exposure to DTT. In table C the absolute number and relative percentage (*between brackets*; in relation to the total number of *A. gossypii* genes) of genes transcriptionally up-regulated (↑), down-regluated (↓) and unchanged (=) by DTT after different exposure times are indicated. The Volcano plots (D) obtained from the LIMMA analysis show the overall significant fold changes in the *A. gossypii* gene expression profiles after 30 min, 1 h and 4 h of DTT adition. Negative fold changes represent genes with increased (↑) expression and positive fold changes genes with reduced (↓) expression after exposure to DTT.

Figure 2.4 – Overall variation in the transcript levels of the genes included in different co-expression clusters after addition of DTT to the cultures of recombinant *A. gossypii* EGI producing strain. DTT was added after 9.5 h of culture (EG_D_0) and expression analyses were performed after 30 min (EG_D_0.5), 1 h (EG_D_1) and 4 h (EG_D_4) of exposure. The image also includes information for the gene expression variation in the non-treated EGI producing strain (EGI) and corresponding empty vector strain (EV). EG_D_0.5 (10 h of culture) can be compared with EG_10 and e_10, and EG_D_4 (13.5 h of culture) can be compared with EG_13.5 and e_14. Dashed squares indicate gene clusters down-regulated by DTT and the other represented clusters were up-regulated by DTT.

Figure 2.5 – Regulatory DNA elements significantly over- (***yellow***) and under-represented (***blue***) ($p<0.05$) in different gene clusters. In panel A are represented the co-expression clusters and in panel B the clusters of genes differentially up- (C1) and down-regulated (C2) by DTT. C0 comprises the genes that were not included in any of the other clusters.

CHAPTER 3

Figure 3.1 – MALDI-TOF mass spectra of neutral *N*-glycans released by PNGase F from *A. gossypii* glycoproteins present in the supernatants of cultivations in defined minimal medium (**A**, **C** and **E**) and complex rich medium (**B**, **D** and **F**) at exponential (**A** and **B**), early stationary (**C** and **D**) and late stationary (**E** and **F**) phases. Glycan species are annotated with proposed monosaccharide compositions derived by mass calculation. H, hexose; N, *N*-acetylhexosamine; *, known contaminants from the culture media; ↓, unknown polyhexose contaminants.

Figure 3.2 – MALDI-TOF mass spectra of acidic *N*-glycans released by PNGase F from *A. gossypii* glycoproteins present in the supernatant of mycelia grown in defined minimal medium at exponential (**A**), early stationary (**B**) and late stationary (**C**) phases. Glycan species are annotated with proposed

monosaccharide compositions derived by mass calculation and based on NMR data. H, hexose; N, *N*-acetylhexosamine; *P*, phosphate; ↓, glycans containing two phosphate residues; □, glycans containing two phosphate residues and only one *N*-acetylhexosamine; *, known contaminants from the culture media.

Figure 3.3 – Expansions of 600 MHz ^{1}H NMR spectrum of a mixture of *N*-glycans released from the total pool of glycoproteins secreted by *A. gossypii* in defined minimal medium by PNGase F digestion. The spectrum was recorded at 600 MHz at 25 ºC and chemical shifts were referenced to internal acetone (2.225 ppm). **A**. A model structure showing the naming of the monosaccharide units used for assignment of the structural reporter groups. ●, mannose; ■, *N*-acetylglucosamine; P, phosphate. **B**. Anomeric region of the spectrum. Assignments are given for the structural reporter group signals identified and **red** identifies terminal position localization. **C**. Methyl signals of the NAc groups of the two reducing end GlcNAc units.

Figure 3.4 – Expansion of ^{1}H-^{31}P HSQC spectrum of a mixture of *N*-glycans released from the total pool of glycoproteins secreted by *A. gossypii* in defined minimal medium by PNGase F digestion showing the cross peaks between Man H-1 and the phosphate groups.

CHAPTER 4

Figure 4.1 – Secreted (●) and cell-associated (○) specific invertase activity during *A. gossypii* ATCC 10895 growth in defined minimal medium containing sucrose as primary carbon source, at pH 6.0, 30°C, 500 rpm with 1 vvm aeration. In panel A, specific invertase activity (*circles*) is read on the primary axis and biomass concentration (*squares*) on the secondary axis. Error bars represent the standard error of the mean of three independent measurements. Panel B shows the sugar concentration in the culture supernatants at each time point the activities in panel A were measured.

Figure 4.2 – Nucleotide and deduced amino acid sequence of the *AgSUC2*-encoded invertase and its 5' untranslated region (UTR). Potential recognition sites for AgMig1p binding and the preceding A-T rich sequence in the *AgSUC2* 5' UTR are blocked in black. The putative secretion signal peptide and the catalytic domains A, D and E are underlined, with the aminoacid of the catalytic site marked in **bold**. Potential *O*-glycosylation (S/T) and *N*-glycosylation sites (N-X-S/T, X any aminoacid except proline) are marked by two lines. *N*-glycosylation sites conserved in ScSuc2p are highlighted in grey and sites with higher probability of being occupied represented in **bold**.

Figure 4.3 – Comparison of specific invertase activity in the culture supernatant (*black circles*), extracellular cell wall-bound (*grey circles*) and total cell-associated (*open circles*) during growth of *A. gossypii* ATCC 10895 in 250 ml flasks containing 50 ml of AFM with sucrose (A) or glucose (B) as primary carbon source. Specific invertase activity in read on the primary axis and sugar concentration on the secondary axis. Error bars represent the standard error of the mean of six independent biological replicates and where not seen were smaller than the symbol.

Figure 4.4 – Specific invertase activity in the culture supernatant (*black circles*) and total cell-associated (*open circles*) during growth of *A. gossypii* ATCC 10895 in 250 ml flasks containing 50 ml of AFM with glycerol as primary carbon source. Specific invertase activity in read on the primary axis and glycerol concentration on the secondary axis. Error bars represent the standard error of the mean of three independent biological replicates and where not seen were smaller than the symbol.

Figure 4.5 – Comparison of biomass production by *A. gossypii* ATCC 10895 (parent strain, *triangles*) and the *A. gossypii Agsuc2* null mutant (*squares*) growing in 250 ml flasks containing 50 ml of SC medium with sucrose (*solid symbols*) or glucose (*open symbols*) as primary carbon source. Error bars represent the standard error of the mean of three independent biological replicates.

Figure 4.6 – Schematic representation of the *AgSUC2* gene deletion strategy (**A**) and PCR confirming the integration of the *GEN3* deletion module in the *AgSUC2* locus and complete removal of this gene from the genome of *A. gossypii* homokaryotic *Agsuc2* null mutants (**B**). In panel B, the first lane is the molecular marker (NZYDNA Ladder III, NZYTech); lanes 1, 3 and 5 are genomic DNA from the *A. gossypii* parent strain amplified with primers V1/V2, V3/V4 and V5/V4, respectively; and lanes 2, 4 and 6 are genomic DNA from *A. gossypii Agsuc2* mutant strain amplified with primers V1/V2, V3/V4 and V5/V4, respectively.

Figure 4.7 – Electrophoretic analysis of *A. gossypii* culture filtrate collected after 32 h of cultivation in a bioreactor with defined minimal medium containing sucrose as primary carbon source. To 8% native polyacrylamide gels were applied: *A. gossypii* culture filtrates concentrated 25 times (lane 1), 0.05 U of commercial *S. cereviae* invertase (lane 2) and 0.005 U of commercial *S. cereviae* invertase (lane 3). According to the manufacturer's indications the molecular weight of the *S. cereviae* invertase is 270 kDa.

CHAPTER 5

Figure 5.1 – Graphic representation of the plasmids constructed for the expression of *A. niger* *β*-galactosidase in *A. gossypii* under the regulation of different promoters: (A) *S. cerevisiae ADH1* promoter, (B) *S. cerevisiae PGK1* promoter, (C) *A. gossypii GPD* promoter, and (D) *A. gossypii TEF* promoter. Images created with PlasmaDNA v1.4.2 (University of Helsinki).

Figure 5.2 – Test tubes for *β*-galactosidase activity screening containing *A. gossypii*ADH1lacA (A), PGK1lacA (B), TEFlacA (C) and GPDlacA (D) grown for 2 days in 5 ml of selective AFM supplemented with 0.04 mg/ml X-gal.

Figure 5.3 – Glucose consumption (A) and specific *β*-galactosidase activity (B) for the *A. gossypii* recombinant strains growing at 30ºC and 200 rpm in 250 ml shake-flasks containing 50 ml of selective AFM: *A. gossypii* TEFlacA (*solid square*), GPDlacA (*solid triangle*), ADH1lacA (*open square*), PGK1lacA (*solid*

circle) and empty vector (*open circle*). Error bars represent the standard deviation of three biological replicates.

Figure 5.4 – *β*-Galactosidase zymogram. *A. niger β*-galactosidase produced by *A. gossypii* TEFlacA (lane a) and *S. cerevisiae* NCYC869-A3/pVK1.1 (lane b). The recombinant *β*-galactosidase samples applied to the gel had an activity of approximately 600 U.

CHAPTER 6

Figure 6.1 – Schematic representation of the strategy used for gene deletion and marker rescue in *A. gossypii* and respective verification PCRs of the generated mutants. **(A)** Construction of the *A. gossypii Agura3* deletion strain. The *AgURA3* gene (*AEL059W*) was replaced by the *lox*P-*GEN3*-*lox*P marker to yield the strain *Agura3GEN3*. This strain was subsequently transformed with the Cre expression plasmid pAgNatCre and excision of the *GEN3* marker occurred after one week of growth in solid AFM containing 10 g/l ethanol as sole carbon source. **(B)** Construction of the *A. gossypii Agade1* deletion strain. The *AgADE1* gene (*AER221W*) was replaced by the *lox*P-*NATPS*-*lox*P marker to yield the strain *Agade1NATPS***.** This strain was transformed with the Cre expression plasmid pAgBleCre and the *NATPS* marker was excised after induction of the Cre expression as described above. **(C)** The *GEN3* marker, which had been excised from the *Agura3* strain, was used again to successfully delete the *AgADE1* gene in this strain, yielding the strain *Agura3ade1GEN3*. **(D)** Verification of the integration of the *lox*P-*GEN3*-*lox*P cassette in the *AgURA3* locus and confirmation of its excision using the primers U1_FW and U2_RV (lanes 1, 2 and 3), U3_FW and U2_RV (lanes 4, 5 and 6) and U4_FW and U2_RV (lanes 7, 8 and 9). Genomic DNA from the *A. gossypii* parent strain (ATCC 10895) (lanes 1, 4 and 7), *Agura3GEN3* strain (lanes 2, 5 and 8) and *Agura3* strain (lanes 3, 6 and 9) were used as template. **(E)** Verification of the integration of the *lox*P-*NATPS*-*lox*P cassette in the *AgADE1* locus and confirmation of its excision using the primers A1_FW and A2_RV (lanes 1, 2 and 3), A1_FW and A4_RV (lanes 4, 5 and 6), and A1_FW and A3_RV (lanes 7, 8 and 9). Genomic DNA from the *A. gossypii* parent strain (ATCC 10895) (lanes 1, 4 and 7), *Agade1NATPS* strain (lanes 2, 5 and 8) and *Agade1* strain (lanes 3, 6 and 9) were used as template. **(F)** Verification of the integration of the *lox*P-*GEN3*-*lox*P cassette in the *AgADE1* locus of the *Agura3* strain using the primers A1_FW and A2_RV (lane 1 and 2), U3_FW and A3_RV (lane 3 and 4), and A1_FW and A3_RV (lane 5 and 6). Genomic DNA from the *A. gossypii* parent strain (ATCC 10895) (lanes 1, 3 and 5) and from the *Agura3ade1GEN3* strain (lanes 2, 4 and 6) were used as template. The positions of the primers and expected PCR amplicon sizes are indicated in the scheme.

Figure 6.2 – Graphic representation of the plasmids pUGGEN3, pUGNATPS and pUGBLE3. Images created with PlasmaDNA v1.4.2 (University of Helsinki).

Figure 6.3 – Characterization of the *A. gossypii Agura3GEN3* mutant growth on agar-solidified SC medium with (URA^+) and without (URA^-) uracil and on AFM supplemented with 5-FOA (1 mg/ml). Plates were incubated for one week at 30ºC and the parental strain (ATCC 10895) used as control.

Figure 6.4 – Characterization of the *A. gossypii Agura3ade1GEN3* mutant growth on agar-solidified SC medium with (URA^+ADE^+) and without (URA^-ADE^-) uracil and adenine and with uracil and without adenine (URA^+ADE^-). Plates were incubated for one week at 30ºC and the parental strain (ATCC 10895) used as control.

Figure 6.5 – *A. gossypii Agade1* mutant colony grown on agar-solidified AFM for one week at 30ºC.

LIST OF TABLES

CHAPTER 2

CHAPTER 3

Table 3.2 – ^{1}H chemical shifts of the structural reporter group protons of the constituent monosaccharides of *A. gossypii N*-glycans. Assignments were made by comparison with NMR data reported in the literature for several *N*-glycan structures and supported by 2-D homonuclear ^{1}H-^{1}H DQF-COSY and ^{1}H-^{1}H TOCSY experiments. The identified signals were consistent with high-mannose type *N*-glycan structures with average monosaccharide composition $Man_9GlcNAc_2$.

Table 3.3 – Putative *A. gossypii* genes predicted to be involved in the *N*-glycan biosynthesis pathway. The table shows all of the *S. cerevisiae* genes involved in the *N*-glycosylation pathway for which a homolog was identified in the *A. gossypii* genome. No homolog genes were identified in *S. cerevisiae* for two putative *A. gossypii* genes predicted to encode glycoside hydrolases with a possible role in *N*-glycosylation.

CHAPTER 4

Table 4.1 – Primers used in this study. Upper case sequences correspond to sequences complementary to the template. Lower case sequences correspond to additions for restriction sites (underlined) or to the disruption cassette guide sequences with homology to the *AgSUC2* flanking regions (**bold**).

CHAPTER 5

Table 5.1 – Primers used in this study. The restriction sites used for the construction of plasmids are highlighted in **bold and underlined**. Highlithed only in **bold** is the multiple cloning site introduced in the the plasmids used.

Table 5.2 – Maximum volumetric β-galactosidase activity secreted by *A. gossypii* and *S. cerevisiae* recombinant strains under the regulation of different promoters. The values represent the average ± standard deviation of the β-galactosidase activity measured at the glucose depletion time-point in the supernatants of three independent shake-flask cultures performed at 30ºC and 200 rpm in 50 ml of selective AFM (for *A. gossypii*) and YPD (for *S. cerevisiae*). The average biomass for the *A. gossypii* recombinant strains was 6.37±0.38 g/l and for the *S. cerevisiae* recombinant strains was 5.75±0.07 g/l.

CHAPTER 6

Table 6.1 – Strains and plasmids used in this study.

Table 6.2 – Oligonucleotide primers used in this study. Underlined letters are nucleotide sequences complementary to the gene deletion regions. Recognition sequences for restriction enzymes are in **bold**.

LIST OF ABBREVIATIONS

AFM	*Ashbya* Full Medium
ARS	Autonomous Replicating Sequences
BSA	Bovine Serum Albumin
bp	base pairs
bZIP	basic-leucine Zipper
cDNA	complementary DNA
gDNA	genomic DNA
rDNA	ribosomal DNA
CRM	Complex Rich Medium
CBHI	Cellobiohydrolase I
DMM	Defined Minimal Medium
DNA	Deoxyribonucleic Acid
DNS	3,5-Dinitrosalicylic Acid
Dol	Dolichol
DQF-COSY	Double Quantum Filtered-Correlation Spectroscopy
DTT	Dithiothreitol
EGI	Endoglucanase I
ENGase	Endo-*β*-*N*-acetylglucosaminidase
ER	Endoplasmic Reticulum
ERAD	ER-Associated Degradation
F	Deoxyhexose
FAD	Flavin Adenine Dinucleotide
FMN	Flavin Mononucleotide
FIRE	Finding Informative Regulatory Elements
GC	Guanine-Cytosine
GFP	Green Fluorescent Protein
GH32	Glycoside Hydrolase family 32
GH85	Glycoside Hydrolase family 85
GH92	Glycoside Hydrolase family 92
GlcNAc	*N*-acetylglucosamine

GO	Gene Ontology
GPI	Glycosylphosphatidylinositol
GRAS	Generally Recognized as Safe
GT	UDPglucose:glycoprotein glucosyltransferase
GTP	Guanosine-5-Triphosphate
ICL	Isocitrate lyase
IEF	Isoelectric focusing
H	Hexose
HPLC	High Performance Liquid Chromatography
HSQC	Heteronuclear Single Quantum Coherence
LB	Luria-Bertani
MALDI-TOF	Matrix-Assisted Laser Desorption-Ionization Time-Of-Flight
Man	Mannose
MCS	Multiple Cloning Site
mRNA	messenger Ribonucleic acid
MULac	methylumbelliferyl-*β*-D-lactoside
MW	Molecular Weight
m/z	mass-to-charge ratio
N	*N*-acetylhexosamine
NAc	*N*-acetyl
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Enhancement Spectroscopy
OD_{600}	Optical Density at 600 nm
ORF	Open Reading Frame
OST	Oligosaccharyltransferase
P	Phosphate
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
pI	Isoelectric point
PMT	*O*-mannosyltransferase
PNGase F	Peptide:*N*-glycosidase F
***p*NPG**	*p*-nitrophenyl-*β*-D-galactopyranoside
PRPP	Phosphoribosyl pyrophosphate
QC	Quality Control

RESS	REpression under Secretion Stress
rESR	Environmental Stress Response
RIDD	Regulated *IRE1*-Dependent mRNA Decay
RMA	Robust Multichip Average
RNA	Ribonucleic acid
SC	Synthetic Complete
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis
SRP	Signal Recognition Particle
TCA	Tricarboxylic Acid
TOCSY	Total Correlation Spectroscopy
TTC	2,3,5-triphenyltetrazolium chloride
UPR	Unfolded Protein Response
UPRE	UPR Element
UTR	Untranslated Region
vvm	volumes of air per volume of liquid per minute
WGD	Whole-Genome Duplication
X-gal	5-bromo-4-chloro-3-indolyl- β-D-galactopyranoside
1-D	One-dimensional
2-D	Two-dimensional
5-FOA	5-Fluoroorotic Acid

MOTIVATION AND OUTLINE OF THE THESIS

Fungi, including both true filamentous fungi and yeast, have been exploited by humans for centuries as sources of metabolites and enzymes used in traditional processes such as food and beverage production. Based on early biotechnological knowledge, the use of fungi for the production of other products outside of their traditional application was expanded. This has led to the identification and development of improved fungal cell factories that allow the production of large amounts of a variety of industrially relevant bio-products in an economically and environmentally sustainable way. In this context, the industrial production of riboflavin (vitamin B2) by the true filamentous fungus *Ashbya gossypii* is an example of the sustainability achieved by some fungal-based biotechnological processes.

The flavinogenic fungus *A. gossypii* gathers several interesting features that make it an attractive host to further explore as a cell factory for the production of other industrially relevant bio-products in addition to riboflavin. The high similarity of its small genome (one of the smallest eukaryotic genomes known) to that of the so well-studied *Saccharomyces cerevisiae* is one of these features. This facilitates the assignment of potential functions to *A. gossypii* genes, allowing a good understanding of its molecular biology. In addition, its efficient homologous recombination system, haploid nuclei and uninucleate spores allow easy genetic manipulation through simple PCR-based gene targeting protocols, facilitating rapid strain construction in *A. gossypii*. The ability of *A. gossypii* to freely replicate plasmids containing *S. cerevisiae* autonomous replicating sequences (ARS) also constitutes an interesting feature, as several plasmids already developed for *S. cerevisiae* can be used in *A. gossypii*. *A. gossypii* also performs well in large-scale fermentation processes, growing to high cell densities in complex media based on agro-industrial residues. Thus, inexpensive environmentally friendly culture media can be designed for growth and metabolite/enzyme production.

Considering these attributes, a research project aimed at exploring the potential of *A. gossypii* as a cell factory organism was initiated in our research group in collaboration with VTT Biotechnology (Finland) – the AshByofactory project. In this context, early investigation of the *A. gossypii* potential as a host for the production of heterologous proteins revealed that the protein secretion abilities of this filamentous fungus seemed to resemble those of the closely related yeast *S. cerevisiae*, even though more efficient protein secretion is usually associated with filamentous growth (Ribeiro et al. 2010). However, proteins produced by *A. gossypii* were apparently less extensively glycosylated than those

produced by *S. cerevisiae* (Ribeiro et al. 2010). Several questions were, thus, raised about the protein secretion abilities of this fungus and the glycosylation patterns it performed, becoming necessary a better characterization of its protein secretory pathway.

The purpose of this thesis was, therefore, to increase our understanding of the biotechnological potential of *A. gossypii* by addressing different topics that aimed at better characterizing its protein secretory pathway and at improving its heterologous protein production capacities. A parallel goal was the development of a molecular tool that would allow the construction of *A. gossypii* strains for industrial use free of exogenous selective markers.

A short summary of the content of each chapter of this thesis follows.

Chapter 1 starts with a general literature review about *A. gossypii* and about the main subjects covered in this thesis.

Chapter 2 presents the results from genome-wide analyses performed in *A. gossypii* with the intent of exploring its secretome and its transcriptional responses to protein secretion stress. These results provide the first insights into the secretion stress response of *A. gossypii*, as well as a basic understanding of its protein secretion potential and protein quality control systems. From the analysis of these results valuable information was gathered that helped to conceive the experiments described in Chapters 4 and 5.

Chapter 3 includes the results from the first comprehensive characterization of the *N*-glycan structures attached to the proteins secreted by *A. gossypii* and provides genomic insights into the *N*-glycosylation pathway of this fungus. In this work, culture medium- and cultivation stage-specific *N*-glycan profiles were assessed, providing an overall understanding of the *N*-glycosylation heterogeneity in *A. gossypii*.

Chapter 4 describes the molecular and functional characterization of an invertase that was predicted to be secreted by *A. gossypii* in Chapter 2. This provides the characterization of only the second hydrolytic enzyme natively secreted by *A. gossypii*, expanding our knowledge about the protein secretion capacities of this fungus. This work confirms the usefulness of the *in silico* approach followed in Chapter 2 for guiding the experimental characterization of genes putatively encoding secreted proteins in *A. gossypii*.

In **Chapter 5** *A. gossypii* was further assessed as a host for the production of secreted heterologous proteins through the expression of the extracellular *β*-galactosidase from the filamentous fungus *Aspergillus niger*. In this work, the strength of different native and foreign constitutive promoters was tested and high levels of secreted *β*-galactosidase were obtained with the native promoters. The results obtained demonstrate that the potential of *A. gossypii* as a recombinant protein production platform is still underexplored and that further optimization of foreign gene expression strategies, culture conditions or strain tailoring could much improve the productivity of this host.

Chapter 6 describes a molecular tool based on the Cre-*lox*P recombination system of the bacteriophage 1 developed to remove and recycle marker genes in *A. gossypii*. This method offers advantages over the existing methods by enabling the creation of *A. gossypii* mutants entirely free of foreign genes and by allowing the reuse of selection markers. The set of disruption cassettes and plasmids constructed greatly expand the possibilities for genetically engineer *A. gossypii*, being suitable for use in both laboratorial and industrial strains.

Finally, **Chapter 7** presents a summary of the main conclusions and some future perspectives.

CHAPTER 1

Introduction

1.1 *ASHBYA GOSSYPII*

1.1.1 Habitat and pathogenicity

The filamentous hemiascomycete commonly known as *Ashbya gossypii* (syn. *Eremothecium gossypii*) was originally isolated from infected cotton bolls (*Gossypium* s.p.) harvested in the British West Indies (Caribbean region) and identified as one of the causative agents of stigamtomycosis (Ashby and Nowell 1926). Subsequently, *A. gossypii* strains have also been isolated from other economically important crops widespread along the tropic and sub-tropic regions of Northern America, Southern America and South Africa, such as coffee, citrus fruits and tomatoes (Pridham and Raper 1950, Batra 1973). In cotton, the most characteristic symptoms of stigmatomycosis are (Pridham and Raper 1950, Batra 1973): premature dropping of the bolls or drying out of those which remain on the plant; seed coats stained brown in spots; and yellow to yellowish-brown lint fibres. In coffee, the disease affects the beans, which become dry rot, black and shrunken (Pridham and Raper 1950). When infection occurs in tomatoes and citrus the disease causes lesions in the fruits and soft brown rotting (Batra 1973).

During the first half of the 20th century this destructive phytopathogen caused extensive crop damages and severe economical losses (Batra 1973). Insect vectors, predominantly those with pierce-sucking mouthparts such as the Heteroptera, are essential for the transmission of this fungus, as the fungus itself is unable to penetrate the outer cells of healthy fruits and seeds unless a mechanical injury occurs (Pridham and Raper 1950, Batra 1973). Since the most important mode of infection requires insects to transport and inject spores or mycelial fragments into the tissue of the plant (Pridham and Raper 1950, Batra 1973), nowadays adequate control of the insect population by insecticides has proven efficient in preventing *A. gossypii* dissemination (Dammer and Ravelo 1990). Nevertheless, *A. gossypii* is still present in the nature and was recently isolated from large milkweed bugs in Florida (Dietrich et al. 2013).

1.1.2 Taxonomic history

A. gossypii was first characterized in 1926 by Ashby and Nowell (Ashby and Nowell 1926), who at the time named it *Nematospora gossypii*, based on the close resemblance of its sporiferous sacs and spores to those of other *Nematospora* species. In

1928, based on cytologic and phylogenetic studies, Guilliermond (1928) placed this species in the class Hemiascomycetes and assigned to it a new genus name, *Ashbya*, because in opposition to other *Nematospora* species, it grew exclusively in a polinucleated mycelial form and multiple nuclei were in the origin of its sporangia. Fragoso and Ciferri (1928) considered this species to belong to the order Saccharomycetales (Endomycetales) and based on Guilliermond (1928) cytological observations also suggested separating it from the *Nematospora* genus, presenting a Latin description for it under the generic name *Ashbia*.

As this species appears to possess characteristics of several families, further attempts to classify it differed along the years. Ascospore morphology (needle-shaped with a whiplike filament at one end) and the presence of hyphal growth were considered to separate the genus *Ashbya*, at least at the family level, from yeast genera belonging to the Saccharomycetales (von Arx and van der Walt 1987). However, the close relationship between *A. gossypii* and *Saccharomyces cerevisiae* started to became obvious when the sequence analysis of the first set of clones of *A. gossypii* genomic DNA (gDNA) revealed an unpredictably high degree of synteny to the gene order of their homologs in *S. cerevisiae* (Steiner and Philippsen 1994, Altmann-Jöhl and Philippsen 1996). Based on phylogenetic analysis of ribosomal DNA (rDNA) sequence divergence, Kurtzman (1995) placed the genera *Ashbya*, *Eremothecium*, *Holleya* and *Nematospora* in a single genus, *Eremothecium*, and introduced the family Eremotheciaceae for this genus. Prillinger et al. (1997) data on cell wall sugar composition, dityrosine content in ascospores, ubiquinone side chains and sequence analysis of the genes coding for the 18S rDNA and ITS region supported the inclusion of these genera within a single genus, as proposed by Kurtzman (1995), but within the Saccharomycetaceae family.

The sequencing and annotation of the *A. gossypii* genome gave a complete picture about the close relation between *A. gossypii* and *S. cerevisiae* (Brachat et al. 2003, Dietrich et al. 2004). Based on the gene order, 91% of the 4776 annotated *A. gossypii* genes are syntenic and only 4% non-syntenic to *S. cerevisiae* genes (Dietrich et al. 2004, Schmitz and Philippsen, 2011). The remaining 5% have no homolog in *S. cerevisiae* (Dietrich et al. 2004). The 9.12 Mb genome of *A. gossypii* also provided compelling evidence for whole-genome duplication (WGD) in the *S. cerevisiae* lineage. This could be established based on the 2:1 mapping of *S. cerevisiae* loci to blocks of *A. gossypii* genes (Dietrich et al. 2004).

With the genome sequence of several fungal species available, fungal phylogenomic studies have been conducted (Fitzpatrick et al. 2006, Wang et al. 2009), congruently placing *A. gossypii* in the order Saccharomycetales, close to other pre-WGD species, like *Kluyveromyces lactis*, *Kluyveromyces waltii* and *Saccharomyces kluyveri* (Figure 1.1). However, the monophyly of these four species is still unsettled (Wang et al. 2009). Kurtzman and Robnett (2003) suggested that these species are paraphyletic, assigning the genus *Eremothecium* to the clade 12 of the *Saccharomyces* complex, *K. lactis* to the clade 11 and *K. waltii* and *S. kluyveri* to the clade 10. In contrast, other studies suggest that the four species are monophyletic, constituting themselves a clade (Fitzpatrick et al. 2006, Wang et al. 2009).

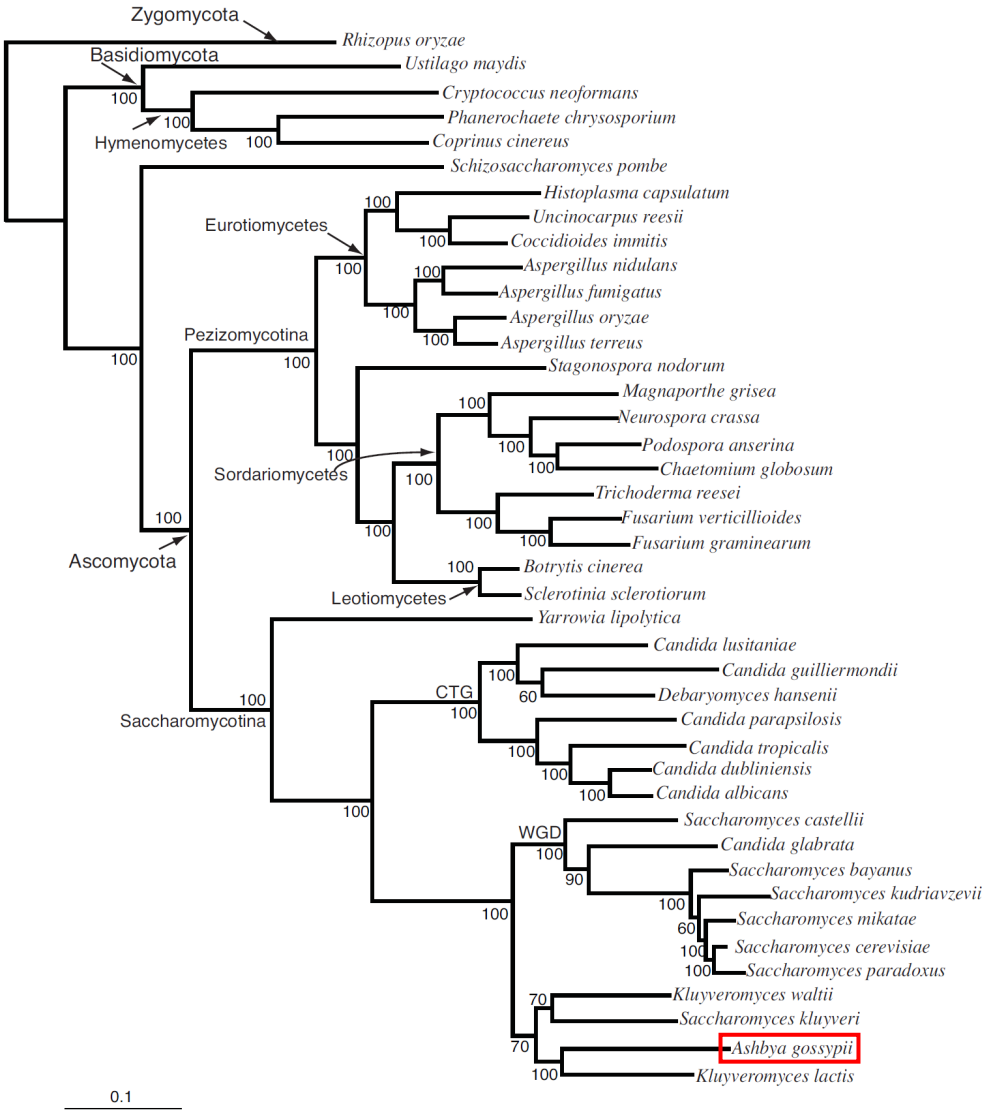


Figure 1.1 – Maximum likelihood phylogeny reconstructed using a concatenated alignment of 153 universally distributed fungal genes from 42 taxa. Adapted from Fitzpatrick et al. (2006).

1.1.3 Life cycle

A. gossypii is one of the simplest filamentous fungal species studied. Its life cycle (Figure 1.2) starts with a short period of isotropic growth that initiates the germination of its haploid uninucleated needle-shaped spores and leads to the formation of a spherical germ bubble at the centre of the spores, which is where the nucleous is positioned (Wendland and Philippsen 2000, Alberti-Segui et al. 2001). Then, actin patches start to accumulate in a region of the germ bubble cortex perpendicular to the axis defined by the spore needle, marking the emergence of the first germ tube and the switch to polarized growth (Knechtle et al. 2003). The tip of the germ tube continues extending and an actin ring at the neck between the germ bubble and the first germ tube initiates the formation of the first septum (Knechtle et al. 2003). At the opposite side of the germ bubble a second germ tube is formed, giving rise to a bipolar germling (Wendland and Philippsen 2000). From these initial hyphae a young mycelium is generated by lateral branching. Hyphal tip growth speed increases during maturation and concomitantly hyphal tips start to undergo apical branching, leading to the formation of Y-shaped hyphal filaments (Ayad-Durieux et al. 2000). In the older parts of the mycelium, once growth has ceased, *A. gossypii* eventually enters the sporulation phase, event that is thought to be associated with the production of riboflavin (vitamin B2), as the production of this metabolite by *A. gossypii* dramatically increases during sporulation (Stahmann et al. 2001). Old hyphae fragment at septal sites to form sporangia that usually contain eight endospores bound together by filaments, which can be set free by lysis. The hyphae around the spores stick together and form synnemata like structures that are exposed into the air (Wendland and Walther 2005). Sporulation in *A. gossypii* seems to be affected by the pheromone response pathway, as the deletion of the transcription factor AgSte12p, a key component of this pathway, led to a hypersporulation phenotype (Wendland et al. 2011).

The *A. gossypii* reference strain ATCC10895 was found to harbor four identical mating-type cassettes in its genome containing *MAT**a*** information, but no *MAT*α orthologs were found (Wendland et al. 2011, Dietrich et al. 2013), which led to question the existence of a sexual cycle in this fungus (Wendland and Walther 2011). However, the recent genome sequencing of a new *A. gossypii* strain (FDAG1) isolated from large milkweed bugs revealed that this wild isolate encodes both *MAT**a*** and *MAT*α sequences, suggesting that *A. gossypii* likely has a sexual cycle (Dietrich et al. 2013).

The relatively simple life cycle of *A. gossypii* in the laboratory, in which uninucleated haploid spores germinate and give rise to multinucleated vegetative hyphae that grow exclusively in a polarized manner, together with the astonishing similarity of its genome with the genome of the unicellular yeast *S. cerevisiae*, have raised interest to this fungus as an attractive biological model to study fungal developmental biology (reviewed in Wendland and Walther 2005, Schmitz and Philippsen 2011).

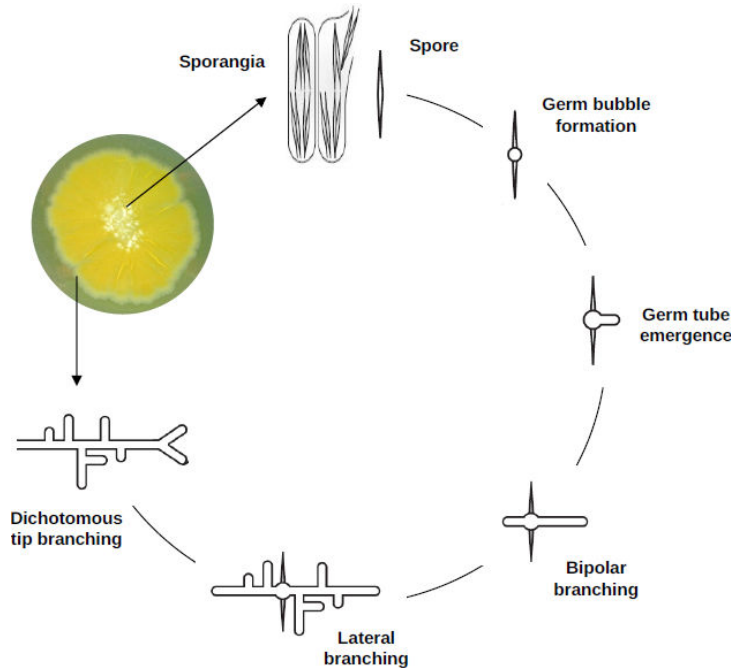


Figure 1.2 – Schematic representation of the *A. gossypii* life cycle.

1.1.4 Metabolism and physiology

A. gossypii is a natural overproducer of riboflavin (vitamin B2), which confers to its mycelia a characteristic yellow colour (Wickerham et al. 1946). Overproduction of this vitamin in *A. gossypii* is known to start at the stationary phase, when the growth rate declines, and to be linked with sporulation (Stahmann et al. 2001, Karos et al. 2004). As a rule, similar kinetics of production have been seen during microbial synthesis of secondary metabolites, but not primary metabolites. Therefore, riboflavin is considered in *A. gossypii* as a pseudo-secondary metabolite (Schlösser et al. 2001).

Riboflavin is synthesized *de novo* from guanosine-5-triphosphate (GTP) and ribulose 5-phosohate through a multi-step pathway controlled by the *RIB* genes (Karos et al. 2004). The riboflavin production phase is characterized by a strong increase in the

expression of some of theses genes (*AgRIB3*, *AgRIB4*, and *AgRIB5*), which are regulated at the transcription level (Shlösser et al. 2007). Overproduction of riboflavin by *A. gossypii* has been shown to be triggered by environmental stresses, such as nutritional and oxidative stress (Shlösser et al. 2007, Kavitha and Chandra 2009, Walther and Wendland 2012). The AgYap1p, a transcription factor that plays a major role in directing oxidative stress responses, was shown to regulate the expression of the *A. gossypii RIB4* gene and to mediate a Yap1p-dependent increase in riboflavin production during oxidative stress (Walther and Wendland 2012). Therefore, the secretion of large amounts of riboflavin was suggested to function as a possible scavenging mechanism against reactive oxygen species to protect *A. gossypii* against plant defences, thus representing an ecological advantage (Walther and Wendland 2012).

The few physiological studies focusing on the influence of some common environmental factors on the *A. gossypii* growth started to appear in the literature as a reflection of the eventual interest in using this fungus for industrial applications. As early as 1930, Farries and Bell (1930) noted the production of a yellow pigment (riboflavin) by certain *A. gossypii* strains when investigating the nitrogen requirements of this species. In their study they found that potassium nitrate and ammonium salts were not utilize as nitrogen sources, observations that were later confirmed by Buston et al. (1938). However, Ribeiro et al. (2011) recently verified that several *A. gossypii* strains actually grow well on chemically defined medium containing ammonium as sole nitrogen source when the medium pH is 6.5 and only poorly at pH 4.5, indicating that the lack or limited growth previously reported probably reflected low medium pH. Supporting their observations is the fact that *A. gossypii* is genetically equipped for ammonium (but not for nitrate) assimilation (Ribeiro et al. 2011).

The sensitivity of *A. gossypii* to low medium pH had been previously observed by Tanner et al. (1949), whom verified that media initially adjusted to pH 4.0 gave little growth and riboflavin production. Between pH 4.5 and 5.5 there was a good and rapid multiplication, but the best yields of riboflavin were obtained when the initial pH of the medium was above pH 5.5, preferably in the range of pH 6.0 to 7.0 (Tanner et al. 1949). The growth temperature was also found to have an impact on riboflavin production. Although *A. gossypii* grows over a rather wide temperature range (20 to 40ºC), the greatest riboflavin accumulation occurs when cultures are incubated at 26 to 28ºC, which is below the optimum for growth (30 to 35ºC) (Tanner et al. 1949, Pfeifer et al. 1950, Özbas and Kutsal 1986).

Myo-inositol and biotin have early been found to be required for *A. gossypii* growth (Buston and Kasinathan 1933, Kögl and Fries 1937), as this fungus lacks the genes necessary for their biosynthesis. Thiamine was also found required for adequate growth (Pridham and Raper 1950), although *A. gossypii* has the genetic capacity for its biosynthesis. When these are added to the medium *A. gossypii* is capable of utilizing simple mixtures of amino acids, ammonium aspartate or asparagine as nitrogen sources (Buston and Kasinathan 1933), the later supporting better growth in defined media (Demain 1972). Peptone and casein constitute, however, better nitrogen sources for *A. gossypii* (Farries and Bell 1930). Yeast extract is also a good source of several factors necessary for *A. gossypii* growth (Wickerhan et al. 1946) and our recent observations have shown that it alone can support *A. gossypii* residual growth in liquid medium without addition of any other nitrogen, vitamin or carbon source. Crude nitrogen and vitamin sources, like animal steep liquor, corn steep liquor and distillers' solubles have, however, been preferably employed in large-scale *A. gossypii* fermentations (Tanner et al. 1949, Pfeifer et al. 1950, Smiley et al. 1951).

Several studies on the *A. gossypii* carbon requirements indicated that glucose, fructose, sucrose, starch, maltose and glycerol support good growth, but plant oils (such as corn oil and soybean oil) are superior and favour riboflavin production (Tanner et al. 1949, Pridham and Raper 1950, Kutsal and Özbas 1989, Ribeiro et al. 2011). Ethanol also supports *A. gossypii* growth, but with more limited growth rates (Pridham and Raper 1950). Polymeric carbohydrates, pentoses (arabinose, xylose), galactose, lactose and inulin are not utilized as carbon sources (Pridham and Raper 1950, Ribeiro et al. 2011), which is explained by the absence of enzymes and/or pathways necessary for the metabolism of these substrates in the *A. gossypii* genome (Slot and Rokas 2010, Ribeiro et al. 2011, Ribeiro 2012).

Under aerobic conditions, glucose is oxidized by *A. gossypii* more or less completely to carbon dioxide and water (Mickelson 1950, Mickelson and Schuler 1953). However, considerable amounts of ethanol and traces of pyruvic, acetic and citric acid are formed during the course of the fermentation (Mickelson 1950, Mickelson and Schuler 1953). Ethanol has been shown to be produced during aerobic growth of *A. gossypii* on glycerol as well (Ribeiro et al. 2011). *A. gossypii* is also able to ferment glucose anaerobically, yielding 2 moles of carbon dioxide per mole of sugar consumed, which approaches the theoretical yield of a typical yeast alcoholic fermentation (Mickelson

1950). However, aerobiosis is required for riboflavin overproduction (Tanner et al. 1949, Pfeifer et al. 1950, Demain 1972).

When plant oils are used as sole carbon source, the glyoxylate cycle plays a fundamental role with respect to *A. gossypii* growth and riboflavin synthesis, as the precursors of riboflavin, GTP and ribulose 5-phosphate, originate from the carbohydrate metabolism (Figure 1.3) (Schmidt et al. 1996). The first enzyme of this pathway, isocitrate lyase (ICL), is essential for catalyzing the cleavage of isocitrate to succinate and glyoxylate, thus diverting the carbon flux from the tricarboxylic acid (TCA) cycle into a carbon-conserving pathway. The final result of this anaplerotic activity is a net conversion of fats to carbohydrates through gluconeogenesis (Schmidt et al. 1996, Maeting et al. 2000). An extracellular lipase secreted by *A. gossypii* also plays an important role in the degradation of plant oils to fatty acids and glycerol before their uptake into intracellular organelles (where β-oxidation takes place) (Stahmann et al. 1997, Maeting et al. 2000). Addition of non-ionic surface active agents to the culture medium (like Tween 80) stabilizes the activity of this enzyme (Stahmann et al. 1997) and enhances riboflavin production (Goodman and Ferrera 1954, Demain et al 1972).

Riboflavin biosynthesis is closely related to other biosynthetic pathways, such as the purine synthesis pathway (Figure 1.3), all of which compete for the same substrates. The riboflavin precursor ribulose 5-phosphate is one of the end products of the oxidative branch of the pentose phosphate pathway, which starts with the conversion of glucose 6-phosphate (obtained through gluconeogenesis from oils) to 6-phosphogluconolactone (Figure 1.3) (Kruger and von Schaewen 2003). In addition of being directly used in the riboflavin biosynthesis pathway, ribulose 5-phosphate can alternatively isomerize to ribose 5-phosphate (Kruger and von Schaewen 2003). GTP, the other riboflavin precursor, is synthesized through the purine pathway, which starts with the formation of phosphoribosyl pyrophosphate (PRPP) from ribose 5-phosphate and ATP (Figure 1.3). PRPP is not only required for the *de novo* and salvage pathways of purine, pyrimidine and pyridine (NAD^+, $NADP^+$) nucleotides, but also as a precursor of histidine and tryptophan byosynthesis (Jiménez et al. 2008). The addition of purines (like hypoxanthine) and, most importantly, glycine to the culture medium has yield-enhancing effects on riboflavin production (Demain 1972, Monschau et al. 1998). Glycine is an important precursor during the *de novo* purine byosynthesis and is obtained through the conversion of threonine by the Gly1p threonine aldolase (Monschau et al. 1998).

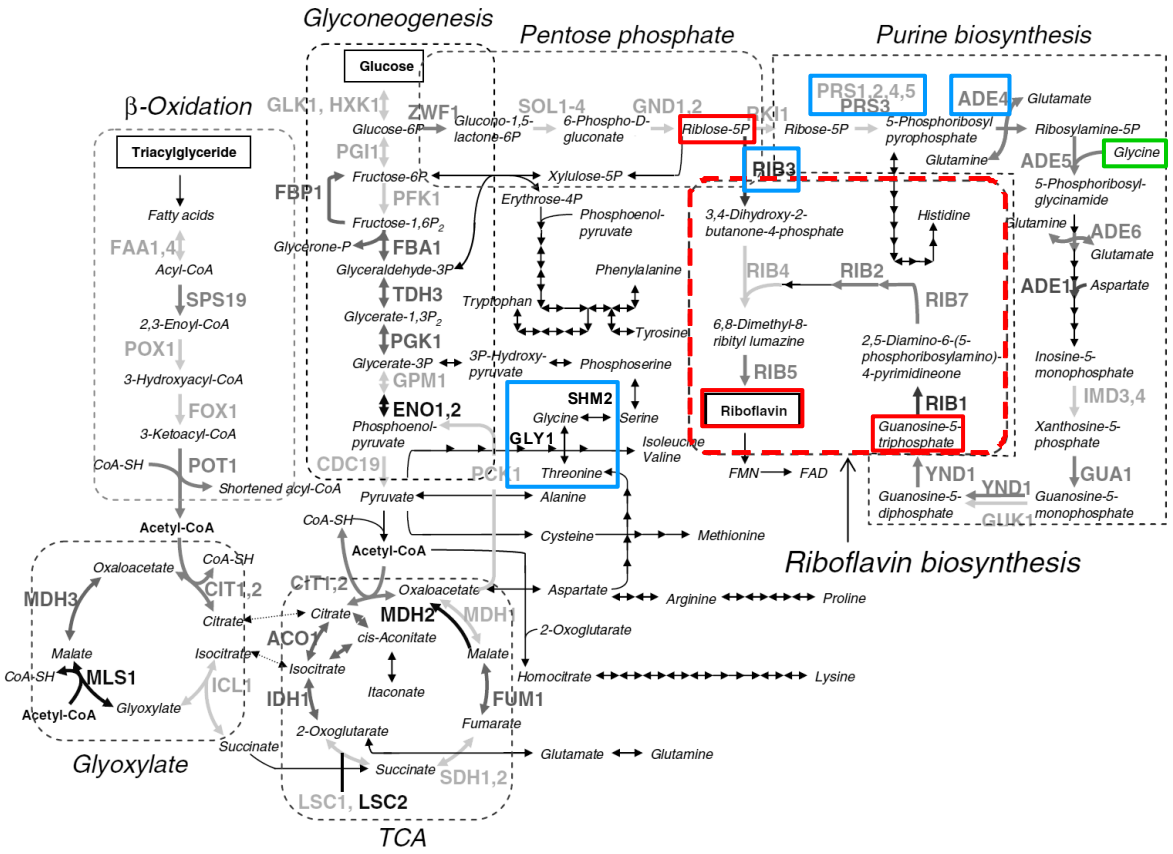


Figure 1.3 – Pathways involved in riboflavin biosynthesis from plant oils. The riboflavin biosynthetic pathway and precursors are highlighted in **red**. Enzymatic steps that were manipulated in *A. gossypii* leading to improvement of riboflavin production are highlighted in **blue**. **Green** indicates the step where glycine is necessary as precursor in the purine biosynthesis pathway. Adapted from Park et al. (2011).

One important aspect on the physiology of *A. gossypii* that needs to be stressed out is the fact that significantly different sporulation efficiencies and growth parameters, such as colonial radial growth rate, have been reported not only for different strains within the species but also for the same strain deposited in different culture collections (Ribeiro et al. 2011). Therefore, one should be cautious when analyzing the available information regarding the metabolism and physiology of this fungus.

1.1.5 Molecular tools

Many different approaches have been used to improve *A. gossypii* strains, mainly for riboflavin production. In the beginning, random mutagenesis and screening by colony colour were the main strategies employed (Bigelis 1989), followed by the use of

anti-metabolites (like itaconate and oxalate, which inhibit ICL activity) in combination with several rounds of mutagenesis (Schmidt et al. 1996, Park et al. 2007, Sugimoto et al. 2010). With the introduction of recombinant DNA technology in *A. gossypii*, molecular manipulations have been added to mutational techniques as a means of further improving strains in a rational way.

Transformation methods for inserting foreign DNA into *A. gossypii* were initially developed by Wright and Philippsen (1991). The surprising result of their study was that freely replicating plasmids are propagated in *A. gossypii* based on *S. cerevisiae* autonomously replicating sequences (ARS) (*ARS1* and 2-micron). However, an ARS element and centromere of the *A. gossypii* chromosome V, which function in *Holleya sinecauda* (a close relative of *A. gossypii*; Kurtzman 1995), were found not to be functional in *S. cerevisiae* (Schade et al. 2003). *S. cerevisiae* centromeres do not function in *A. gossypii* as well (Wendland and Walther 2005). Plasmids containing *S. cerevisiae ARS* elements have since been widely used to express homologous and heterologous genes in *A. gossypii* (Altmann-Jöhl and Philippsen 1996, Monschau et al. 1998, Ayad-Durieux et al. 2000, Kato and Park 2006, Ribeiro et al. 2010).

Strong, constitutive promoters for driving homologous and heterologous gene expression in *A. gossypii* have been described, such as those of the *A. gossypii* translation elongation factor 1α (*AgTEF*) or glyceraldehyde-3-phosphate dehydrogenase (*AgGPD*) (Steiner and Philippsen 1994, Revuelta et al. 1999). The *S. cerevisiae TEF2*, *PDC1*, *PGK1* and *LEU2* constitutive promoters have been shown to drive the expression of heterologous genes in *A. gossypii* as well (Wendland et al. 2000, Ribeiro et al. 2010, Kaufmann 2009). Regulatable promoters, such as the *A. gossypii* and *S. cerevisiae MET3* and *THI13* promoters have also been described to regulate the expression of genes in *A. gossypii* (Dünkler and Wendland 2007, Kaufmann 2009).

Another interesting characteristic of *A. gossypii* is that, in contrast to other filamentous fungi and similarly to *S. cerevisiae*, it has a highly efficient homologous recombinantion system (Steiner et al. 1995). This has facilitated the establishment of simple PCR-based gene targeting techniques in *A. gossypii* (Wendland et al. 2000). Integrative transformation of *A. gossypii* results in primary heterokaryotic transformants, which contain both wild type and mutant nuclei, and therefore show wild type-like phenotype. These heterokaryotic transformants generate uninucleate spores that when grown under selective conditions allow the isolation of homokaryotic mutants (Steiner et

al. 1995). Therefore, compared with other filamentous fungi, gene analysis in *A. gossypii* is particularly straightforward.

A. gossypii strains auxotrophic for leucine (*Agleu2*Δ) and threonine (*Agthr4*Δ) have been developed for use in molecular genetic studies (Altmann-Jöhl and Philippsen 1996, Mohr 1997). These auxotrophies are complemented by the corresponding *S. cerevisiae* homolog genes (Altmann-Jöhl and Philippsen 1996, Mohr 1997). Several heterologous selectable marker cassettes for dominant selection have also been developed, such as the *GEN3*, *NATPS* and *BLE3* marker cassettes, which confer resistance to G418/geneticin, clonNAT/nourseothricin and phleomycin, respectively (Wendland et al. 2000, Hoepfner *in* Kaufmann 2009, Ribeiro et al. 2013). A hygromycin resistance (*Hyg*^{r}) marker gene has also been used to select *A. gossypii* transformants (Mateos et al. 2006, Jiménez et al. 2008).

Reporter proteins, like the *Escherichia coli* β-galactosidase and the green fluorescent protein (GFP), have been widely used in *A. gossypii* to study gene expression and protein localization (Monschau et al. 1998, Ayad-Durieux et al. 2000, Dünkler and Wendland 2007, Walther and Wendland 2012). In this context, Kaufmann (2009) developed a plasmid collection for PCR-based gene targeting in *A. gossypii* that covers a wide selection of modules for fluorescent protein tagging combined with different heterologous selection markers. Additionally, several fluorescent dyes can be used in *A. gossypii* to stain various cellular organelles and cytoskeleton structures (Wendland and Philippsen 2001, Walther and Wendland 2004).

Important molecular tools for genetic manipulation of *A. gossypii* have been established, but optimal strain design for industrial applications requires a transformation system that allows the removal of exogenous genes, especially those conferring resistance to antibiotics, but such a system was still not available for *A. gossypii*. A methodology that has been widely used in several organisms to remove and recycle marker genes is based on the Cre-*lox*P recombination system (Sauer 1987, Steensma and Ter Linde 2001, Forment et al. 2006, Steiger et al. 2011). The Cre-*lox*P system makes use of the site-specific recombinase Cre from the bacteriophage P1 that catalyzes the recombination between two of its DNA recognition sites, called *lox*P (van Duyne 2001). The *lox*P site is a 34 bp consensus sequence composed of two 14 bp recombinase binding elements arranged as inverted repeats around a central 6-bp crossover region (van Duyne 2001). The Cre-induced recombination between *lox*P sites allows the excision of DNA fragments that are flanked by *lox*P sequences. In Chapter 6, a Cre-*lox*P based system developed in the

scope of this thesis for generating *A. gossypii* strains free of exogenous marker genes is described.

1.1.6 Biotechnological applications

A. gossypii constitutes a paradigm of the sustainable "white" biotechnology with regard to industrial riboflavin production. Riboflavin is a precursor of the flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), which are essential co-factors for numerous enzymes (including dehydrogenases, oxidases, oxidoreductases) that participate in a range of redox reactions critical for major biological processes. This vitamin is commercially used as a yellow colorant and animal food additive, and for many years it was mainly obtained by chemical synthesis until its industrial production shifted to microbial fermentation (Stahmann et al. 2000, Meyer zu Berstenhorst et al. 2009). As an outcome of classical improvement of naturally overproducing strains, metabolic engineering and culture conditions optimization, an industrial fermentation process using *A. gossypii* for riboflavin production is among the few biotechnological processes that replaced the chemical process in use due to its higher economical competitiveness (Stahmann et al. 2000, Meyer zu Berstenhorst et al. 2009).

Since the 1990s, when BASF (Germany) launched a riboflavin production plant with *A. gossypii*, the performance of *A. gossypii* strains could be much improved by rational metabolic design. The deep understanding of the *A. gossypii* riboflavin biosynthetic pathway, which has been acquired along the last two decades, allowed this (reviewed in Kato and Park 2012). Among the rational strategies tested, the majority focused on increasing the flux of the riboflavin biosynthetic pathway by (Figure 1.3): (**1**) overexpressing the *AgRIB3* gene, which encodes the first enzyme in the ribulose-5-phosphate branch of the riboflavin biosynthetic pathway (Meyer zu Berstenhorst et al. 2009); (**2**) overexpressing the *AgGLY1* gene, involved in the conversion of threonine to glycine (Monschau et al. 1998); (**3**) expressing the *S. cerevisiae AGX1* gene, which is not present in *A. gossypii* and encodes an alanine:glyoxylate aminotransferase that converts glyoxylate into glycine (Kato and Park 2006); (**4**) disrupting the *AgSHM2* gene, which encodes a cytosolic serine hydroxymethyltransferase that converts glycine into serine (Schlüpen et al. 2003); and (**5**) overexpressing several genes involved in the purine biosynthetic pathway (*AgADE4*, *AgPRS2,4* and *AgPRS3*) (Jiménez et al. 2005, Jiménez et al. 2008). Another strategy focused on increasing the excretion of

riboflavin into the medium by disrupting the gene encoding the vacuolar ATPase subunit A, *AgVMA1* (Förster et al 1999).

The development of novel biotechnological applications for *A. gossypii* has been motivated by several attractive features of this filamentous fungus, such as good performance in large-scale fermentation processes, ability to grow in inexpensive media to high cell densities (Park et al. 2007; Tajima et al. 2009), and small haploid genome with high similarity to that of the well studied *S. cerevisiae*. These features, combined with the *A. gossypii* natural ability to overproduce riboflavin and with the availability of several tools for its easy genetic manipulation, have raised attention to this fungus as a potential cell factory organism, which could be tailor-made to produce other metabolites an/or proteins. In this context, the genome scale-metabolic model of *A. gossypii* has been constructed, being now in its experimental validation phase (Gomes et al. personal communication). This tool will help design novel rational metabolic engineering strategies for further improving riboflavin productivity and/or for adding new desired genes and pathways that could lead to the production of other commercially interesting compounds.

The potential of *A. gossypii* as a host for the production of heterologous proteins has also been investigated through the expression of two cellulases from *Trichoderma reesei*, endoglucanase I (EGI) and cellobiohydrolase I (CBHI) (Ribeiro et al. 2010). *A. gossypii* possesses the ability to secrete native and heterologous enzymes to the extracellular medium and to recognize signal peptides of other organism as secretion signals (Stahmann et al. 1997, Ribeiro et al. 2010). This is a desired property for cost-efficient downstream processing of low- and medium-value enzymes. Moreover, it is also able to perform protein post-translation modifications, such as *N*-glycosylation and other modifications required for biological activity and stability of proteins (Ribeiro et al. 2010). Compared to the closely related yeast *S. cerevisiae*, one of the most commonly used fungal hosts for the production of heterologous proteins (Porro et al. 2005, Demain and Vaishnav 2009), *A. gossypii* seems to have the tendency to hyperglycosylate secreted glycoproteins less extensively, which is advantageous for the production of proteins whose properties may be adversely affected by extensive glycosylation (Ribeiro et al. 2010). However, the production levels of active EGI and CBHI secreted by *A. gossypii* were comparable to those secreted by *S. cerevisiae*. EGI and CBHI have been reported as challenging enzymes to be produced by recombinant microbial hosts, both as functional proteins and at reasonable yields (Ribeiro et al. 2010). Therefore, further exploitation of *A.*

gossypii as a recombinant protein producer should be done with proteins that have been previously secreted biologically-active and at high levels by other model hosts.

The expression and secretion of proteins by *A. gossypii* have been scarcely explored and little is also known about the protein secretory pathway of this fungus. In Chapters 2 and 3, valuable insights into the *A. gossypii* protein secretory pathway are provided. Chapters 4 and 5 further explore the natural and recombinant protein secretion abilities of this biotechnologically relevant fungus.

1.2 PROTEIN SECRETION IN FUNGI

Protein secretion is a cell translocation process of major biological and technological significance. Cell communication, intercellular signaling and growth of eukaryotic cells depend on the protein secretory pathway. Within the fungal kingdom, filamentous fungi exhibit an extraordinary capacity to secrete large amounts of extracellular proteins, which are necessary for hyphal extension, pathogenicity and degradation of substrates in natural ecosystems. This high secretion capacity of filamentous fungi has been widely exploited by the biotechnology industry for the commercial production of proteins (Nevalainen et al. 2005, Demain and Vaishnav 2009, Ward 2012). The protein secretion capacity of yeast is considerably lower than that of filamentous fungi and the different morphologies of these organisms is thought to contribute to some of the differences observed in the organization of their protein secretory pathway (reviewed in Conesa et al. 2001).

In contrast to budding yeast, filamentous fungi have a highly polarized cell growth at the hyphal typ. To deal with the high demand for novel cell surface at the growing typ, filamentous fungi require a more complex organization, communication and efficient secretory vesicle transport across long septated hyphae (Schmitz and Philippsen 2011). Unique for filamentous fungi (*A. gossypii* included; Köhli et al. 2008) is the presence of a Spitzenkörper at the hyphal tip, which is a complex multi-component structure dominated by vesicles (reviewed by Harris et al. 2005, Steinberg 2007). This structure is considered important for the efficient delivery of proteins required for hyphal elongation and for efficient secretion of extracellular proteins necessary for lysis of substrates or synthesis of the cell wall (Steinberg 2007). Protein secretion in filamentous fungi is thus thought to occur preferentially (but not exclusively) at hyphal apices (Read 2011).

The high levels of extracellular proteins secreted by filamentous fungi have been generally accepted to be linked to their efficient secretion machinery. However, despite presenting sustained polarized hyphal growth, the extracellular protein secretion capacity of *A. gossypii* seems to resemble that of its closely related yeast species (Ribeiro et al. 2010, Ribeiro et al. 2013). The characteristics of its natural habitat (rich in reducing sugars and oils; Caskey and Gallup 1931) and its insect-dependent infection mode justify modest secretion of extracellular proteins by *A. gossypii*, as it does not need them for substrate degradation. However, a better knowledge about the range of proteins natively secreted by this fungus and a better understanding of its protein secretory pathway would contribute to fully elucidate its secretion abilities. Chapters 2 and 3 of this thesis address these topics.

1.2.1 The secretory pathway

Proteins destined to be secreted to the extracellular space are trafficked through the secretory pathway, a part of the endomembrane system of the cell. Along with secreted proteins, proteins destined for other cellular locations such as the plasma membrane, vacuole/lysosome or the Golgi complex, are also transported by the endomembrane system (reviewed on Conesa et al. 2001, Shoji et al. 2008; Saloheimo and Pakula 2012).

Secretory proteins enter the general secretory pathway in the endoplasmic reticulum (ER), directed by a targeting signal commonly referred as signal sequence/peptide. This is usually located at the N-terminal side of the proteins and consists of a continuous stretch of hydrophobic residues which mediates their translocation into the ER (Martoglio and Dobberstein 1998). Two protein translocation pathways across the ER membrane have been described in *S. cerevisiae*: the signal recognition particle (SRP)-dependent pathway, in which translocation occurs co-translationally, and the SRP-independent pathway, which occurs post-translationally (Conesa et al. 2001).

In the ER, newly synthesized proteins acquire their proper conformation and undergo initial post-translational modifications, such as disulfide bridge formation, glycosylation, phosphorylation and subunit assembly (Conesa et al. 2001). The ER has a sophisticated quality control system that monitors the folding state of the cargo proteins and only allows the exit of fully folded ones. Misfolded or aggregated proteins are removed from the ER and degraded by the ER-associated protein degradation (ERAD) system (Stolz and Wolf 2010). An increased burden of improperly folded proteins activates

a signalling cascade known as the unfolded protein response (UPR), which induces a number of events aimed at decreasing the ER load (reviewed in Mori 2000, Walter and Ron 2011). These events include an increased expression of folding chaperones, an attenuation of translation and/or removal of parts of the ER containing aggregated unfolded proteins by autophagy (Mori 2000, Yorimitsu et al. 2006, Walter and Ron 2011). If stress cannot be relieved, the UPR can also serve as an apoptotic executor (Walter and Ron 2011). In this regard, protein folding in the ER constitutes a bottleneck in heterologous protein production by fungi, as expression of heterologous proteins often activates the UPR (Gasser et al. 2008, Arvas et al. 2006, Saloheimo and Pakula 2012).

After folding and core glycosylation of the secreted proteins, their trafficking towards the cell exterior begins via the Golgi complex inside membrane vesicles coated by specific protein coats. Alongside with its crucial role in the sorting of secreted and vacuolar proteins to their final destinations, the Golgi also plays a key role in the further post-translational processing of secreted proteins, such as glycosylation and phosphorylation (Shoji et al. 2008). In filamentous fungi, secretory vesicles are transported to their destinations in a cytoskeleton-assisted process (Schmitz and Philippsen 2011, Saloheimo and Pakula 2012). Concerning the transport to the plasma membrane, a large protein complex called the exocyst is important for the localization of the fusion event on the membrane (Saloheimo and Pakula 2012). In *A. gossypii*, when fused to GFP, exocyst components localize as a cortical cap to the tip, marking the vesicle fusion zone, and accumulate in a Spitzenkörper-like spherical body in fast hyphae (Köhli et al. 2008). These observations suggest that *A. gossypii* seems to have dominant secretion at hyphal tips.

1.2.1.1 Protein glycosylation

During the past few years, it has become increasingly evident that glycosylation is vital for cell wall synthesis and thus vital for growth and morphology of filamentous fungi (Jin 2012). Moreover, in the context of heterologous protein production, correct glycosylation is important for the stability and biological activity of the produced protein (Nevalainen et al. 2005, Ward 2012).

All eukaryotes possess three major types of protein glycosylation, *N*-glycosylation of asparagine residues, *O*-glycosylation of threonine and serine residues, and glycosylphosphatidylinositol-anchoring (GPI-anchoring) of the C-terminus of some

proteins. The transfer of these carbohydrate structures is initiated during, or soon after, the translocation of nascent polypeptide chains into the ER lumen. The core glycans are further elaborated by the actions of glycosidases and glycosyltransferases in the ER and Golgi as the protein traverses the secretory pathway, yielding a diverse array of structures.

***N*-glycosylation** begins with the assembly of a dolichol (Dol)-linked oligosaccharide precursor at the cytoplasmic leaflet of the ER membrane (Figure 1.4; reviewed in Weerapana and Imperiali 2006, Jigami 2008). In *S. cerevisiae*, the glycosylphosphotransferase Alg7p first transfers *N*-acetylglucosamine (GlcNAc)-P from UDP-GlcNAc to membrane bound Dol-P to form Dol-PP-GlcNAc. A second GlcNAc residue is added by the dimmer Alg13p/Alg14p and the structure is further modified by a series of mannosyltransferases (Alg1p, Alg2p and Alg11p) that add five mannose (Man) residues derived from GDP-Man. The resulting Dol-PP-$GlcNAc_2Man_5$ molecule is then translocated by the flippase Rft1p to the lumenal leaflet of the ER and synthesis of the *N*-glycan precursor continues with the transfer of four Man residues from Dol-P-Man to the oligosaccharide core by the mannosyltransferases Alg3p, Alg12p and Alg9p. Finally, the Dol-PP-$GlcNAc_2Man_9$ structure is capped with three glucose residues (from Dol-P-Glc) via the actions of the glucosyltransferases Alg6p, Alg8p and Alg10p. The resulting Dol-PP-$GlcNAc_2Man_9Glc_3$ precursor is subsequently transferred to an asparagine residue within an Asn-X-Ser/Thr (where X is any amino acid except for Pro) consensus sequence in the nascent peptide. This transfer is catalyzed by the oligosaccharyltransferase (OST) complex. In *S. cerevisiae*, the OST complex consists of at least eight different subunits, including Ost1p, Ost2p, Wbp1, Stt3p, Swp1p, Ost4p, Ost5p, and Ost3p/Ost6p. Although the function of each subunit is still unclear, Stt3p is believed to be the catalytic subunit, and its homologs are found in almost all eukaryotes (Lennarz 2007).

Once $Glc_3Man_9GlcNAc_2$ is transferred to proteins, the *N*-glycan is processed sequentially in the ER and Golgi. *N*-Glycan processing is initiated by the removal of the glucose residues catalyzed by the ER α-glucosidases I (encoded by *CWH41*) and II (encoded by *ROT2*). In *S. cerevisiae*, the resulting structure is further trimmed by the mannosidase Mns1p, which removes a single α1,2-Man residue (Jigami 2008). The folded protein, with its $GlcNAc_2Man_8$ modification, is then transferred from the ER to the Golgi.

In mammalian cells, *N*-glycans plays a decisive role as a quality control (QC) of the folding of secretory proteins, which is composed of calnexin, calreticulin, UDPglucose:glycoprotein glucosyltransferase (GT) and glucosidase II (GII) (Helenius and

Aebi 2004). Recently, some studies have suggested the existence of an *N*-glycan-dependent QC system in filamentous fungi as well (Jin 2012). Unlike mammalian cells, *S. cerevisiae* lacks a calnexin cycle and GT and only has an effective mannosidase I-dependent ERAD system. A mannosidase-like protein, Mnl1p (Htm1p), recognizes $Man_8GlcNAc_2$ oligosaccharides to prevent the secretion of a misfolded glycoprotein, delivering it to the ERAD pathway (Jigami 2008).

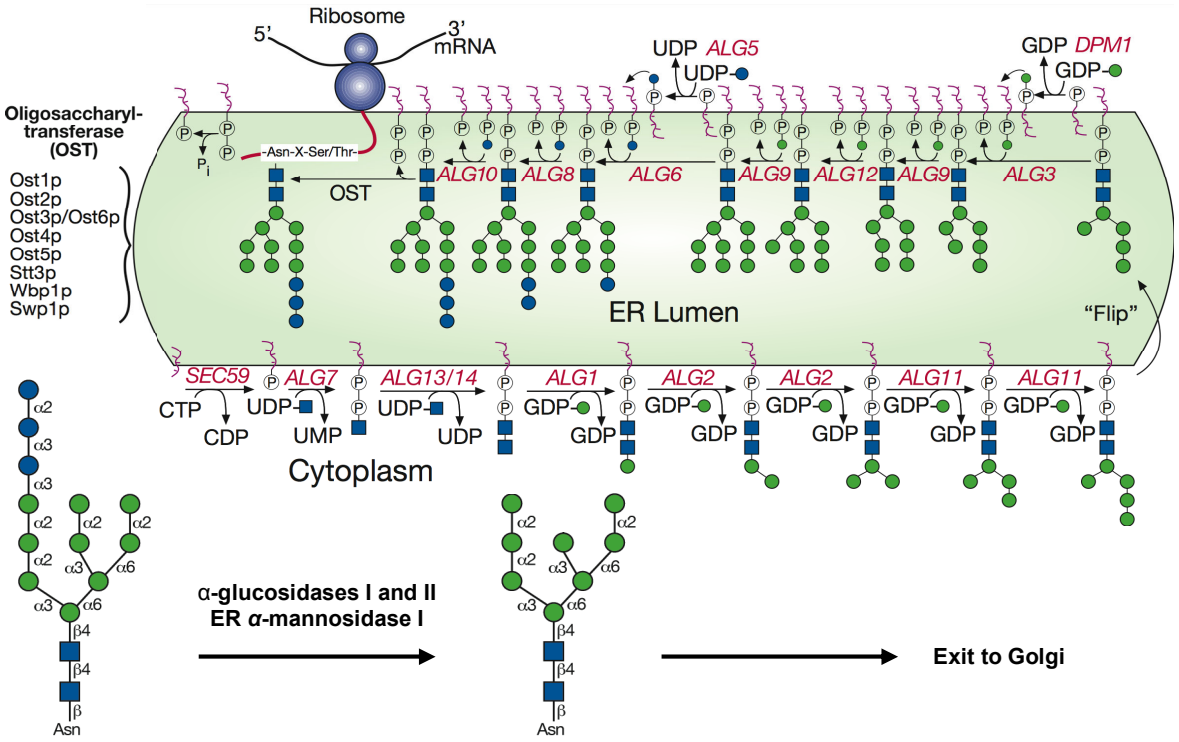


Figure 1.4 – Biosynthesis and transfer of the precursor *N*-glycan to Asn-X-Ser/Thr residues in newly synthesized glycoproteins in the ER. ■ *N*-acetylhexosamine; ●, mannose; ● glucose. Adapted from Cummings and Doering (2009).

Fungi do not generate 'complex-type' *N*-glycans typical of mammalian systems. Instead, the core oligosaccharide structures of *N*-glycans are either minimally modified ('core-type') or receive extensive modifications ('hypermannosylated'). In *S. cerevisiae*, protein modifications involve Man addition exclusively, although other yeast species may also add GlcNAc (*K. lactis*) or galactose (*Schizosaccharomyces pombe*) (Gemmill and Trimble 1999). When a $Man_8GlcNAc_2$ modified polypeptide arrives in the Golgi apparatus, Och1p adds a single α1,6-Man to the structure. A glycoprotein that will retain a core-type *N*-glycan structure is further modified at this new Man by an α1,2-mannosyltransferase (the protein responsible has not yet been identified) that adds a

single Man and by subsequent capping of secondary branches by the α1,3-mannosyltransferase Mnn1p (Figure 1.5; Jigami 2008). In contrast to these modest alterations, hypermannosylated *N*-glycans receive extensive modifications beyond the actions of Och1p (Figure 1.5; Jigami 2008). First, the enzyme complex mannan polymerase I (M-Pol I composed by Mnn9p and Van1p) modifies the new Man with a linear branch of α1,6-Man with approximately ten residues of length. A second mannan polymerase complex (M-Pol II composed by Mnn9p, Anp1p, Mnn10p, Mnn11p and Hoc1p) elongates this branch with up to fifty α1,6-Man residues more. This α1,6-Man backbone may be elaborated with α1,2-Man by the action of Mnn2p or Mnn5p. Mannosylphosphate (Man-P) may also be transferred to mannan by Mnn6p (Ktr6p) under the regulation of Mnn4p. Finally, the secondary branches of these extensive mannan chains are capped with an α1,3-Man residue by Mnn1p (Jigami 2008).

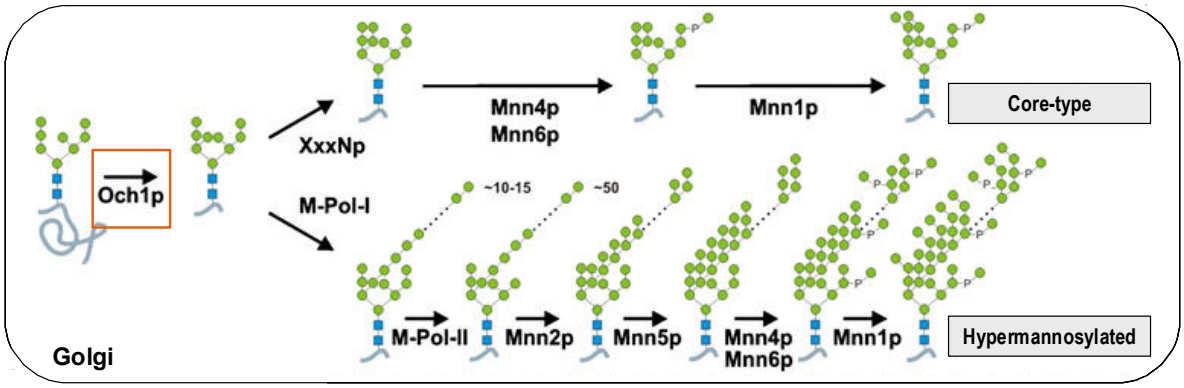


Figure 1.5 – Pathways of *N*-glycan modification in the yeast Golgi. The α1,2-mannosyltransferase responsible for the first addition in the pathway to the core-type structure is currently unknown (XxxNp). ■ *N*-acetylhexosamine; ●, mannose. Adapted from De Pourcq et al. (2010).

The *N*-glycosylation patterns from filamentous fungi are more similar to those of mammals than the patterns observed in common yeast hosts (Nevalainen et al. 2005, Ward 2012). Although hypermannosylation may occur in some filamentous fungi, most often they synthesize small high-mannose type *N*-glycans. They have two distinct α1,2-mannosydases, one of which is similar to the mammalian Golgi α1,2-mannosydases that trim three Man residues from $Man_8GlcNAc_2$ to form $Man_5GlcNAc_2$ as substrate for GlcNAc transferase 1. The other is a distinct fungal α1,2-mannosydase (De Pourcq et al. 2010). *N*-glycans from filamentous fungi may also have terminal altered substituents such as glucose, galactose or phosphoesters (Deshpand et al. 2008).

***O*-mannosylation** is initiated by a family of *O*-mannosyltransferases (PMTs) that are evolutionarily conserved from yeast to human. In *S. cerevisiae*, seven PMTs have been identified (Pmt1-7p). These fall into three major groups of homology: the PMT1 family (Pmt1p and Pmt5), the PMT2 family (Pmt2p, Pmt3p and Pmt6p) and the PMT4 family (Pmt4p). Filamentous fungi generally contain only three *PMT* genes that belong to the PMT1, PMT2 and PMT4 families, respectively (Jin 2012).

Similarly to the progress of *N*-glycosylation, the core Man of *O*-glycans that is added to proteins in the ER is further elongated in the Golgi, where one to six Man residues derived from GDP-Man may be added to extend the linear chain (Jigami 2008). Mannose may be added in α1,2-linkages through the actions of the KTR family (Ktr1p, Ktr3p and Kre2p) or in α1,3-linkages by the MNN1 family (Mnn1p, Mnt2p and Mnt3p) (Figure 1.6). As in *N*-glycans, Man-P may also be added by *S. cerevisiae* to some Man residues in *O*-glycan structures through the action of Mnn6p (Ktr6p) (Jigami 2008). Simple *O*-glycans containing only α1,2-Man residues are found in *Candida albicans* and *S. pombe* adds galactose residues to the non-reducing end of short α1,2-Man chains (Figure 1.6) (Gemmill and Trimble 1999). In addition to linear chains of oligosaccharides, branched *O*-glycans with Man residues linked by α1,6-, α1,2- and α1,3-linkages have been identified in some filamentous fungi (Deshpand et al. 2008). The presence of glucose, galactopyranose, galactofuranose, sulfate or phosphate residues in *O*-glycans has also been documented for some filamentous fungi species (Deshpand et al. 2008).

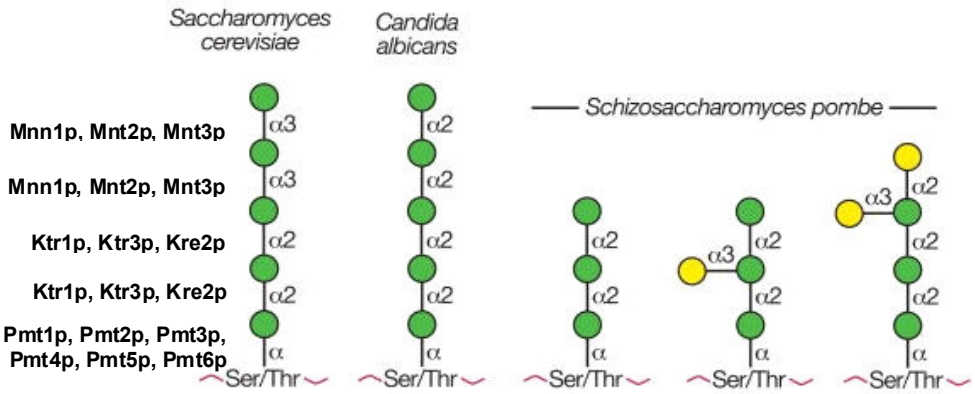


Figure 1.6 – Yeast *O*-glycan structures. The enzymes involved in the addition of the different Man residues to *S. cerevisiae O*-glycans are indicated. ●, mannose; ●, galactose. Adapted from Cummings and Doering (2009).

GPI-anchoring, by which proteins are attached to the plasma membrane, is conserved in all eukaryotic cells (Jigami 2008). GPI anchor synthesis is essential for *S. cerevisiae* viability and plays an important role in the biosynthesis and organization of the cell wall. Synthesis of the GPI anchor begins in the cytoplasmic leaflet of the ER membrane, where GlcNAc is transferred from UDP-GlcNAc to phosphatidylinositol (PI). This process is catalysed in *S. cerevisiae* by the transmembrane protein Gpi3p in association with five other polypeptides (Gpi1p, Gpi2p, Gpi15p, Gpi19p and Eri1p) that form the GPI-GlcNAc transferase complex (Pittet and Conzelmann, 2007). The resulting GlcNAc-PI is then deacetylated by Gpi12p to yield GlcN-PI and transferred to the lumenal leaflet of the ER membrane by an unidentified flippase. Once in the lumen, the inositol moiety is palmitoylated by Gwt1p, followed by the addition of up to four Man residues (from Dol-P-Man) and three phosphoethanolamine (EtnP) moieties. The presence of EtnP on the third Man is absolutely required for the association of the GPI anchor with a protein. Therefore, both Gpi13p (which adds this EtnP) and Smp3p (which adds the fourth Man, whose presence is required for the actions of Gpi13p) are essential in *S. cerevisiae* (Jigami 2008). The complete GPI precursor is then transferred to a newly synthesized protein with a C-terminal hydrophobic motive by the GPI anchor transamidase complex (Gpi8p, Gaa1p, Gpi16p, Gpi17p and Gab1p). After transfer to the protein, the inositol portion of the GPI is deacylated by Bst1p. The lipid moiety of GPI is often changed (lipid remodelling) to another form, such as diacylglycerol, with longer fatty acids or ceramides (Pittet and Conzelmann, 2007). The glycan portion of GPIs is also sometimes modified by the addition of another α1,2- or α1,3-Man to the fourth Man of the anchor by an unknown mannosyltransferase (Jin 2012).

1.2.2 Heterologous protein secretion

Many fungal species have been developed as cell factories for the production of secreted heterologous proteins. Contrary to prokaryotic hosts, which are simpler and easier to manipulate, fungi have the ability to secrete soluble proteins at high yields and to process eukaryotic post-translational protein modifications, such as protein glycosylation, proteolytic maturation and disulfide bond formation. Comparing to other eukaryotic systems, fungi can be more readily manipulated and grow relatively rapidly to high cell densities in low-cost media. Moreover, several fungal species have long been used in food

fermentation processes and are Generally Recognized as Safe (GRAS) (Demain and Vaishnav 2009).

Although heterologous genes from many different natural sources have been successfully expressed in fungal systems, the secreted yields of heterologous proteins by fungal species have been better when the gene to express is of fungal origin (Nevalainen et al. 2005, Porro et al. 2005, Demain and Vaishnav 2009, Ward 2012). Product yield is a key factor in using fungi as cell factories and several strategies developed to increase protein expression and translation efficiency have included (Porro et al. 2005, Demain and Vaishnav 2009, Ward 2012): (**1**) use of strong (preferably homologous) promoters; (**2**) integration of genes at a transcriptionally active locus; (**2**) optimization (not necessarily maximization) of gene copy number; (**3**) codon optimization of the heterologous gene; (**4**) addition of secretion sequences or fusion of the heterologous gene with a gene encoding a naturally well secreted protein.

There are, however, several pertinent areas which remain under-investigated. For example, there is a limited range of promoters in current use and the genome sequences now available for many fungi should provide other options quite readily, even if they will require testing. Moreover, promoters have not usually been modified for optimised activity by, for example, improving core promoter activity. The range of signal sequences available for targeting the secretion of recombinant proteins is also limited and the screening of signal peptides in native proteins highly secreted by each host should provide more suitable options.

Increasing gene expression and translational efficiency beyond a critical point may however become ineffective, as increased secretion of at least some heterologous proteins induces the UPR (Arvas et al. 2006, Gasser et al. 2008, Saloheimo and Pakula 2012). Studies in yeast and filamentous fungi with chemicals that function as secretion blockers, such as dithiothreitol (DTT) and tunicamycin, also induce the UPR, with the DTT response being very strong. Therefore, these chemicals have been widely used to study the UPR in fungi (Travers et al. 2000, Saloheimo et al. 2003, Graf et al. 2008, Wimalasena et al. 2008, Sims et al. 2005, Guillemette et al. 2007).

The most conserved and best understood UPR signalling pathway in lower eukaryotes is controlled by the ER-resident sensor Ire1p and the downstream transcription factor Hac1p. The chaperone immunoglobulin heavy-chain binding protein (Kar2p in yeast and BiP in mammals) binds to Ire1p when not bound by unfolded proteins and keeps it in an inactive form, but when the load of unfolded proteins increases they compete for this

chaperone, favouring its dissociation from Ire1p, thus allowing its activation (reviewed in Kimata and Kohono 2011). However, Ire1p can also sense secretion stress by binding directly to unfolded proteins and subsequently be activated (Walter and Ron 2011). Activated Ire1p promotes splicing of *HAC1* messenger RNA (mRNA) by an unconventional mechanism and only then Hac1p is actively translated and capable of inducing its downstream target genes, such as *KAR2/BIP* and *PDI1*, by binding to an UPR Element (UPRE) existent in their promoter regions (Mori et al. 1996, Saloheimo et al. 1999). Apart from activating the UPR in a *HAC1*-dependent manner, Ire1p signaling may also cause regulated *IRE1*-dependent mRNA decay (RIDD) to induce the degradation of mRNAs encoding secretory proteins (Walter and Ron 2011; Miyazaki et al. 2013). Under secretion stress conditions, some fungi also exhibit transcriptional down regulation of genes encoding secreted proteins (REpression under Secretion Stress, RESS), such as *T. reesei*, *Aspergillus niger* and *S. cerevisiae* (Pakula et al. 2003, Al-Sheikh et al. 2004, Kimata et al. 2006). The mechanism of RESS is unknown, but its regulation is likely to occur at the transcription level. A stress response induced by misfolded cytosolic proteins which do not enter the secretory pathway, called UPR-Cyto, has also been reported to induce in *S. cerevisiae* the production of several cytosolic chaperones and co-chaperones in an apparently Hsf1p-mediated way (Metzger and Michaelis 2009, Geiler-Samerotte et al. 2010).

By manipulating levels of ER resident chaperones and transcription factors involved in UPR, improvements in the secretion of a number of proteins by different fungi have been achieved. For instance, simultaneous or independent overexpression of *KAR2/BIP* and *PDI1* have been shown to enhance the secreted yields of various heterologous proteins in some fungi, such as *S. cerevisiae*, *Pichia pastoris* and *S. pombe* (Robinson et al. 1994, Shusta et al. 1998, Damasceno et al. 2012), but not in others, such as some *A. niger* species (Ward 2012). Constitutive activation of the UPR by overexpressing the active form of *HAC1* also increased in *S. cerevisiae*, *Pichia pastoris* and *A. niger* the production of some secreted heterologous proteins (Valkonen et al. 2003a,b, Damasceno et al. 2012).

Product stability and authenticity are other key factors in using fungi as cell factories. In this regard, the engineering of humanized glycosylation pathways in yeast and fungi has been widely explored to generate strains capable of producing the correct glycoform of therapeutic proteins, as foreign glycosylation may induce an immune response in the patient, reducing the treatment efficacy (reviewed in De Pourcq et al.

2010). Even when the heterologous protein to be produced is of fungal origin, correct gycosylation of the protein assumes particular importance in its stability and biological activity, as there are significant differences among fungi in the types of sugars added and structures formed during glycosylation.

The reduction of the extracellular proteolytic activity of fungal strains, mainly of filamentous fungi strains (which secrete large amounts of these hydrolytic enzymes), through the construction of protease-deficient hosts has also received much attention (Ward 2012). Fungal proteases severely hamper heterologous protein production, as they degrade the proteins secreted into the medium before its downstream processing. Moreover, large amounts of natively secreted proteins also contaminate the final product and difficult its purification. Knowledge of proteases from genomic information may prove to be very useful in strain construction.

Whereas rational optimization of the protein secretory pathway of fungal strains has been resulting in increased production of secreted heterologous proteins, better results are obtained when this is accompanied by process optimization. Manipulation of fungal culture pH away from the optimal pH for protease activity and implementation of cultivation strategies which prevent release of intracellular proteases via cellular lysis have been shown to reduce proteolysis of secreted recombinant proteins (Ward 2012). Different culture media and growth rates have also been shown to result in different protein glycosylation processing (Stals et al. 2004, Schenk et al. 2008). Therefore, the optimization and control of the fermentation processes may by itself have yield-enhancing effects.

The results presented along the different Chapters of this thesis provide valuable insights into different aspects of protein secretion in *A. gossypii*, namely in what concerns native and heterologous protein secretion, protein secretion stress and protein glycosylation, leading to a wider understanding of the biotechnological potential of this fungus.

CHAPTER 2

Genome-wide study of protein secretion and secretion stress in *Ashbya gossypii*

ABSTRACT

Protein secretion is a cellular process of great biological and technological importance. To gain insight into the protein secretory pathway of the biotechnologically relevant fungus *Ashbya gossypii*, we undertook genome-wide analyses to explore its secretome and its transcriptional responses to secretion stress. A computational pipeline was used to predict the inventory of proteins putatively secreted by *A. gossypii* via the general secretory pathway. The proteins actually secreted by this filamentous fungus into the supernatants of submerged cultures in minimal and rich medium were mapped by two-dimensional gel electrophoresis, revealing that most of the *A. gossypii* secreted proteins have an isoelectric point between 4 and 6 and a molecular mass above 25 kDa. These analyses also indicated that 1-4% of the *A. gossypii* proteins are likely to be secreted, of which less than 33% are putative hydrolases. Whole-genome transcriptomic analyses carried out in *A. gossypii* cells under recombinant protein production conditions and DTT-induced secretion stress unexpectedly revealed that a conventional unfolded protein response (UPR) was not activated in none of the conditions, as the expression levels of several well-known UPR target genes (e.g. *IRE1*, *KAR2*, *HAC1*, *PDI1* and *EUG1* homologs) remained unaffected. However, several other genes involved in protein unfolding, ERAD, proteolysis, vesicle trafficking, vacuolar protein sorting, secretion and mRNA degradation were up-regulated by DTT-induced secretion stress. Conversely, the transcription of several genes encoding secretory proteins, such as components of the glycosylation pathway, was severely repressed. These transcriptional responses suggest that alternative mechanisms exist in *A. gossypii* to cope with secretion stress.

2.1 INTRODUCTION

The protein secretory pathway constitutes an important area in the field of fungal research, as the secretion of proteins by fungal cells is of major biological and commercial significance. *Ashbya gossypii* (syn. *Eremothecium gossypii*), a well known industrial producer of riboflavin (Stahmann et al. 2000), is a filamentous fungus that has been recently considered as a host for the production of recombinant proteins (Ribeiro et al. 2010). However, its protein secretory pathway and the spectrum of proteins natively secreted by this fungus to the extracellular space remain virtually unexplored. *A. gossypii* is phylogenetically closer to yeast than to other filamentous fungi (Prillinger et al. 1997) and has one of the smallest eukaryotic genomes known, which shares a high degree of gene homology and gene order conservation with that of the budding yeast *Saccharomyces cerevisiae* (Dietrich et al. 2004). Although efficient protein secretion is generally associated with filamentous growth, the secretion levels of the heterologous proteins EGI (endoglucanase I) and CBHI (cellobiohydrolase I) from *Trichoderma reesei* in *A. gossypii* were comparable to those obtained with *S. cerevisiae* (Ribeiro et al. 2010). The production of heterologous proteins by fungal species is usually much less efficient than the production of native proteins and several steps ing the secretory pathway (e.g. translation, translocation, folding, transport and secretion) are potential bottlenecks for heterologous protein production (Gouka et al. 1997, Sims et al. 2005, Resina et al. 2009, Yoon et al. 2010).

In eukaryotes, newly synthesised proteins are typically targeted for entry into the general secretory pathway by the presence of an N-terminal signal sequence that typically has a length between 15 to 30 amino acids and comprises a central hydrophobic region flanked by hydrophilic N- and C- terminal regions (Martoglio and Dobberstein 1998). Some proteins lacking this sorting signal may also reach the extracellular space, but the vast majority of fungal secreted proteins are believed to contain classical N-terminal signal peptides, which are cleaved off during translocation across the endoplasmatic reticulum (ER) membrane (Lum and Min 2011). Taking advantaged of the characteristics of these signal peptides and other sorting signals, several computational tools have been developed to predictict the subcellular location of proteins such as the extracellular space (Emanuelsson et al. 2007, Horton et al. 2007, Käll et al. 2007). These have been widely applied to the genome-wide prediction of putative fungal secretomes (Lum and Min 2011). As the translocation of proteins into the ER is determined by the secretion signal, the

correct processing of signal peptides, together with the proper folding of proteins within the ER, are important in recombinant protein production and secretion (Gasser et al. 2008, Damasceno et al. 2012).

The ER serves as the first station of the secretory pathway. Its lumen provides a unique oxidizing environment in which a highly active folding machinery, including molecular chaperones and foldases, facilitates and promotes the folding, assembling, modification and maturation of proteins. To ensure that only properly folded proteins move onward through the secretory pathway, the ER also contains stringent quality control mechanisms that retain malfolded proteins and ultimately retrotranslocate them into the cytosol for proteasomal degradation through a process called ER-associated degradation (ERAD) (reviewed in Stolz and Wolf 2010). Environmental and physiological demands (e.g. cell differentiation, pH and temperature, nutrient limitation, expression of heterologous proteins, etc) can lead to an imbalance between the protein folding load and the protein folding capacity in the ER lumen, resulting in an accumulation of misfolded proteins, i.e. ER stress (Kauffman et al. 2002, Rutkowski and Kaufman 2004, Matsumoto et al. 2005). In response to ER stress, eukaryotic cells have evolved signalling pathways that induce the unfolded protein response (UPR), which helps to restore the ER-protein folding homeostasis by enhancing the ER protein folding capacity and ERAD, and reducing the translation and entry of new proteins into the ER (reviewed in Mori 2000, Walter and Ron 2011).

Cellular responses to the accumulation of malfolded proteins in the ER have been described for yeast, filamentous fungi and higher eukaryotes, and were shown to play a significant role in the stress response to production of secreted recombinant proteins (Cudna and Dickson 2003, Gasser et al. 2008). The inositol-requiring enzyme 1 (*IRE1*) gene encodes the protein that controls the most conserved and best understood UPR signalling pathway in lower eukaryotes (Mori 2009). Ire1p has a luminal sensing domain coupled to cytosolic kinase and endoribonuclease (RNase) domains (Cox et al. 1993). The accumulation of unfolded proteins in the ER lumen leads to the oligomerization of the Ire1p luminal domain and thereby to the activation of its kinase and RNase functions (Walter and Ron 2011). The unfolded protein load in the ER regulates Ire1p activation by two known processes. One involves the ER chaperone immunoglobulin heavy-chain binding protein (Kar2p in yeast and BiP in mammals), which usually is bound to Ire1p, maintaining it in a monomeric, inactive state. When the load of unfolded proteins increases they compete for this chaperone, favouring its dissociation from Ire1p, thus allowing Ire1p

oligomerization (reviewed in Kimata and Kohono 2011). In the other process, interaction of unfolded proteins with the stress-sensing luminal domain of Ire1p directly activates its oligomerization (Gardner and Walter 2011, Walter and Ron 2011).

Upon Ire1p activation, the Ire1p RNase initiates the splicing of a non-conventional intron from *HAC1* messenger RNA (mRNA), thus allowing the translation of an active Hac1p basic-leucine zipper (bZIP) transcription factor that specifically binds to UPR elements (UPREs) existent in the promoter region of UPR target genes, thereby up-regulating their transcription (Cox and Walter 1996, Mori et al. 1996). In response to strong ER stress, Ire1p signaling may also cause regulated *IRE1*-dependent mRNA decay (RIDD) to reduce the ER load by inducing the degradation of mRNAs encoding secretory proteins (Hollien et al. 2006, Walter and Ron 2011, Miyazaki et al. 2013). Several UPR target genes have been identified, among which are those encoding ER chaperones and protein folding enzymes, ER structural and transport proteins, members of the ERAD machinery and components that mediate autophagy (Kaufman 1999, Travers et al. 2000, Yorimitsu et al. 2006). The bZIP transcription factor Gcn4p, a major controller of the amino acid starvation response, has been shown to also play an essential role in the induction of a large subset of these target genes during ER stress, by directly interacting with Hac1p and modulating its activity in an *IRE1*-independent way (Leber et al. 2004, Patil et al. 2004).

The characterization of protein secretory pathway components and of the regulatory range of secretion stress responses in yeast and filamentous fungi have often relied on inducing ER stress with chemical secretion blockers such as the folding inhibitor dithiothreitol (DTT), the glycosylation inhibitor tunicamycin and the protein trafficking inhibitor brefeldin A. In the present study, we have analysed the events taking place at the transcription level in *A. gossypii* under recombinant protein production conditions and also under DTT-induced secretion stress. Moreover, to explore the native proteins putatively secreted by this fungus, we also analysed its predicted secretome by combining comparative *in silico* predictions for classically secreted proteins with experimental data derived from two-dimensional (2-D) gel electrophoresis.

2.2 MATERIALS AND METHODS

2.2.1 Strains and culture conditions

A. gossypii ATCC 10895, kindly provided by Prof. P. Philippsen (Biozentrum, University of Basel, Switzerland) and here referred as the parental strain, was used for proteomic analyses. A recombinant *A. gossypii* EGI producing strain (VTT D-101398) and its corresponding empty vector described in Ribeiro et al. (2010) were used for transcriptomic analyses. Stock cultures were maintained as spores suspended in 20% (v/v) glycerol, 0.8% (w/v) NaCl with 0.025% (v/v) Tween 20 at -80°C.

Pre-cultures to inoculate bioreactors were grown in 250 ml Erlenmeyer flasks containing 50 ml of AFM (1% (w/v) yeast extract, 1% (w/v) tryptone, 2% (w/v) glucose and 0.1% (w/v) myo-inositol), which was supplemented with 200 µg/ml G418 (Sigma) for maintenance of the recombinant strains. *A. gossyppii* pre-cultures were inoculated with 10^6 spores and grown for 14-17 h at 30ºC and 200 rpm.

For cultivation of the recombinant strains, Biostat® CT bioreactors, maximum working volume of 2.5 l (B. Braun Biotech International, Sartorius AG), containing 1.5 l or 2 l of AFM plus 200 µg/ml G418 were used. Biostat® B-DCU bioreactors, maximum working volume of 2 l (B. Braun Biotech International, Sartorius AG), were used for cultivation of the parental strain in 1 l of either modified AFM or defined minimal medium (DMM; Verduyn et al. 1992), both containing 2% (w/v) sucrose as primary carbon source instead of glucose. Bioreactors were inoculated to an initial biomass of 0.13 ± 0.08 g/l, for recombinant strains, or 0.60 ± 0.05 g/l, for the parental strain. Cultures were grown at 30°C and 500 rpm, with 1.0 volume of gas per volume of culture per minute (vvm) aeration. Culture pH was kept at 6.0 ± 0.1 by the addition of 1 M KOH or 1 M H_3PO_4. Polypropylene glycol (mixed molecular weights; Wiebe et al. 2001) was added to prevent foaming. Gas concentration (CO_2, O_2, N_2 and Ar) was analyzed continuously in an Omnistar quadrupole mass spectrometer (Balzers AG), calibrated with 3% CO_2 in Ar.

For dry weight determination, culture samples were filtered through pre-dried and pre-weighted Whatman GF/B glass fibre filters, washed with at least two sample volumes of double-distilled water and dried to a constant weight at 105ºC. Aliquots of the culture filtrates were stored at -20°C.

Residual sugars and produced metabolites in the culture filtrates were quantified by high performance liquid chromatography (HPLC) as previously described (Toivari et al. 2010). Total protein concentration in the cell-free broth was measured using the Thermo

Scientific Pierce Coomassie (Bradford) Protein Assay kit, with bovine serum albumin (BSA) as standard. The activity of secreted EGI in the culture filtrates was determined as described in Ribeiro et al. (2010), using 4-methylumbelliferyl-β-D-lactoside (MULac) (Sigma) as substrate. Volumetric EGI enzyme activity was defined as micromoles of 4-methylumbelliferone (MU) (Sigma) formed per minute and per litre of culture (μmol min^{-1} l^{-1}) under the assay conditions.

For gene expression analysis, mycelial samples from the recombinant strains were collected from duplicate bioreactor cultivations after 4 h, 7 h, 10 h and 13.5 h after inoculation. After 9.5 h, DTT was added to two out of four *A. gossypii* VTT D-101398 cultures at a final concentration of 10 mM. Samples from DTT-treated cultures were collected 30 min, 1 h and 4 h after DTT addition. Mycelium was rapidly separated from the culture supernatant by filtration through Whatman GF/B glass fibre filters, washed with two sample volumes of 0.9% (w/v) NaCl, frozen immediately in liquid nitrogen and stored at -80ºC.

2.2.2 *In silico* secretome prediction

A computational approach similar to those described to predict the secretomes of *Candida albicans* (Lee et al. 2003), *Kluyveromyces lactis* (Swaim et al. 2008), *Pichia pastoris* (Mattanovich et al. 2009) and *Trichoderma* species (Druzhinina et al. 2012) was used to analyze the putative protein sequences of the 4776 open reading frames (ORFs) annotated in the *A. gossypii* genome (ftp://ftp.ncbi.nlm.nih.gov/genomes/Fungi/Eremothecium_gossypii_uid10623/, accessed in January 2013). In this protocol, SignalP version 3 (Bendtsen et al. 2004) was used to identify the presence of N-terminal signal peptides and TMHMM version 2 (Krogh et al. 2001) was used to identify transmembrane regions in proteins with putative signal peptides. Only proteins with signal peptides predicted by both SignalP models were included. Proteins with 1 predicted transmembrane spanning region were kept in the dataset if it was located in the N-terminal region before the predicted signal peptide cleavage site. Sequences with more than 1 transmembrane spanning region were excluded. TargetP version 1.1 (Emanuelsson et al. 2007) and the fungal version of big-PI (Eisenhaber et al. 2004) were then used to eliminate proteins predicted to be targeted to the mitochondrion and to contain a GPI anchor, respectively. Finally, WoLF PSORT (Horton et al. 2007) was used for sub-cellular localization

prediction. The default value for the total number or nearest neighbors (*k*) was 27 and only proteins with a $k > 13$ for extracellular location were included in the secretome.

The NetNGlyc version 1.0 server (www.cbs.dtu.dk/services/NetNGlyc/) was used to predict *N*-glycosylation sites and the databases UniProt (www.uniprot.org; UniProt Consortium 2013), CAZy (www.cazy.org; Cantarel et al. 2009) and *MEROPS* (http://merops.sanger.ac.uk; Rawlings et al. 2012) were used to retrieve predicted functions for *A. gossypii* putative proteins.

2.2.3 Gel electrophoresis of secreted proteins

The proteins present in 15 µl of culture supernatant were analyzed by 12% (w/v) sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) followed by silver staining.

For 2-D gel electrophoresis, the total proteins in parental strain culture filtrates collected at the beginning of the stationary phase were precipitated overnight at -20ºC in 10% (w/v) trichloroacetic acid (TCA) and 66.6% (v/v) acetone. The pellet was washed with ice-cold acetone, dried and resuspended (15 min at room temperature) in 2-D sample solution (8 M urea, 10 mM DTT, 2% (v/v) Pharmalyte 3-10 (GE Healthcare) and 2% (v/v) Triton X-100). Insoluble material was removed by centrifugation and the protein concentration in the samples was determined using the 2-D Quant Kit (GE Healthcare). Each culture sample was independently prepared and used for duplicate 2-D electrophoresis analyses.

Equal amounts of total extracellular protein (100 µg) were cup-loaded in Immobiline DryStrip gel strips pH 3-10, 18 cm (GE Healthcare) previously rehydrated and subjected to isoelectric focusing using an Ettan IPGphor II equipment (GE Healthcare) according to the instructions of the manufacturer. The first-dimension isoelectric focusing (IEF) was followed by second-dimension 11% (w/v) SDS-PAGE using an Ettan DALT electrophoresis system (GE Healthcare). After electrophoresis, the gels were fixed for 30 min with 30% (v/v) ethanol and 0.5% (v/v) acetic acid in water and subsequently stained with SYPRO Ruby (Bio-Rad) according to the instructions of the manufacturer. The 2-D gels were scanned in a Typhoon 8610 variable mode imager (GE Healthcare) at 300 dpi resolution and the gel images analysed with the Melanie software version 7.0 (Geneva Bioinformatics (GeneBio) SA). After automatic spot detection, artefacts were manually

removed and the weaker spots (< 0.1% of the whole gel volume) were eliminated. The remaining spots were then linked to allow comparison between samples.

2.2.4 RNA extraction and microarray data acquisition

Total RNA extraction from frozen mycelium was carried out using the RNeasy Plant Mini kit (QIAGEN) according to the manufacturer's instructions for isolation of total RNA from filamentous fungi. RNA concentration and purity were determined using a NanoDrop ND-1000 (NanoDrop Technologies) and integrity of RNA was analyzed using an Agilent 2100 Bioanalyzer (Agilent Technologies).

For microarray analysis, custom-made *A. gossypii* gene expression 12x135K arrays were designed and manufactured by Roche NimbleGen. Each slide contained 12 independent arrays, each comprising four replicates of 33,364 probes covering 4767 Open Reading Frames (ORFs) of *A. gossypii* (7 probes/target ORF). Ten micrograms of total RNA were used for reverse transcription and synthesis of cDNA using the SuperScript II Double-Stranded cDNA Synthesis Kit (Invitrogen) according to the Roche NimbleGen Arrays User's Guide: Gene Expression Arrays v5.0, available from the NimbleGen website (www.nimblegen.com). The cDNA was quantified in a NanoDrop ND-1000 (NanoDrop Technologies) and its integrity analyzed using an Agilent 2100 Bioanalyzer (Agilent Technologies). The double-stranded cDNA was labelled with Cy3 fluorescent dye, hybridized to the custom-made microarray slides (Roche NimbleGen) and scanned using a NimbleGen MS 200 Microarray Scanner (Roche NimbleGen) according to the instructions of the manufacturer.

2.2.5 Microarray data analysis

The raw array data obtained from NimbleScan software version 2.5.26 (Roche NimbleGen) was preprocessed with the Robust Multichip Average (RMA) method (Irizarry et al. 2003). Array data quality was controlled with arrayQualityMetrics (Kauffmann et al. 2009) and sample wise Principal Component Analysis (PCA) of raw, RMA preprocessed and repeat averaged data. Linear Models for Microarrays (LIMMA) (Smyth 2004) was subsequently used to select significantly changing genes with a cut-off of pvalue < 0.01 (which corresponds to a false discovery rate (FDR) of 1% in this analysis) and fold-change > 1.5. For each gene its Pearson correlation with EGI activity

(μmol min^{-1} l^{-1}) was calculated. The FDR of these correlations (4.4% for absolute correlation > 0.7) was estimated from the Q-value (Storey et al. 2004) using the R package 'qvalue'.

Each gene's repeat averaged expression values over all the samples i.e. expression profiles were clustered with R-package 'Mfuzz' (Futschik and Carlisle 2005) with parameter m set to 1.35 and number of clusters to 9. Parameters were selected based on visual evaluation of cluster profiles. Genes with cluster membership > 0.7 were assigned to a co-expression cluster for further analysis.

A. gossypii gene mappings to *S. cerevisiae* genes from the *Ashbya* Genome Database (http://agd.vital-it.ch; Gattiker et al. 2007) were used to map the array results of differential gene expression, gene expression correlation with EGI activity (μmol min^{-1} l^{-1}) and co-expression clusters to *S. cerevisiae*. Gene Ontology (GO) class analyses used *S. cerevisiae* GO annotations. The R-package 'GSA' (Efron and Tibshirani 2007) was used for the GO analysis of differentially expressed genes, the R-package 'GOstats' (Falcon and Gentleman 2007) was used to analyse gene expression correlation and co-expression clusters and REVIGO (Supek et al. 2011) was used to summarize the GO term lists.

Promoter sequences were analysed with the tool Finding Informative Regulatory Elements (FIRE) (Elemento et al. 2007), with default settings using the *A. gossypii* genome. Co-expression clusters and clusters of differentially expressed genes were used as groups of genes from which to find common promoter elements. Genes which were not assigned to any cluster were assigned to group number 0 for this analysis. As promoter we used 1500 bases upstream from the putative ORF of each gene. The identified promoter elements were mapped with FIRE to the known promoter elements described in Gasch et al. (2004) in order to annotate them.

2.3 RESULTS

2.3.1 The *A. gossypii* predicted secretome

Of the 4776 ORFs annotated in the *A. gossypii* genome, 333 (7%) were predicted to contain an N-terminal signal peptide, and thus to enter the general secretory pathway. However, only 54 (1%) of these proteins were predicted to be secreted extracellularly, the others being targeted to different cellular compartments. Of the 54 proteins in the *A. gossypii* predicted secretome, 7 have no homologs in *S. cerevisiae*. 65% were predicted to

contain at least one *N*-glycosylation site (Table 2.1). Enzymes predicted to have hydrolytic activity on poly- and oligosaccharides (carbohydrate active enzymes or CAZymes), proteins, lipids, ester bounds and carbon-nitrogen bounds other than in peptides comprised 33% of the *A. gossypii* predicted secretome (Table 2.1).

A. gossypii ATCC 10895 strain produced 5.7 ± 0.2 g/l dry biomass in defined minimal medium (DMM) and 8.1 ± 0.3 g/L in rich medium (AFM) with sucrose as primary carbon source. At the beginning of the stationary phase the supernatant of the culture growing in minimal medium contained a total protein content of 130 mg/l and that of the culture growing in rich medium contained 218 mg/l.

SDS-PAGE showed 12 distinct protein bands in the culture supernatant of both DMM and AFM cultures, ranging from 7 kDa to 209 kDa (Figure 2.1B). On 2-D gels, 18 protein spots common to both DMM and AFM culture supernatants (from a total of 101 spots commom to both) were visible at higher abundance. Of these at least 2 were obviously isoforms of other protein spots with identical molecular weight (MW) but different isoelectric points (pIs) (Figure 2.1A and C). As shown in Figures 2.1A and C more protein spots were detected in AFM (182) than in DMM (157) culture supernatants. In addition, the distribution of the protein spots on the 2-D gels indicated that most of the *A. gossypii* secreted proteins have a pI between 4 and 6 and a MW above 25 kDa, although some proteins present only in AFM cultures ran at slightly higher pIs (6-8).

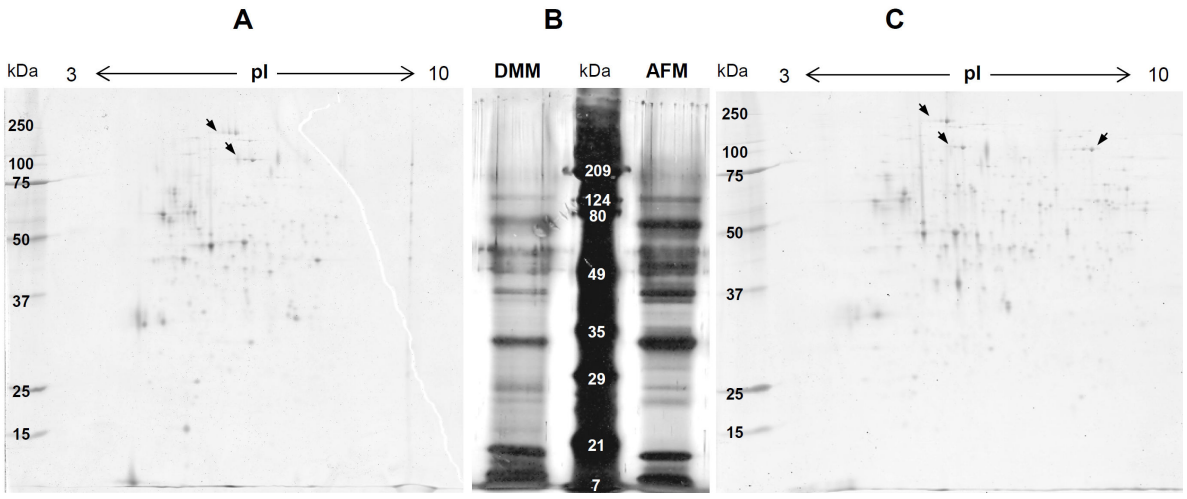


Figure 2.1 – Electrophoretic profiles of the proteins secreted by *A. gossypii* into the culture supernatant. Panels A and C show representative 2-D electrophoresis gels of *A. gossypii* culture supernatants derived from bioreactor batch cultures in minimal (DMM) and rich (AFM) medium, respectively. Pannel B shows the SDS-PAGE gel of the same supernatants. The arrows indicate spots that are most likely diffent isoforms of the same protein.

Table 2.1 – *A. gossypii* predicted secretome. The table shows the *A. gossypii* ORFs encoding the putative proteins predicted to be secreted to the extracellular space and the corresponding *S. cerevisiae* homologs. Predicted functions, number of *N*-glycosylation sites, theoretical isoelectric points (pIs) and molecular weights (MWs) are also indicated.

***A. gossypii* ORF**	***S. cerevisiae* homolog(s)**	**Predicted protein function**	**Theoretical pI / MW (kDa)**	**Predicted *N*-glycosylation sites**
Chromosome I				
AAL179W	*PRY1*	Probable Ca^{2+}-chelating fucntion	39 / 5.0	0
AAL178W	*PRY1*, *PRY2*	Probable Ca^{2+}-chelating fucntion	22 / 8.1	3
AAL092C	*SVS1*, *SRL1*	Cell wall and vacuolar mannoprotein	20 / 4.9	2
AAL056C	*CPR5*, *CPR2*	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	22 / 4.8	2
AAR129C	*YDR415C*	Metallo protease M28 (Aminopeptidase)	41 / 5.6	4
AAR163C	*MF(ALPHA)*1, *MF(ALPHA)2*	Mating pheromone α-factor	10 / 6.0	1
Chromosome II				
ABL189W	*YDL237W*	Altered inheritance of mitochondria protein 6 (AIM6); probable phosphoric diester hydrolase activity	44 / 4.9	2
ABL123C	*PEP4*	Aspartyl protease A1 (CtsD peptidase (Aspergillus-type))	54 / 6.7	6
ABR026C	*CWP1*	Cell wall mannoprotein	16 / 4.5	0
Chromosome III				
ACL202W	*DFG5*	α-1,6-mannanase	51 / 6.5	7
ACR143W	*PEP4*	Aspartyl protease A1	43 / 4.9	0
ACR144W	*PEP4*	Aspartyl protease A1 (Saccharopepsin)	44 / 4.7	1
ACR272C	*CWP1*	Cell wall mannoprotein	15 / 4.4	1
ACR273W	*CWP1*	Cell wall mannoprotein	16 / 6.1	0
Chromosome IV				
ADL398C	*CWP1*	Cell wall mannoprotein	16 / 4.8	0
ADL397C-B	No homolog	FMN-binding protein with probable oxidoreductase activity	22 / 7.4	1
ADL369C	*MRL1*	Mannose 6-phosphate receptor-like protein 1	30 / 9.9	2
ADR020C	*MNT3*	α-mannosyltransferase	67 / 10.0	1
ADR056W	*TOS1*	Cell wall protein of unknown function	53 / 6.7	1
ADR123W	*APE3*	Metallo protease M28 (Aminopeptidase Y)	57 / 4.4	5
ADR238C	No homolog	Acid phosphatase	54 / 4.9	8
ADR322W	*UTH1*, *NCA3*	Member of the SUN family	53 / 4.2	6
Chromosome V				
AEL312C	*YMR244W*	Member of the SUN family	34 / 5.3	2
AEL197C	*MID1*	Membrane protein required for Ca^{2+} influx stimulated by pheromone	59 / 4.9	9
AEL134W	*KRE9*	Glycoprotein involved in cell wall β-glucan assembly	32 / 8.2	2
AEL110W	*PIR1, HSP150*	Cell wall glycoprotein	26 / 4.2	0
AEL103W	*CIS3*, *PIR1*, *PIR3*	Cell wall glycoprotein	20 / 4.6	1
AEL040W	*YCR015C*	Haloacid dehydrogenase (HAD)-superfamily hydrolase	17 / 9.2	0
AER022W	*YBR139W*	Serine carboxypeptidase Y (S10 family)	59 / 6.4	3

Table 2.1 – *Continued.*

A. gossypii ORF	*S. cerevisiae* homolog(s)	Predicted protein function	Theoretical pI / MW (kDa)	Predicted *N*-glycosylation sites
Chromosome V				
AER039W	*MPD2*	Member of the protein disulfide isomerase (PDI) family	28 / 5.0	0
AER372C	*SUN4*, *SIM1*	Member of the SUN family	46 / 6.6	0
AER428W	*OM45*	Unknown	54 / 7.3	3
AER454C	*YJR107W*	Triglyceride lipase	36 / 6.4	1
Chromosome VI				
AFL173C	No homolog	Unknown	39 / 8.3	1
AFL062W	*MF(ALPHA)*1, *MF(ALPHA)2*	Mating pheromone α-factor	13 / 5.2	1
AFR029C	*DSE2*	Daughter cell-specific secreted protein with similarity to glucanases	16 / 5.0	1
AFR038W	*YHR138C*	Unknown	10 / 5.8	0
AFR048W	No homolog	Ureohydrolase	42 / 6.4	1
AFR298C	*YIL156W-B*	Unknown	8 / 8.2	0
AFR324W	*VPS62*, *TDA6*	Vacuolar protein sorting (VPS) protein	52 / 6.7	0
AFR367W	*FRD1*, *OSM1*	Succinate dehydrogenase/fumarate reductase (NADH) flavoprotein	54 / 9.2	0
AFR473W	No homolog	Unknown	89 / 4.6	3
AFR476C	*AGA2*	Adhesion subunit of α-agglutinin	27 / 7.3	1
AFR529W	*SUC2*	Invertase	65 / 4.6	7
AFR755W	No homolog	FMN-binding protein with probable oxidoreductase activity	22 / 7.4	1
AFR756W	*CWP1*	Cell wall mannoprotein	16 / 4.8	0
Chromosome VII				
AGL354C	*SCW4*, *SCW10*	Glucan 1,3-β-glucosidase	47 / 6.2	0
AGL343C	*BGL2*	Glucan 1,3-β-glucosidase	34 / 4.5	0
AGL326W	*CPS1*	Metallo carboxypeptidase S (M20 family)	63 / 4.7	4
AGR072W	*LCL2*	Long chronological lifspan protein 2	14 / 7.7	0
AGR178W	*GTB1*	Glucosidase II β subunit	75 / 5.2	2
AGR225W-A	No homolog	Unknown	10 / 4.4	0
AGR240W	*BAR1*	Aspartyl protease A1 (Pepsin A)	48 / 4.7	3
AGR407C	*PEP4*	Aspartyl protease A1 (Pepsin A)	42 / 5.7	4

Comparing the number of protein spots observed in the 2-D gels with the *A. gossypii* predicted secretome, slightly more spots were observed than putative proteins predicted to be secreted through the plasma membrane. The existence of different isoforms of the same proteins (e.g. different glycoforms) may have contributed to this observation. Additionally, some proteins may be secreted via an alternative secretion pathway and, therefore, failed to be predicted as secreted with the computational methodology used. The possible contribution of intracellular proteins for some of the weakest spots detected cannot also be ruled out. Nevertheless, the results from both experimental and computational analyses indicated that 1-4% of the *A. gossypii* proteins are secreted.

2.3.2 Effect of DTT on *A. gossypii* growth

The ability of *A. gossypii* cells to cope with secretion stress was assessed by monitoring cell growth in the presence and absence of a well known secretion stress inducer, DTT, which inhibits the formation of disulfide bonds, thus interfering with protein folding in the ER. As previously reported (Ribeiro et al. 2010), the *A. gossypii* EGI producing strain and corresponding empty vector control strain exhibited similar growth rates during bioreactor batch cultures in selective rich medium (cf. Figure 2.2 and Ribeiro et al. 2010). The addition of 10 mM final concentration DTT to cultures after 9.5 h caused, however, a substantial and immediate reduction in the specific growth rate of the *A. gossypii* EGI producing cells (Figure 2.2). When the same concentration of DTT was added after only 6 h, the *A. gossypii* cells immediately stopped growing (data not shown).

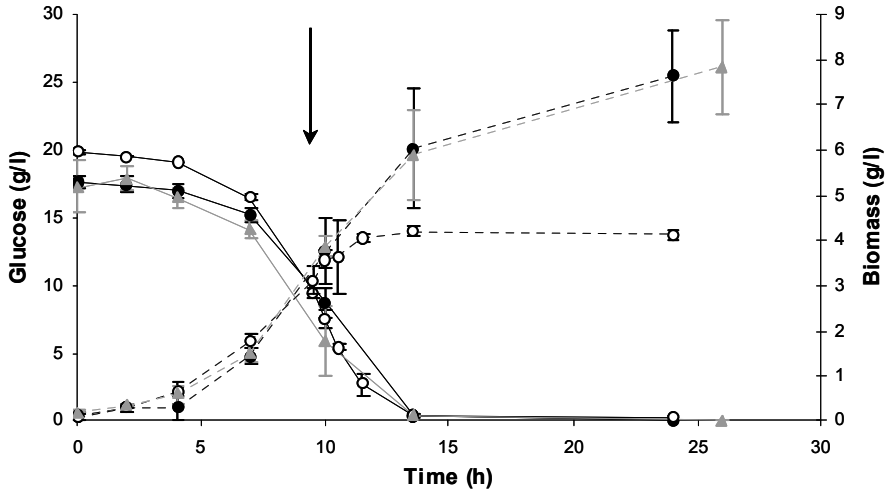


Figure 2.2 – Glucose consumption (*solid lines*) and growth (*dashed lines*) curves of recombinant *A. gossypii* in batch cultures with selective AFM at pH 6.0, 30ºC and 500 rpm, with 1.0 vvm aeration. (●) EGI producing strain, (▲) empty vector strain and (○) DTT-stressed cells. Values represent the average ± standard deviation of two independent bioreactor cultures. The arrow indicates the time when DTT was added.

2.3.3 Effect of recombinant protein production on the *A. gossypii* transcriptome

The overall variation in the gene expression profiles between *A. gossypii* cells secreting recombinant EGI and those which did not was surprisingly small, such that significant differences between the profiles were only detected when the microarray data analysis took into account the correlation between expression of each gene and EGI production. As the correlation analysis included the variation in replicates of single

conditions (unlike LIMMA), it was thus expected to be more sensitive than LIMMA. Only 21 genes were found to be differentially expressed (FDR of 4.4%) using this approach, of which 16 were up-regulated and 5 down-regulated in the strain producing recombinant protein (Table 2.2). GO enrichment analyses for this set of genes hinted at translation down-regulation and ion and amino acid transmembrane transport up-regulation having occurred during EGI production.

Table 2.2 – Genes differentially regulated in *A. gossypii* under recombinant protein production conditions (FDR of 4.4%). The corresponding *S. cerevisiae* homologs are indicated, as well as predicted functions. The biological processes enriched ($p < 0.01$) in the up- and down-regulated gene clusters are also indicated.

	***A. gossypii* gene**	***S. cerevisiae* homolog(s)**	**Predicted protein function**	**Biological processes**
Up-regulated	*AAR030W*	*CTR1*	High-affinity copper transporter	
	AAR080W	No homolog	Unknown	
	ADL123C	*PHO4*	Transcription factor that activates transcription cooperatively with Pho2p in response to phosphate limitation	
	ADL153W	*RRI2*	Subunit of the COP9 signalosome complex	
	ADR080W	*FRE1*	Ferric reductase and cupric reductase	
	AEL294C	*FTR1*	High-affinity iron permease	
	AER428W	*OM45*	Major constituent of the mitochondrial outer membrane	Amino acid transport
	AFL135W	*YMR181C, YPL229W*	Unknown	Iron transmembrane transport
	AFR156W	*PUT4*	High-affinity proline permease	Transmembrane transport
	AFR442C	*PHO84*	High-affinity inorganic phosphate transporter and low-affinity manganese transporter	Ion transport
	AFR529W	*SUC2*	Invertase	
	AFR595W	*MCH1*	Protein with similarity to mammalian monocarboxylate permeases	
	AFR668W	*CAN1, ALP1*	Plasma membrane arginine permease	
	AFR739C	No homolog	Unknown	
	AGL097C	*ENA2, ENA5, ENA1*	P-type ATPase sodium pump, involved in Na^+ and Li^+ efflux to allow salt tolerance	
	AGR304W	*MTH1, STD1*	Protein involved in the control of glucose-regulated gene expression	
Down-regulated	*ABL065W*	*RPG1*	Subunit of the core complex of translation initiation factor 3 (eIF3)	Regulation of translation
	ABL174C	*SSB2, SSB1*	Cytoplasmic ATPase that is a ribosome-associated molecular chaperone; may be involved in the folding of newly-synthesized polypeptide chains; member of the HSP70 family	Posttranscriptional regulation of gene expression
	AEL032W	*GCN20*	Positive regulator of the Gcn2 kinase activity	Regulation of cellular protein metabolic process
	AER366W	*FLX1*	Protein required for transport of FAD across the mitochondrial membrane	
	AGR261W	*RPS28B, RPS28A*	Protein component of the small ribosomal subunit	Regulation of translational elongation

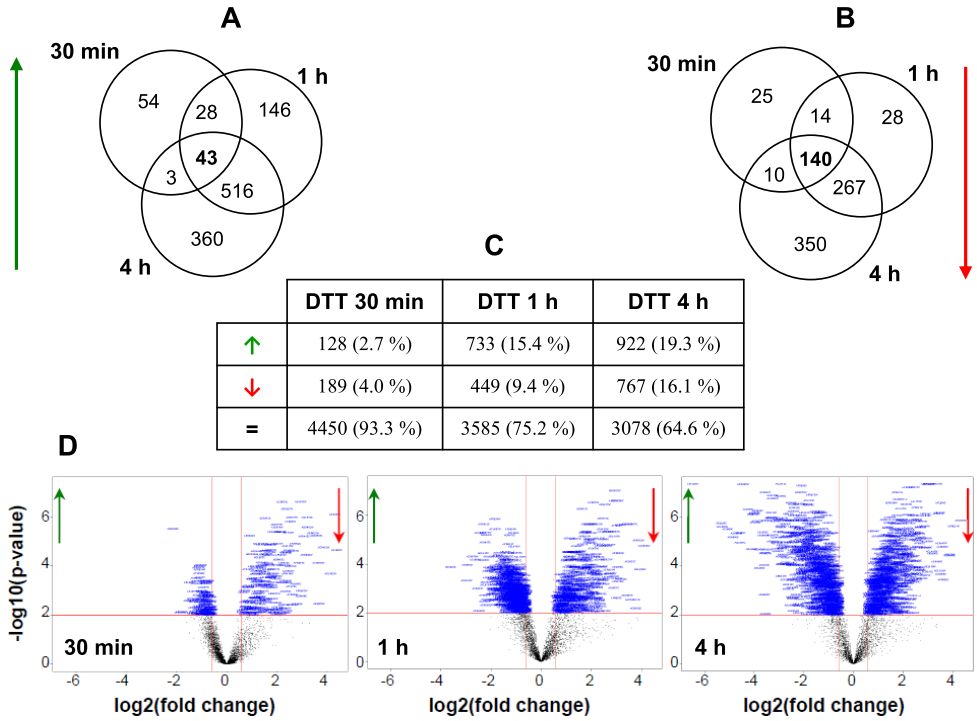
Analysis of the *EGI* expression levels in the EGI producing cells revealed that they were within the average of overall gene expression and around 5 fold higher than the background of this gene in the non-producing cells, which contained the empty vector. However, they were approximately 14 fold lower than the expression levels of *TEF*, a gene that is reported to be highly expressed in *A. gossypii* (Steiner and Phillippsen 1994). Therefore, *EGI* was not being highly overexpressed in the EGI producing cells.

2.3.4 Effect of DTT-induced stress on the *A. gossypii* transcriptome

Upon addition of DTT into the cultures of *A. gossypii* EGI producing cells during logarithmic growth, the gene expression profile changed significantly, has had the growth rate (Figure 2.3). When comparing the transcript levels of all genes at the time immediately before the addition of DTT (9.5 h after inoculation and defined as time zero for DTT addition) with those at 30 min, 1 h and 4 h after DTT addition, 128 genes were already up-regulated and 189 down-regulated after 30 min of exposure to DTT (Figure 2.3C). The up-regulation of 43 of these genes was sustained up to 4 h of treatment (Figure 2.3A), and the same was observed for 140 of the 189 down-regulated genes (Figure 2.3B). As can be seen from Figure 2.3C, DTT induced more genes than it repressed. However, down-regulation was greater than up-regulation after 30 min and 1 h of DTT treatment, as higher fold changes were observed in the transcript levels of down-regulated than up-regulated genes (Figure 2.3D). 32% of the genes whose transcription was significantly decreased 30 min after DTT addition (p-value < 0.01 for a fold-change > log2(1.5)) were predicted to encode secretory proteins. This decreased to 27% after 1 h and to 19% after 4 h of treatment. Less than 5% of the significantly up-regulated genes were predicted to encode secretory proteins.

DTT caused wide-ranging effects on the *A. gossypii* gene expression profiles (Tables 2.3, 2.4 and 2.5). Some of the changes probably correlate with the observed reduction in growth rate (Figure 2.2). For instance, within 30 min, the transcription of genes involved in filamentous growth (Table 2.4), glycosylation, lipoprotein (Cluster 1 in Figure 2.4, Tables 2.3 and 2.4) and cell wall biosynthesis were already down-regulated by DTT. Analysis of the co-expression clusters also revealed that genes encoding ribosomal proteins were more down-regulated in the DTT-treated cells than in the control, particularly after 1 h of exposure to DTT (Cluster 1 in Figure 2.4, Table 2.3). A

down-regulation of ribosomal protein-encoding genes has been correlated with reductions in the growth rate (Hayes et al. 2002; Lim et al. 2003).



	DTT 30 min	DTT 1 h	DTT 4 h
↑	128 (2.7 %)	733 (15.4 %)	922 (19.3 %)
↓	189 (4.0 %)	449 (9.4 %)	767 (16.1 %)
=	4450 (93.3 %)	3585 (75.2 %)	3078 (64.6 %)

Figure 2.3 – Overall variations in the *A. gossypii* transcriptome after addition of DTT. The data refers to the outcome of the LIMMA analysis for differentially expressed genes within 30 min, 1 h and 4 h of exposure to DTT, when compared to the time immediately before DTT addition (fold change > log2(1.5) and p-value < 0.01). The Venn diagrams indicate the number of genes that had increased (A) and reduced (B) expression after 30 min, 1 h and 4 h of exposure to DTT. In table C the absolute number and relative percentage (*between brackets*; in relation to the total number of *A. gossypii* genes) of genes transcriptionally up-regulated (↑), down-regluated (↓) and unchanged (=) by DTT after different exposure times are indicated. The Volcano plots (D) obtained from the LIMMA analysis show the overall significant fold changes in the *A. gossypii* gene expression profiles after 30 min, 1 h and 4 h of DTT adition. Negative fold changes represent genes with increased (↑) expression and positive fold changes genes with reduced (↓) expression after exposure to DTT.

Genes involved in the response to stress, transcription, protein unfolding, proteasome assembly, proteolysis, vesicle trafficking, vacuolar protein sorting, secretion, trehalose biosynthesis and DNA repair were induced by DTT (Clusters 2, 7 and 9 in Figure 2.4, Tables 2.3 and 2.5), hinting at the activation of a secretion stress response by DTT. However, contrary to what has been commonly observed in several yeast and filamentous fungi (Travers et al. 2000; Sims et al. 2005; Arvas et al. 2006; Guillemette et al. 2007; Graf et al. 2008), the expression levels of classical UPR targets such as *IRE1*, *KAR2*, *HAC1*,

PDI1 and *EUG1* homologs were not altered in *A. gossypii* cells treated with 10 mM DTT. Moreover, the *A. gossypii ERO1* and *LHS1* homologs, two other classical UPR targets whose transcription is usually induced by secretion stress (Travers et al. 2000; Guillemette et al. 2007; Graf et al. 2008), were in *A. gossypii* transcriptionally down-regulated (Table 2.6), as was the *GCN4* homolog, which has been previously shown to be induced by ER stress (Patil et al. 2004; Arvas et al. 2006).

Treatment of *A. gossypii* cells with DTT led to a rapid and major repression of the protein glycosylation pathway (in particular of the *N*-glycosylation pathway) (Tables 2.3, 2.4 and 2.6). The observed repression was not only at the ER, but also at the Golgi processing level (Table 2.6), indicating that a major accumulation of unglycosylated proteins may have occurred. In a similar way, the *A. gossypii YOS9* and *HTM1*/*MNL1* homologs, which in yeast encode two proteins that are required for the ERAD of misfolded glycoproteins (Kim et al. 2005), were also repressed by DTT. Several genes involved in the COPI retrograde transport of proteins from the Golgi back to the ER were induced by DTT within 1 h of exposure (Tables 2.6), which might have favoured the recycling of proteins. Some genes involved in the ER to Golgi protein trafficking were also induced by DTT, but the COPII vesicle-mediated transport was more strongly repressed by DTT (Tables 2.6).

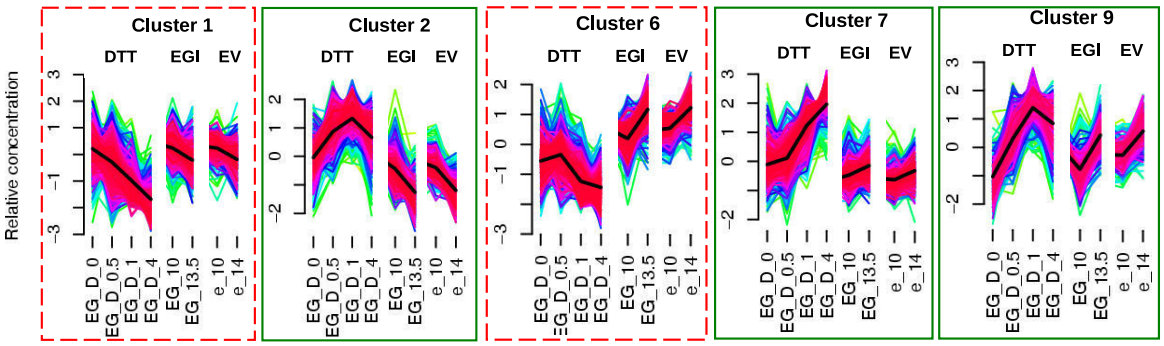


Figure 2.4 – Overall variation in the transcript levels of the genes included in different co-expression clusters after addition of DTT to the cultures of recombinant *A. gossypii* EGI producing strain. DTT was added after 9.5 h of culture (EG_D_0) and expression analyses were performed after 30 min (EG_D_0.5), 1 h (EG_D_1) and 4 h (EG_D_4) of exposure. The image also includes information for the gene expression variation in the non-treated EGI producing strain (EGI) and corresponding empty vector strain (EV). EG_D_0.5 (10 h of culture) can be compared with EG_10 and e_10, and EG_D_4 (13.5 h of culture) can be compared with EG_13.5 and e_14. Dashed squares indicate gene clusters down-regulated by DTT and the other represented clusters were up-regulated by DTT.

Table 2.3 – Biological processes enriched ($p < 0.01$) in the co-expression clusters whose transcription varied differently with culture time in the *A. gossypii* cells treated with DTT and in the non-treated cells. Clusters 1 and 6 were down-regulated by DTT, whereas clusters 2, 7 and 9 were up-regulated by DTT (Figure 2.4).

Cluster 1	Cluster 6
Ion transport	Reproduction
Ribosomal large subunit biogenesis	Response to pheromone
Glycosylation	Ascospore wall assembly
Glycoprotein metabolic process	Mrna splicing, via spliceosome
Macromolecule modification	Conjugation
Cellular iron ion homeostasis	M phase
Attachment of GPI anchor to protein	Peptide transport
Lipoprotein metabolic process	Dna recombination
Translational elongation	External encapsulating structure organization
Sulfur amino acid transport	Regulation of microtubule polymerization or depolymerization
Peptidyl-diphthamide biosynthetic process from peptidyl-histidine	Dna metabolic process
	Cell wall assembly
	RNA splicing
Cluster 2	**Cluster 7**
Reproduction	Endocytosis
Organelle organization	Response to biotic stimulus
Response to pheromone	Positive regulation of homeostatic process
Growth	Proteasome assembly
Multi-organism process	Purine ribonucleoside catabolic process
Biological regulation	Cell division
Transcription elongation from RNA polymerase II promoter	NAD biosynthesis via nicotinamide riboside salvage pathway
Conjugation	Amide biosynthetic process
Gene expression	Response to singlet oxygen
Regulation of transcription during mitosis	Membrane invagination
Mitotic cell cycle	Response to osmotic stress
Actin filament-based process	Response to abiotic stimulus
Isoleucyl-trna aminoacylation	
Regulation of protein catabolic process	**Cluster 9**
Protein localization to organelle	Developmental process involved in reproduction
Chromosome segregation	Vacuolar transport
Small gtpase mediated signal transduction	Negative regulation of biological process
Cellular localization	Response to stimulus
Regulation of localization	Autophagy
Cell cycle	Dna-dependent transcription initiation
DNA-dependent transcription elongation	Cellular membrane fusion
Proteolysis	Vacuolar protein processing
Transcription from RNA polymerase II promoter	Post-translational protein modification
Macromolecule localization	Cellular response to stress
Cellular macromolecule biosynthetic process	Transcription initiation from rna polymerase iii promoter
Regulation of biological process	Macromolecule localization
Nucleus organization	Cytokinesis
Regulation of cell size	Dna metabolic process
Transmembrane transport	Vesicle-mediated transport
Mrna-binding (hnrnp) protein import into nucleus	Negative regulation of metabolic process
Nuclear pore organization	Vacuole organization
Nucleosome disassembly	Cell communication
	Meiotic mismatch repair
	Response to extracellular stimulus

Table 2.4 – Biological processes enriched ($p < 0.001$) in the gene clusters significantly down-regulated (fold change > log2(1.5) and p-value < 0.01) after 30 min, 1 h and 4 h of DTT treatment, in comparison with the time immediately before addition of DTT. GO terms highlighted in bold were overrepresented at two (**grey**) or all (**black**) DTT exposure times.

30 min	1 h	4 h
Thiamine transport	'De novo' IMP biosynthetic process	Adenine salvage
Filamentous growth	**Cadmium ion transport**	**Cadmium ion transport**
Response to copper ion	**Response to copper ion**	**Response to copper ion**
Phytochelatin biosynthetic process	Cell-cell adhesion	Organophosphate metabolic process
Lipid metabolic process	Iron assimilation	Polyphosphate metabolic process
Barrier septum assembly	Polyphosphate metabolic process	Cellular oligosaccharide metabolic process
Positive regulation of catabolic process	***N*-glycan processing**	Group II intron splicing
Carbohydrate metabolic process	Peptide biosynthetic process	Triglyceride biosynthetic process
Hydrogen peroxide metabolic process	Glutamine metabolic process	Mannoprotein biosynthetic process
Glycoprotein metabolic process	**Acylglycerol biosynthetic process**	**Thiamine transport**
Alditol biosynthetic process	**Organic alcohol transport**	Cellular biosynthetic process
Purine ribonucleoside monophosphate metabolic process	**Thiamine transport**	Regulation of translational fidelity
Acylglycerol biosynthetic process	Endonucleolytic cleavage to generate mature 3'-end of ssu-rRNA from (ssu-rRNA, 5.8s rRNA, lsu-rRNA)	**Vitamin transport**
Organic alcohol transport	**Lipid storage**	Glycoprotein metabolic process
Nucleoside transport	**Organic ether metabolic process**	**Lipid storage**
Lipid storage	**Vitamin transport**	Septin checkpoint
Cadmium ion transport	**Nucleoside monophosphate metabolic process**	Protein metabolic process
Alcohol metabolic process	Ctp metabolic process	Box c/d snoRNA metabolic process
Vitamin transport	**Nucleobase-containing compound biosynthetic process**	Nucleoside metabolic process
Nucleoside monophosphate metabolic process	Glycerol ether metabolic process	**Nucleoside monophosphate metabolic process**
Glycosylation		**Glycosylation**
Organic ether metabolic process		Regulation of mating-type specific transcription, DNA-dependent
Lipoprotein biosynthetic process		Cell wall glycoprotein biosynthetic process
Purine base biosynthetic process		Purine nucleoside metabolic process
Lipoprotein metabolic process		
Deadenylation-dependent decapping of nuclear-transcribed mRNA		
Nucleobase metabolic process		
Glycerol ether metabolic process		
Cellular carbohydrate metabolic process		
***N*-glycan processing**		
Coenzyme a biosynthetic process		
Nucleobase-containing compound biosynthetic process		

Translation was also down-regulated by DTT, but only after 4 h of treatment (Cluster 1 in Figure 2.4, Tables 2.3 and 2.4). DTT also repressed the transcription of the *A. gossypii SSH1* homolog, which in *S. cerevisiae* is involved in the co-translational translocation of proteins into the ER (Robb and Brown 2001).

On the other hand, genes involved in mRNA degradation, such as *DOM34*, *KEM1*, *SKI2*, *SKI3*, *SKI7*, *LSM2* and *NMD2* homologs, were up-regulated by DTT within 1 h of treatment.

Table 2.5 – Biological processes enriched ($p < 0.001$) in the gene clusters significantly up-regulated (fold change > log2(1.5) and p-value < 0.01) after 30 min, 1 h and 4 h of DTT treatment, in comparison with the time immediately before addition of DTT. GO terms highlighted in bold were overrepresented at two (**grey**) or all (**black**) DTT exposure times.

30 min	1 h	4 h
Spermine biosynthetic process	**Protein unfolding**	DNA dealkylation involved in dna repair
Interspecies interaction between organisms	Bipolar cellular bud site selection	Carbon utilization
S-adenosylmethionine transport	**Early endosome to golgi transport**	Asexual reproduction
Traversing start control point of mitotic cell cycle	Membrane docking	Response to stimulus
Macromolecule metabolic process	**Trehalose biosynthetic process**	**Negative regulation of transferase activity**
Glucose 1-phosphate metabolic process	Phosphorus metabolic process	**Early endosome to golgi transport**
Cis assembly of pre-catalytic spliceosome	Regulation of DNA repair	Cellular aldehyde metabolic process
Macromolecule catabolic process	Fructose transport	NAD biosynthesis via nicotinamide riboside salvage pathway
Protein unfolding	Asymmetric protein localization	**Trehalose biosynthetic process**
Golgi localization	Vesicle-mediated transport	**Protein unfolding**
Leading strand elongation	mRNA polyadenylation	Nuclear mrna 5'-splice site recognition
Cellular macromolecule metabolic process	Cellular localization	Asymmetric protein localization
Negative regulation of transcription from rna polymerase II promoter during mitosis	Cell wall macromolecule catabolic process	**Secretion**
Proteolysis	**Secretion**	Sodium ion transport
Transcription initiation, DNA-dependent	Macromolecule localization	**Negative regulation of developmental process**
Stress-activated protein kinase signaling cascade	Cofactor transport	**Oligosaccharide metabolic process**
Negative regulation of transferase activity	**Oligosaccharide metabolic process**	Tetrapyrrole catabolic process
Trehalose biosynthetic process	**Macromolecule catabolic process**	Intron homing
	Positive regulation of lipid metabolic process	**Negative regulation of catalytic activity**
	Transcription from RNA polymerase II promoter	**Regulation of hydrolase activity**
	Negative regulation of catalytic activity	**Regulation of transferase activity**
	Regulation of hydrolase activity	Glycerophospholipid catabolic process
	Regulation of transferase activity	
	Proteolysis	
	Regulation of response to stimulus	
	Negative regulation of developmental process	

The expression of several genes involved in protein folding was significantly induced by DTT after 4 h of exposure, including *JEM1*, *SIL1*, *SSA2*, *STI1*, *SIS1*, *FES1*, *HSP104*, *HSC82*, *HSP82*, *AHA1*, *MDJ1*, *HSP78*, *APJ1* and *HSP26* homologs (Table 2.6). Some genes involved in the ERAD (*HRD1*, *USA1*, *UBX2* and *HLJ1* homologs) and proteasome degradation (*DOA4*, *UBP5*, *RPT3*, *UBA1*, *UBP2*, *CDC48*, *UFD1*, *DOA1* and *SHP1* homologs) were also significantly induced by DTT within 1 h of treatment, suggesting that *A. gossypii* cells were targeting proteins for proteolytic degradation. Up-regulation of vacuolar protein sorting (Cluster 9 in Figure 2.4, Tables 2.3, 2.5 and 2.6) and vesicle trafficking was also observed in the DTT-treated cells (Tables 2.5 and 2.6).

A search for common regulatory DNA motifs in the promoter region of the DTT-regulated genes identified only seven motifs common within one or more gene

clusters (Figures 2.5), none of which matched known consensus binding sites for the transcriptional factors Hac1p or Gcn4p. Among the common promoter elements found, four were similar to known binding sequences for the *S. cerevisiae* transcription factors Rap1p, Adr1p and Hcm1p. The binding site for Rap1p, a positive transcriptional regulator for multiple growth related genes such as ribosomal protein genes (Lieb et al. 2001), was over-represented in the DTT-induced gene cluster (Figure 2.5B). The consensus sequence for Adr1p binding was the only motif over-represented in the *A. gossypii* DTT down-regulated genes. Adr1p is a carbon source responsive zinc-finger transcription factor that is required for transcription of the glucose-repressed gene *ADH2*, of peroxisomal protein genes and of genes required for ethanol, glycerol, and fatty acid utilization (Denis and Young 1983). Another motif had similaity to a Ribosomal RNA Processing Element (RRPE; AAAAATTT), to which the *S. cerevisiae* Stb3p has been demonstrated to bind (Liko et al. 2007) and that was identified by Gasch et al. (2000) as being a common element found in the promoter region of several genes repressed during Environmental Stress Response (rESR). This motif was over-represented in the cluster of *A. gossypii* genes induced by DTT (Figure 2.5B) and also in the co-expression clusters 3 and 4 (Figure 2.5A), which are enriched in various functions associated with the RNA metabolism.

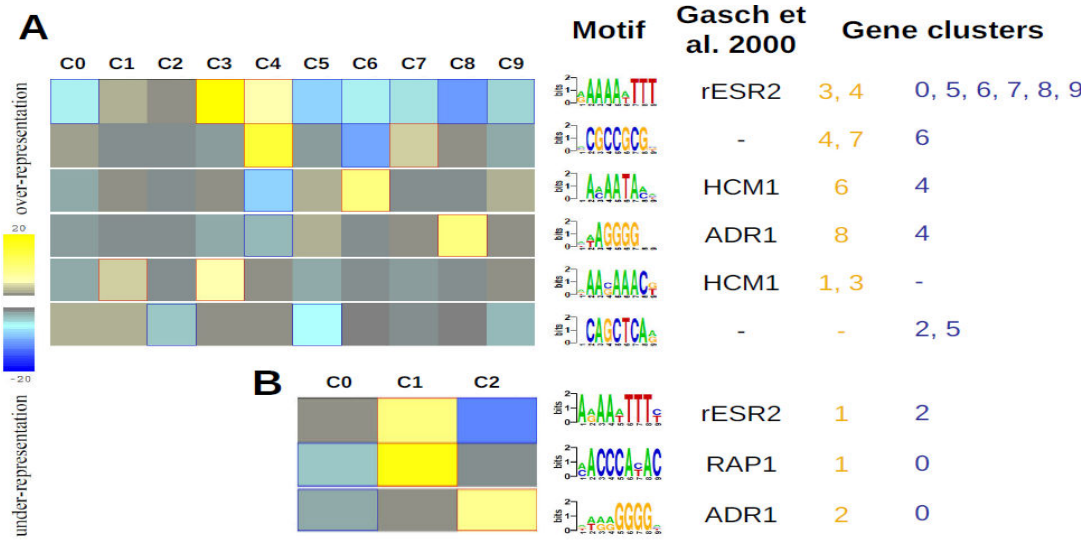


Figure 2.5 – Regulatory DNA elements significantly over- (*yellow*) and under-represented (*blue*) ($p<0.05$) in different gene clusters. In panel A are represented the co-expression clusters and in panel B the clusters of genes differentially up- (C1) and down-regulated (C2) by DTT. C0 comprises the genes that were not included in any of the other clusters.

Table 2.6 – Secretion-related *A. gossypii* genes with significant fold changes ([a]p-value < 0.01) in their transcript levels after treatment with DTT for 30 min, 1 h and 4 h. The corresponding *S. cerevisiae* homologs are indicated, as well as predicted functions. Green corresponds to the up-regulated genes and red to the down-regulated genes.

Predicted function	*A. gossypii* gene	*S. cerevisiae* homolog(s)	Fold Change[a]		
			30 min	1 h	4 h
Translocation					
Subunit of the Ssh1p translocon complex	*AFR613C*	*SSH1*	-2,28	-2,02	-2,61
Protein folding					
DnaJ-like chaperone, localizes to the ER membrane; enhancer of the ATPase activity of Kar2p	*ADR124C*	*JEM1*		1,90	2,12
Molecular chaperone of the ER lumen involved in polypeptide translocation and folding	*ADR128C*	*LHS1*	-0,43		-1,65
Thiol oxidase required for oxidative protein folding in the ER	*ADL348W*	*ERO1*		-1,58	
Member of a transmembrane complex required for efficient folding of proteins in the ER	*AFR720W*	*ECM1*	-3,31	-2,28	-2,25
Member of a transmembrane complex required for efficient folding of proteins in the ER	*AFR429C*	*ECM3*		-1,67	-1,81
Member of a transmembrane complex required for efficient folding of proteins in the ER	*ADL092W*	*EMC5*		-2,60	-2,70
Putative protein with a suggest a role in folding of ER membrane proteins	*AFR176W*	*YER140W*	-2,20	-1,77	-2,48
Nucleotide exchange factor for the ER lumenal Hsp70 chaperone Kar2p	*AEL340W*	*SIL1*		1,56	
ATP binding protein, member of the HSP70 family	*AFR114W*	*SSA2*			2,02
Nucleotide exchange factor for the Hsp70 Ssa1p, cytosolic homolog of Sil1p	*ADL319W*	*FES1*		2,15	3,06
Heat shock protein that cooperates with Ydj1p (Hsp40) and Ssa1p (Hsp70) in protein refolding	*AGL036C*	*HSP104*		2,21	3,89
Hsp90 co-chaperone, interacts with the Ssa group of the cytosolic Hsp70 chaperones	*AEL224W*	*STI1*			2,28
Cytoplasmic chaperone of the Hsp90 family	*AFL148C*	*HSC82, HSP82*			2,12
Co-chaperone that binds to Hsp82p and activates its ATPase activity	*AGL090W*	*AHA1*			2,79
Type II HSP40 co-chaperone that interacts with the HSP70 Ssa1p	*AAL008W*	*SIS1*		2,80	3,68
Type I HSP40 co-chaperone for the HSP70 Ssc1p in the mitochodrial matrix	*AFR507W*	*MDJ1*	1,50		2,38
Mitochondrial chaperone that cooperates with Ssc1p, prevents misfolded proteins aggregation	*AEL223C*	*HSP78*		4,40	8,23
Putative chaperone of the HSP40 (DNAJ) family	*AER427W*	*APJ1*			2,82
Small heat shock protein with chaperone activity, suppresses unfolded proteins aggregation	*ADL397C*	*HSP26*			4,80
Glycosylation/modification					
ER *N*/*O*-linked glycosylation					
Dolichyl-phosphoglucose-dependent α1,2 glucosyltransferase	*ADL138C*	*DIE2*		2,10	
UDP-N-acetyl-glucosamine-1-P transferase	*AFL037W*	*ALG7*	-3,65	-3,30	-3,50
Mannosyltransferase catalyzes two consecutive steps in the *N*-linked glycosylation pathway	*AFL098W*	*ALG2*	2,14	2,43	
Dolichol-P-Man dependent α 1,3 mannosyltransferase	*AGL299C*	*ALG3*	-2,27	-3,16	-5,24
Mannosyltransferase involved in N-linked glycosylation	*ABL185C*	*ALG9*	-2,61	-1,75	-2,05
α1,3-glucosyltransferase involved in *N*-linked protein glycosylation in the ER	*ACR004W*	*ALG6*		1,87	2,04
Dolichyl pyrophosphate (Dol-P-P) phosphatase	*ACL186W*	*CAX4 (CWH8)*		-1,59	-1,96
Dolichol kinase required for protein *N*-glycosylation	*ABR051C*	*SEC59*	-2,05	-2,46	-5,40
Gamma subunit of the oligosaccharyltransferase complex of the ER lumen	*AER413C*	*OST3*	-2,89	-3,11	-3,40
Subunit of the oligosaccharyltransferase complex of the ER lumen	*AFR491W*	*STT3*	-5,02	-5,10	-6,60
α-glucosidase I, role in assembly of cell wall β1,6-glucan and *N*-linked protein glycosylation	*AFR483C*	*CWH41*	-2,43		

Table 2.6 – *Continued.*

Predicted function	*A. gossypii* gene	*S. cerevisiae* homolog(s)	Fold Change[a]		
			30 min	1 h	4 h
Glycosylation/modification					
ER *N*/*O*-linked glycosylation					
Glucosidase II α subunit, involved in *N*-linked protein glycosylation in the ER	*AAR173C*	*ROT2*	-3,11	-2,13	-2,31
Glucosidase II β subunit, involved in *N*-linked protein glycosylation in the ER	*AGR178W*	*GTB1*	-3,35	-2,48	-2,47
Protein *O*-mannosyltransferase	*ABL085W*	*PMT4*		1,74	1,94
ER-localized and oxidants sensitive protein involved in *N*-glycosylation	*AER423C*	*EOS1*		-1,63	-2,12
Golgi *N*/*O*-linked glycosylation					
Guanosine diphosphatase involved in the transport of GDP-mannose into the Golgi lumen	*AFR362C*	*GDA1*	-2,59	-1,90	-1,86
Subunit of the Golgi mannosyltransferase complex	*AGL259C*	*MNN9*	-5,05	-3,53	
α1,6-mannosyltransferase of the Golgi mannosyltransferase complex	*AFR454W*	*MNN10*	-5,31	-5,33	-5,35
α1,6-mannosyltransferase of the Golgi mannosyltransferase complex	*AEL142W*	*MNN11*	-2,49	-2,21	-3,82
α1,2-mannosyltransferase, responsible for addition of the first α1,2-linked mannose	*AER018C*	*MNN2*	-4,84	-4,11	-3,30
α1,2-mannosyltransferase, responsible for addition of the second α1,2-linked mannose	*AEL148W*	*MNN5*	-2,97	-2,17	-2,63
α1,2-mannosyltransferase of the Golgi involved in *O*- and *N*-linked protein mannosylation	*AGL267C*	*KRE2, KTR6*	-2,46	-2,28	-1,90
α1,2-mannosyltransferase involved in *O*- and *N*-linked protein glycosylation	*ADL265W-A*	*KTR1*	-2,52	-2,22	-3,64
Mannosyltransferase involved in protein *N*-glycosylation in the Golgi	*ACR251C*	*KTR2, YUR1*	-2,56	-3,14	-3,72
Putative KRE2/MNT1 mannosyltransferase family protein involved in protein glycosylation	*AEL243W*	*KTR4*		7,92	26,45
Glycosylphosphatidylinositol (GPI) anchoring					
UDP-GlcNAc-binding and catalytic subunit of enzyme that mediates the first step in GPI biosynthesis	*AAL108C*	*SPT14*		1,96	1,66
Subunit of GPI-GlcNAc transferase involved in synthesis of the first intermediate in GPI biosynthesis	*AER333C*	*GPI19*			-3,95
Putative dolichyl-phosphate mannosyltransferase polypeptide 2, regulatory subunit	*ACR063W-A*	*YIL102C-A*			-2,08
Protein involved in the inositol acylation of glucosaminyl phosphatidylinositol (GlcN-PI)	*AFR094C*	*GWT1*		1,79	2,03
Mannosyltransferase that transfers the second mannose in GPI biosynthesis	*AER005C*	*GPI18*		2,42	4,13
ER membrane protein involved in a late step of GPI anchor assembly	*AAL059W*	*GPI11*	1,78	2,04	1,69
Subunit of the GPI transamidase complex that adds GPIs to newly synthesized proteins	*ADR299W*	*GPI8*		2,07	2,17
Subunit of the GPI transamidase complex that adds GPIs to newly synthesized proteins	*ADR333C*	*GPI16*	-2,25	-2,16	-2,35
Subunit of the GPI transamidase complex that adds GPIs to newly synthesized proteins	*ADR180C*	*GPI17*	-2,29	-2,40	-2,36
Integral plasma membrane protein involved in the synthesis of the GPI core structure	*AEL166C*	*LAS21*	-2,02	-2,09	-3,14
α1,2-mannosyltransferase involved in GPI biosynthesis	*AFR395C*	*SMP3*			-2,36
Protein degradation					
ER-associated degradation (ERAD)					
Ubiquitin-protein ligase required for ERAD of misfolded proteins	*ACL019C*	*HRD1*		2,12	2,71
Protein involved in ERAD; component of the Hrd1p complex	*AER311W*	*USA1*		2,77	2,63
Protein of the ER membrane that plays a central role in ERAD; forms HRD complex with Hrd1p	*ACL160C*	*HRD3*	-2,67	-1,90	-1,87
ER quality-control lectin	*AER171W*	*YOS9*	-2,10	-1,75	
Alpha mannosidase-like protein of the ER required for degradation of glycoproteins	*ADL390W*	*MNL1 (HTM1)*			-2,85
Protein proposed to coordinate the assembly of proteins involved in ERAD	*ADR182W*	*UBX2*		1,78	1,86

Table 2.6 – *Continued.*

Predicted function	*A. gossypii* gene	*S. cerevisiae* homolog(s)	Fold Change[a]		
			30 min	1 h	4 h
Protein degradation					
ER-associated degradation (ERAD)					
Co-chaperone for Hsp40p that promotes ERAD of integral membrane substrates	*AFL190C*	*HLJ1*		1,59	1,92
ER membrane protein that recruits the ubiquitin-conjugating enzyme Ubc7p to the ER	*AER279W*	*CUE1, CUE4*		-1,94	-2,48
Ubiquitin/proteasome					
Ubiquitin isopeptidase	*AFR007W*	*DOA4, UBP5*	1,72	2,67	3,39
ATPases of the26S proteasome involved in the degradation of ubiquitinated substrates	*AFR394W*	*RPT3*	1,72	1,63	
Ubiquitin activating enzyme (E1), involved in ubiquitin-mediated protein degradation	*AFR433C*	*UBA1*		2,11	
Ubiquitin-specific protease required for MVB sorting of membrane proteins	*ACL164C*	*UBP2*		2,21	2,82
ATPase involved in the transport of ubiquitinated proteins for degradation by the proteasome	*AFR158W*	*CDC48*		1,82	1,76
Ubiquitin fusion degradation protein that interacts with Cdc48p and Npl4p	*AFR662C*	*UFD1*		2,18	
Putative ubiquitin fusion degradation protein that interacts with Cdc48p and Npl4p	*ADL320C*	No homolog		3,93	5,61
Protein required for ubiquitin-mediated protein degradation, forms complex with Cdc48p	*AER448W*	*DOA1*		2,40	
Ubiquitin regulatory X domain-containing protein that interacts with Cdc48p	*ABR211C*	*SHP1*		1,89	2,22
Vacuolar protein sorting					
Sec1-like protein essential for vacuolar protein sorting (Golgi to endosome)	*AFL053W*	*VPS45*		4,68	6,30
Subunit of the vacuole protein sorting complex (Golgi to endosome and endosome to vacuole)	*AGR149W*	*VPS41*		2,05	1,74
Subunit of the vacuole protein sorting complex (Golgi to endosome and endosome to vacuole)	*AGL252W*	*VPS16*		2,55	2,06
Component of the Golgi-associated retrograde protein complex (endosome to late Golgi)	*ADR112W*	*VPS52*		2,35	3,99
Component of the Golgi-associated retrograde protein complex (endosome to late Golgi)	*ABL093W*	*VPS54*		2,80	3,41
Nexin-1 homolog involved in the recycling of proteins (endosome to late Golgi)	*ADR211W*	*VPS5, YKR078W*		2,42	2,41
Vacuolar sorting protein (cytoplasm to vacuole)	*AFR324W*	*VPS62, YPR157W*		1,58	3,04
Vacuolar sorting protein (cytoplasm to vacuole)	*AGL106C*	*VPS64, FAR10*	1,84	1,91	1,74
Protein required for sorting proteins to the vacuole	*ACL014C*	*MVP1*		2,59	2,11
Htz1p-binding component of the SWR1 complex required for vacuolar protein sorting	*ADR114C*	*VPS72*		2,01	2,38
Endosomal protein required for recycling Golgi proteins and sorting ubiquitinated proteins	*AFL176C*	*VPS27*		2,72	2,32
Cytoplasmic class E vacuolar protein sorting factor	*AGR204W*	*BRO1 (VPS31)*		2,18	2,74
Vesicle trafficking/transport					
Ras-like small GTPase, involved in the ER to Golgi step of the secretory pathway	*ABR220W*	*YPT1*		2,12	
Alpha subunit of Type II geranylgeranyltransferase required for vesicular transport (ER to Golgi)	*ACR042C*	*BET4*		1,68	1,93
GTPase-activating protein for the Ypt1p, involved in protein trafficking (ER to Golgi)	*AFL161C*	*GYP5, GYL1*		1,98	1,69
GTPase-activating protein for the Ypt1p, involved in the regulation of vesicle transport (ER to Golgi)	*AFR154C*	*GYP8*		2,39	
ATPase required for ER to Golgi transport	*AER169C*	*SEC18*		2,18	
GPI inositol deacylase of the ER that negatively regulates COPII vesicle formation (ER to Golgi)	*AFR521W*	*BST1*	-4,00	-3,72	-5,84
Integral membrane component of ER-derived COPII-coated vesicles (ER to Golgi)	*ABR243W*	*EMP47, EMP46*	-3,23	-3,31	-2,41

Table 2.6 – *Continued.*

Predicted function	*A. gossypii* gene	*S. cerevisiae* homolog(s)	Fold Change[a]		
			30 min	1 h	4 h
Vesicle trafficking/transport					
Protein localized to COPII-coated vesicles, forms a complex with Erv46p (ER to Golgi)	*AER136W*	*ERV41*		-2,52	-3,19
ER chaperone, required for incorporation of amino acid permeases into COPII coated vesicles	*ABL137W*	*SHR3*		-2,55	-2,18
Integral membrane protein of the early Golgi and ER, involved in COP II vesicle transport (ER to Golgi)	*AFR127W*	*SVP26*			-2,17
Protein member of the p24 family involved in ER to Golgi transport	*ADR185W*	*ERV25*		-3,72	
Conserved phosphoesterase domain-containing protein that acts in cargo exit from the ER	*ADL110W*	*TED1*	-6,20	-7,48	-6,24
GTPase, Ras-like GTP binding protein involved in the secretory pathway (Golgi to vacuole)	*AGR257C*	*YPT6*			-2,07
GTPase-activating protein for the Ypt6p, involved in vesicle mediated protein transport	*ACR057C*	*GYP6*		2,81	3,96
Target membrane receptor (t-SNARE) for vesicular intermediates (Golgi to vacuole)	*ACR092C*	*PEP12 (VPS6)*		2,06	3,95
Alpha subunit of COPI vesicle coatomer complex (Golgi to ER)	*ADR077C*	*COP1*		1,89	
Essential beta'-coat protein of the COPI coatomer (Golgi to ER)	*AFL118W*	*SEC27*		1,67	1,62
Delta subunit of the coatomer complex COPI (Golgi to ER)	*AFR274C*	*RET2*		1,51	1,58
Peripheral membrane protein required for retrograde traffic (Golgi to ER)	*AFR652W*	*DSL1*		2,28	2,53
t-SNARE required for ER membrane fusion and vesicular traffic, part of Dsl1p complex (Golgi to ER)	*ADR220W*	*UFE1*		1,79	1,55
Protein of unknown function proposed to be involved in protein secretion, interacts with Dsl1p	*AER129C*	*SEC39*	1,75	2,51	2,68
Peripheral membrane protein required for fusion of COPI vesicles with the ER (Golgi to ER)	*ADL286W*	*TIP20*		2,35	1,87
v-SNARE binding protein that facilitates specific protein retrieval from a late endosome to the Golgi	*AFR325W*	*BTN2, YPR158W*		2,97	4,52
Protein presumably functioning within the endosomal-vacuolar plasma membrane proteins trafficking	*AER006W*	*RCR1, RCR2*		2,55	5,79
Non-essential subunit of the exocyst complex (distal secretion)	*ADR012C*	*SEC3*		2,35	2,58
Essential 107kDa subunit of the exocyst complex (distal secretion)	*AGL158C*	*SEC5*		3,13	4,27
Essential 121kDa subunit of the exocyst complex (distal secretion)	*ADL317C*	*SEC8*		2,15	2,72
Essential 113kDa subunit of the exocyst complex (distal secretion)	*AFR251C*	*SEC15*		2,80	3,14
Essential 100kDa subunit of the exocyst complex (distal secretion)	*AGL130C*	*SEC10*		2,95	3,47
Subunit of the exocyst complex (distal secretion)	*AFR100W*	*EXO70*		3,10	4,34
Essential protein dual roles in spliceosome assembly and exocytosis (distal secretion)	*ADL321W*	*EXO84*		2,44	2,50
Alpha-adaptin involved in vesicle mediated transport	*ADL302W*	*APL3*		2,07	2,07
Guanine nucleotide exchange factor for Arf proteins involved in vesicular transport	*ABR102W*	*SYT1*		2,55	2,82

2.4 DISCUSSION

The secretion of proteins by filamentous fungi is important for hyphal extension, degradation of substrates in natural ecosystems, pathogenicity and for biotechnological exploitation. Thus, many filamentous fungi have evolved to secrete high amounts of proteins. Previous observations have suggested that the secretion abilities of *A. gossypii* were more similar to those of closely related yeast species than to those of other filamentous fungi (Ribeiro et al. 2010). Here, we observed that the total protein

concentration in the supernatants of *A. gossypii* submerged cultures with sucrose as primary carbon source was indeed relatively low (< 218 mg/l). *In silico* analysis of the *A. gossypii* secretome predicted that it should represent around 1% of the total proteome, a percentage closer to that predicted for yeast secretomes (2-4%) (Lee et al. 2003, Swaim et al. 2008, Brustolini et al. 2009, Mattanovich et al. 2009, Lum et al. 2011) than to that predicted for the secretomes of filamentous fungi (5-8%) (Tsang et al. 2009, Lum et al. 2011, Druzhinina et al. 2012). Nevertheless, in 2-D electrophoresis gel maps, up to 101 protein spots could be detected in the supernatants of *A. gossypii* grown in different media, indicating that, although in low amount, this fungus secretes a variety of proteins through the plasma membrane.

Of the 54 proteins that were predicted to comprise the *A. gossypii* secretome, less than 33% were putative enzymes with hydrolytic activity. This is in line with the limited range of carbon sources which *A. gossypii* utilises (Pridham and Raper 1950, Ribeiro et al. 2011). Extracellular lipase (Stahmann et al. 1997), amylase (Ribeiro et al. 2013) and *β*-glucosidase (Ribeiro et. al 2013) activities have previously been found in *A. gossypii* culture supernatants. In agreement with these observations, one putative lipase (AER454C) and two putative *β*-glucosidases (AGL354C and AGL343C) were predicted to be secreted by *A. gossypii*. However, neither of the putative *A. gossypii* amylases (AEL044W and AEL276C) were predicted to contain an N-terminal signal peptide and, thus, would not be expected to be secreted via the general secretory pathway.

Although extracellular protease activity in *A. gossypii* supernatants has been reported as negligible (Ribeiro et al. 2010), nine putative proteases (AAR129C, ABL123C, ACR143W, ACR144W, ADR123W, AER022W, AGL326W, AGR240W and AGR407C) were predicted to be secreted, the majority of which would probably be most active at acidic pH (Rawlings et al. 2012). Given that the optimum pH range for *A. gossypii* is 6-7 (Ribeiro et al. 2011) and that only low concentration of proteins are secreted by this fungus, extracellular protease activity would indeed be expected to be low and undetectable. An invertase (AFR529W) was also predicted to be secreted by this fungus and subsequent experimental characterization of this protein confirmed its function and secretion into the culture supernatant (Aguiar et al. Chapter 4). Additionally, there was a putative acid phosphatase, a putative ureohydrolase and two putative FMN-binding proteins with probable oxidoreductase activity among the predicted secretome proteins with no homolog in *S. cerevisiae*, but having homologs in *K. lactis*. Well-known

extracellular proteins like α-factor mating pheromones (AAR163C and AFL062W) were also predicted to be secreted.

The rather small number of proteins secreted by *A. gossypii*, together with the low concentration in which they were produced and with the negligible extracellular protease activity would be major advantages of this organism as a recombinant protein production host, as secreted products are less likely to be contaminated or degraded by its native proteins. Moreover, comprehensive analysis of the *A. gossypii* native secreted proteins can have a positive impact on product purification and quality control.

Despite these several advantages, a major drawback of *A. gossypii* for protein production is still its low productivity. Results from this transcriptomic study of recombinant EGI production suggest that stronger promoters and/or better expression strategies (like genome integration) need to be employed to obtain higher expression levels of the recombinant genes. In fact, changing the vector used to express EGI has already resulted in a 2 fold increase in the extracellular EGI activity (Ribeiro et al. 2013). High extracellular production levels of recombinant β-galactosidase from *Aspergillus niger* have also been achieved by using the *A. gossipii* native *TEF* promoter (Magalhães et al. Chapter 5). Enhancement of the protein translation efficiency could also increase the production of secreted proteins, as indicated by the down-regulation of some genes involved in regulation of protein translation during EGI production.

Previous studies have shown that secretion of recombinant proteins can lead to secretion stress and trigger the UPR, which modulates both general and protein-specific transcriptional responses (Sims et al. 2005, Arvas et al. 2006, Guillemette et al. 2007, Gasser et al. 2008). This did not occur in *A. gossypii*, but may be explained by the fact that *EGI* was not highly expressed and, thus, its production did not constitute a major burden to the cells. Carvalho et al (2011) have recently shown that, in *A. niger*, the induction of the UPR pathway is dependent on the level of heterologous gene expression. Under relative low-expressing conditions, the basal protein folding and quality control machinery of the *A. niger* ER was apparently adequate, but under high-expressing conditions ER stress was induced.

We induced secretion stress in *A. gossypii* cells with DTT, since it was not induced by EGI production. The use of DTT to induce secretion stress has been widespread in investigations of the UPR (Travers et al. 2000, Saloheimo et al. 2003, Sims et al. 2005, Arvas et al. 2006, Guillemette et al. 2007, Graf et al. 2008, Wimalasena et al. 2008), but DTT also affects the transcription of other genes which may not be closely related to the

UPR. A transcriptomic comparison of *Aspergillus nidulans* cells stressed by DTT treatment or by recombinant chymosin secretion showed similar changes in the expression of some genes, but not others (Sims et al. 2005). Different stresses elicit slightly different responses. However, DTT has consistently induced the UPR in various yeast and filamentous fungi (Travers et al. 2000, Saloheimo et al. 2003, Sims et al. 2005, Arvas et al. 2006, Guillemette et al. 2007, Graf et al. 2008, Wimalasena et al. 2008).

DTT did not trigger a conventional UPR in *A. gossypii*: the expression levels of several well-known UPR target genes (such as *IRE1*, *HAC1*, *KAR2*, *PDI1* and *EUG1*) remained unchanged and no UPRE-like motif was overrepresented in the gene clusters up-regulated by DTT. The amount of DTT used in this study (10 mM final concentration) was comparable to that used in other studies with yeast and filamentous fungus (2-10 mM; Travers et al. 2000, Saloheimo et al. 2003, Graf et al. 2008, Wimalasena et al. 2008). To our knowledge, only in studies of *A. nidulans* and *A. niger* has DTT been added to the cultures at a higher concentration (20 mM; Sims et al. 2005, Guillemette et al. 2007).

Although a classical UPR was not induced, there was evidence of secretion stress induced by DTT in *A. gossypii* cells. Expression of several genes involved in protein unfolding, ERAD, proteasome degradation, proteolysis, vesicle trafficking, vacuolar protein sorting and secretion significantly increased within 1 h of DTT treatment. In fungi, the UPR was thought to be exclusively dependent on Ire1p-mediated splicing of *HAC1* mRNA. However, an *IRE1*-, *HAC1*- and UPRE-independent pathway for transcriptional activation upon ER stress exists in *S. cerevisiae*, which may activate a core promoter through stimulation of RNA polymerase II holoenzyme activity (Schröder et al. 2003). Miyazaki et al. (2013) have also demonstrated that *Candida glabrata* has lost the classic Ire1p-Hac1p UPR, but instead possesses an alternative mechanism, RIDD. In *A. gossypii*, the expression of several genes involved in mRNA degradation was induced by DTT, which suggests that a RIDD-like mechanism may exist in *A. gossypii* to reduce the ER load when there is secretion stress. Another mechanism to alleviate the load of proteins in the ER in some fungi is the transcriptional down-regulation of genes encoding secreted proteins in response to secretion stress (RESS). This down-regulation mechanism has been described in *T. reesei* (Pakula et al. 2003), *A. niger* (Al-Sheikh et al. 2004) and *S. cereviaise* (Kimata et al. 2006). In *A. gossypii*, DTT also repressed the transcription of a large number of genes encoding putative secretory proteins.

Two genes encoding for a subunit of the translocon complex (Ssh1p) and for a chaperone involved in the translocation of newly synthesised proteins into the ER (Lhs1p)

were also repressed by DTT, which may have contributed to accumulation of unfolded proteins in the cytosol. Intriguingly, most the genes involved in protein folding that were up-regulated by DTT encoded cytosolic chaperones, co-chaperones and nucleotide exchange factors. A stress response induced by misfolded cytosolic proteins that do not enter the secretory pathway, called UPR-Cyto, has been preliminarily characterized in *S. cerevisiae* (Metzger and Michaelis 2009, Geiler-Samerotte et al. 2010). This induces the production of several cytosolic chaperones and co-chaperones. The UPR-Cyto response appears to be a specific *HSF1*-mediated module of the eukaryotic heat shock response (Metzger and Michaelis 2009, Geiler-Samerotte et al. 2010). The transcript level of the *HSF1* homolog in DTT-stressed *A. gossypii* cells was only slightly increased. Moreover, no Hsf1p-like consensus binding sequence was overrepresented in the gene clusters analyzed. However, the homologs of several *A. gossypii* genes that were up-regulated by DTT have consensus sequences for Hsf1p binding in their promoter region in *S. cerevisiae* (Teixeira et al. 2006). These include all genes encoding cytosolic folding elements which were up-regulated in *A. gossypii* in response to DTT. Thus, an UPR-Cyto may have been activated in *A. gossypii* in response to secretion stress induced by DTT, but not by Hsf1p.

Another striking difference in the *A. gossypii* transcriptional responses to DTT-induced stress compared to *S. cerevisiae*, *A. niger* or *T. reesei* was the rapid and severe down-regulation of the protein glycosylation pathway, an effect that at similar extent has only been described for treatments with tunicamycin. This could lead to an accumulation of improperly glycosylated proteins. In mammals, calnexin provides chaperone activity to retain incompletely glycosylated proteins in the ER, functioning as a component of the glycoprotein quality control system in the ER (Caramelo and Parodi 2008). The *S. cerevisiae* homolog, Cne1p, also binds specifically to monoglucosylated oligosaccharides (Xu et al. 2004). However, no homolog for the *CNE1* was found in the *A. gossypii* genome. An alternative mechanism of quality control should, therefore, exist to balance this absence.

The best understood mechanism of ERAD substrate recognition is the glycan-dependent pathway of ERAD-L (lumen). The α-1,2-mannosidase Mns1p removes a mannose residue from the middle branch (B) in misfolded *N*-glycosylated substrates, leading to an acceleration of its ERAD (Stolz and Wolf 2010). The mannosidase Mnl1p/Htm1p is also needed for substrate binding and further mannose trimming (Stolz and Wolf 2010). The glycan structure is then recognized by the lectin Yos9p as being missfolded. Structurally, Yos9p is part of the Hrd1p complex through a direct interaction

with the large luminal domain of Hrd3p (Carvalho et al. 2006). The Hrd1p/Hrd3p ligase forms a near stoichiometric membrane complex by binding to Der1p via the linker protein Usa1p (Carvalho et al. 2006). In *A. gossypii*, the expression of the *MNL1*, *YOS9* and *HDR3* homologs were repressed by DTT, while the *HRD1* and *USA1* homologs were transcriptionally induced. Two other ERAD quality control pathways have been described as well: ERAD-M (membrane) and ERAD-C (cytosol) (Carvalho et al. 2006). ERAD-M substrates may be directly recognized by the Hrd1p E3 ligase and ERAD-C substrates are recognized by the cytoplasmic Hsp70p/Hsp40p chaperones Ssa1p, Ydj1p and Hlj1p, and by the Doa10 E3 ligase (Nakatsukasa and Brodsky 2008). The degradation pathway of ERAD-L and ERAD-C substrates merges at the Cdc48p-Ufd1p-Npl4p machinery, which further delivers the misfolded polyubiquitylated proteins to the proteasome for degradation (Stolz and Wolf 2010). The Cdc48p/Ufd1p/Npl4p ATPase complex and the linker protein Ubx2p were up-regulated by DTT in *A. gossypii*, as was the Hlj1p chaperone.

Like the ER, the Golgi complex may also be involved in conformation-based disposal of abnormal proteins that were targeted for degradation (Arvan et al. 2002). Here we show that several post-ER pathways for protein disposal were up-regulated upon DTT treatment in *A. gossypii*. These included both the retrograde transport of proteins back to the ER for ERAD (via COPI-vesicle mediated transport) and protein transport via the endosomal system for degradation. COPII-vesicle mediated export of proteins from the ER to the Golgi was, however, down-regulated by DTT.

Our results provide the first insights into the secretion stress response of *A. gossypii*, as well as a basic understanding of its protein secretion potential and protein quality control systems. We have shown that the lack of an active conventional UPR in *A. gossypii* may be compensated by alternative pathways, probably working simultaneously, to relieve the cells from secretion stress. The fact that *A. gossypii* has one of the smallest eukaryotic genomes known and, consequently, a reduced genetic machinery, may have contributed to the differences between its transcriptional responses to secretion stress and those reported for other fungal species. The absence of a calnexin homolog in *A. gossypii* indicates that it lacks some of the ER quality control mechanisms of other fungi. Despite the high genetic similarity it shares with *S. cerevisiae*, the regulation of the protein secretory pathway of *A. gossypii* and *S. cerevisiae* differed considerably. It must not be forgotten that significant sequence similarity between *A. gossypii* and *S. cerevisiae* is restricted to coding regions (Brachat et al. 2003). Therefore, several differences at the gene regulation level are expected to exist between these organisms.

CHAPTER 3

Characterization of the *Ashbya gossypii* secreted *N*-glycome and genomic insights into its *N*-glycosylation pathway

ABSTRACT

The riboflavin producer *Ashbya gossypii* is a filamentous hemiascomycete closely related to the yeast *Saccharomyces cerevisiae* that has been used as a model organism to study fungal developmental biology. It has also been explored as a host for the expression of recombinant proteins. However, although *N*-glycosylation plays important roles in protein secretion, morphogenesis and development of multicellular organisms, the *N*-glycan structures synthesised by *A. gossypii* had not been elucidated. In this study, we report the first characterization of *A. gossypii N*-glycans and provide valuable insights into their biosynthetic pathway. By combined matrix-assisted laser desorption-ionization time-of-flight (MALDI-TOF) mass spectrometry profiling and nuclear magnetic resonance (NMR) spectroscopy we determined that the *A. gossypii* secreted *N*-glycome is characterized by high-mannose type structures in the range $Man_{4-18}GlcNAc_2$, mostly containing neutral core-type *N*-glycans with 8 to 10 mannoses. Cultivation in defined minimal media induced the production of acidic mannosylphosphorylated *N*-glycans, generally more elongated than the neutral *N*-glycans. Truncated neutral *N*-glycan structures similar to those found in other filamentous fungi ($Man_{4-7}GlcNAc_2$) were detected, suggesting the possible existence of trimming activity in *A. gossypii*. Homologs for all of the *S. cerevisiae* genes known to be involved in the endoplasmatic reticulum and Golgi *N*-glycan processing were found in the *A. gossypii* genome. However, processing of *N*-glycans by *A. gossypii* differs considerably from that by *S. cerevisiae*, allowing much shorter *N*-glycans. Genes for two putative *N*-glycan processing enzymes were identified, that did not have homologs in *S. cerevisiae*.

3.1 INTRODUCTION

Ashbya gossypii is a pre-whole genome duplication hemiascomycete (Dietrich et al. 2004) that has long been known in the scientific and industrial communities, first as a cotton pathogen and subsequently as a riboflavin overproducer (Stahmann et al. 2000; Wendland and Walther 2005). It grows exclusively in a filamentous way, but has the smallest free living eukaryotic genome known, which shares a high degree of gene homology and gene order conservation with that of *Saccharomyces cerevisiae* (Dietrich et al. 2004).

The similarities of *A. gossypii* to both yeast and filamentous fungi, its small genome, haploid nuclei and ease of genetic manipulation led to expanded interest in this organism beyond riboflavin production. Over the past two decades *A. gossypi* has been used as a model organism to study the biological processes and regulatory mechanisms that govern the differences between filamentous and yeast growth (Wendland and Walther 2005; Schmitz and Philippsen 2011). Recently, it has also been considered as a host for the expression of recombinant proteins (Ribeiro et al. 2010).

Sustained polarized hyphal growth and generally high secretion capacity are typical features of filamentous fungi that distinguish them from yeast. Interestingly, although its secretory pathway remains poorly characterized, there is evidence that the level of protein secretion in *A. gossypii* is more similar to that of yeast than to that of other filamentous fungi (Ribeiro et al. 2010).

In eukaryotes, proteins destined for the secretory pathway are targeted to the endoplasmatic reticulum (ER), where they are folded and can undergo post-translational modifications. They are then transferred to the Golgi apparatus, where they undergo further processing before being delivered to the plasma membrane for exocytosis by vesicles of the distal secretory system (Conesa et al. 2001; Shoji et al. 2008). In filamentous fungi this is a highly polarized process, which takes place mainly at hyphal tips through intricate mechanisms as yet not fully understood (Conesa et al. 2001; Shoji et al. 2008; Hayakawa et al. 2011). Recently, the importance of protein glycosylation in the polarized biosynthetic sorting of proteins and filamentous growth has become increasingly evident (Potter et al. 2006; Kotz et al. 2010; Motteram et al. 2011; Jin 2012).

Glycosylation is a post-translational modification that strongly influences protein folding, secretion, cellular localization and biological activity (Helenius and Aebi 2004). It involves the covalent attachment of one or more carbohydrates (glycans) to a protein,

providing a means to enhance its structural and functional diversity. In most cases, the attachment is to an asparagine residue (within the sequence Asn-X-Ser/Thr, where X denotes any aminoacid except proline) in a process termed *N*-glycosylation. The early stages of *N*-glycan processing within the ER are highly conserved among eukaryotes, whereas the further trimming and elongation of the core structures made in the ER differ remarkably between species and even between individual proteins within a single species or cell type (Stolz and Munro 2002).

N-glycosylation has been well studied in several biotechnologically important yeast and filamentous fungi, mostly in those with potential as recombinant protein producers, as the type of sugars that compose the *N*-glycans, along with the extent and structure of these polymers greatly affects the production, stability and bioactivity of recombinant glycoproteins (Maras et al. 1999; De Pourcq et al. 2010). Fungal *N*-glycans are mainly composed of mannosyl residues (high-mannose type), although additional residues, such as *N*-acetylglucosamine, glucose, galactose, xylose, fucose, pyruvate, or phosphate may also be found in some species (Gemmill and Trimble 1999; Maras et al. 1999; Deshpande et al. 2008; Buser et al. 2010; Grass et al. 2011). Yeast often specialize their high-mannose *N*-glycans further by extending them into large hypermannosylated outer chains (Gemmill and Trimble 1999). In contrast, although there are reports pointing to the occurrence of hypermannosylated *N*-glycans in specific strains, filamentous fungi most often synthesize small high-mannose type *N*-glycans (Maras et al. 1999; Deshpande et al. 2008). Although no information is available on the composition and structural characteristics of *A. gossypii* *N*-glycans, Ribeiro et al. (2010) suggested that they may be less extensively glycosylated than *S. cerevisiae* glycans.

The mannosyltransferase Och1 protein is the key enzyme for the synthesis of hypermannosylated outer chains in yeast and filamentous fungi. Recent findings suggest a link between filamentous growth and Och1 proteins having a signal peptide (Kotz et al. 2010). Of particular interest is the putative Och1 protein of *A. gossypii* that, although being highly homologous to its *S. cerevisiae* counterpart, was predicted to have an N-terminal signal sequence like that found in other filamentous fungi rather than an N-terminal membrane anchor, as predicted for yeast Och1 proteins (Kotz et al. 2010).

Considering the unique characteristics of *A. gossypii* and that, to our knowledge, no information was available on the *N*-glycan structures synthesized by this fungus, we undertook a global analysis of the *N*-glycans attached to its secreted glycoproteins by MALDI-TOF mass spectrometric profiling and NMR spectroscopy. Here, we report the

first characterization of *A. gossypii N*-glycan structures, including culture medium- and cultivation stage-specific *N*-glycan profiles, to provide an enriched context for assessing overall *N*-glycosylation heterogeneity in *A. gossypii*.

3.2 MATERIALS AND METHODS

3.2.1 Strain, media and culture conditions

The strain used in this study was *A. gossypii* ATCC 10895, kindly provided by Prof. P. Philippsen (Biozentrum, University of Basel, Switzerland). Stock cultures were maintained as spores suspended in 20% (v/v) glycerol, 0.8% (w/v) NaCl with 0.025% (v/v) Tween 20 at -80°C. For bioreactor inoculation, 50 ml pre-cultures, inoculated with 10^6 spores, were cultivated in 250 ml Erlenmeyer flasks at 30°C and 200 rpm. Pre-cultures were grown in Ashbya Full Medium (Altmann-Jöhl and Philippsen 1996) for 17 h before being harvested by filtration through disks of sterile disposable cleaning cloth (X-tra, Inex Partners Oy, Finland) and resuspended in an appropriate volume of sterile distilled water. Bioreactors were inoculated to an initial biomass of 0.55 g/l. Batch cultivations were carried out in B. Braun Biotech International (Sartorius AG) Biostat® B-DCU (2.0 l working volume) bioreactors containing 1.2 l of either defined minimal medium (DMM; Verduyn et al. 1992) or complex rich medium (CRM; 1% (w/v) tryptone, 1% (w/v) yeast extract, 0.1% (w/v) myo-inositol), both with 2% (w/v) sucrose as carbon source. Polypropylene glycol (mixed molecular weights (Wiebe et al. 2001)) was added to a final concentration of 0.2% (v/v) to control foam production. Cultures were grown at 30°C, with 600 rpm agitation and aeration of 2.4 volumes of air per volume of liquid per minute (vvm). Culture pH was maintained at 6.0 ± 0.1 by the addition of sterile 1 M KOH or 1 M H_3PO_4. *A. gossypii* maximum specific growth rate in DMM was 0.22 ± 0.01 h^{-1} and in CRM 0.36 ± 0.01 h^{-1}. Culture supernatants were collected by filtration and stored at -20°C until further use. For dry weight determination, mycelia were harvested from 5 ml of fermentation broth by filtration through disks of disposable cleaning cloth (X-tra), washed with 2 sample volumes of distilled water and dried to a constant weight at 100°C. Triplicate samples were taken for all biomass measurements.

3.2.2 Preparation of *N*-glycans for MALDI-TOF mass spectrometry

The protein content of the culture supernatants collected during exponential (Biomass: 1.9 g/l for DMM and 2.9 g/l for CRM), early stationary (Biomass: 5.5 g/l for DMM and 7.6 g/l for CRM) and late stationary (Biomass: 5.7 g/l for DMM and 8.1 g/l for CRM) phases was determined by the Bradford protein assay (Bio-Rad) and 500 µl of each sample (5-45 µg of total protein) were taken for mass spectrometry *N*-glycan profiling. Samples of the culture media alone were also analysed. *N*-glycans were detached from the secreted glycoproteins by *Flavobacterium meningosepticum* peptide:*N*-glycosidase F (PNGase F; Prozyme) digestion and purified by organic extraction-precipitation and miniaturized solid-phase extraction steps as described in Hemmoranta et al. (2007).

3.2.3 Mass spectrometry

MALDI-TOF mass spectrometry was performed on a Bruker Ultraflex III TOF/TOF instrument (Bruker Daltonics) as previously described (Hemmoranta et al. 2007). Neutral *N*-glycans were detected in positive ion reflector mode as $[M + Na]^+$ ions and acidic *N*-glycans were detected in negative ion reflector mode as $[M - H]^-$ ions. The software FlexAnalysis 3.3 (Bruker Daltonics) was used for analysis of the mass spectra. Relative molar abundances of neutral and acidic glycan components were assigned based on their relative signal intensities in the mass spectra when analyzed separately as the neutral and acidic *N*-glycan fractions. The presented glycan profiles were extracted from the resulting signal lists by removing the effect of isotopic pattern overlapping, multiple alkali metal adduct signals, products of elimination of water from the reducing oligosaccharides, and other interfering mass spectrometric signals not arising from the original glycans in the sample. The resulting glycan signals in the presented glycan profiles were normalized to 100% to allow comparison between samples (Hemmoranta et al. 2007).

3.2.4 Preparation of *N*-glycans for NMR analysis

Approximately 750 ml of DMM culture supernatant was collected during late stationary phase by filtration and concentrated under vacuum in a RapidVap® N2/48 Evaporation System (Labconco Corporation) to a final volume of 40 ml. Subsequently, the proteins were precipitated overnight at -20°C with 2 volumes of ice-cold acetone

containing 10% (v/v) trichloroacetic acid. The pellet was washed twice with ice-cold acetone, dried at room temperature and redissolved with 8 ml of 20 mM sodium phosphate buffer, pH 7.3. Proteins (100 mg) were denaturated in 50 ml polypropylene tubes by heating at 100°C for 5 min in the presence of 0.1% (w/v) sodium dodecyl sulfate and 1% (v/v) 2-mercaptoethanol. After addition of 0.75% (v/v) Nonidet P-40 and 2 U/ml of PNGase F (Roche), *N*-glycans were released from the glycoproteins by digestion at 37°C for 48 h. Glycans were extracted by selective organic procedures essentially as previously described (Verostek et al. 2000). Briefly, proteins and glycans were precipitated with 4 volumes of ice-cold acetone at -20°C and the glycans extracted twice from the pellets with 60% (v/v) ice-cold methanol at -20°C. The two extracts were combined and dried. Prior to NMR experiments, the isolated glycans were lyophilized twice from 99.8% methanol-d4 (MeOD; Sigma-Aldrich), repeatedly exchanged with 99.9% deuterium oxide (D_2O; Sigma-Aldrich), with intermediate lyophilization, and finally dissolved in 600 µl D_2O.

3.2.5 NMR spectroscopy

The ^{1}H NMR spectra were recorded at 10ºC or 25ºC on a 600 MHz Bruker Avance III NMR spectrometer (Bruker Daltonics) equipped with a QCI Cryoprobe. One-dimensional (1-D) ^{1}H spectra were recorded with 4 s presaturation of the residual water signal using the 1-D NOESY presaturation sequence. Two-dimensional (2-D) DQF-COSY and TOCSY spectra were recorded using standard Bruker pulse sequences. In TOCSY, the mixing time used was 80 ms. The ^{31}P decoupled ^{1}H spectrum was recorded with GARP4 ^{31}P decoupling (Shaka et al. 1985) during the acquisition time. The ^{1}H-^{31}P HSQC spectrum was recorded with standard echo/antiecho-TPPI gradient selection pulse sequence and GARP4 ^{31}P decoupling (Shaka et al. 1985) during the acquisition time. The ^{1}H-^{31}P coupling constant was estimated to 8 Hz. The ^{1}H chemical shifts were referenced to internal acetone (δ 2.225 in $^{2}H_2O$) and the ^{31}P chemical shifts were referenced to external 85% phosphoric acid (0 ppm).

3.2.6 *In silico* analysis of the *A. gossypii N*-glycosylation pathway

The Ashbya Genome Database (http://agd.vital-it.ch) (Gattiker et al. 2007) was used to map a set of *S. cerevisiae* genes involved in the various steps of the *N*-glycosylation pathway to *A. gossypii* genes. To search for putative *A. gossypii* genes

encoding glycoside hydrolases and glycosyltransferases, with a possible role in *N*-glycosylation and for which the homolog was absent in *S. cerevisiae*, the Carbohydrate-Active enZYmes (CAZy) database (www.cazy.org/e189.html) (Cantarel et al. 2009) was used. Prediction of signal peptides or signal anchors on putative proteins was done using SignalP 3.0 (www.cbs.dtu.dk/services/SignalP-3.0/) and MultiLoc2-HighRes (Fungal) (http://www-bs.informatik.uni-tuebingen.de/Services/MultiLoc2) was used to predict subcellular protein localization.

3.3 RESULTS

3.3.1 *A. gossypii* neutral *N*-glycan profiles

Neutral and acidic *N*-glycans derived from *A. gossypii* secreted proteins were analysed separately by MALDI-TOF mass spectrometry. Using a mass matching approach, monosaccharide compositions were proposed for different glycan signals (Table 3.1) and annotated in the mass spectra profiles (Figures 3.1 and 3.2).

In the positive-ion mode mass spectra of the neutral *N*-glycans (Figure 3.1), a series of peaks predominated with mass-to-charge ratio (*m/z*) values that correlated with the $[M + Na]^+$ adduct ions of high-mannose type *N*-glycans ranging in size from H4N2 to H18N2 (Table 3.1). The high-mannose type *N*-glycans with composition H8N2, H9N2 and H10N2 were enriched in all of the neutral *N*-glycan pools, and only minor relative amounts of *N*-glycans containing more than 11 hexoses were detected. *N*-glycans containing 7 hexoses (H7N2) were also enriched in the neutral *N*-glycan pools during exponential growth in complex rich medium (Figure 3.1B), decreasing during stationary phase (Figures 3.1D and 3.1F). In agreement with these results, chemical shifts typical of a mixture of high-mannose type *N*-glycan structures with an average of nine mannoses per molecule were observed by ^{1}H NMR spectroscopy of the total pool of unfractionated *N*-glycans (Figures 3.3 and Table 3.2).

As is evident from Figure 3.1, the size distribution and relative amounts of neutral *N*-glycans were generally similar during cultivation in both defined minimal and complex rich medium, although a higher amount of small (H4N2 to H7N2) and high (H16N2 to H18N2) molecular weight glycan structures was observed when complex medium was used. In minimal medium, three pseudomolar ion signals observed at *m/z* 1241.3, 1386.3 and 1531.2 were among the most abundant signals observed during exponential growth

(Figure 3.1A). Of these, only the signal at *m/z* 1241.3 corresponded to the molecular weight of a known *N*-glycan, with the monosaccharide composition H4N2F1. However, no evidence was obtained from our NMR data for the presence of fucosylated *N*-glycans in *A. gossypii* (see section 3.3.3), nor were genes putatively encoding for fucosyltransferases found in the *A. gossypii* genome. The molecular weight of the peak at *m/z* 1531.2 correlated with a series of ion signals that were observed later in the cultivations at 162 *m/z* intervals (indicated with arrows in Figure 3.1), which is the molecular weight of a hexose. Considering that these signals did not correspond to known *N*-glycans nor to contaminants from the media, they are referred to here as unknown polyhexose contaminants.

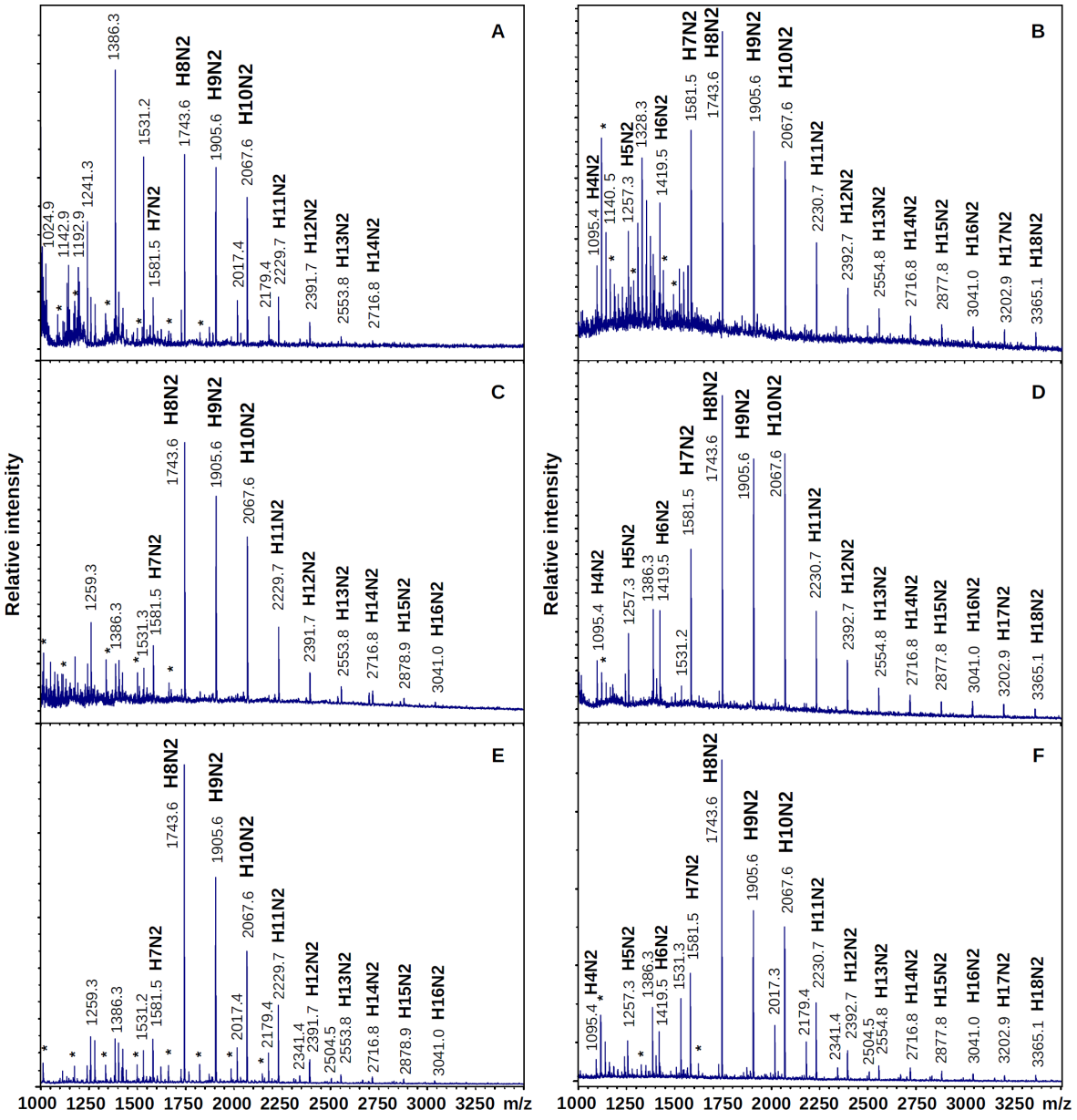


Figure 3.1 – MALDI-TOF mass spectra of neutral *N*-glycans released by PNGase F from *A. gossypii*

glycoproteins present in the supernatants of cultivations in defined minimal medium (**A**, **C** and **E**) and complex rich medium (**B**, **D** and **F**) at exponential (**A** and **B**), early stationary (**C** and **D**) and late stationary (**E** and **F**) phases. Glycan species are annotated with proposed monosaccharide compositions derived by mass calculation. H, hexose; N, *N*-acetylhexosamine; *, known contaminants from the culture media; ↓, unknown polyhexose contaminants.

3.3.2 *A. gossypii* acidic *N*-glycan profiles

Acidic *N*-glycans were only detected when *A. gossypii* was grown in defined minimal medium (Figure 3.2 and Table 3.1), as all ion signals observed in the negative ion mode mass spectra from cultivation in complex rich medium corresponded to known contaminants from the culture medium (not shown).

In minimal medium, a series of peaks predominated in the acidic *N*-glycan profiles (Figure 3.2) at *m/z* values that correlated with the $[M - H]^-$ adduct ions of monophosphorylated/monosulphated high-mannose type *N*-glycans ranging in size from *P*1H5N2 to *P*1H17N2 (Table 3.1). Consistent with these results, NMR spectroscopy revealed the presence of mannosylphosporylated *N*-glycans, but no evidence for the presence of sulphate (see section 3.3.3).

N-Glycans with composition *P*1H5N2, *P*1H13N2, *P*1H14N2 and *P*1H15N2 were enriched in the acidic *N*-glycan pools during stationary phase, while *P*1H5N2 was also enriched during exponential growth. In general, the major acidic *N*-glycan structures present during the stationary phase were about five hexose units larger than the major neutral *N*-glycans.

Similar to what was observed in the neutral *N*-glycan profiles, the negative ion mode mass spectra from cultivation in defined minimal medium also included a series of peaks at 160 *m/z* intervals (*m/z* values indicated with boxes in Figure 3.2), that did not correspond to known contaminant signals from the media, and which were among the major signals detected during the exponential growth. The *m/z* values of these peaks correlated with the $[M - H]^-$ adduct ions of high-mannose type *N*-glycans containing two phosphate/sulphate residues and only one *N*-acetylhexosamine (*P*2H5N1 to *P*2H8N1). The *m/z* values of another series of minor peaks observed at 160 *m/z* intervals (indicated with arrows in Figure 3.2) corresponded to diphosphorylated/disulphated *N*-glycans in the range *P*2H3N2 to *P*2H16N2.

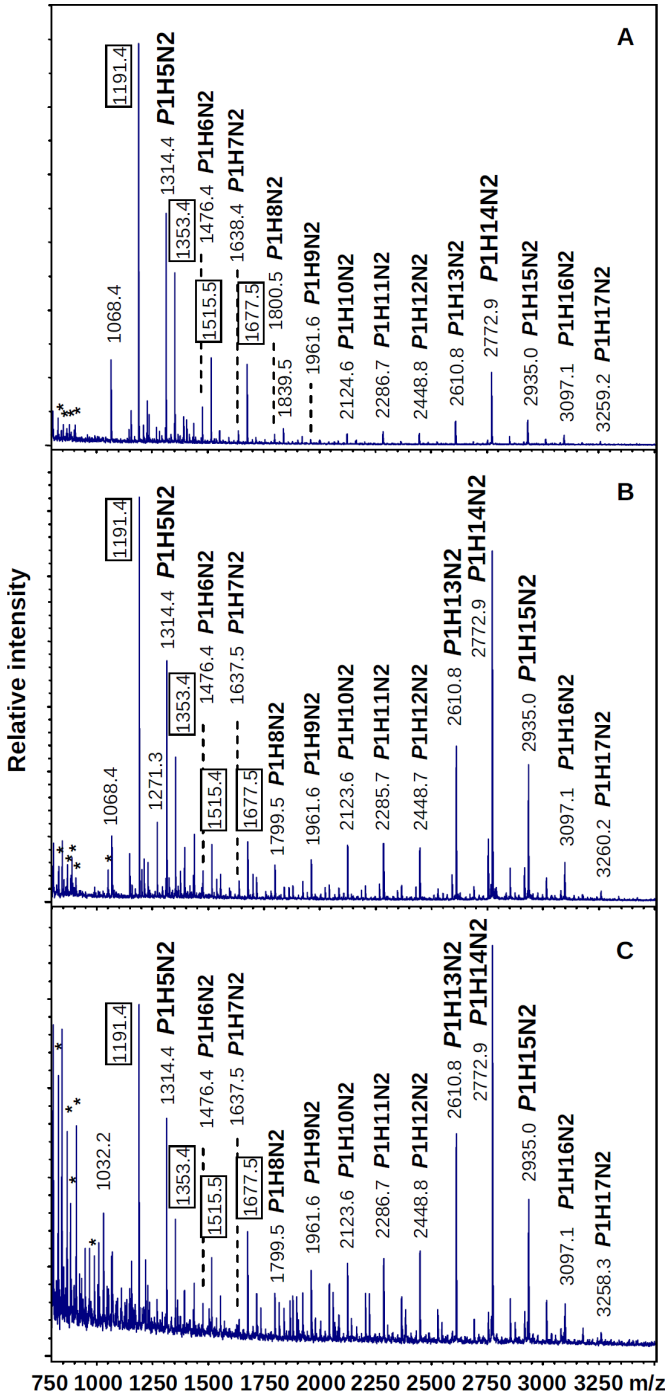


Figure 3.2 – MALDI-TOF mass spectra of acidic *N*-glycans released by PNGase F from *A. gossypii* glycoproteins present in the supernatant of mycelia grown in defined minimal medium at exponential (**A**), early stationary (**B**) and late stationary (**C**) phases. Glycan species are annotated with proposed monosaccharide compositions derived by mass calculation and based on NMR data. H, hexose; N, *N*-acetylhexosamine; *P*, phosphate; ↓, glycans containing two phosphate residues; □, glycans containing two phosphate residues and only one *N*-acetylhexosamine; *, known contaminants from the culture media.

Table 3.1 – Putative composition of underivatized *N*-glycans from *A. gossypii* secreted glycoproteins detected by MALDI-TOF mass spectrometry. Proposed compositions were calculated on the basis of mass matching and annotated as follows: H, hexose; N, *N*-acetylhexosamine; F, deoxyhexose; *P*, phosphate or sulfate.

Putative composition	Observed	Calculated
Neutral	**m/z $[M + Na]^+$**	
H4N2	1095.4	1095.4
H4N2F1	1241.3	1241.4
H5N2	1257.3-1257.4	1257.4
Unknown	1386.3	
H6N2	1419.5	1419.5
Unknown	1531.2-1531.3	
H7N2	1581.5	1581.5
H8N2	1743.6	1743.6
H9N2	1905.6	1905.6
Unknown	2017.3-2018.4	
H10N2	2067.6-2067.7	2067.7
Unknown	2179.4	
H11N2	2229.7-2230.7	2229.7
Unknown	2341.4	
H12N2	2391.7-2392.7	2391.8
Unknown	2504.5	
H13N2	2553.8-2554.8	2553.9
H14N2	2716.8	2715.9
H15N2	2877.8-2878.9	2878.0
H16N2	3039.9-3041.0	3040.0
H17N2	3202.9-3203.0	3202.1
H18N2	3365.0-3365.1	3364.1
Acidic	**m/z $[M - H]^-$**	
*P*2H3N2	1068.4	1069.3
*P*2H5N1	1191.4	1190.3
*P*2H4N2	1230.4	1231.3
*P1*H5N2	1314.4	1313.4
*P*2H6N1	1353.4	1352.3
*P*2H5N2	1394.3	1393.4
*P*1H6N2	1476.4	1475.4
*P*2H7N1	1515.4-1515.5	1514.4
*P*2H6N2	1554.5-1555.5	1555.4
*P*1H7N2	1637.5-1638.4	1637.5
*P*2H8N1	1677.5	1676.4
*P*2H7N2	1716.5-1717.5	1717.5
*P*1H8N2	1799.5-1800.5	1799.6
*P*2H8N2	1879.5-1879.6	1879.5
*P*1H9N2	1961.6	1961.6
*P*2H9N2	2041.6-2042.6	2041.6
*P*1H10N2	2123.6-2124.6	2123.7
*P*2H10N2	2203.6-2204.6	2203.6
*P*1H11N2	2285.7-2286.7	2285.7
*P*2H11N2	2365.7-2366.7	2365.7
*P*1H12N2	2448.7-2448.8	2447.8
*P*2H12N2	2527.7-2528.7	2527.7
*P*1H13N2	2610.8	2609.8
*P*2H13N2	2689.8-2690.9	2689.8
*P*1H14N2	2772.9	2771.9
*P*2H14N2	2852.9	2851.8
*P*1H15N2	2935.0	2933.9
*P*2H15N2	3015.0	3013.9
*P*1H16N2	3097.1	3096.0
*P*2H16N2	3176.1-3177.1	3175.9
*P*1H17N2	3258.3-3260.2	3258.0

3.3.3 Structural analysis of *A. gossypii N*-glycans by NMR spectroscopy

A mixture of *N*-glycans concentrated from the supernatant of mycelia grown in defined minimal medium until late stationary phase was analysed by NMR spectroscopy, which yielded substantial data about the most abundant *N*-glycan structures present. Figures 3.3B and 3.3C respectively show the anomeric region and the methyl signals of the *N*-acetyl (NAc) groups of the 1-D ^{1}H NMR spectrum obtained, and Table 3.2 indicates the chemical shift values of the various structural reporter group signals that, despite the complexity of the mixture, remained individually observable.

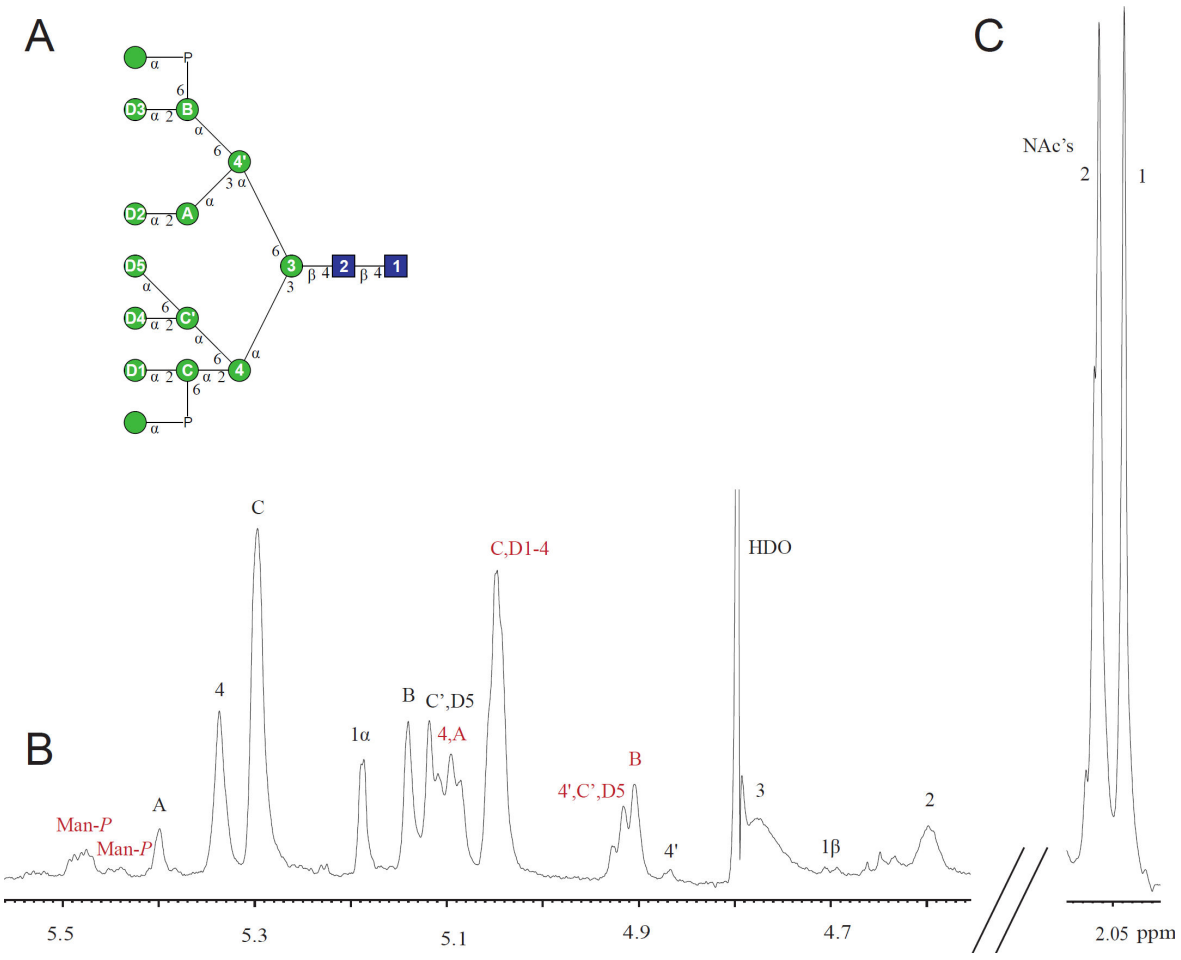


Figure 3.3 – Expansions of 600 MHz ^{1}H NMR spectrum of a mixture of *N*-glycans released from the total pool of glycoproteins secreted by *A. gossypii* in defined minimal medium by PNGase F digestion. The spectrum was recorded at 600 MHz at 25 ºC and chemical shifts were referenced to internal acetone (2.225 ppm). **A**. A model structure showing the naming of the monosaccharide units used for assignment of the structural reporter groups. ●, mannose; ■, *N*-acetylglucosamine; P, phosphate. **B**. Anomeric region of the spectrum. Assignments are given for the structural reporter group signals identified and **red** identifies terminal position localization. **C**. Methyl signals of the NAc groups of the two reducing end GlcNAc units.

Table 3.2 – ^{1}H chemical shifts of the structural reporter group protons of the constituent monosaccharides of *A. gossypii N*-glycans. Assignments were made by comparison with NMR data reported in the literature for several *N*-glycan structures and supported by 2-D homonuclear ^{1}H-^{1}H DQF-COSY and ^{1}H-^{1}H TOCSY experiments. The identified signals were consistent with high-mannose type *N*-glycan structures with average monosaccharide composition $Man_9GlcNAc_2$.

Glycan residue			^{1}H NMR chemical shift
Residue	**Linkage**	**Proton**	**Observed**
GlcNAc-**1**		H-1α	5.187
		H-1β	4.697
		NAc	2.038
GlcNAc-**2**	4	H-1	4.600
		NAc	2.064, 2.068 and 2.077
Man-**3**	4,4	H-1	4.774
		H-2	n.d.
Man-**4'**	6,4,4	H-1	4.917* / 4.87
		H-2	3.99* / 4.144
Man-**B**	6,6,4,4	H-1	4.905* / 5.141
		H-2	3.99* / 4.02
Man-**A**	3,6,4,4	H-1	5.08-5.11* / 5.400
		H-2	4.01* / 4.009
Man-**4**	3,4,4	H-1	5.08-5.11* / 5.337
		H-2	4.01* / 4.089
Man-**C**	2,3,4,4	H-1	5.05* / 5.30
		H-2	4.072* / 4.112
Man-**D1**, Man-**D2**, Man-**D3** or Man-**D4**	2,2,3,4,4, 2,3,6,4,4, 2,6,6,4,4 or 2,6,3,4,4	H-1	5.05*
		H-2	4.072*
Man-**C'**	6,3,4,4	H-1	4.929* / 5.120
		H-2	3.99* / 4.070
Man-**D5**	6,6,3,4,4	H-1	4.929* / 4.917*/ 5.120
		H-2	3.99* / 4.070
Man-***P***	n.d.	H-1	5.486, 5.448 and 5.416

* corresponding to residue in a terminal position
n.d. not determined

The 1-D ^{1}H NMR spectrum revealed proton signals consistent with a mixture of high-mannose type *N*-glycans with an average of eight α-mannoses per molecule plus the β-mannose linked to the *N,N'*-diacetylchitobiose core ($Man_9GlcNAc_2$), as estimated by integration of the structural reporter group indicator signals. The following glycosidic linkages were found from the spectral interpretation in *A. gossypii N*-glycans: Man-**4**(α1,3)[Man-**4'**(α1,6)]Man-**3**(β1,4)GlcNAc-**2**(β1,4)GlcNAc-**1,** Man-**A**(α1,3)[Man-**B**(α1,6)]Man-**4'**, Man-**D2**(α1,2)Man-**A**, Man-**D3**(α1,2)Man-**B**, Man-**C**(α1,2)Man-**4**, Man-**C'**(α1,6)Man-**4**, Man-**D1**(α1,2)Man-**C**, Man-**D4**(α1,2)/**D5**(α1,6)Man-**C'** and Man(α1-***P***-6).

The H-1 signals of the two anomers of the reducing end GlcNAc-**1** were found at 5.187 (α) and 4.697 (β) ppm, and the methyl proton resonance of its NAc group at 2.038 ppm. The characteristic H-1 signal of the second GlcNAc-**2** was detected at 4.60 ppm and its NAc signals at 2.064, 2.068 and 2.077 ppm, reflecting the structural heterogeneity of

the sample. At 10ºC, the H-1 signal for the mannose β1,4-linked to GlcNAc (Man-**3)** was found at the expected position 4.774 ppm, overlapped with the residual water signal at 25ºC.

All other major peaks observed in the ^{1}H NMR spectrum on the left side of the water signal were assigned to H-1 proton signals of α-mannoses. These assignments were confirmed by 2-D NMR experiments, where scalar couplings to characteristic H-2 protons of the mannoses could be observed (not shown). The set of chemical shifts typical of the H-l protons of the α1,6-linked Man-**4'** in the trimannosyl *N,N'*-diacetylchitobiose core gave signals in two areas. The first group of H-l signals could be found at 4.87 ppm, which is indicative of a substitution of Man-**4'** by an α1,3-linked Man-**A**, or possibly by an α1,6-linked Man-**B.** In addition, a group of three H-1 signals arising from α1,6-linked mannoses was found at 4.905, 4.917 and 4.929 ppm. The 4.917 ppm signal was consistent with the presence of Man-**4'** (Vliegenthart et al. 1983; Jongen et al. 2007) and/or Man-**D5** (Blanchard et al. 2006) at a terminal position, and the 4.905 ppm signal with a terminal Man-**B** in α1,6-linkage to Man-**4'** (Vliegenthart et al. 1983; Jongen et al. 2007). The minor H-1 signal at 4.929 ppm was consistent with Man-**C'** and/or Man-**D5** occurring in a terminal position (Trimble et al. 1991; Blanchard et al. 2006). H-1 signals consistent with internal Man-**B** substituted by a terminal α1,2-linked Man-**D3** were found at 5.141 ppm. Proton signals from Man-**A** α1,3-linked to Man-**4'** were also found. The signal at 5.400 ppm indicated the substitution of Man-**A** with an α1,2-linked Man-**D2**. The group of H-1 signals at 5.08-5.11 ppm resonate in an area that is characteristic of an α1,3-linked Man-**A** or Man-**4** occurring in a terminal position (Vliegenthart et al. 1983; Trimble and Atkinson 1986; Trimble et al. 1991; Jongen et al. 2007).

Chemical shifts typical of the H-l protons from α1,3-linked Man-**4** from the trimannosyl *N,N'*-diacetylchitobiose core were also observed at 5.337 ppm, indicating the presence of Man-**C** α1,2-linked to Man-**4**. A major signal found at 5.30 ppm corresponded to a Man-**C** residue bearing an α1,2-linked Man-**D1** residue. The H-1 and H-2 resonances of all terminal α1,2-linked mannoses were found in the same area (H-1 5.05 ppm and H-2 4.072 ppm), being impossible to distinguish between their signals. The major H-1 signal at 5.120 ppm was consistent with the presence of internal Man-**C'** and/or Man-**D5** substituted by an α1,2- and/or α1,6-linked mannose (Trimble et al. 1991; Blanchard et al. 2006).

The signals found at 5.486 and 5.448 ppm were consistent with the presence of α-mannosylphosphate (Man-***P***) groups (Hernandez et al. 1989; Blanchard et al. 2006). The ^{31}P decoupled ^{1}H spectrum (not shown) clearly showed that these signals were affected by

the decoupling, confirming that they represented protons at the sites of phosphorylation. The remaining homonuclear coupling constant of these signals was smaller than 3 Hz, confirming that they originate from mannoses. The phosphorylation was further confirmed by a ^{1}H-^{31}P HSQC experiment (Figure 4), where cross peaks were observed between these proton signals and ^{31}P atoms resonating at two different frequences. However, no cross peaks were detected at these ^{31}P chemical shifts to any other protons. Therefore, the site of phosphorylation could not be determined. In addition to these two proton signals, a week signal at 5.416 ppm was also affected by the ^{31}P decoupling, but was not detected in the HSQC spectrum, probably due to low abundance. In agreement with the MALDI-TOF results, the relative low intensities of the Man-***P*** H-1 signals support the presence of phosphorylated *N*-glycan structures predominantly containing only one phosphate residue.

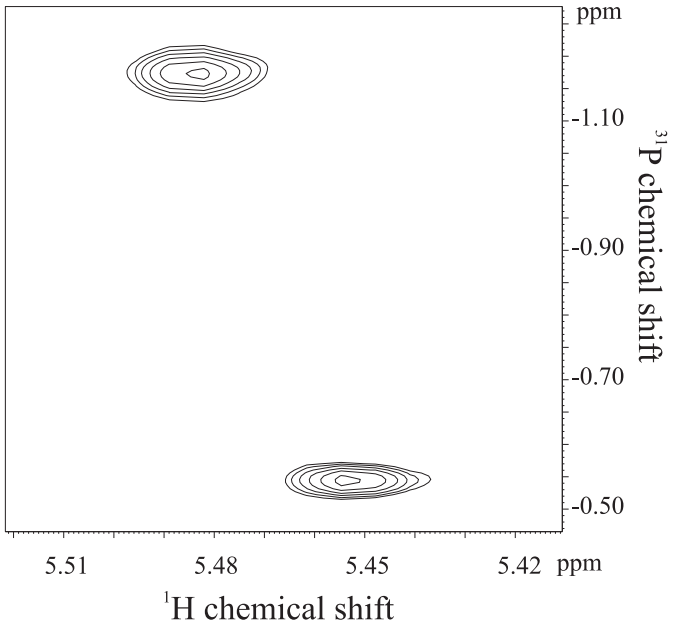


Figure 3.4 – Expansion of ^{1}H-^{31}P HSQC spectrum of a mixture of *N*-glycans released from the total pool of glycoproteins secreted by *A. gossypii* in defined minimal medium by PNGase F digestion showing the cross peaks between Man H-1 and the phosphate groups.

Chemical shifts typical of the H-l and H-2 protons of α1,3-mannose extensions at Man-**D1** and Man-**D3**, like reported for *S. cerevisiae N*-glycans (Trimble and Atkinson 1986), were not detected, as no H-1/H-2 correlation was found at 5.140-5.144/4.224-4.226 ppm (Trimble and Atkinson 1986; Blanchard et al. 2006). No evidence was obtained for the presence of sulphate. Sulphatation usually leads to large shifts of the ^{1}H NMR signals, and such shifts were absent. Fucose in either α1,3-, α1,4- or α1,6-linkage to GlcNAc

residues was also undetectable, by virtue of the absence of fucose methyl groups in the region of the 1.17-1.28 ppm (Vliegenthart et al. 1983; Manzi et al. 2000). Nevertheless, due to the low detection limit of NMR, the presence of small amounts of glycan structures containing these or other residues cannot be ruled out.

3.4 DISCUSSION

The data presented in this study represent the first characterization of the *N*-glycans synthesised by *A. gossypii*. Here, we demonstrate that *A. gossypii* secreted proteins carry high-mannose type *N*-glycan structures in the range $Man_{4-18}GlcNAc_2$, predominantly with neutral core-type $Man_{8-10}GlcNAc_2$ *N*-glycans not terminally capped by α1,3-linked mannose residues. *A. gossypii* also produces acidic mannosylphosphorylated *N*-glycan structures that, depending on the growth conditions, can be predominantly longer than the neutral structures ($Man_{13-14}PGlcNAc_2$).

In eukaryotes, the *N*-glycan biosynthetic pathway begins with the synthesis of the precursor *N*-glycan $Glc_3Man_9GlcNAc_2$ and its transfer by the oligosaccharyltransferase complex to the amino group of asparagine residues of nascent polypeptides entering the ER (Helenius and Aebi 2004). Subsequent trimming of the precursor *N*-glycan by glucosidases I and II and a specific ER-residing α1,2-mannosidase lead to the formation of a core glycan ($Man_8GlcNAc_2$) (Gemmill and Trimble 1999). Glycoproteins containing $Man_8GlcNAc_2$ (or $Man_9GlcNAc_2$, in *Schizosaccharomyces pombe* and *Kluyveromyces lactis*) are then transported to the Golgi apparatus (Gemmill and Trimble 1999; De Pourcq et al. 2010).

Most yeast do not trim the precursor $Glc_3Man_9GlcNAc_2$ shorter than the core $Man_8GlcNAc_2$, unlike mammalian cells and filamentous fungi. However, small amounts of *N*-glycans smaller than $Man_8GlcNAc_2$ have also been found in some wild-type yeast. *N*-Glycans in the range $Man_{5-7}GlcNAc_2$ were found in *Cryptococcus neoformans* cell wall and secreted proteins (Park et al. 2012), $Man_{6-7}GlcNAc_2$ *N*-glycans were detected in recombinant aspartic protease produced by *Pichia pastoris* (Montesino et al. 1999), and *N*-glycans with composition $Man_7GlcNAc_2$ were detected in *Hansenula polymorpha* (Kim et al. 2004) and *Yarrowia lipolytica* (Song et al. 2007) secreted proteins. Filamentous fungi, on the other hand, predominantly produce mammalian high-mannose type ($Man_{5-9}GlcNAc_2$) *N*-glycan structures (Stals et al. 2004; Kainz et al. 2008). They usually

possess two types of α1,2-mannosidases: one residing in the ER, like that found in yeast, and which trims $Man_9GlcNAc_2$ to $Man_8GlcNAc_2$, and another that resembles the mammalian Golgi α1,2-mannosidases, which trims $Man_8GlcNAc_2$ to $Man_5GlcNAc_2$ (De Pourcq et al. 2010).

N-Glycans smaller than the core $Man_8GlcNAc_2$ were also found in *A. gossypii*, especially when it was grown in complex rich medium. This suggests that trimming of the core $Man_8GlcNAc_2$ *N*-glycan may occur in *A. gossypii*. Genomic analysis indicated that *A. gossypii* has a glycosylation machinery very similar to that of *S. cerevisiae* (Table 3.3), but also revealed a putative Golgi-localized α-mannosidase of the glycoside hydrolase family 92 (GH92), encoded by the *A. gossypii ADR339C* gene, for which there is no homolog in *S. cerevisae*. The potential role of this putative enzyme on *N*-glycan trimming in *A. gossypii* is a matter of further study.

In yeast, processing in the Golgi involves the addition of an α1,6-mannose residue by the Och1 protein. This branch is then elongated by the stepwise addition of mainly mannose residues, leading to the formation of small core-type ($Man_{8\text{-}14}GlcNAc_2$) or large hypermannosylated mannan-type ($Man_{>15}GlcNAc_2$) outer chains (Gemmill and Trimble 1999). Hypermannosylation is not a typical feature for filamentous fungi (Maras et al. 1999), although it may occur if Och1 activity is present (Wallis et al. 2001; Kotz et al. 2010). *A. gossypii* produced small amounts of *N*-glycan structures containing more than 15 hexoses (Figures 3.1 and 3.2). Moreover, proton signals consistent with the presence of α1,6-mannose extensions from the α1,3-mannose residue of the trimannosyl *N,N'*-diacetylchitobiose core were detected by NMR, indicating that outer chain processing of *N*-glycans occurred in *A. gossypii*. Indeed homologs are found in its genome for all *S. cerevisiae* genes known to be involved in outer chain biosynthesis (Table 3.3). However, our results indicate that outer chain processing differs considerably in *A. gossypii* and in *S. cerevisiae*, since shorter *N*-glycans were predominantly produced by *A. gossypii* when comparing to *S. cerevisiae* (Trimble and Atkinson 1986; Ballou et al. 1990; Montesino et al. 1999). These were more similar in extent to those produced by nonconventional yeast species such as the methylotrophic species *P. pastoris* (Montesino et al. 1999; Choi et al. 2003; Blanchard et al. 2006) and *H. polymorpha* (Kim et al. 2004), and the dimorphic yeast *Y. lipolytica* (Song et al. 2007).

Table 3.3 – Putative *A. gossypii* genes predicted to be involved in the *N*-glycan biosynthesis pathway. The table shows all of the *S. cerevisiae* genes involved in the *N*-glycosylation pathway for which a homolog was identified in the *A. gossypii* genome. No homolog genes were identified in *S. cerevisiae* for two putative *A. gossypii* genes predicted to encode glycoside hydrolases with a possible role in *N*-glycosylation.

Predicted protein function	*S. cerevisiae* gene	*A. gossypii* gene
Synthesis and transport of sugar nucleotide donors		
UDP-GlcNAc		
Glutamine-fructose-6-phosphate amidotransferase	*GFA1, YMR084W, YMR085W*	*ABL036C*
Glucosamine-6-phosphate *N*-acetyltransferase	*GNA1*	*AFR510W*
N-acetylglucosamine-phosphate mutase	*PCM1*	*ACR015W*
UDP-*N*-acetylglucosamine pyrophosphorylase	*QRI1/UAP1*	*AFL192C*
GDP-Man		
Mannose-6-phosphate isomerase	*PMI40*	*ADR003C*
Phosphomannomutase	*SEC53*	*ABR236W*
Phosphoglucomutase	*PGM1, PGM2*	*ABL029W*
Phosphoglucomutase	*PGM3*	*ABL126W*
Mannose-1-phosphate guanyltransferase	*PSA1/MPG1*	*AFR599W*
Guanosine diphosphatase	*GDA1*	*AFR362C*
Apyrase with wide substrate specificity	*YND1*	*ADR006W*
GDP-mannose transporter	*VRG4, HVG1, YER039C-A*	*AFR236C*
Putative nucleotide sugar transporter similar to Vrg4p	*YMD8*	*AFR462C*
UDP-Glc		
Glucokinase	*GLK1, EMI2*	*AFR716C*
Hexokinase	*HXK1, HXK2*	*AFR279C*
UDP-glucose pyrophosphorylase	*UGP1, YHL012W*	*AGL082W*
Phosphoglucose isomerase	*PGI1*	*AEL249C*
Synthesis of the dolicholphosphate linked *N*-glycan precursor $Glc_3Man_9GlcNAc_2$		
Dolichol kinase	*SEC59*	*ABR051C*
Dolichol-phosphate-mannose (Dol-P-mannose) synthase	*DPM1*	*AEL275C*
Putative Dol-P-mannose biosynthesis regulatory (DPM2) protein	*YIL102C-A*	*ACR063W-A*
UDP-*N*-acetylglucosamine-1-phosphate transferase	*ALG7*	*AFL037W*
UDP-*N*-acetylglucosamine transferase: catalytic subunit	*ALG13*	*AGL042W*
UDP-*N*-acetylglucosamine transferase: transmembrane part	*ALG14*	*AGL202W*
GDP-mannose:$GlcNAc_2$-PP-Dol β1,4-mannosyltransferase	*ALG1*	*ADL338C*
GDP-mannose:$Man_{1/2}GlcNAc_2$-PP-Dol α1,3/1,6-mannosyltransferase	*ALG2*	*AFL098W*
GDP-mannose:$Man_{3/4}GlcNAc_2$-PP-Dol α1,2-mannosyltransferase	*ALG11*	*ADL235W*
Protein required for translocation of $Man_5GlcNac_2$-PP-Dol to the lumenal side of the ER	*RFT1*	*AFR015W*
Dol-P-mannose:$Man_5GlcNAc_2$-PP-Dol α1,3-mannosyltransferase	*ALG3*	*AGL299C*
Dol-P-mannose:$Man_{6/8}GlcNAc_2$-PP-Dol α1,2-mannosyltransferase	*ALG9*	*ABL185C*
Dol-P-mannose:$Man_7GlcNAc_2$-PP-Dol α1,6-mannosyltransferase	*ALG12*	*AAR043C*
Dolichol-phosphate β-glucosyltransferase	*ALG5*	*AFL133C*
Dol-P-glucose:$Man_9GlcNAc_2$-PP-Dol α1,3-glucosyltransferase	*ALG6*	*ACR004W*
Dol-P-glucose:$Glc_1Man_9GlcNAc_2$-PP-Dol α1,3-glucosyltransferase	*ALG8*	*ADR210C*
Dol-P-glucose:$Glc_2Man_9GlcNAc_2$-PP-Dol α1,2-glucosyltransferase	*DIE2/ALG10*	*ADL138C*
Dolichol pyrophosphate phosphatase	*CAX4/CWH8*	*ACL186W*

Table 3.3 – *Continued.*

Predicted protein function	S. cerevisiae gene	A. gossypii gene
Oligosaccharyltransferase (OST) subunits		
Subunit of the OST complex	*OST4*	*ABL170C*
Beta subunit of the OST complex	*WBP1*	*ADL003C*
Zeta subunit of the OST complex	*OST5*	*AFR241W*
Alpha subunit of the OST complex - Ribophorin I	*OST1*	*AAL170W*
Subunit of the OST complex	*OST6*	*AER330W*
Delta subunit of the OST complex - Ribophorin II	*SWP1*	*ADR268W*
Gamma subunit of the OST complex	*OST3*	*AER413C*
Epsilon subunit of the OST complex	*OST2*	*ADL261C*
Subunit of the OST complex	*STT3*	*AFR491W*
Processing of the *N*-glycan precursor $Glc_3Man_9GlcNAc_2$ after transfer to a polypeptide		
ER		
Processing alpha glucosidase I	*CWH41*	*AFR483C*
Glucosidase II - alpha subunit	*ROT2*	*AAR173C*
Glucosidase II - beta subunit	*GTB1*	*AGR178W*
α1,2-mannosidase	*MNS1*	*AER165W*
Putative α-mannosidase → GH92 family member	**No homolog**	***ADR339C***
α1,2-mannosidase-like protein (ER degradation enhancer)	*MNL1/HTM1*	*ADL390W*
Putative mannosidase involved in ER-associated protein degradation	*YLR057W/MNL2*	*ACR213W*
Golgi		
Initiating α1,6-mannosyltransferase	*OCH1*	*AFR465C*
Subunit of the Golgi α1,6-mannosyltransferase complexes M-Pol I and M-Pol II	*MNN9*	*AGL259C*
Subunit of the Golgi α1,6-mannosyltransferase complex M-Pol I	*VAN1*	*ABL124W*
Subunit of the Golgi α1,6-mannosyltransferase complex M-Pol II	*MNN10*	*AFR454W*
Subunit of the Golgi α1,6-mannosyltransferase complex M-Pol II	*MNN11*	*AEL142W*
Subunit of the Golgi α1,6-mannosyltransferase complex M-Pol II	*ANP1*	*AFR357W*
Subunit of the Golgi α1,6-mannosyltransferase complex M-Pol II	*HOC1*	*AFR435W*
α1,2-mannosyltransferase	*MNN2*	*AER018C, AEL082W*
α1,2-mannosyltransferase	*MNN5*	*AEL148W*
Putative α1,2-mannosyltransferase and/or mannosylphosphate transferase	*KRE2/MNT1, KTR6/MNN6*	*AGL267C*
Putative positive regulator of the mannosylphosphate transferase	*MNN4, YJR061W*	*AFR416C*
Golgi α1,3-mannosyltransferase	*MNT3, MNN1*	*AFL235W*
α1,2-mannosyltransferase	*KTR1*	*ADL265W-A*
α1,2-mannosyltransferase	*KTR2, YUR1*	*ACR251C*
Putative α1,2-mannosyltransferase	*KTR3*	*ADR048W*
Putative mannosyltransferase	*KTR4*	*AEL243W*
Cytosol and vacuole		
Vacuolar α-mannosidase, involved in free oligosaccharide degradation	*AMS1*	*AGR385C*
Unknown		
Putative endo-β-*N*-acetylglucosaminidase → GH85 family member	**No homolog**	***AFR597W***

S. cerevisiae usually adds terminal α1,3-linked mannoses to its *N*-glycans through the action of Mnn1p (Gemmill and Trimble 1999). The *A. gossypii* gene *AGL267C* is a

nonsyntenic homolog of the *S. cerevisiae MNT3* and *MNN1* genes, with a higher degree of homology to *MNT3* than to *MNN1*. Both genes encode α1,3-mannosyltransferases in *S. cerevisiae*, but Mnt3p has only been described as involved in *O*-glycosylation (Romero et al. 1999). In this study, no terminal α1,3-mannose caps were detected in *A gossypii* *N*-glycans by NMR. Therefore, given that only the most abundant sugar species can be detected by this technique, extensive α1,3-mannose capping did not occur in *A. gossypii*.

N-glycans may be further modified by the addition of acidic phosphate groups in yeast and other fungi (Gemmill and Trimble 1999; Maras et al. 1999). *A. gossypii* also added phosphate groups to its *N*-glycans (Figures 2 and 3). In *S. cerevisiae* (Hernandez et al. 1989), *P. pastoris* (Blanchard et al. 2006) and *Trichoderma reesei* (Maras et al. 1997) glycoproteins, Man-***P*** extensions occur on *O*-6 of Man-**B** and/or Man-**C** (in *T. reesei* only on Man-**B**) within the *N*-glycan chain. In this study, two rather strong proton signals with different chemical shifts representing phosphorylation sites were detected by NMR in *A. gossypii* *N*-glycans (Figure 3), however the site(s) of phosphorylation could not be determined.

The extent of mannosylphosphorylation in *P. pastoris* (Montesino et al. 1999) and *T. reesei* (Stals et al. 2004) depends on culture conditions, including medium composition and pH. In *S. cerevisiae*, mannosylphosphorylation is related to a cellular stress response, increasing under conditions of high osmolarity and in late exponential and stationary phases of growth (Odani et al. 1997; Jigami et al. 1999). In this study, the culture conditions and cultivation stage also affected the extent of phosphorylation and extension of the *A. gossypii* *N*-glycans, highlighting the importance of these factors on the regulation of the *N*-glycosylation pathway in this fungus. Growth in defined minimal medium resulted in a lower specific growth rate and promoted more *N*-glycan extension and mannosylphosphorylation than growth in complex rich medium, especially during the stationary phase, as observed in *S. cerevisiae* (Odani et al. 1997). Similarly, *T. reesei* prominently mannosylphosphorylated cellobiohydrolase I (CBHI) under minimal growth conditions, but only at low pH (Stals et al. 2004). This modification was proposed to be related to a stress response of *T. reesei* to the low pH conditions of the minimal medium (Stals et al. 2004). Here, the culture pH was kept constant at 6 ± 0.1 in all *A. gossypii* cultivations, so this should not have been the case in *A. gossypii*. However, in light of the observations made in *S. cerevisiae*, where mannosylphosphorylation was observed to occur in response to high osmolarity conditions (Odani et al. 1997), a lower osmoprotection offered by the minimal medium to *A. gossypii* cells in comparison to the rich medium,

which contained higher amount of known osmoprotectors (derived from the tryptone and yeast extract), may help explain why extensive mannosylphosphorylation only occurred under minimal growth conditions in *A. gossypii*.

When defined medium was used, diphosphorylated *N*-glycans containing only one *N*-acetylglucosamine were detected by MALDI-TOF mass spectrometry, suggesting the presence of endo-*β*-*N*-acetylglucosaminidase (ENGase) activity in *A. gossypii* under these conditions. Single *N*-acetylglucosamine residues are often found at the *N*-glycosylation sites of *T. reesei* proteins (Stals et al. 2004). These have been attributed to the action of a GH18 family protein with extracellular ENGase-type activity (Stals et al. 2004; Stals et al. 2010). *Mucor hiemalis* also produces an ENGase, but of the GH85 family (Endo M), which catalyses the release of *N*-glycans with only one *N*-acetylglucosamine residue at their reducing end, leaving one *N*-acetylglucosamine attached to the protein (Fujita et al. 2004). The *A. gossypii AFR597W* gene, for which there is no homolog in *S. cerevisiae*, is predicted to encode a putative intracellular GH85 family ENGase, and we have observed that this gene is transcribed (unpublished data), indicating that it is probably expressed in *A. gossypii*. Thus, AFR597W is a good candidate ENGase for catalyzing the liberation of *N*-glycans containing only one *N*-acetylglucosamine residue and details regarding its functionality warrant further investigation.

The *N*-glycan structures on *A. gossypii* secreted glycoproteins and the genomic information regarding their biosynthetic pathway indicate that protein glycosylation in *A. gossypii* is intermediate to that of *S. cerevisiae* and other filamentous fungi species. *A. gossypii* possibly trims *N*-glycans as other filamentous fungi do, although it generally produces yeast-like *N*-glycan structures slightly longer than those typical of filamentous fungi. Hypermannosylation occurs, but is more limited than in *S. cerevisiae*. Given these observations, it would be intriguing to elucidate in detail the *A. gossypii N*-glycosylation pathway and investigate the role of glycosylation in the development of *A. gossypii* multicellular hyphae.

CHAPTER 4

Molecular and functional characterization of the *Ashbya gossypii* invertase

ABSTRACT

The repertoire of hydrolytic enzymes natively secreted by the filamentous fungus *Ashbya gossypii* has been poorly explored. Here, an invertase secreted by this flavinogenic fungus was for the first time molecularly and functionally characterized. Invertase activity was detected in *A. gossypii* culture filtrates and cell-associated fractions, being predominantly extracellular. Hydrolytic activity towards sucrose was approximately 10 times higher than towards raffinose. Inulin and levan were not hydrolyzed. The invertase activity in the different fractions was slightly inhibited by the presence of glucose at 7 ± 2 to 17 ± 1 g/l. The presence of fermentable sugars in the medium negatively regulated the production of secreted invertase. The *A. gossypii* invertase was demonstrated to be encoded by the *AFR529W* (*AgSUC2*) gene, which is highly homologous to the *Saccharomyces cerevisiae SUC2* (*ScSUC2*) gene. *Agsuc2* null mutants were unable to hydrolyze sucrose, proving that invertase is encoded by a single gene in *A. gossypii*. This mutation was functionally complemented by the *ScSUC2* and *AgSUC2* genes, when expressed in a 2 micron plasmid. The signal sequences of both AgSuc2p and ScSuc2p were able to direct the secretion of invertase into the culture medium in *A. gossypii*.

4.1 INTRODUCTION

The filamentous hemiascomycete *Ashbya gossypii* (syn. *Eremothecium gossypii*), a well known riboflavin overproducer (Demain 1972), was sequenced in 2004 (Dietrich et al. 2004). The remarkably high degree of gene homology and gene order conservation existent between its genome and the genome of the baker's yeast *Saccharomyces cerevisiae* (Brachat et al. 2003; Dietrich et al. 2004) has facilitated the assignment of potential functions to *A. gossypii* open reading frames (ORFs). However, up to now only a small percentage of ORFs have been experimentally characterized in *A. gossypii*. Furthermore, although this fungus has been recently considered as a host for the expression and secretion of recombinant proteins (Ribeiro et al. 2010), only one native protein secreted by *A. gossypii*, a lipase, has been characterized so far (Stahmann et al. 1997). Nevertheless, extracellular amylase (Ribeiro et al. 2011, Ribeiro et. al 2013) and *β*-glucosidase (Ribeiro et. al 2013) enzymatic activities have been detected in *A. gossypii* ATCC 10895 culture filtrates. The proteins responsible for the detected activities have not yet been characterized.

Invertase, or *β*-fructofuranosidase, (EC 3.2.1.26) is an industrially important enzyme secreted by many fungi that has wide applications in the food, pharmaceutical and bioethanol production sectors. It catalyzes the release of terminal *β*-fructose residues from various β-$_D$-fructofuranoside substrates, such as sucrose and raffinose. Fungal invertases have been widely studied in yeast (Perlman et al. 1984; Reddy and Maley 1996; Zhang and Chi 2004; Linde et al. 2009) and filamentous fungi (Vainsteein and Peberdy 1991; Heyer and Wendenburg 2001; Guimarães et al. 2007; Orikasa and Oda 2012), and based on the homology of their aminoacid sequences they have been classified within the family 32 of the glycoside hydrolases (GH32) (Cantarel et al. 2009).

A. gossypii is able to utilize sucrose as carbon source (Mickelson 1950; Pridham and Raper 1950), which should reflect the presence of either intra- or extracellular invertase activity in this fungus. A putative invertase-enconding gene, homologous to the *S. cerevisiae SUC2* (*ScSUC2*) gene, can be identified in its genome. The deduced amino acid sequence of the protein it presumably encodes also shares high identity with that of the *S. cerevisiae* Suc2 invertase (ScSuc2p).

The *ScSUC2* gene encodes two different invertase isoforms: a constitutively expressed cytoplasmic form, which is non-glycosylated, and a glucose-repressible glycosylated form, which is secreted into the periplasmic space (Carlson and Botstein

1982; Lutfiyya and Johnston 1996). Several factors affect invertase secretion in different fungi, and the carbon source used has been found to play an important role as repressor or inducer of its synthesis (Gancedo 1998; Rubio and Navarro 2006).

Here, we describe for the first time the molecular and functional characterization of the *A. gossypii* invertase. Studies of its secreted and cell-associated hydrolytic activities were performed, and the functionality of the *A. gossypii ScSUC2* homolog gene (*AgSUC2*) was demonstrated through its deletion and complementation by recombinant expression.

4.2 MATERIALS AND METHODS

4.2.1 Strains

A. gossypii ATCC 10895 was obtained from Prof. Peter Philipsen (University of Basel). *A. gossypii* strains *Agsuc2* (*Agsuc2::GEN3*), *Agsuc2*pTAgSUC (*Agsuc2::GEN3*, pTAgSUC) and *Agsuc2*pTScSUC2 (*Agsuc2::GEN3*, pTScSUC2) were generated in this study using the ATCC 10895 strain as the parent. *Escherichia coli* TOP10 (Invitrogen) was used as the recipient for all cloning steps.

4.2.2 Media and culture conditions

E. coli was grown at 37ºC in LB medium supplemented with 100 µg/ml of ampicillin (Sigma-Aldrich) for selection. *A. gossypii* was maintained on agar-solidified (15 g/l agar) *Ashbya* full medium (AFM; 10 g/l yeast extract, 10 g/l tryptone, 1 g/l myo-inositol, 20 g/l glucose). Spore suspensions were prepared and stored as described by Ribeiro et al. (2010), except that the mycelium was digested using 4.5 mg/ml Lysing Enzymes from *Trichoderma harzianum* (Sigma-Aldrich). For selection of *A. gossypii* transformants, the antibiotics geneticin (G418) (Sigma-Aldrich) or nourseothricin (clonNAT) (WERNER BioAgents) were used at a final concentration of 200 and 100 µg/ml, respectively. In liquid medium, selection was maintained with 50 µg/ml of clonNAT.

Submerged cultures were inoculated with 10^6 spores and grown at 30ºC and 200 rpm in 250 ml Erlenmeyer flasks containing 50 ml of AFM or synthetic complete (SC) medium (Sherman et al. 1986) with 20 g/l of glucose, sucrose or glycerol as carbon source.

0.1 M sodium-phosphate at pH 7.0 was used to buffer SC medium. For bioreactor cultivations, pre-cultures were grown in liquid AFM for 17 h before the mycelia were harvested by filtration through disks of sterile disposable cleaning cloth (X-tra, Inex Partners Oy, Finland) and washed with sterile distilled water. Bioreactors (Sartorius AG, 2 l Biostat® B-DCU, 1.0 l working volume) containing defined minimal medium (Verduyn et al. 1992) with 20 g/l of sucrose as carbon source were inoculated to an initial biomass of 0.62 ± 0.05 g/l. Polypropylene glycol (mixed molecular weights (Wiebe et al. 2001)) was added to a final concentration of 0.2% (v/v) to prevent foam production. Cultures were grown at 30°C, with 500 rpm agitation and aeration of 1 volume of air per volume of liquid per minute (vvm). Culture pH was maintained at 6.0 ± 0.1 by the addition of sterile 1 M KOH or 1 M H_3PO_4.

4.2.3 Analyses

Culture optical density at 600 nm (OD_{600}) was used to monitor growth of submerged cultures. Biomass concentration was determined by dry cell weight. Mycelium was collected by filtration through disks of disposable cleaning cloth (X-tra, Inex Partners Oy) or filter paper (Advantec qualitative grade 2), washed with two sample volumes of distilled water and dried to a constant weight at 105°C.

Supernatant was obtained by filtration of culture broth through 0.2 μm cellulose acetate filters (Whatman or Advantec) for analysis of enzymatic activities, substrates and metabolites (sucrose, glucose, fructose, glycerol and ethanol). Substrates and metabolites were quantified by high performance liquid chromatography (HPLC) using a Fast Acid Analysis Column (100 mm x 7.8 mm, Bio-Rad) linked to an Aminex HPX-87H Organic Acid Analysis Column (300 mm x 7.8 mm, Bio-Rad) or a MetaCarb 87H column (300 x 7.8 mm; Varian) at 35 ºC, with 5 mM H_2SO_4 as mobile phase and a flow rate of 0.5 ml/min.

To assess biomass associated invertase activity, biomass was collected by filtration, washed with cold 10 mM sodium azide in 0.05 M phosphate-citrate buffer (pH 6.0) and then resuspended in 0.5 ml of the same solution.

4.2.3.1 Invertase assays

To determine the secreted and extracellular cell wall-bound invertase activity, 50 µl of filtered supernatant or treated mycelial suspensions were mixed with 100 µl of 0.1 M phosphate-citrate buffer (pH 6.0) and 50 µl of 0.2 M sucrose. For total cell-associated invertase activity (including both intracellular and extracellular cell wall-bound invertase), 250 µl of resuspended mycelia were first disrupted with 0.25 g of 20 mm diameter glass beads (Sigma-Aldrich) using a FastPrep FP120 (Qbiogene; 4 cycles at speed 6 for 10 s, with ice cooling between cycles). Resuspended lysed mycelia (50 µl) were then used for the assay. The reactions were carried out at 40°C for 90 min and stopped by boiling for 10 min. As a blank, each reaction mixture was incubated without the sample, which was only added to the mix immediately before stopping the reaction. Substrate specificity assays were performed with culture filtrates as described above, but using 50 µl of 0.2 M raffinose, 1% (w/v) levan or 1% (w/v) inulin as substrate instead of sucrose.

The amount of reducing sugars formed was determined by the 3,5-dinitrosalicylic acid (DNS) method (Miller 1959), using glucose as standard. One unit (U) of enzyme activity was defined as the amount of enzyme that hydrolyzed sucrose to yield 1 µmol of glucose (or reducing sugars, for the other substrates) per minute under the stated conditions. Specific enzyme activities are expressed as U per gram of dried biomass (U/g).

4.2.3.2 Electrophoretic analyses

The proteins secreted by *A. gossypii* to the cultures supernatants were concentrated using Amicon® Ultra-15 10,000 MWCO centrifugal filter devices (Millipore) according to the manufacturer's instructions. For buffer exchange, the concentrated fractions were repeatedly diluted with 0.05 M phosphate-citrate buffer (pH 6.0) and concentrated again. All procedures were carried out at 4°C.

Invertase activity was detected *in situ* after native polyacrylamide gel electrophoresis (PAGE) on 8% (w/v) polyacrylamide gels. After incubation for 24 h at 30ºC in 0.1 M phosphate-citrate buffer (pH 6.0) containing 0.2 M sucrose, the reducing sugars in the gels were stained with 0.1% (w/v) 2,3,5-triphenyltetrazolium chloride (TTC; Sigma-Aldrich) in 0.25 M NaOH. Sucrolytic activity was reflected by the formation of red formazan bands resulting from the reaction of TTC with the reducing sugars. The reaction

was stopped with 7.5% (v/v) acetic acid. Invertase from *S. cerevisiae* (Sigma-Aldrich # I4504) was used as a positive control.

4.2.3 Plasmid constructions

Expression plasmids containing the *S. cerevisiae* 2 micron replication origin were generated as follows. The *NATPS* cassette conferring resistance to clonNAT was amplified from pUC19NATPS (Hoepfner *in* Kaufmann 2009) with the primers N1/N2 (Table 4.1). The *kan*MX and the *S. cerevisiae URA3* expression modules in pMI516 (Ribeiro et al. 2010) were exchanged with the *NATPS* module using the *Bgl*II/*Bgl*II sites, generating plasmid pMINATPS. Based on the annotated sequence for the *AFR529W* (*AgSUC2*) ORF in the AGD database (http://agd.vital-it.ch; Gattiker et al. 2007) and *YIL162W* (*ScSUC2*) ORF in the SGD database (www.yeastgenome.org), the complete coding regions of *AgSUC2* (1719 bp) and *ScSUC2* (1599 bp) were amplified by PCR from *A. gossypii* ATCC 10895 and *S. cerevisiae* CEN.PK 113-7D genomic DNA using the primers I1/I2 and S1/S2 (Table 4.1), respectively. These fragments were cloned into the *Eco*RI site of pMINATPS, between the *ScPGK1* promoter and terminator sequences. The resulting plasmids were named pTAgSUC and pTScSUC2, respectively. The sequence and orientation of the inserts in the plasmids were confirmed by sequencing (Eurofins MWG Operon) with primers P1, P2, I1 and I2 (Table 4.1).

4.2.4 *A. gossypii* transformation

For the deletion of the complete coding region of the *AgSUC2* gene, a disruption cassette containing the *GEN3* expression module (conferring resistance to G418) flanked by 40 bp sequences with homology to the upstream and downstream regions of the *AgSUC2* ORF was obtained by PCR with primers G1/G2 (Table1), using pGEN3 (Wendland et al. 2000) as template. The amplified cassette was purified with the QIAquick PCR purification Kit (Qiagen) and used to transform *A. gossypii* by electroporation, as described in Wendland et al. (2000). Homokaryotic *Agsuc2* null mutants were obtained through the isolation of single spores from primary heterokaryotic transformants, which were germinated under selective conditions. Diagnostic PCR with primers V1/V2, V3/V4

and V5/V4 (Table 4.1) was used to verify the correct integration of the disruption cassette and absence of the target gene in homokaryotic deletion mutants.

A. gossypii Agsuc2 null mutants were transformed by electroporation (Wendland et al. 2000) with plasmids pTAgSUC, pTScSUC2 and pMINATPS (empty vector) and positive clones were selected in AFM containing clonNAT. Additionally, the ability of the clones to use sucrose as sole carbon source was assessed in flask cultivations with AFM containing either glucose or sucrose. Two strains, one expressing the homologous AgSuc2 invertase and other expressing the heterologous ScSuc2 invertase, were selected for further study.

Table 4.1 – Primers used in this study. Upper case sequences correspond to sequences complementary to the template. Lower case sequences correspond to additions for restriction sites (underlined) or to the disruption cassette guide sequences with homology to the *AgSUC2* flanking regions (**bold**).

Primer	Sequence (5' – 3')
N1 - NATPS-BglII_FW	gaagatcttcCCTGCAGAACCGTTACGGTA
N2 - NATPS-BglII_RV	gaagatcttcCCTGCAGCCAAACAGTGTT
I1 - INV-EcoRI_FW	cggaattccgTGCATCACTTAACATCAATCAGCA
I2 - INV-EcoRI_RV	cggaattccGCGCACGTATTGTGCTTTACTAG
S1 - SUC2-EcoRI_FW	cggaattccgAAAGCTTTTCTTTTCACTAACG
S2 - SUC2-EcoRI_RV	cggaattccgCCTTTAGAATGGCTTTTGAA
P1 - PGK_FW	GTTTAGTAGAACCTCGTGAAAC
P2 - PGK_RV	GGCATTAAAAGAGGAGCG
G1 - GEN3_FW	**cgactgcgataagagatgcatcacttaacatcaatcagca**GCTAGGGATAACAGGGTAAT
G2 - GEN3_ RV	**ggaacactctacgtaggcgcacgtattgtgctttactaga**AGGCATGCAAGCTTAGATCT
V1 - INV_FW	GAGGCCGTTGTCGTGTAGAG
V2 - Kan_RV	GTTTAGTCTGACCATCTCATCTG
V3 - Kan_FW	TCGCAGACCGATACCAGGATC
V4 - INV_RV	TCCGGAACATCACATAAGCA
V5 - INV_FW	GTTCCAGGGATTCTACAACA

4.2.5 Bioinformatic analyses

The deduced amino acid sequence encoded by the putative invertase gene from *A. gossypii* (GenBank accession no. AAS53900.1) was analyzed by using the BLAST server against the *S. cerevisiae* S288c invertase (GenBank accession no. DAA08390.1), and the

two sequences were aligned using the Clustal Omega (Sievers et al. 2011) server (http://www.ebi.ac.uk/Tools/msa/clustalo/; default parameters).

Prediction of signal peptides was done using the SignalP 3.0 (Bendtsen et al. 2004) server (www.cbs.dtu.dk/services/SignalP-3.0/) and prediction of putative *O*- and *N*-glycosylation sites was done using the NetOGlyc 3.1 (Julenius et al. 2005) (www.cbs.dtu.dk/services/NetOGlyc/) and NetNGlyc 1.0 (Gupta et al. 2004) (www.cbs.dtu.dk/services/NetNGlyc/) servers, respectively.

4.3 Results and Discussion

4.3.1 Invertase production by *A. gossypii*

The filamentous fungus *A. gossypii* is able to consume sucrose as carbon source (Mickelson 1950; Pridham and Raper 1950), but invertase activity had not yet been described from this organism. We observed invertase activity in the culture supernatant, as well as associated with cell biomass (Fig. 1). Substrate specificity assays revealed that *A. gossypii* culture filtrates displayed approximately 10 times more hydrolytic activity towards sucrose (213 ± 4 mU/l) than towards raffinose (21 ± 1 mU/l), whereas inulin and levan were not hydrolyzed. Sucrose was subsequently used as substrate in all enzymatic assays.

In bioreactor cultivations of *A. gossypii* mycelium in defined minimal medium containing sucrose as primary carbon source, maximum specific invertase activities (36 ± 1 to 48 ± 1 U/g) were detected in the culture supernatant at the stationary phase (Figure 4.1A), when sucrose, glucose and fructose had been depleted from the medium (Figure 4.1B). These levels of activity were about 3-fold higher than those quantified when sugars were still present in the medium and were maintained for about 25 h of culture before slowly decreasing. In a similar way, the invertase activity of the total cell-associated fraction (which included both intracellular and extracellular cell wall-bound invertase) also reached the maximum levels during this period (76 ± 6 to 91 ± 12 U/g), being approximately 2-fold higher than the activity detected in the culture supernatant (Figure 4.1A). A prominent increase in the invertase activity from both culture supernatant and total cell-associated fraction was observed between the 15 h and the 20 h of culture (Figure 4.1A). This coincided with the depletion of the glucose and fructose from the medium

(Figure 4.1B). Similar observations were made from bioreactor cultivations in AFM with sucrose as the primary carbon source for cell-associated invertase activity. In the culture supernatant only low invertase activity (< 5 U/g) was observed (data not shown).

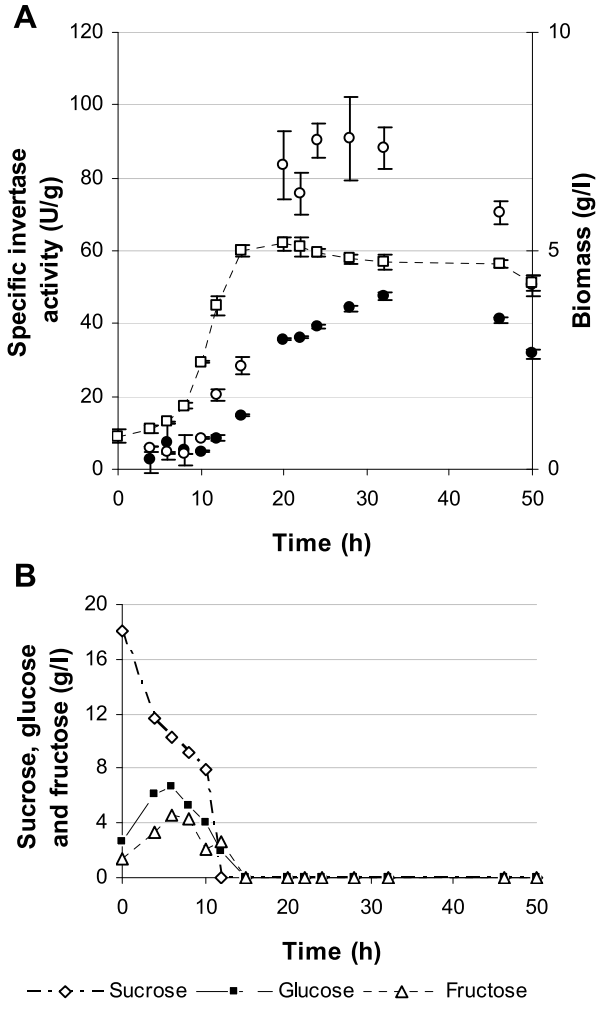


Figure 4.1 – Secreted (●) and cell-associated (○) specific invertase activity during *A. gossypii* ATCC 10895 growth in defined minimal medium containing sucrose as primary carbon source, at pH 6.0, 30°C, 500 rpm with 1 vvm aeration. In panel A, specific invertase activity (*circles*) is read on the primary axis and biomass concentration (*squares*) on the secondary axis. Error bars represent the standard error of the mean of three independent measurements. Panel B shows the sugar concentration in the culture supernatants at each time point the activities in panel A were measured.

The patterns of secreted and cell-associated invertase activity observed during *A. gossypii* bioreactor cultivations in both defined and complex medium suggested that the presence of glucose in the medium may repress invertase production in *A. gossypii*, as observed in *S. cerevisiae* (Gancedo 1998), *Schizosaccharomyces pombe* (Zhang and Chi

2004), *Aspergillus nidulans* (Vainstein and Peberdy 1991) and *A. niger* (Rubio and Navarro 2006). Glucose repression in fungi is mediated by the binding of regulatory proteins (Mig1p and Mig2p in *S. cerevisiae* (Lutfiyya and Johnston 1996, Gancedo 1998), CreA/Cre1 in *A. nidulans* and *Trichoderma reesei* (Kulmburg et al. 1993, Cubero and Scazzocchio 1994, Mach et al. 1996, Takashima et al. 1996) to GC-rich consensus sequences ([G/C][C/T]GGGG in *S. cerevisiae*, [G/C][C/T]GG[A/G]G in *A. nidulans* and *T. reesei*) in the promoter region of the regulated gene. In *S. cerevisiae*, the presence of an AT-rich sequence 5' to the recognition site of Mig1p is essential for its binding (Lundin et al. 1994). *In silico* analysis of the deduced 5' untranslated region (UTR) of the *A. gossypii* homolog to the *ScSUC2* gene (*AFR529W*) revealed the presence of four putative binding sites for a catabolite repressor ([G/C][C/T]GGGG) preceded by an AT-rich sequence (Figure 4.2). These should be recognition sites for binding of the *A. gossypii* Mig1p homolog (AFR471C), as the zync-finger domains of the AgMig1p and ScMig1p are highly similar. Interestingly, in bioreactor cultivations using AFM with glucose as primary carbon source, the level of the *AgSUC2* mRNA transcript increased 1.79-fold ($p < 0.0018$) after glucose was depleted from the culture medium, when compared with the transcript levels measured when glucose was present at least from 17 ± 1 to 7 ± 2 g/l (our microarray data), which also suggested that glucose represses *AgSUC2* transcription and that sucrose is not needed as an inducer.

In flask cultivations of *A. gossypii* in AFM containing glucose as primary carbon source, invertase activity was found in the supernatant and cell-associated fractions during growth when glucose was present in concentrations as high as 15 g/l (Figures 4.3A and B). The level of activity was approximately 2-fold lower than in AFM with sucrose as the primary carbon source, even after glucose had been consumed (Figures 4.3A and B). In these conditions, invertase activity was mostly extracellular, as the total cell-associated activity measured was not considerably different from the activity quantified in the extracellular cell wall-bound fraction (Figures 4.3A and B). The specific invertase activity of the cell-associated fraction in flask cultivations with AFM containing glycerol as primary carbon source was similar to that observed in sucrose growths (Figure 4.4). However, the maximum specific invertase activity in the supernatant of these cultures was approximately 10-fold higher than that quantified in sucrose cultures (Figure 4.4). These observations reinforce the idea that the production of secreted invertase by *A. gossypii* is repressed by the presence of fermentable sugars in the medium and that sucrose is not needed as an inducer.

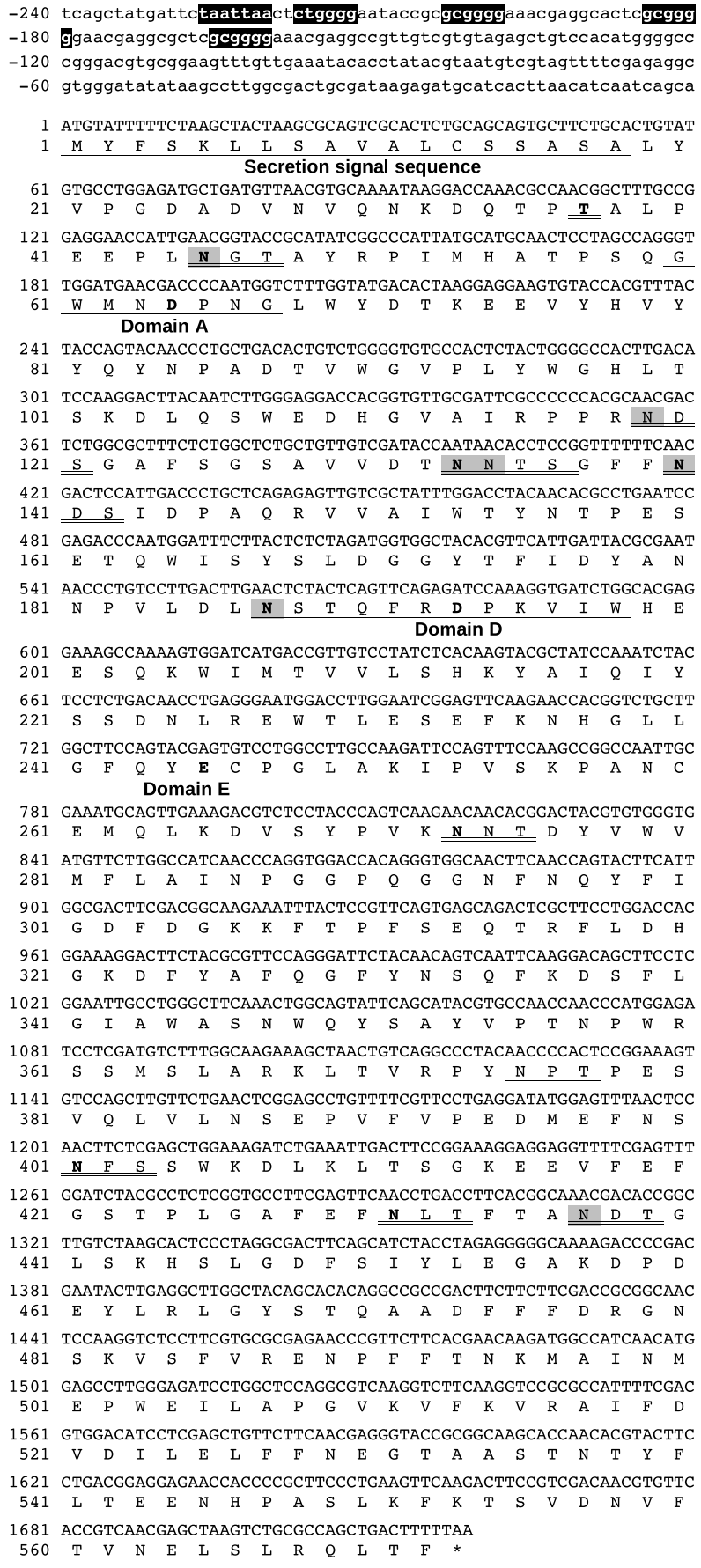


Figure 4.2 – Nucleotide and deduced amino acid sequence of the *AgSUC2*-encoded invertase and its

5' untranslated region (UTR). Potential recognition sites for AgMig1p binding and the preceding A-T rich sequence in the *AgSUC2* 5' UTR are blocked in black. The putative secretion signal peptide and the catalytic domains A, D and E are underlined, with the aminoacid of the catalytic site marked in **bold**. Potential *O*-glycosylation (S/T) and *N*-glycosylation sites (N-X-S/T, X any aminoacid except proline) are marked by two lines. *N*-glycosylation sites conserved in ScSuc2p are highlighted in grey and sites with higher probability of being occupied represented in **bold**.

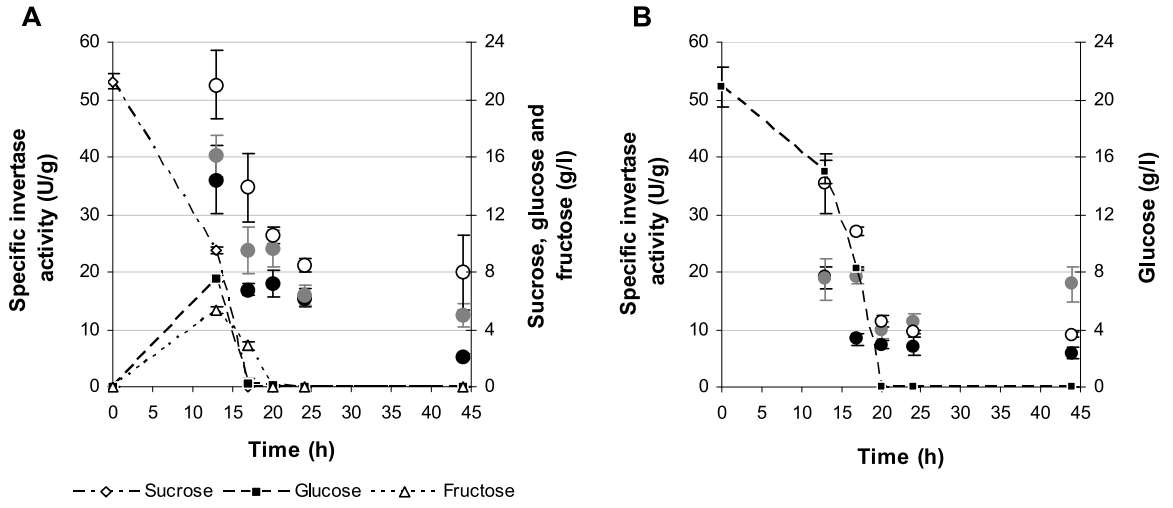


Figure 4.3 – Comparison of specific invertase activity in the culture supernatant (*black circles*), extracellular cell wall-bound (*grey circles*) and total cell-associated (*open circles*) during growth of *A. gossypii* ATCC 10895 in 250 ml flasks containing 50 ml of AFM with sucrose (A) or glucose (B) as primary carbon source. Specific invertase activity in read on the primary axis and sugar concentration on the secondary axis. Error bars represent the standard error of the mean of six independent biological replicates and where not seen were smaller than the symbol.

In contrast to the high specific invertase activity observed during the deceleration and stationary phase in bioreactor cultivations (Figure 4.1), specific activity in flask cultivations was highest during the growth phase and decreased during stationary phase. Thus, the extent to which invertase should be considered to be growth associated in *A. gossypii* is unclear, and may be partly dependent on factors such as aeration and pH, which would differ in flask and bioreactor cultures.

Unexpectedly, when *A. gossypii* spores (derived from the same stocks) were grown in SC medium containing glucose as carbon source, the lag phase was much longer than when sucrose was used (Figure 4.5). No difference in the lag phase for cultures growing on glucose or sucrose has been observed in AFM (cf. Figure 4.3 and Ribeiro et al. 2010). The long lag phase in SC may reflect differences in the osmolarity of SC medium and AFM, or

increased sodium toxicity from the sodium phosphate buffer which was used to buffer SC medium but not AFM, since both high osmolarity and sodium toxicity have been observed to increase the lag phase of *A. gossypii* (Förster et al. 1998). Regardless, sucrose was beneficial for the initial stages of *A. gossypii* growth in SC medium, possibly because it contributed less to the osmotic pressure of the medium than did the glucose.

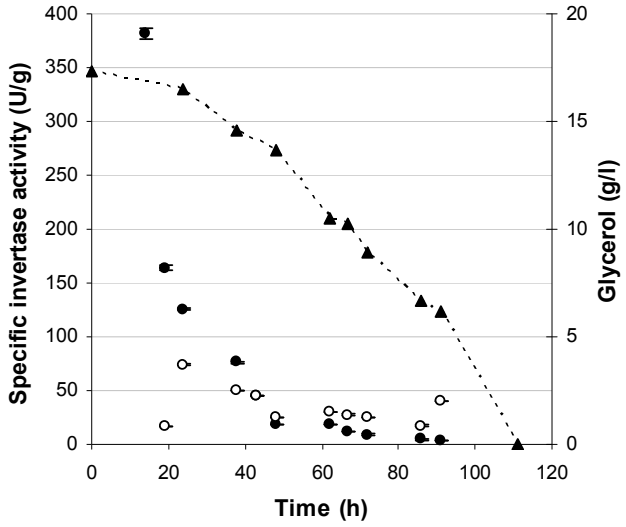


Figure 4.4 – Specific invertase activity in the culture supernatant (*black circles*) and total cell-associated (*open circles*) during growth of *A. gossypii* ATCC 10895 in 250 ml flasks containing 50 ml of AFM with glycerol as primary carbon source. Specific invertase activity in read on the primary axis and glycerol concentration on the secondary axis. Error bars represent the standard error of the mean of three independent biological replicates and where not seen were smaller than the symbol.

4.3.2 Molecular and functional characterization of the *A. gossypii* invertase

In silico analysis of the *A. gossypii AFR529W* ORF (*AgSUC2*) showed that its 1719 bp DNA sequence putatively encodes a protein with 572 amino acids and a calculated molecular mass of 64.99 kDa, which shares 53% of amino acid identity with the ScSuc2p. Moreover, the conserved domains A, D and E, which contain highly conserved acidic residues located in the active site of GH32 members (Alméciga-Díaz et al. 2011), were all present in AgSuc2p (Figure 4.2). Thus, we expected that this protein would be responsible for the hydrolyis of sucrose to fructose and glucose in *A. gossypii*.

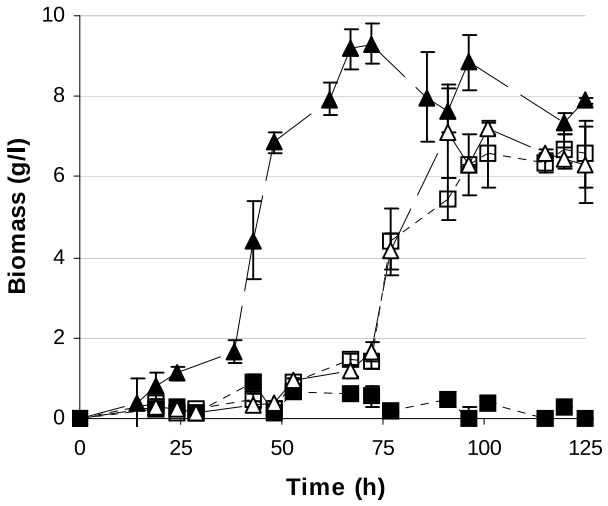


Figure 4.5 – Comparison of biomass production by *A. gossypii* ATCC 10895 (parent strain, *triangles*) and the *A. gossypii Agsuc2* null mutant (*squares*) growing in 250 ml flasks containing 50 ml of SC medium with sucrose (*solid symbols*) or glucose (*open symbols*) as primary carbon source. Error bars represent the standard error of the mean of three independent biological replicates.

To investigate the function of the *AgSUC2* gene, its complete coding region was deleted from the *A. gossypii* genome through a one-step PCR-based gene targeting approach (Wendland et al. 2000) and the generated *Agsuc2* null mutants were physiologically characterized (Figure 4.6). These mutants lost their ability to grow on sucrose as sole carbon source, but growth on glucose was unaffected (Figure 4.5). Moreover, invertase activity was not detected in either the secreted or cell-associated fractions. Transformation of these mutants with the plasmids pTAgSUC and pTScSUC2, expressing the homologous *AgSUC2* and the heterolougous *ScSUC2* genes under the control of the *S. cerevisiae PGK1* constitutive promoter, respectively, restored their ability to grow on sucrose and to secrete invertase. Invertase activity was detected in the culture filtrates of recombinant *A. gossypii* expressing both AgSuc2p and ScSuc2p. The mutants transformed with the empty vector (pMINATPS) remained unable to hydrolyze sucrose. These data demonstrate that the *AgSUC2* gene encodes all active invertase isoforms in *A. gossypii* and that it is functionally complemented by the *ScSUC2* gene. In addition, the signal peptide of the ScSuc2p was efficiently recognized by *A. gossypii* as a secretion signal.

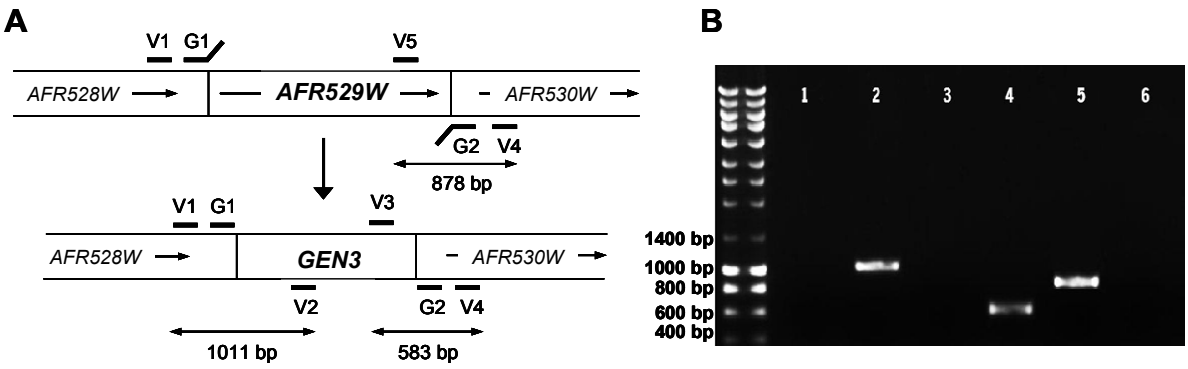


Figure 4.6 – Schematic representation of the *AgSUC2* gene deletion strategy (**A**) and PCR confirming the integration of the *GEN3* deletion module in the *AgSUC2* locus and complete removal of this gene from the genome of *A. gossypii* homokaryotic *Agsuc2* null mutants (**B**). In panel B, the first lane is the molecular marker (NZYDNA Ladder III, NZYTech); lanes 1, 3 and 5 are genomic DNA from the *A. gossypii* parent strain amplified with primers V1/V2, V3/V4 and V5/V4, respectively; and lanes 2, 4 and 6 are genomic DNA from *A. gossypii Agsuc2* mutant strain amplified with primers V1/V2, V3/V4 and V5/V4, respectively.

In silico analysis of the deduced amino acid sequence of the AgSuc2p provided evidence of a potential secretion signal and a predicted cleavage site between positions 18 and 19 (SiganlP 3.0; Bendtsen et al. 2004). In addition, it was predicted to contain 1 putative *O*-glycosylation site (NetOGlyc 3.1; Julenius et al. 2005) and 11 putative *N*-glycosylation sites (NetNGlyc 1.0; Gupta et al. 2004), 7 of which are conserved in ScSuc2p (Clustal Omega; Sievers et al. 2011) (Figure 4.2).

Analysis of a native gel stained for invertase activity, which had been loaded with *A. gossypii* concentrated culture filtrates, revealed two smeared bands (Figure 4.7, lane 1), one with a molecular mass similar to that of a purified extracellular *S. cerevisiae* invertase (270 kDa) (Figure 4.7, lanes 2 and 3) and another with a lower molecular mass, indicating that there were two extracellular invertase isoforms. These are most likely glycosylated, as this fungus has been previously shown to glycosylate its secreted proteins (Ribeiro et al. 2010). The commercial invertase from *S. cerevisiae* which was used as a control produced only one band (Figure 4.6, lanes 2 and 3), which according to the manufacturer's information corresponds to the extracellular glycosylated homodimer described by Gascón et al. (1968).

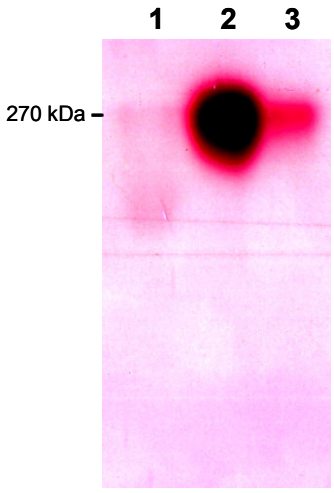


Figure 4.7 – Electrophoretic analysis of *A. gossypii* culture filtrate collected after 32 h of cultivation in a bioreactor with defined minimal medium containing sucrose as primary carbon source. To 8% native polyacrylamide gels were applied: *A. gossypii* culture filtrates concentrated 25 times (lane 1), 0.05 U of commercial *S. cereviae* invertase (lane 2) and 0.005 U of commercial *S. cereviae* invertase (lane 3). According to the manufacturer's indications the molecular weight of the *S. cereviae* invertase is 270 kDa.

The existence of different invertase isoforms is not unusual in fungi. In *S. cerevisiae*, invertase occurs in two homodimeric isoforms, one intracellular and non-glycosylated (135 kDa) and other extracellular and glycosylated (270 kDa) (Gascón et al. 1968). The *S. pombe* invertase natively occurs as a dimeric (205 kDa) or multimeric (1070–1200 kDa) enzyme (Moreno et al. 1990). Invertases from the yeast *Arxula adeninivorans* (Böer et al. 2004) and *Rhodotorula glutinis* (Rubio et al. 2002), respectively exist as hexameric (600 kDa) and dimeric (100 kDa) structures. Invertases from *A. niger* (Rubio and Maldonado 1995) and *A. ochraceus* (Guimarães et al. 2007) are homodimeric proteins with apparent molecular masses of 95 kDa and 135 kDa, respectively. Invertase from *Fusarium oxysporum* culture filtrates indicated that two secreted isoforms with apparent molecular weight of 60 and 120 kDa were present (Wolska-Mitaszko et al. 2007). High and low molecular weight isoforms (S- and F-forms differing in dimerisation and glycosylation) of secreted invertase have also been described in *A. nidulans* (Chen et al. 1996). The sugar composition of the growth media has been shown to deeply influence the different invertase isoforms created (Wolska-Mitaszko et al. 2007). The existence of multiple isoenzymes under different growth conditions may represent a physiological advantage for microorganisms, by allowing them greater flexibility in the regulation of carbohydrate metabolism (Wolska-Mitaszko et al. 2007).

4.4 CONCLUSION

Based on the analysis of the *A. gossypii* genome sequence we have isolated and characterized the *AgSUC2* gene (*AFR529W* ORF) and proved that it encodes a secreted invertase that is responsible in *A. gossypii* for hydrolysing sucrose to the readily assimilated sugars glucose and fructose. Similarly to the invertases of other fungi, the *A. gossypii* invertase natively exists in at least two isoforms and its synthesis may be regulated by the sugars present in the growth medium. This provides the characterization of the second hydrolytic enzyme natively secreted by *A. gossypii*, expanding our knowledge about the secretion capacities of this fungus. Moreover, this work enlarges the pool of experimentally characterized *A. gossypii* ORFs.

CHAPTER 5

High-level expression of *Aspergillus niger* β-galactosidase in *Ashbya gossypii*

ABSTRACT

Ashbya gossypii has been recently considered as a host for the expression of recombinant proteins. The production levels achieved thus far were similar to those obtained with *Saccharomyces cerevisiae* for the same proteins. Here, the β-galactosidase from *Aspergillus niger* was successfully expressed and secreted by *A.gossypii* from 2-micron plasmids carrying the native signal sequence at higher levels than those secreted by *S. cerevisiae* laboratorial strains. Four different constitutive promoters were used to regulate the expression of β-galactosidase: *A. gossypii AgTEF* and *AgGPD* promoters, and *S. cerevisiae ScADH1* and *ScPGK1* promoters. The native *AgTEF* promoter drove the highest expression levels of recombinant β-galactosidase in *A. gossypii*, leading to 2 and 7-fold higher extracellular activity than the *AgGPD* promoter and the heterologous promoters, respectively. In similar production conditions, the levels of active β-galactosidase secreted by *A. gossypii* were up to 37 times higher than those secreted by recombinant *S. cerevisiae* CEN.PK 113-7D and approximately 2.5 times higher than those previously reported for the β-galactosidase-highproducing *S. cerevisiae* NCYC869-A3/pVK1.1. The substitution of glucose by glycerol in the production medium led to an 1.5-fold icrease in the secretion of active β-galactosidase by *A. gossypii*. Recombinant β-galactosidase secreted by *A. gossypii* was extensively glycosylated, as are the native *A. niger* β-galactosidase and recombinant β-galactosidase produced by yeast. These results highlight the potential of *A. gossypii* as a recombinant protein producer and open new perspectives to further optimize recombinant protein secretion in this fungus.

5.1 INTRODUCTION

Ashbya gossypii is a filamentous hemiascomycete (Ashby and Nowell 1926) of important biotechnological interest due to its natural ability to overproduce riboflavin/vitamin B2 (Wickerham et al. 1946). Additionally, *A. gossypii* presents several unique characteristics that have expanded the interest in this organism beyond riboflavin production, such as one of the smallest eukaryotic genomes sequenced, which shares high homology and gene order conservation with that of the yeast *Saccharomyces cerevisiae* (Dietrich et al. 2004), haploid nuclei, high homologous recombination efficiency, allowing the manipulation of genes by simple PCR-based gene targeting techniques (Steiner et al., 1995 Wendland et al., 2000), and capacity to freely replicate plasmids harboring *S. cerevisiae* autonomous replicating sequence (ARS) elements (Wright and Philippsen 1991). These features, combined with the *A. gossypii* ability to grow in inexpensive media to high cell densities (Park et al. 2007, Tajima et al. 2009), to perform protein post-translation modifications (Ribeiro et al. 2010) and to natively secrete few proteins to the culture medium (Stahmann et al. 1997, Ribeiro et al. 2010, Ribeiro et al. 2011) have raised attention to this fungus as a potential host for the production of heterologous proteins.

The expression and secretion of recombinant proteins by *A. gossypii* have been scarcely explored. Two cellulases from *Trichoderma reesei*, endoglucanase I (EGI) and cellobiohydrolase I (CBHI), were successfully secreted by *A. gossypii* into the culture medium (Ribeiro et al. 2010). Recombinant EGI was expressed as a functional enzyme and at similar levels to those obtained with the closely related host *S. cerevisiae*, while recombinant CBHI was presumably inactive and only detectable by Western blot using monoclonal antibodies (Ribeiro et al. 2010). These cellulases have been reported as challenging enzymes to be produced by recombinant microbial hosts, both as functional proteins and at reasonable yields, especially CBHI (Ribeiro et al. 2010). Therefore, to further assess *A. gossypii* as a recombinant host, the expression in this organism of recombinant proteins which have been produced biologically-active and at high levels by other model hosts is of interest.

The β-galactosidase from the filamentous fungus *Aspergillus niger* is an extracellular enzyme of easy detection that has been efficiently produced by recombinant *S. cerevisiae* strains in high amounts (Domingues et al. 2002, Oliveira et al. 2007). Furthermore, this β-galactosidase has a high biotechnological interest, namely for the

hydrolysis of lactose from acid cheese whey (Oliveira et al. 2011). For these reasons, this enzyme was chosen to be recombinantly expressed in *A. gossypii*, aiming at further investigating the potential of this fungus as a recombinant protein producer.

Several approaches can be carried out to improve the productivity of recombinant hosts. Among them, random mutagenesis and targeted modification of the cell wall permeability have been previously attempted in *A. gossypii*, but with no relevant effect on the production levels of the recombinant protein (Ribeiro et al. 2013). Another strategy commonly used involves finding an appropriate promoter to drive the overexpression of recombinant proteins. Thus, this work also aimed at studying the influence of different homologous (*A. gossypii AgGPD* and *AgTEF*) and heterologous (*S. cerevisiae ScPGK1* and *ScADH1*) constitutive promoters on the expression of recombinant *β*-galactosidase by *A. gossypii*.

5.2 MATERIALS AND METHODS

5.2.1 Strains

A. gossypii ADH1lacA (pFMAlacA), PGK1lacA (pFMPlacA), GPDlacA (pFMGlacA) and TEFlacA (pFMTlacA) strains were obtained by transformation of *A. gossypii* ATCC10895 (provided by Prof. P. Philippsen from Basel University) as described below. *S. cerevisiae* CEN.PK 113-7D (MATα, MAL2-8^c SUC2) was used as host to obtain comparable *S. cerevisiae* transformant strains: *S. cerevisiae* ADH1lacA (pFMAlacA), PGK1lacA (pFMPlacA), GPDlacA (pFMGlacA) and TEFlacA (pFMTlacA). *S. cerevisiae* NCYC869-A3/pVK1.1 (Domingues et al. 2002) was used to produce recombinant *β*-galactosidase for zymogram analysis. *Escherichia coli* NZ5α (NZYTech; fhuA2Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17) was used as the recipient for all cloning experiments.

5.2.2 Culture media

A. gossypii strains were maintained on agar-solidified (15 g/l agar) *Ashbya* full medium (AFM; 10 g/l yeast extract, 10 g/l tryptone, 1 g/l myo-inositol, 20 g/l glucose), supplemented with 150 µg/ml G418 (Sigma-Aldrich) for selection. When indicated, 20 g/l

glycerol were used instead of glucose. Spore suspensions were prepared and stored as described by Ribeiro et al. (2010), except that the mycelium was digested using 4.5 mg/ml Lysing Enzymes from *Trichoderma harzianum* (Sigma-Aldrich). *S. cerevisiae* strains were maintained on agar-solidified (15 g/l agar) YPD (20 g/l glucose, 20 g/l peptone, 10 g/l yeast extract), supplemented with 150 µg/ml G418 for selection. Submerged cultures of *A. gossypii* and *S. cerevisiae* transformants were made in selective AFM and YPD, respectively. LB medium (10 g/l sodium chloride, 10 g/l yeast extract, 5 g/l tryptone, pH adjusted to 7.5) supplemented with 15 g/l agar and 100 µg/ml ampicillin was used for *E. coli* cultivations.

5.2.3 Plasmids construction

The plasmid pMI516 used by Ribeiro et al. (2010) to express recombinant proteins in *A. gossypii* was modified as follows. A multiple cloning site (MCS) containing restriction sites for *Eco*RI, *Bam*HI, *Nhe*I, *Eag*I, *Not*I, *Sac*II, *Kpn*I, *Sal*I and *Xho*I was introduced between the *ScPGK1* promoter and terminator, generating plasmid pMI516MCS. This was accomplished by amplifying the *ScPGK1* promoter (*ScPGK1p*) from pMI516 with primers pMI-MCS_FW and pMI-MCS_RV (Table 5.1), digesting the amplified fragment with *Sac*I and *Xho*I (New England Biolabs) and cloning it between the *Sac*I and *Xho*I sites in pMI516. The *kanMX* and *S. cerevisiae URA3* expression modules in pMI516MCS were then exchanged with the *GEN3* module from pGEN3 (Wendland et al. 2000) using the *Bgl*II/*Bgl*II sites, generating plasmid pMIGEN3.

Plasmids pFMA, pFMG and pFMT were obtained by exchanging the *ScPGK1p* in pMIGEN3 with the *ScADH1*, *AgGPD* and *AgTEF* promoters, respectively. The 700 bp *ScADH1p* (Ruohonen et al. 1995) was amplified by PCR from *S. cerevisiae* CEN.PK 113-7D genomic DNA using the primers ScADH1p-SacI_FW and ScADH1p-EcoRI_RV (Table 5.1), the *AgGPDp* (Revuelta et al. 1999) was amplified from *A. gossypii* ATCC10895 genomic DNA with primers AgGPDp-SacI_FW and AgGPDp-EcoRI_RV (Table 5.1) and the *AgTEFp* (Steiner and Philippsen 1994) was amplified by PCR from pUG66 (Gueldener et al. 2002) with primers AgTEFp-SacI_FW and AgTEFp-EcoRI_RV (Table 5.1). The amplified promoters were digested with *Sac*I and *Eco*RI (New England Biolabs) and cloned between the *Sac*I and *Eco*RI sites in pMIGEN3, replacing the *ScPGK1p*. The *A. niger lacA* gene was amplified from pVK1.1 vector (Kumar et al. 1992) with primers LacA_KpnI_FW and LacA_XhoI_RV (Table 5.1), digested with *Kpn*I and

*Xho*I (New England Biolabs) and inserted into the *Kpn*I/*Xho*I sites of pFMA, pMIGEN3, pFMG and pFMT, in frame with the different promoters. The resulting plasmids were named pFMAlacA (Figure 5.1A), pFMPlacA (Figure 5.1B), pFMGlacA (Figure 5.1C) and pFMTlacA (Figure 5.1D), respectively. The constructions were verified by sequencing.

All PCR reactions were carried out with Taq DNA polymerase 5 U/µl (NZYTech), with the exception of *lacA* gene amplification, which was done using Phusion® High-fidelity DNA polymerase 2 U/µl (Finnzymes). PCR conditions for amplification with Taq DNA polymerase were: 5 min at 95°C, 35 cycles of 45 s at 95°C, 45 s at 50°C (53ºC for amplification of the MCS) and 1 min at 72°C, with a final extension at 72°C for 10 min. For *lacA* amplification PCR conditions were: 1 min at 98°C, 35 cycles of 10 s at 98°C, 30 s at 68°C and 1 min 30 s at 72°C, with a final extension at 72°C for 10 min.

Table 5.1 – Primers used in this study. The restriction sites used for the construction of plasmids are highlighted in **bold and underlined**. Highlithed only in **bold** is the multiple cloning site introduced in the the plasmids used.

Primer	Sequence (5' – 3')
pMI-MCS_FW	CCAGTGAATTGGCCGATGC
pMI-MCS_RV	GCC**CTCGAGGTCGACGGTACCCCGCGGGCGGCCGCGCTAGCGGA TCCGAATTC**GGCGCGTAAGTT
ScADH1p-SacI_FW	TCTATC**GAGCTC**CGGGTGTACAATATGGACT
ScADH1p-EcoRI _RV	CG**GAATTC**TGAGATAGTTGATTGTATGCTTGGT
AgGPDp-SacI_FW	TCTATC**GAGCTC**TGTCTGGGTGCACGACAC
AgGPDp-EcoRI_RV	CG**GAATTC**GGTGTGTATGTGTGGACTGAA
AgTEFp-SacI_FW	TCTATC**GAGCTC**GATCTGTTTAGCTTGCCTCG
AgTEFp-EcoRI_RV	CG**GAATTC**TTGTTTATGTTCGGATGTGATG
LacA_KpnI_FW	CGG**GGTACC**ATGAAGCTTTCCTCCGCTTG
LacA_XhoI_RV	CCG**CTCGAG**CTAGTATGCACCCTTCCGCTTC

5.2.4 *A. gossypii* transformation and screening of transformants

A. gossypii ATCC10895 was transformed by electroporation as described by Wendland et al. (2000) with 5 µg of plasmid DNA (pFMAlacA, pFMPlacA, pFMGlacA, pFMTlacA and pFMT as a negative control). Transformant colonies were selected on selective AFM plates. For the initial β-galactosidase activity screenings a small patch of mycelium from the edge of several transformant colonies was used to inoculate 10 ml test

tubes containing 5 ml of selective AFM. After 2 days of growth at 30ºC and 200 rpm, β-galactosidase activity was measured from culture supernatants and 5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-gal; NZYTech) was added to the tubes to a final concentration of 0.04 mg/ml. Hydrolysis of the X-gal by the produced β-galactosidase turned the supernatants blue. Based on the β-galactosidase assays and on the intensity of the blue colour in the different tubes, four transformant colonies were selected for further study: *A. gossypii* ADH1lacA, *A. gossypii* PGK1lacA, *A. gossypii* GPDlacA and *A. gossypii* TEFlacA.

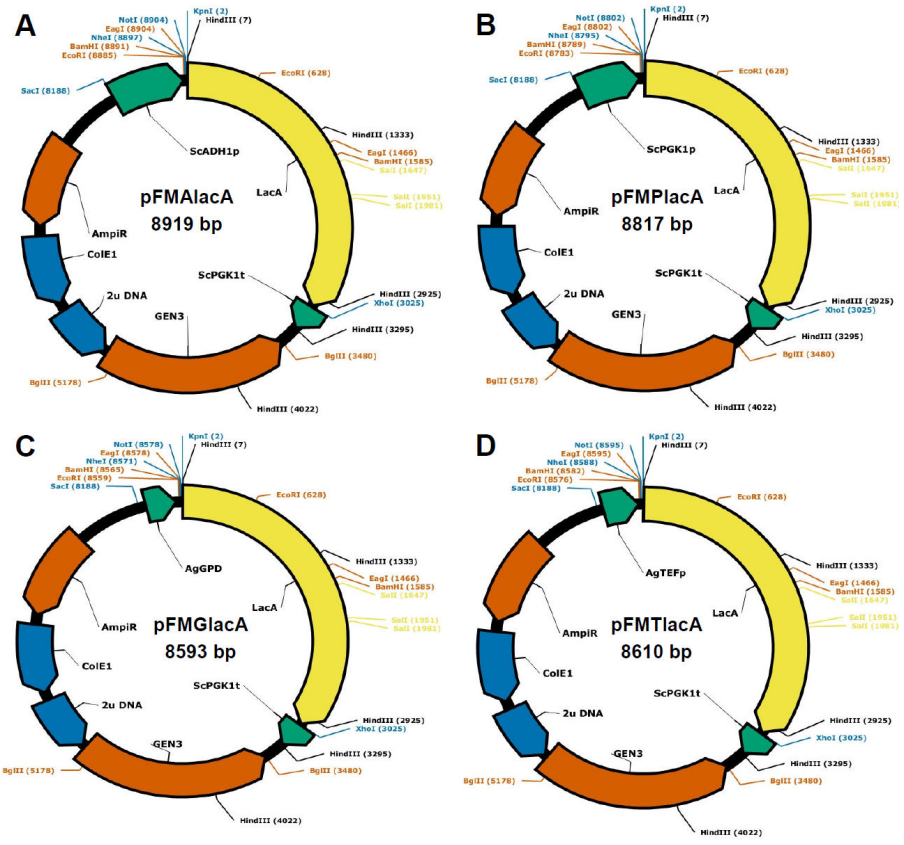


Figure 5.1 – Graphic representation of the plasmids constructed for the expression of *A. niger* β-galactosidase in *A. gossypii* under the regulation of different promoters: (A) *S. cerevisiae ADH1* promoter, (B) *S. cerevisiae PGK1* promoter, (C) *A. gossypii GPD* promoter, and (D) *A. gossypii TEF* promoter. Images created with PlasmaDNA v1.4.2 (University of Helsinki).

5.2.5 *S. cerevisiae* transformation and selection of transformants

S. cerevisiae CEN.PK 113-7D was transformed by the Lithium Acetate method (Gietz et al. 1995) with the same plasmids as *A. gossypii*. Transformant colonies were

selected on selective YPD medium and expression of β-galactosidase was confirmed in X-gal plates. After 3 days of growth at 30ºC four transformant colonies were randomly chosen for further study: *S. cerevisiae* ADH1lacA, *S. cerevisiae* PGK1lacA, *S. cerevisiae* GPDlacA and *S. cerevisiae* TEFlacA.

5.2.6 Expression of recombinant β-galactosidase

Recombinant *A. gossypii* and *S. cerevisiae* strains were grown at 30ºC and 200 rpm in 250 ml Erlenmeyer flasks containing 50 ml of selective AFM and YPD, respectively. An inoculum of 10^7 spores was used for *A. gossypii* growths. For *S. cerevisiae*, pre-cultures grown for 12 h were used to inoculate production cultures to an initial absorbance at 600 nm (OD_{600}) of 0.1. Samples were taken to determine cell concentration (OD_{600}) and to collect culture supernatants by centrifugation at 5000 rpm for 10 min, at 4ºC. The OD_{600} was converted to dry weight (DW) using a standard curve. Total reducing sugars concentration in the culture supernatants was determined by the dinitrosalicylic acid method (Gonçalves et al. 2010).

5.2.7 β-galactosidase assay

β-Galactosidase activity was measured from culture supernatants through the quantification of *p*-nitrophenol released from *p*-nitrophenyl-β-D-galactopyranoside (*p*NPG), as described in Domingues et al. (2004). Briefly, the samples were incubated for 10 min at 65ºC in 1.7 mM of *p*NPG in 0.075 M acetate buffer (pH 4.5). The reaction was stopped by raising the pH to 10 with 1 M Na_2CO_3 and the absorbance was measured at 405 nm. One enzyme unit (U) was defined as the amount of enzyme that hydrolysed 1 nmol of *p*NPG per min at 65ºC.

5.2.8 β-galactosidase zymogram

The supernatant of a shake-flask culture of recombinant *A. gossypii* TEFlacA and *S. cerevisiae* NCYC869-A3/pVK1.1 was concentrated in Amicon® Ultra-15 30,000 MWCO centrifugal filter devices (Millipore) according to the manufacturer's instructions. For *in situ* detection of β-galactosidase activity, the concentrated samples (with a normalized

β-galactosidase activity of approximately 40,000 U/ml) were separated by polyacrylamide gel electrophoresis (PAGE) in an 8% native polyacrylamide gel. *β*-Galactosidase activity staining was carried out as described in O'Connell and Walsh (2010) by incubating the native gel in a solution of 0.02% (w/v) X-gal in 0.2 M sodium acetate buffer (pH 4.5) for 5 h at 45ºC, with agitation (100 rpm).

5.3 RESULTS

5.3.1 Expression of *A. niger* *β*-galactosidase in *A. gossypii* under the regulation of different promoters

Four different promoters were used to drive the expression of *A. niger* *β*-galactosidase in *A. gossypii*: two *A. gossypii* promoters (*AgGPDp* and *AgTEFp*) and two *S. cerevisiae* promoters (*ScADH1p* and *ScPGK1p*). As described in the materials and methods section, each promoter was cloned *in frame* with the *lac*A gene (coding for secreted *A. niger* *β*-galactosidase) in the modified pMI516 plasmid (Ribeiro et al. 2010), resulting in four expression vectors: pFMAlacA, pFMPlacA, pFMGlacA and pFMTlacA (Figure 5.1). *A. gossypii* mycelium was transformed by electroporation with each expression vector constructed, with transformation efficiencies of 50 to 100 colonies per µg of plasmid DNA. In the screening assays, it was already possible to see that the native promoters (*AgTEFp* and *AgGPDp*) led to a stronger expression of recombinant *β*-galactosidase than the *S. cerevisiae* promoters (*ScADH1p* and *ScPGK1p*) (Figure 5.2).

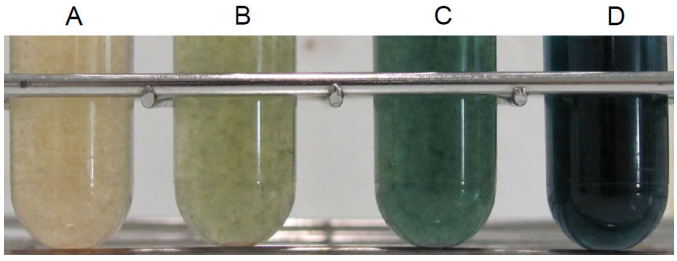


Figure 5.2 – Test tubes for *β*-galactosidase activity screening containing *A. gossypii*ADH1lacA (A), PGK1lacA (B), TEFlacA (C) and GPDlacA (D) grown for 2 days in 5 ml of selective AFM supplemented with 0.04 mg/ml X-gal.

The β-galactosidase activity secreted by the selected recombinant *A. gossypii* strains under the control of each promoter was measured in the culture supernatants of shake-flasks fermentations. No differences in the specific growth rate of the recombinant strains were observed when compared to the parental strain transformed with the empty vector, indicating that the secretion of recombinant β-galactosidase did not impose significant metabolic load in this host (not shown). In all strains, the specific activity of recombinant β-galactosidase increased along the fermentation time and reached the maximum values when all glucose was consumed (at about 87 h) (Figure 5.3). Similar activity levels were obtained with both *S. cerevisiae* promoters (*ScADH1p* and *ScPGK1p*), as observed in the preliminary screening, and higher levels were obtained with the native promoters. The native *AgTEFp* led to 2-fold higher recombinant β-galactosidase activity than the native *AgGPDp* and 7-fold higher than the *S. cerevisiae* promoters (Table 5.2). Maximum specific extracellular β-galactosidase activities (148 ± 7 U/mg, which corresponded to a volumetric activity of 1039 ± 51 U/ml) were obtained when the *A. gossypii* TEFlacA strain was grown in modified AFM containing glycerol instead of glucose. These are approximately 1.5-fold higher than those obtained with the same strain growing in standard AFM with glucose as primary carbon source.

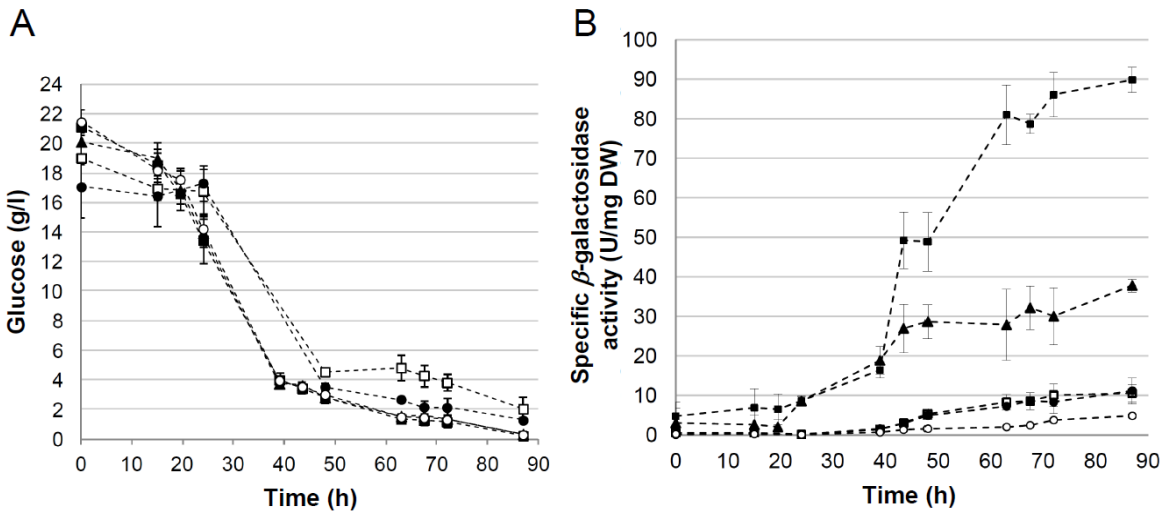


Figure 5.3 – Glucose consumption (A) and specific β-galactosidase activity (B) for the *A. gossypii* recombinant strains growing at 30ºC and 200 rpm in 250 ml shake-flasks containing 50 ml of selective AFM: *A. gossypii* TEFlacA (*solid square*), GPDlacA (*solid triangle*), ADH1lacA (*open square*), PGK1lacA (*solid circle*) and empty vector (*open circle*). Error bars represent the standard deviation of three biological replicates.

5.3.2 Expression of *A. niger* β-galactosidase in *S. cerevisiae* under the regulation of different promoters

In order to compare the levels of recombinant β-galactosidase activity obtained in *A. gossypii* with a yeast host, the *S. cerevisiae* CEN.PK 113-7D strain was transformed with the same expression vectors used to transform *A. gossypii*. The β-galactosidase activity secreted by the selected yeast recombinants was determined as mentioned above for the *A. gossypii* recombinants. The levels of recombinant β-galactosidase activity secreted by *S. cerevisiae* were considerably lower than those obtained in *A. gossypii*, independently of the promoter used (Table 5.2). The yeast secreted nearly 40 times less recombinant β-galactosidase activity than *A. gossypii* in similar production conditions (Table 5.2). Furthermore, no relevant differences were observed between the levels of recombinant β-galactosidase activity obtained in *S. cerevisiae* with their own native promoters and with the *A. gossypii* promoters (Table 5.2).

Table 5.2 – Maximum volumetric β-galactosidase activity secreted by *A. gossypii* and *S. cerevisiae* recombinant strains under the regulation of different promoters. The values represent the average ± standard deviation of the β-galactosidase activity measured at the glucose depletion time-point in the supernatants of three independent shake-flask cultures performed at 30ºC and 200 rpm in 50 ml of selective AFM (for *A. gossypii*) and YPD (for *S. cerevisiae*). The average biomass for the *A. gossypii* recombinant strains was 6.37±0.38 g/l and for the *S. cerevisiae* recombinant strains was 5.75±0.07 g/l.

	Maximum β-galactosidase activity (U/ml)	
Promoter	***A. gossypii***	***S. cerevisiae***
ScADH1	72 ± 18	2 ± 1
ScPGK1	70 ± 16	8 ± 1
AgGPD	248 ± 14	10 ± 2
AgTEF	523 ± 30	14 ± 2

5.3.3 Partial characterization of the recombinant β-galactosidase secreted by *A. gossypii*

The recombinant β-galactosidase secreted by *A. gossypii* was analyzed in a NATIVE-PAGE gel stained for β-galactosidase activity with X-gal (Figure 5.4). The *A. niger* β-galactosidase produced by recombinant *S. cerevisiae* (Domingues et al. 2002) was included for comparison. *A. gossypii* recombinant β-galactosidase is seen in the gel as a

diffuse high molecular weight band with a size ranging from 135 kDa to more than 245 kDa (Figure 5.4). Such molecular weights are higher than the predicted weight based on the amino acid sequence of β-galactosidase (116 kDa) (Figure 5.4), indicating that this recombinant β-galactosidase is probably glycosylated, like the β-galactosidase secreted by *A. niger* and recombinant *S. cerevisiae* (Kumar et al. 1992). It can also be seen in the zymogram that the recombinant β-galactosidase produced by *A. gossypii* is more heterogeneous than the one produced in *S. cerevisiae*, judging by the larger smeared band corresponding to the *A. gossypii* recombinant enzyme (Figure 5.4).

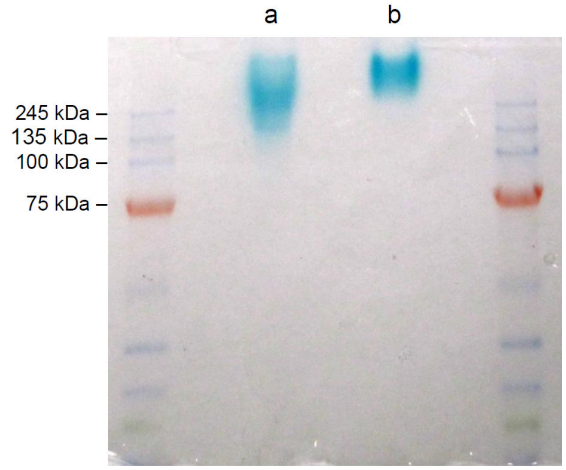


Figure 5.4 – β-Galactosidase zymogram. *A. niger* β-galactosidase produced by *A. gossypii* TEFlacA (lane a) and *S. cerevisiae* NCYC869-A3/pVK1.1 (lane b). The recombinant β-galactosidase samples applied to the gel had an activity of approximately 600 U.

5.4 DISCUSSION

A. gossypii presents several features that suggest it can be a good host for recombinant protein production, but until now only two heterologous proteins had been reported to be secreted by this fungus and at modest levels (Ribeiro et al. 2010). In this work, the β-galactosidase from *A. niger* was successfully expressed and secreted by the *A. gossypii* ATCC10895 strain from 2-micron plasmids under the control of the native promoters *AgGPDp* and *AgTEFp*, and of the *S. cerevisiae* promoters *ScPGK1p* and *ScADH1p* (Figures 5.2 and 5.3). This is the first time that *A. niger* β-galactosidase is expressed in this fungus, expanding the number of recombinant proteins of different origins to be effectively secreted by *A. gossypii*.

The level of recombinant β-galactosidase secreted by *A. gossypii* was strictly dependent on the promoter used. The highest extracellular β-galactosidase activity was obtained with the native *AgTEFp*, followed by the other native promoter, the *AgGPDp* (Table 5.2). The *AgTEFp* and *AgGPDp* are the most commonly used promoters in *A. gossypii* but their strengths have never been directly compared. These promoters have been used in *A. gossypii* for the overexpression of genes associated with riboflavin biosynthesis, resulting in increased riboflavin production (Monschau et al. 1998, Jiménez at al. 2005, Jiménez et al. 2008), and for the expression of intracellular heterologous proteins (Kato and Park 2006). In this work, the *AgTEFp* presented the double driving force of the *AgGPDp*. The *S. cerevisiae* promoters *ScPGK1p* and *ScADH1p* functioned in a similar way in *A. gossypii*, but led to lower protein expression levels than the native promoters, as commonly happens in other filamentous fungi (Nevalainen et al. 2005) (Table 5.2). A better performance of native promoters over *S. cerevisiae* promoters in *A. gossypii* has also been previously reported, namely using the *AgTEFp* (Kaufmann 2009). The *ScPGK1p* has already been employed once in the expression of heterologous proteins in *A. gossypii* (Ribeiro et al. 2010), but to our knowledge, this is the first time that the *ScADH1p* has been used. The plasmids constructed in this work, especially that carrying the native *AgTEFp*, thus revealed to be suitable protein expression vehicles for *A. gossypii*.

The laboratorial *S. cerevisiae* CEN.PK 113-7D strain transformed with the same plasmids secreted lower levels of recombinant β-galactosidase than *A. gossypii* (up to 14 U/ml; Table 5.2). This result contrasts with previous reports, where the levels of recombinant proteins secreted by both hosts were comparable, using a 2-micron based plasmid identical to that used in this work (Ribeiro et al. 2010). However, this does not directly imply that *A. gossypii* is a better *A. niger* β-galactosidase producer than *S. cerevisiae*, since the efficient expression of this enzyme in *S. cerevisiae* seems to be strain and/or plasmid dependent. In the first attempts to produce *A. niger* β-galactosidase in *S. cerevisiae* very low activity levels were obtained. A Mauri distiller's yeast transformed with plasmid pVK1.1, harbouring the *URA3* selective marker and the *A. niger* β-galactosidase under the control of the *S. cerevisiae ADH1* promoter and terminator (*lacA* cassette), expressed 10 U/ml of extracellular recombinant β-galactosidase (Ramakrishnan and Hartley 1993). Similarly, a flocculent brewer's yeast transformed with a plasmid having as selective marker the yeast *CUP1* gene, conferring resistance to copper (pET13.1 plus *lacA* cassette), expressed 17 U/ml of extracellular recombinant β-galactosidase

(Domingues et al. 2000). These activity levels are of the same order of magnitude as those obtained in this work for the *S. cerevisiae* CEN.PK 113-7D strain transformed with the plasmids here described, even though different media and carbon sources were used (in this work glucose was the primary carbon source used, while lactose has been used in the previous studies). However, very high yields of recombinant β-galactosidase activity were reported for the *S. cerevisiae* NCYC869-A3 strain transformed with the plasmid pVK1.1 (Domingues et al. 2002). Unfortunately, it is not possible to transform *A. gossypii* with the pVK1.1 plasmid for direct comparison with *S. cerevisiae*, because the *S. cerevisiae URA3* does not complement uracil auxotrophy in *A. gossypii Agura3* auxotrophes (Pompejus et al. 1999) and also the *ScADH1* terminator shows ARS activity in *A. gossypii* (Knechtle et al. 2003). However, in similar culture conditions, the recombinant *A. gossypii* strain carrying the plasmid pFMTlacA (where the *lacA* gene is under the regulation of the native *AgTEFp*) secreted approximately 2.5 times higher levels of recombinant β-galactosidase activity than those reported for the recombinant high-level producer *S. cerevisiae* NCYC869-A3/pVK1.1 strain (212 U/ml; Oliveira et al. 2007). This result highlights the suitability of *A. gossypii* as an alternative recombinant β-galactosidase producer with still plenty of room for improvement. A change in the primary carbon source used (from glucose to glycerol) already improved approximately 1.5-fold the productivity of the best recombinant *A. gossypii* β-galactosidase-producing strain. In *S. cerevisiae*, optimization of the operation and culture conditions resulted in a 21-fold increase in the recombinant β-galactosidase production, when compared with the values described above for flask cultivations, approaching the extracellular activity levels obtained in *A. niger* (Oliveira et al. 2011).

Besides good expression capacity, *A. gossypii* also presented in this work other desirable features for a recombinant protein producer. The native signal peptide of the *A. niger* β-galactosidase was recognised by *A. gossypii* as a secretion signal, directing the production of this protein into the culture medium, which facilitates its downstream processing. *A. gossypii* seems to be able to recognize a wide range of signal sequences, which is advantageous in a recombinant host. In previous works, *A. gossypii* recognized the native signal sequences of *T. reesei* cellulases EGI and CBHI (Ribeiro et al. 2010) and *S. cerevisiae* invertase (*SUC2* gene) (our unpublished results). Furthermore, recombinant β-galactosidase secreted by *A. gossypii* was apparently glycosylated, as indicated by its high molecular weight in a native gel (more than 135 kDa) (Figure 5.4). The native *A.*

niger β-galactosidase is a highly glycosylated enzyme, with several potential sites for *N*-linked (12 sites) and *O*-linked glycosylation (Kumar et al. 1992). This enzyme presented an apparent size of 130 kDa in Western blots, migrating with a diffuse aspect, which was attributed to heterogeneous glycosylation (Kumar et al. 1992). Heterogeneous and extensive *N*-glycosylation has also been reported for recombinant β-galactosidase produced in *S. cerevisiae* (Kumar et al. 1992). Here, the β-galactosidase secreted by *A. gossypii* displayed a more diffuse pattern in the native gel than that secreted by *S. cerevisiae*, possibly as a result of higher glycosylation heterogeneity (Figure 5.4). Previously, recombinant *T. reesei* EGI and CBHI secreted by *A. gossypii* were observed to be less extensively glycosylated than the same cellulases secreted by *S. cerevisiae* (Ribeiro et al. 2010). Despite its higher heterogeneity, the same tendency to a less extensive hyperglycosylation of the recombinant β-galactosidase produced by *A. gossypii* in comparison to that produced by *S. cerevisiae* was also apparent in this work (Figure 5.4).

In conclusion, β-galactosidase from *A. niger* was for the first time produced and secreted by *A. gossypii* at higher levels than those obtained with recombinant *S. cerevisiae* laboratorial strains producing this enzyme. Moreover, the *AgTEFp* revealed to be the best promoter for its overexpression. The results obtained herein demonstrate the potential of *A. gossypii* as a new recombinant protein production platform. Further optimization of expression plasmids, culture conditions or strain tailoring can be now envisaged, aiming at improving the productivity of the non-engineered laboratorial strain here used.

CHAPTER 6

Cre-*lox*P based system for removal and reuse of selection markers in *Ashbya gossypii* targeted engineering

ABSTRACT

The filamentous ascomycete *Ashbya gossypii* is amenable to genetic manipulation and is an excellent model system for studying eukaryotic cell biology. However, the number of selection markers that can be used for both targeted gene integration and disruption are very limited. Therefore, the Cre-*lox*P recombination system was adapted for use in *A. gossypii* and its effectiveness in recycling marker genes demonstrated by constructing both single and double deleted *Agura3* and *Agade1* auxotrophic strains free of exogenous markers. In spite of its wide use, this is the first report in which the Cre-*lox*P system was applied to *A. gossypii*, opening new perspectives for targeted engineering of this fungus with several promising biotechnological applications.

6.1 INTRODUCTION

The ability to remove and recycle selection markers is an important tool when conducting genetic engineering projects, not only because it eliminates the need for an unlimited number of selection markers when studying gene expression and gene characterization in the same strain, as it hampers possible disturbing effects of the markers on the cellular processes of the engineered organisms. Removal of selection markers is particularly important when genetically modified strains are intended for industrial or commercial use because of the many regulatory requirements regarding the presence of foreign genes in these organisms, especially exogenous genes conferring resistance to antibiotics.

The Cre-*lox*P recombination system of the bacteriophage P1 has been shown to mediate efficient recombination between directly repeated *lox*P sites flanking selectable markers in several organisms, resulting in the excision of the marker gene (Sternberg and Hamilton 1981, Sauer and Henderson 1988, Orban et al. 1992, Zhang et al. 2009). Cre-*lox*P-based systems have been widely used to study gene function and to remove selection markers in fungi, among which *Saccharomyces cerevisiae* (Sauer 1987, Güldener et al. 1996), *Kluyveromyces lactis* (Steensma and Ter Linde 2001), *Kluyveromyces marxianus* (Ribeiro et al. 2007), *Schizosaccharomyces pombe* (Iwaki and Takegawa 2004), *Yarrowia lipolytica* (Fickers et al. 2003), *Hansenula polymorpha* (Qian et al. 2009), *Cryptococcus neoformans* (Patel et al. 2010), *Aspergillus nidulans* (Forment et al. 2006), *Trichoderma reesei* (Steiger et al. 2011), *Neotyphodium coenophialum*, *N. uncinatum* and *Epichloë festucae* (Florea et al. 2009).

Ashbya gossypii (syn. *Eremothecium gossypii*) is a flavinogenic filamentous ascomycete (Ashby and Novell 1926, Kurtzman 1995) used for the industrial production of riboflavin (Stahmann et al. 2000), which shares remarkable genomic similarities with the budding yeast *S. cerevisiae* (Brachat et al. 2003, Dietrich et al. 2004), being therefore considered an excellent model to study filamentous growth (Wendland and Whalter 2005, Schmitz and Philippsen 2011). The high degree of gene homology and gene order conservation (synteny) between the *A. gossypii* genome and the *S. cerevisiae* genome have facilitated the assignment of potential functions to *A. gossypii* genes (Brachat et al. 2003, Dietrich et al. 2004). Moreover, other *A. gossypii* features have made the genetic manipulation of this fungus to be quite straightforward when comparing to other filamentous fungi, among which: highly efficient homologous recombination, which

allows the use of PCR-based gene targeting tools (Steiner et al. 1995, Wendland et al. 2000); haploid nuclei and unicucleated spores (Steiner et al. 1995), which facilitate the isolation of stable homokaryotic strains; and ability to freely replicate plasmids containing *S. cerevisiae* autonomously replicating sequences (ARS) (Wright and Philippsen 1991) and to regulate gene expression using *S. cerevisiae* promoters (Wendland et al. 2000, Kaufmann 2009, Ribeiro et al. 2010), making possible the use of several tools already available to manipulate *S. cerevisiae*.

Although a wide range of molecular tools exist to genetically engineer *A. gossypii*, the number of selection markers available is limited. Only four marker genes are available for PCR-based gene targeting in *A. gossypii*, which has hampered the disruption and/or integration of multiple genes in the same strains. The existent markers are: (1) the *S. cerevisiae LEU2* gene fused with its own promoter and terminator (Alberti-Segui et al. 2001), which can only be used as auxotrophic marker for gene manipulations in stable *Agleu2* strains, (2) the *Escherichia coli kan*^{r} gene fused to the *S. cerevisiae TEF2* (*ScTEF2*) promoter and terminator (*GEN3*) (Wendland et al. 2000); (3) the *Streptomyces noursei nat*^{r} gene flanked by the *S. cerevisiae PDC1* (*ScPDC1*) promoter and terminator (*NATPS*) (Hoepfner *in* Kaufmann 2009); (4) and a *hyg*^{r} gene fused to the promoter and terminator sequences of the *A. gossypii LEU2* (*AgLEU2*) gene (Revuelta *in* Schlüpen 2003; Jiménez et al. 2008). Because this last marker cassette contains sizeable homology regions to the *AgLEU2* locus, when used for targeted engineering it can direct integration into the *AgLEU2* locus rather than into the desirable target site.

The ability to rescue drug marker genes from the *A. gossypii* genome would therefore offer new perspectives in what concerns *A. gossypii* manipulation. Namely, it would allow further improvement of riboflavin production in industrial strains through rational engineering of different metabolic pathways. Hence, the aim of this work was to develop a Cre-*lox*P-based system for removing and reuse selection markers in *A. gossypii.* To achieve this goal, several marker cassettes and Cre expression plasmids were constructed and successfully used to excise the selection markers used to delete two model genes, the *AgURA3* (*AEL059W*) gene and the predicted *AgADE1* (*AER221W*) open reading frame (ORF). These genes were chosen because their deletion not only should provide easy phenotype identification as would result in new auxotrophic strains, further expanding the molecular toolbox available for *A. gossypii*.

6.2 MATERIALS AND METHODS

6.2.1 Strains and media

A. gossypii ATCC 10895 strain, kindly provided by Prof. P. Philippsen (University of Basel), was used as the background for all deletions. Strains generated in this study are listed in Table 6.1. *A. gossypii* strains were grown at 30ºC on agar-solidified (15 g/l agar) *Ashbya* full medium (AFM; 10 g/l tryptone, 10 g/l yeast extract, 20 g/l glucose, 1 g/l myo-inositol), containing 150 µg/ml geneticin (G418) (Sigma-Aldrich), 50 µg/ml nourseothricin (clonNAT) (WERNER BioAgents) or 5 µg/ml phleomycin (InvivoGen) for selection and maintenance of transformants. *A. gossypii* spores were prepared and stored as described by Ribeiro et al. (2010), with the exception that the mycelium was digested using 4 mg/ml Lysing Enzymes from *Trichoderma harzianum* (Sigma-Aldrich). Solid Synthetic Complete (SC) medium (Sherman et al. 1986) buffered with 0.1 M sodium-phosphate buffer (pH 7.0) and lacking uracil and/or adenine was used for phenotype characterization. When indicated, 5-fluoroorotic acid (5-FOA) (Sigma-Aldrich) was added to AFM at a final concentration of 1 mg/ml.

Escherichia coli TOP10 (Invitrogen) served as plasmid host and was grown on agar solidified (15 g/l agar) LB medium (Sambrook and Russell 2001) containing 100 µg/ml ampicillin (Sigma-Aldrich). Plasmids were maintained in a permanent culture at -80ºC in 2 ml aliquots with 15% (w/w) glycerol.

6.2.2 Molecular methods

DNA manipulations were carried out using standard procedures (Sambrook and Russell 2001). Plasmid DNA was isolated from *E. coli* using the GenElute Plasmid Miniprep Kit (Sigma-Aldrich). PCR reactions were performed using NZYTaq DNA polymerase (NZYTech) or Phusion High-Fidelity DNA polymerase (Finnzymes). All primers used in this work are listed in Table 6.2 and were obtained from Frilabo (Portugal). PCR fragments were purified using the QIAquick PCR Purification Kit (QIAGEN) and DNA fragments were recovered from agarose gels using the QIAquick Gel Extraction Kit (QIAGEN). The restriction enzymes used in this study were purchased from New England BioLabs (USA) or NZYTech (Portugal). T4 DNA ligase was purchased from Promega (USA).

Table 6.1 – Strains and plasmids used in this study.

Strain or Plasmid	Relevant feature(s)	Reference
***A. gossypii* strains**		
ATCC 10895	Parent strain	Prof. P. Philippsen
Agura3GEN3	Derived from ATCC 10895 *Agura3*Δ::*lox*P-*GEN3*-*lox*P	This work
Agade1NATPS	Derived from ATCC 10895 *Agade1*Δ::*lox*P-*NATPS*-*lox*P	This work
Agura3	Derived from *Agura3GEN3 Agura3*Δ::*lox*P	This work
Agade1	Derived from *Agade1NATPS Agura3*Δ::*lox*P	This work
Agura3ade1GEN3	Derived from *Agura3 Agura3*Δ::*lox*P / *Agade1*Δ::*lox*P-*GEN3*-*lox*P	This work
Plasmids		
pGEN3	*GEN3* cassette conferring resistance to geneticin/G418: comprises the *E. coli kan*^{r} gene from transposon Tn*903* under the control of the *S. cerevisiae TEF2* promoter and terminator; ampicilin resistance	Wendland et al. 2000
pUG66	*bleMX* cassette conferring resistance to phleomycin: comprises the bacterial *ble*^{r} gene from transposon Tn*5* under the control of the *A. gossypii TEF* promoter and terminator flanked by two *lox*P sequences; ampicillin resistance	Gueldener et al. 2002
pUC19NATPS	*NATPS* cassette conferring resistance to nourseothricin (clonNAT): comprises the *Streptomyces noursei nat1* gene flanked by the *S. cerevisiae PDC1* promoter and terminator; ampicillin resistance	D. Hoepfner *in* Kaufmann 2009
pUGGEN3	Derivative of pUG66 with the *GEN3* cassette from pGEN3 instead of the *bleMX* cassette	This work
pUGNATPS	Derivative of pUG66 with the *NATPS* cassette from pUC19NATPS instead of the *bleMX* cassette	This work
pUGBLE3	Derivative of pUGGEN3 with the *ble*^{r} gene from pUG66 under the control of the *S. cerevisiae TEF2* promoter and terminator	Ribeiro et al. 2013
pNatCre	*natMX* marker conferring resistance to nourseothricin (clonNAT); comprises the nat1 gene under the control of the *A. gossypii TEF* promoter and terminator; bacteriophage P1 *cre* gene under the control of the *S. cerevisiae GAL1* promoter and CYC1 terminator; *ARSH-CEN6* replication sequences from *S. cerevisiae*; ampicillin resistance	Steensma and Ter Linde 2001
pAgNatCre	Derivative of pNatCre with the *cre* gene under the control of the *S. cerevisiae ADH2* promoter	This work
pAgBleCre	Derivative of pAgNatCre with the *bleMX* marker from pUG66 in place of the *natMX* cassette	This work

Transformation of *A. gossypii* was carried out as described by Wendland et al. (2000), using 30 µg of DNA for integration of deletion cassettes, or 5 µg of DNA for plasmid insertion. *A. gossypii* genomic DNA was extracted as follows. A piece of mycelium collected from the border of *A. gossypii* colonies was transferred to microcentrifuge tubes containing 200 µl of autoclaved DNA extraction buffer (50 mM NaCl, 1 mM EDTA pH 8.0, 10 mM Tris-HCl pH 8.0, and 0.5% (v/v) Triton X-100) and quickly mixed by vortexing. 100 µl of phenol:chloroform:isoamyl alcohol (25:24:1) were

subsequently added, followed by four vortexing cycles of 10 s each, with intercalary cooling on ice. After centrifugation for 15 min at maximum speed, 20 µl of the upper aqueous phase were diluted with 90 µl of water. 1 to 2 µl of these DNA samples were used as template for PCR amplification.

6.2.3 DNA constructs

The deletion cassettes *lox*P-*GEN3*-*lox*P and *lox*P-*NATPS*-*lox*P were constructed as follows. The *GEN3* cassette was amplified from p*GEN3* (Wendland et al. 2000) using primers G1_FW and G2_RV (Table 6.2) and the *NATPS* cassette amplified from pUC19NATPS (Hoepfner *in* Kaufmann 2009) with primers N1_FW and N2_RV (Table 6.2). Both fragments were digested with *Bgl*II and *Xho*I, and cloned between the *lox*P sequences of *Bgl*II/*Xho*I digested pUG66 (Gueldener et al. 2002) to create pUGGEN3 and pUGNATPS, respectivelly. A *lox*P-*BLE3*-*lox*P cassette was also generated as described in Ribeiro et al. (2013). Briefly, the entire pUGGEN3 sequence, with the exception of the *kan*^{r} gene, was amplified with the primers ScTEFt_EcoRI_FW and ScTEFp_NcoI_RV (Table 6.2) and the *ble*^{r} gene obtained by PCR from plasmid pUG66 with the primers Ble_FW and Ble_EcoRI_RV (Table 6.2). Both PCR fragments were digested with *Eco*RI and *Nco*I and ligated to create pUGBLE3.

For deletion of the *AgURA3* (*AEL059W*) gene, the *lox*P-*GEN3*-*lox*P cassette was obtained by PCR from pUGGEN3 with primers G3_FW and G4_RV (Table 6.2). For deletion of the *AgADE1* (*AER221W*) ORF, either *lox*P-*GEN3*-*lox*P or *lox*P-*NATPS*-*lox*P cassettes were obtained by PCR from pUGGEN3 or pUGNATPS, respectively, with primers N3_FW and N4_RV (Table 6.2). The amplified deletion modules comprised the resistance cassettes flanked by *lox*P sequences and 80 bp guide sequences with homology to the 5' and 3' untranslated regions (UTRs) of the target loci, to ensure recombination in the correct location and direction in the genome.

Plasmids pAgNatCre and pAgBleCre were constructed from pNatCre (Steensma and Ter Linde 2001), a shuttle vector that harbours *ARSH*-*CEN6* elements from *S. cerevisiae*, the *natMX* selective marker and the bacteriophage P1 *cre* gene flanked by the *S. cerevisiae GAL1* (*ScGAL1*) promoter and *CYC1* (*ScCYC1*) terminator. For that, a 599 bp fragment containing the *S. cerevisiae ADH2* (*ScADH2*) promoter, from position 872,712 to 873,310 on *S. cerevisiae* chromosome XIII, was amplified from the strain CEN.PK 113-7D genomic DNA with primers AH1_FW and AH2_RV (Table 6.2). The amplified fragment

was digested with the restriction enzymes *Sac*I and *Eco*RI and cloned between the *Sac*I/*Eco*RI sites of pNatCre in place of the *ScGAL1* promoter. This new plasmid was named pAgNatCre. The pAgBleCre plasmid was obtained by the substitution of the *natMX* cassette from pAgNAtCre with the *bleMX* cassette from pUG66 with the restriction enzymes *Bgl*II and *Sac*I. All plasmids used and generated in this study are listed in Table 6.1.

Table 6.2 – Oligonucleotide primers used in this study. Underlined letters are nucleotide sequences complementary to the gene deletion regions. Recognition sequences for restriction enzymes are in **bold**.

Name	Sequence (5' – 3')
G1_FW	GCTAGGGATAACAGGGTAAT
G2_RV	CCG**CTCGAG**GATCTGATGAGGCCGTCT
G3_FW	GATGGTGTAGGCTCCACATCACAGTAAGCATTTGTATAAGGCTGATCACAT AGGGTGCTACCGACCCTAGCCATTGCCACGCAGGTCGACAACCCTTAAT
G4_RV	GGGCATATAATTACAAAAAGGTTGCTGAGATGGTAATCGGCGGTGGAATCC GCCTATATACTGTGGGCGCCAGCGATAGAGCATAGGCCACTAGTGGATC
U1_FW	CCATTGCCACATGTCAACGAAATC
U2_RV	GTGTTGGATGCATGCTACTGAAC
G5_FW	TCGCAGACCGATACCAGGATC
U4_FW	AATCGCTCGAGCAACTCATTGG
N1_FW	GA**AGATCT**TCCCTGCAGAACCGTTACGGTA
N2_RV	CCG**CTCGAG**CCTGCAGCCAAACAGTGTT
N3_FW	CCAGGCAGCGCGAGCTTGACTCTGGTCTATGGCCAATTTAACAGCCCCATA TAGACAAGATATATATAGAGAGATAGCAGGCAGGTCGACAACCCTTAAT
N4_RV	CACATCTCCATTTATCACTTAAAATTATTTAACATTTTAAAAGTACATCATT TTACCATTTCGCGCCTGTAAGTGCTTCGGCATAGGCCACTAGTGGATC
A1_FW	CAATGCGTCTCGAGGAATTTGCTCG
A2_RV	TCGGGTGACGTTAGCGCATCA
A3_RV	AGAAGCAAGCGAGCGACTACG
N5_RV	GATTCGTCGTCCGATTCGTC
AH1_FW	C**GAGCTC**GCAAAGGGGCAAAACGTAG
AH2_RV	CG**GAATTC**CGTGAGTTTCTGGAATAGACATTG
ScTEFt_EcoRI_FW	CG**GAATT**CTCGATGAGTTTTTCTAAGAG
ScTEFp_NcoI_RV	TCTATT**CCATG**GTAATTATAGTTCGTTGACCG
Ble_FW	GCAGGTCGACAACCCTTAAT
Ble_EcoRI_RV	CG**GAATTC**TCATGAGATGCCTGCAAG

6.2.4 Isolation of homokaryotic mutants and diagnostic PCR

Clonal selection of *A. gossypii* homokaryotic mutants was performed through the isolation of single spores from primary heterokaryotic transformants. For that, different

dilutions of spores harvested from primary transformants were spread onto selective AFM plates with a guide line drawn in the bottom. After germination for approximately 24 h, plates were visually analyzed with a Nikon Diaphot 300 inverted microscope (Sony) and single germelings were marked, isolated with a sterile scalpel and transferred to new selective AFM plates. The absence of the gene of interest was tested via diagnostic PCR.

PCR reactions were carried out in 25 µl reaction mixtures with 1-2 µl DNA template, 0.35 µM of each primer, 200 µM dNTP mix, 0.05 U/µl NZYTaq and NZYTaq buffer with 1.5 mM $MgCl_2$ provided by the manufacturer (NZYTech). PCR conditions were: 5 min at 95ºC, 35 cycles of 95ºC for 45 s, 50ºC for 2 min (with increments of 0.3ºC per cycle), 72ºC for 1 min 30 s, and a final extension at 72ºC for 5 min. Correct integration of the *lox*P-*GEN3*-*lox*P cassette in the *AgURA3* and *AgADE1* loci was verified using the primers U1_FW, U2_RV, G5_FW and U3_FW (Table 6.2), and A1_FW, A2_RV, A3_RV and G5_FW (Table 6.2), respectively. Integration of the *lox*P-*NATPS*-*lox*P cassette in the *AgADE1* locus was confirmed using the primers A1_FW, A2_RV, A3_RV and N5_RV (Table 6.2). The positions of the primers and expected PCR amplicon sizes are given in Figure 6.1.

6.2.5 Induction of Cre recombinase and marker cassette excision

For excision of the *lox*P-flanked marker cassettes, the mutant *Agura3GEN3* was transformed with the Cre-expressing plasmid pAgNatCre and the *Agade1NATPS* mutant with the plasmid pAgBleCre. The transformants were selected on AFM containing 50 g/l glucose (to repress the *ScADH2* promoter) and 50 µg/ml of nourseothricin or 5 µg/ml of phleomycin, depending on the resistance conferred by the plasmid. To induce the expression of the Cre recombinase, spores harvested from positive transformant colonies were grown for one week on solid AFM containing 10 g/l ethanol as sole carbon source plus the corresponding antibiotic selection. Simultaneously, the same transformants were also grown on selective AFM containing 50 g/l glucose, as a control for the induction protocol used. Subsequently, spores harvested from these plates were spread onto solid AFM without selection, for plasmid loss. Small patches from the edge of each new mycelial colony were transferred to AFM containing G418 and/or clonNAT (depending on the resistance conferred by the deletion cassette used) and AFM without antibiotic. No growth in plates containing antibiotic indicated successful excision of the marker cassettes through recombination of the *lox*P sequences. If the fungus still had the ability to grow on

G418 and/or clonNAT-containing medium, the induction process was repeated. The removal of the *lox*P-flanked marker cassettes was further confirmed by PCR and stable homokaryotic mutants free of exogenous markers were isolated as described above.

6.3 RESULTS

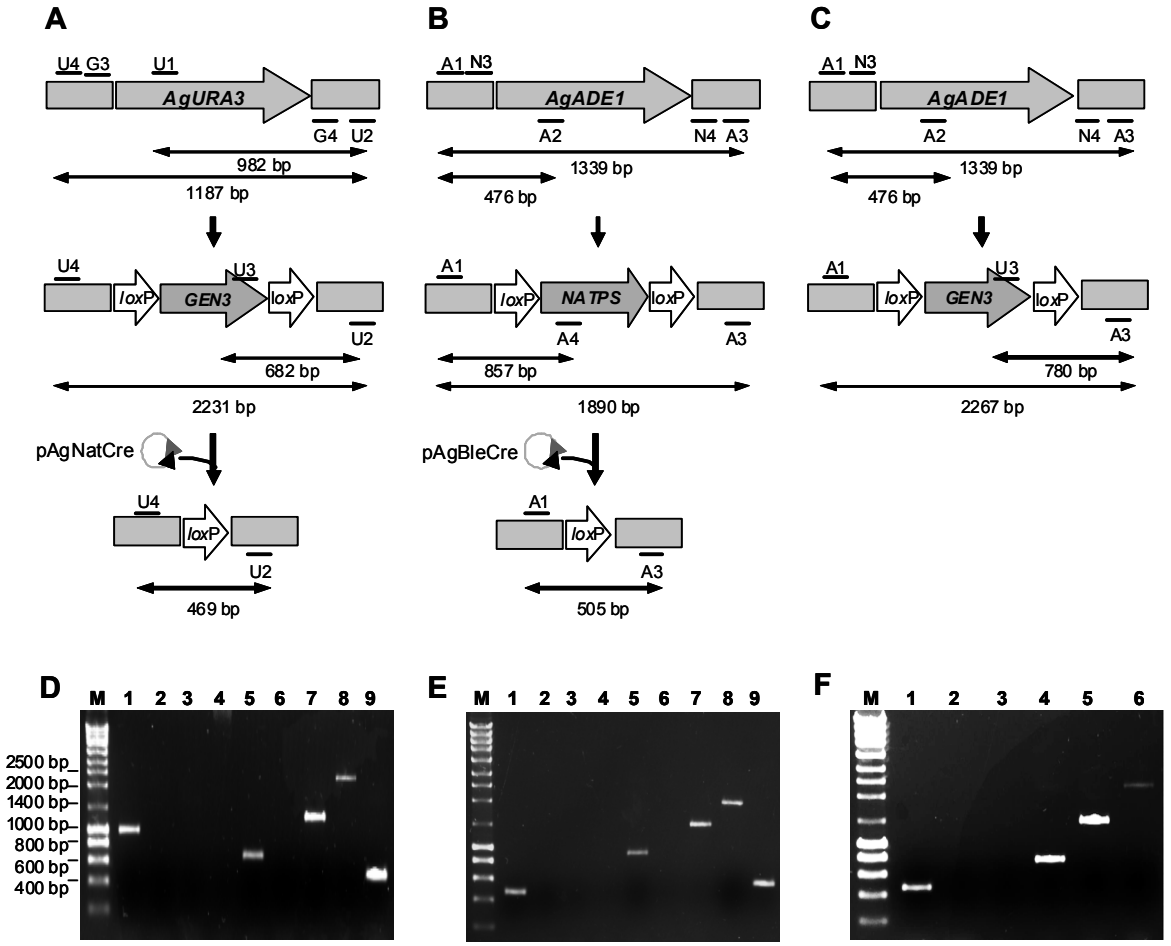


Figure 6.1 – Schematic representation of the strategy used for gene deletion and marker rescue in *A. gossypii* and respective verification PCRs of the generated mutants. **(A)** Construction of the *A. gossypii Agura3* deletion strain. The *AgURA3* gene (*AEL059W*) was replaced by the *lox*P-*GEN3*-*lox*P marker to yield the strain *Agura3GEN3*. This strain was subsequently transformed with the Cre expression plasmid pAgNatCre and excision of the *GEN3* marker occurred after one week of growth in solid AFM containing 10 g/l ethanol as sole carbon source. **(B)** Construction of the *A. gossypii Agade1* deletion strain. The *AgADE1* gene (*AER221W*) was replaced by the *lox*P-*NATPS*-*lox*P marker to yield the strain *Agade1NATPS***.** This strain was transformed with the Cre expression plasmid pAgBleCre and the *NATPS* marker was excised after induction of the Cre expression as described above. **(C)** The *GEN3* marker, which had been excised from the *Agura3* strain, was used again to successfully delete the *AgADE1* gene in this strain, yielding the strain *Agura3ade1GEN3*. **(D)** Verification of the integration of the *lox*P-*GEN3*-*lox*P cassette in the *AgURA3* locus

and confirmation of its excision using the primers U1_FW and U2_RV (lanes 1, 2 and 3), U3_FW and U2_RV (lanes 4, 5 and 6) and U4_FW and U2_RV (lanes 7, 8 and 9). Genomic DNA from the *A. gossypii* parent strain (ATCC 10895) (lanes 1, 4 and 7), *Agura3GEN3* strain (lanes 2, 5 and 8) and *Agura3* strain (lanes 3, 6 and 9) were used as template. **(E)** Verification of the integration of the *lox*P-*NATPS*-*lox*P cassette in the *AgADE1* locus and confirmation of its excision using the primers A1_FW and A2_RV (lanes 1, 2 and 3), A1_FW and A4_RV (lanes 4, 5 and 6), and A1_FW and A3_RV (lanes 7, 8 and 9). Genomic DNA from the *A. gossypii* parent strain (ATCC 10895) (lanes 1, 4 and 7), *Agade1NATPS* strain (lanes 2, 5 and 8) and *Agade1* strain (lanes 3, 6 and 9) were used as template. **(F)** Verification of the integration of the *lox*P-*GEN3*-*lox*P cassette in the *AgADE1* locus of the *Agura3* strain using the primers A1_FW and A2_RV (lane 1 and 2), U3_FW and A3_RV (lane 3 and 4), and A1_FW and A3_RV (lane 5 and 6). Genomic DNA from the *A. gossypii* parent strain (ATCC 10895) (lanes 1, 3 and 5) and from the *Agura3ade1GEN3* strain (lanes 2, 4 and 6) were used as template. The positions of the primers and expected PCR amplicon sizes are indicated in the scheme.

6.3.1 Constructs generated for marker recycling in *A. gossypii*

New disruption cassettes for use in *A. gossypii*, which could be excised after being used to screen for a genetic modification and then reused to produce multiple gene deletions in the same strain, were constructed using as backbone the plasmid pUG66 (Gueldener et al. 2002). The *lox*P–*GEN3*–*lox*P and *lox*P-*NAPS*-*lox*P cassettes were obtained by replacing the *bleMX* unit located between the two *lox*P sites in plasmid pUG66 with the *GEN3* and *NATPS* units from plasmids pGEN3 (Wendland et al. 2000) and pUC19NATPS (Hoepfner *in* Kaufmann 2009), respectively, yielding plasmids pUGGEN3 and pUGNATPS (Figure 6.2). The *ble*^{r} gene from transposon Tn*5*, which confers resistance to phlemoycin (Gatignol et al. 1987), was used to create the new heterologous dominant drug resistance marker cassette *lox*P–*BLE3*–*lox*P, as described in Ribeiro et al. (2013). The *ble*^{r} gene was isolated by PCR from plasmid pUG66 and placed between the *ScTEF2* promoter and terminator in plasmid pUGGEN3, yielding plasmid pUGBLE3 (Figure 6.2).

New plasmids for transient expression of the Cre recombinase in *A. gossypii* were constructed using as backbone the replicative *ARSH-CEN6* plasmid pNatCre (Steensma and Ter Linde 2001). This plasmid harbours a Cre expression cassette where the *cre* gene is flanked by the *ScGAL1* promoter and *ScCYC1* terminator sequences. Several *GAL* pathway genes are absent in the *A. gossypii* genome (Hittinger et al. 2004), including the *ScGAL1* homolog. Hence, the *ScADH2* promoter was used to regulate the Cre recombinase expression in *A. gossypii*. In *S. cerevisiae*, the *ScADH2* promoter is repressed several

hundred-fold in the presence of glucose and induced once glucose in the medium is exhausted (Lee and DaSilva 2005). Glucose has also been observed to have a repressive effect over the transcription of the *ScADH2* homolog gene in *A. gossypii* (*AAR084W*) when present in a concentration as low as 15 g/l, whereas the presence of ethanol in the medium induced it transcription (our microarray data). Therefore, in the plasmid pAgNatCre, the 467 bp *ScGAL1* promoter was substituted with the 599 bp *ScADH2* promoter. This promoter allowed the expression of the Cre recombinase in *A. gossypii* on rich medium containing ethanol as carbon source, as will be further demonstrated. Plasmid pAgBleCre was obtained by substituting the *natMX* drug marker in pAgNatCre with the *bleMX* expression cassette from plasmid pUG66. The usable range of phleomycin concentrations for selection of *A. gossypii* phleomycin-resistant colonies was determined and the optimal concentration for selection was found to be 5 µg/ml phleomycin (in agar-solidified AFM).

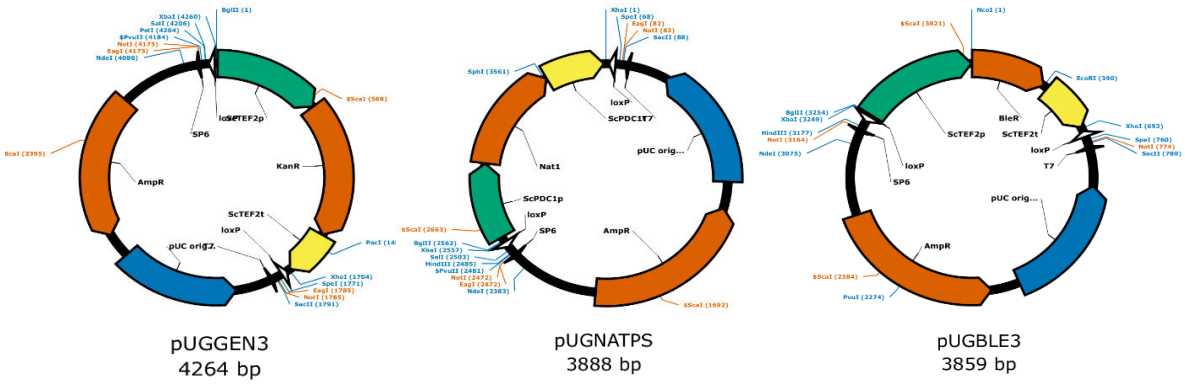


Figure 6.2 – Graphic representation of the plasmids pUGGEN3, pUGNATPS and pUGBLE3. Images created with PlasmaDNA v1.4.2 (University of Helsinki).

6.3.2 Deletion of the *AgURA3* gene and generation of an uracil auxotrophic strain

As a model to test the functionality of the Cre-*lox*P system in *A. gossypii,* the entire coding region of the *AgURA3* (*AEL059W*) gene was deleted from the *A. gossypii* genome using the *lox*P-*GEN3*-*lox*P cassette flanked by 80 bp guide sequences with homology to the *AgURA3* 5' and 3' UTRs. Homologous integration of this deletion module within the *AgURA3* locus resulted in the replacement of *AgURA3* by the *lox*P-*GEN3*-*lox*P cassette with its transcriptional orientation corresponding to that of the deleted gene (Figure 6.1A). All the primary transformants randomly screened were heterokaryotic, containing nuclei

with both wild type and null alleles of *AgURA3* (data not shown). After a single spore purification step, diagnostic PCR indicated the sole presence of the mutant allele in the *Agura3GEN3* strain (Fig. 6.1D).

Deletion of the *AgURA3* gene caused uracil auxotrophy, being the homokaryotic *Agura3GEN3* strain unable to grow on defined medium without uracil (Figure 6.3). Moreover, contrary to the parent strain, the *Agura3GEN3* strain was also able to grow on solid AFM containing 1 mg/ml 5-FOA (Figure 6.3).

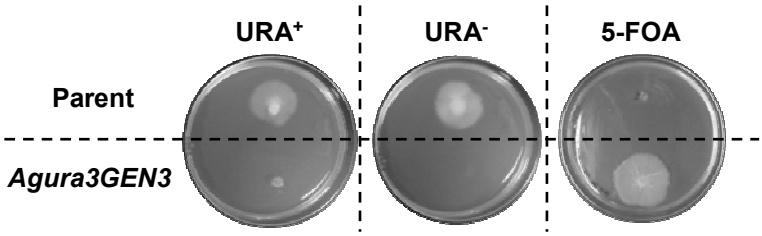


Figure 6.3 – Characterization of the *A. gossypii Agura3GEN3* mutant growth on agar-solidified SC medium with (URA^+) and without (URA^-) uracil and on AFM supplemented with 5-FOA (1 mg/ml). Plates were incubated for one week at 30ºC and the parental strain (ATCC 10895) used as control.

6.3.3 Transient expression of Cre recombinase and marker rescue

The *Agura3GEN3* strain was transformed with the plasmid pAgNatCre to allow the transient expression of Cre recombinase. Transformants were selected on rich AFM containing clonNAT and 50 g/l glucose, to assure the repression of the *ScADH2* promoter and, thus, of the Cre recombinase expression. Spores harvested from a randomly selected transformant colony were grown for one week on solid AFM containing clonNAT and 10 g/l ethanol as sole carbon source to induce the Cre expression. The spores collected after this induction period were spread onto non-selective AFM, to promote plasmid loss, and the sensitivity of the resulting colonies to clonNAT and G418 tested. Of the 10 mycelial colonies screened none was able to grow on AFM containing clonNAT, indicating that all had lost the plasmid that contained the *natMX* resistance marker after only one passage in non-selective conditions. One colony was also unable to grow on AFM containing G418, suggesting that the *GEN3* marker had been looped out and lost. After a single spore purification step, the excision of the *GEN3* cassette from the *Agura3* locus was confirmed

by PCR (Figure 6.1D). This homokaryotic strain free of exogenous marker genes was named *Agura3*.

To test weather the excision of the *GEN3* cassette occurred without the induction step in medium containing ethanol, spores from the randomly selected transformant colony were simultaneously grown on solid AFM containing clonNAT and 50 g/l glucose. In this case, no strain lost its G418 resistance. Altogether, these data suggest that the *ScADH2* promoter was able to differentially regulate the Cre recombinase expression in *A. gossypii*.

6.3.4 Demonstration of marker reuse by deletion of the *AgADE1* gene from the genome of the *Agura3* strain

To demonstrate the successful reuse of the *GEN3* marker, the predicted *AgADE1* (*AER221W*) ORF was deleted from the genome of the *Agura3* strain using the *lox*P-*GEN3*-*lox*P cassette previously used to delete the *AgURA3* ORF, which was flanked by 80 bp guide sequences with homology to the *AgADE1* 5' and 3' UTRs. The correct integration of the deletion cassette in the *AgADE1* locus was verified by PCR as shown in Figure 6.1F. Homokaryotic *Agura3ade1GEN3* mutants were obtained after a single spore purification step (Figure 6.1F).

Physiological characterization of this mutant revealed that it was auxotrophic for both uracil and adenine, as it was unable to grow in defined medium lacking one or both of these nucleobases (Figure 6.4). This strain also continued to be resistant to 5-FOA (1 mg/ml).

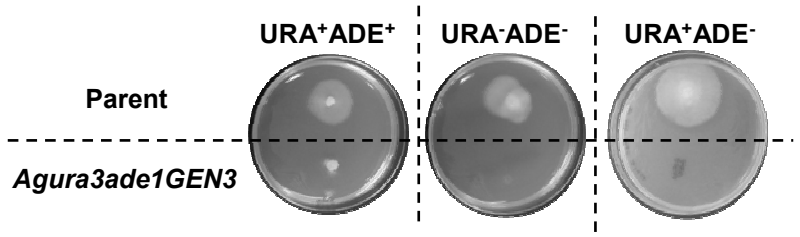


Figure 6.4 – Characterization of the *A. gossypii Agura3ade1GEN3* mutant growth on agar-solidified SC medium with (URA^+ADE^+) and without (URA^-ADE^-) uracil and adenine and with uracil and without adenine (URA^+ADE^-). Plates were incubated for one week at 30ºC and the parental strain (ATCC 10895) used as control.

6.3.5 Deletion of the *AgADE1* gene and generation of a stable adenine auxotrophic strain

To demonstrate the versatility of the developed Cre-*lox*P system, the *AgADE1* ORF was also deleted using the *lox*P-*NATPS*-*lox*P cassette. The correct integration of this deletion cassette was verified by PCR, as shown in Figure 6.1E, and the resulting homokaryotic strain was named *Agade1NATPS*.

This strain was then transformed with the Cre expression plasmid pAgBleCre, transformants were selected on AFM containing phleomycin and 50 g/l glucose, and Cre recombinase was transiently expressed as described above. Following this strategy, the *lox*P-flanked *NATPS* cassette could be looped out of the chromosomal DNA of *Agade1NATPS* transformant colonies. After single spore purification, a homokaryotic *Agade1* strain free of marker genes (Figure 6.1E) and auxotrophic for adenine was obtained.

6.4 DISCUSSION

Here, the Cre-*lox*P recombination system of the bacteriophage P1 was adapted to genetically engineer *A. gossypii* in a targeted way and allow the removal and reuse of the selection markers used during the process. Three heterolougous recyclable drug marker cassettes and two Cre recombinase expression vectors for use in *A. gossypii* were created and successfully used to generate stable *Agura3* and *Agade1* auxotrophic strains free of exogenous selectable markers.

The *AgURA3* gene encodes the orotidine-5'-phosphate decarboxylase involved in the biosynthesis of pyrimidines and, as previously reported (Pompejus et al. 1999), its deletion caused uracil auxotrophy and conferred resistance to 5-FOA. In the parent strain, growth was hampered by the presence of 5-FOA in the medium, which indicates that the AgUra3p converts 5-FOA into 5-fluorouracil, a toxic compound that causes cell death (Boeke et al. 1984). Therefore, the use of the *AgURA3* gene as a bidirectionally selectable marker (confering uracil prototrophy and 5-FOA sensitivity) may be considered in future *A. gossypii* engineering projects. However, it is worth noting that in previous reports neither the *S. cerevisiae URA3* nor the *E. coli pyrF* genes were able to complement the uracil auxotrophy in *A. gossypii Agura3* mutants (Pompejus et al. 1999). Therefore, efforts

are now being made to develop an expression cassette that can complement the *AgURA3* auxotrophy.

Based on the analysis of the annotated genome sequence of *A. gossypii* (http://agd.vital-it.ch; Gattiker et al. 2007), the *AER221W* ORF (*AgADE1*) was found to be homologous to the *S. cerevisiae ADE1* gene. The generated *Agade1* null mutant was auxotrophic for adenine, as described for *S. cerevisiae ade1* null mutants (Fisher 1969). The deletion of the *AgADE4* gene has been also previously shown to cause adenine auxotrophy (Jiménez et al. 2005). In *S. cerevisiae*, the *ADE1* gene encodes the n-succinyl-5-aminoimidazole-4-carboxamide ribotide (SAICAR) synthetase involved in the purine biosynthetic pathway and its deletion causes an accumulation of a red pigment in mutant cells deprived of adenine (Fisher 1969). This was also observed in the *A. gossypii Agade1* null mutants (Figure 6.5).

Figure 6.5 – *A. gossypii Agade1* mutant colony grown on agar-solidified AFM for one week at 30ºC.

For *S. cerevisiae*, several recyclable marker cassettes and Cre expression plasmids are available (Güldener et al. 1996; Steensma and Ter Linde 2001; Gueldener et al. 2002). Unfortunately, these cannot be used in *A. gossypii* for several reasons. First, most of the available cassettes are not heterologous to the *A. gossypii* genome, as they contain the *A. gossypii TEF* promoter and terminator sequences. Second, the heterologous cassettes available are auxotrophic markers, therefore requiring the existence of auxotrophic strains. Third, the *cre* gene present in the available plasmids is under the regulation of the *ScGAL1* promoter (Steensma and Ter Linde 2001; Gueldener et al. 2002), which cannot be used in *A. gossypii* because this fungus lacks pathways for the utilization of galactose in its genome.

Therefore, the heterologous *GEN3* and *NATPS* markers widely used for PCR-based gene targeting in *A. gossypii* were adapted to contain *lox*P flanking sequences and successfully used to select *Agura3* and *Agade1* mutants. A new *lox*P-flanked heterologous

drug resistance cassette was also constructed, which comprises the *ble*^{r} gene from transposon Tn*5* fused to the *ScTEF2* promoter and terminator (*BLE3*). This marker cassette renders resistance to phleomycin, a copper-containing antibiotic that acts by interfering with DNA synthesis (Gatignol et al. 1987). Although the *BLE3* marker was not used to delete any of the genes presented here, it was successfully used to generate and select *A. gossypii Aggas1agas1b* null mutants (Ribeiro et al. 2013).

In the generated Cre expression plasmids, the *ScGAL1* promoter was substituted by the *ScADH2* promoter to allow transient expression of the Cre recombinase in *A. gossypii*. In *S. cerevisiae*, the regulation of the *ScADH2* promoter is well studied, being its expression repressed several hundred-fold in the presence of glucose and activated when the glucose is depleted from the medium (Lee and DaSilva 2005). In *A. gossypii*, as this promoter had never been used before little was known about its functionality. However, previous observations had shown that the transcription of the *A. gossypii ScADH2* homolog (*AAR084W*) was also repressed by the presence of glucose in the medium and induced by ethanol (our microarray data). Hence, it was expected that the *ScADH2* promoter would be similarly regulated in *A. gossypii* and that a straightforward way to induce the Cre recombinase expression would be to transfer the *A. gossypii* mutants transformed with the newly constructed Cre expression plasmids from medium containing high concentration glucose to medium containing ethanol. This strategy proved efficient in generating mutant strains free of exogenous markers, which indicates that the *ScADH2* promoter is functional in *A. gossypii*. Moreover, as no antibiotic sensitive colonies were recovered without induction, the *ScADH2* promoter also proved to regulate the expression of the Cre recombinase in *A. gossypii*.

Other promoters have been already used to regulate gene expression in *A. gossypii*: *A. gossypii* and *S. cerevisiae MET3* promoters (Dünkler and Wendland 2007), and *A. gossypii* and *S. cerevisiae THI13* promoters (Kaufmann 2009). However, their regulation requires the use of synthetic medium and *A. gossypii* grows much faster in complex medium. Therefore, the *ScADH2* promoter offers advantage over these promoters because it can be induced in complex medium, just by shifting the cells from medium containing glucose (where it is repressed) to medium containing ethanol (where it is induced).

In summary, the method here presented for multiple gene disruption and/or integration in *A. gossypii* offers advantages over the existing methods by enabling the creation of mutants entirely free of foreign genes and allowing the reuse of selection markers. The set of disruption cassettes and plasmids constructed greatly expand the

possibilities for genetically engineer *A. gossypii*, being these suitable for use in both laboratorial and industrial strains, as they do not required any predetermined genetic background. Furthermore, since the number of stable *A. gossypii* auxotrophic strains available is limited, the strains generated here are envisaged to be useful for future genetic engineering projects. Efforts are now being made to develop a selection system based on the complementation of these auxotrophic mutants.

CHAPTER 7

Conclusions and future perspectives

The filamentous fungus *A. gossypii* has recently emerged as an interesting organism to explore as a cell factory. Not only has it been safely and successfully used for more than two decades in the industrial production of riboflavin (vitamin B2), as it combines several attributes that make it an attractive host to produce other yet unexploited bio-products in addition to riboflavin, such as heterologous proteins. This thesis focused on the characterization of the *A. gossypii* protein secretory pathway at the genomic, transcriptomic and proteomic levels, and on the investigation of strategies to improve its heterologous protein secretion capacities.

The genome-wide study of the *A. gossypii* predicted secretome (i.e. repertoire of native proteins extracellularly secreted) and the experimental mapping of the proteins natively secreted by this fungus under two different culture conditions by 2-D electrophoresis revealed that only 1-4% of its predicted proteome is secreted. This percentage is in the range of what has been described for yeast secretomes, which are known to be smaller than filamentous fungi secretomes. The low concentration of proteins secreted by *A. gossypii* to the culture supernatants is also in line with what is generally observed in yeast.

Despite its complex way of growing as a multinucleated mycelium, given the minimal set of genes of this organism (which has the smallest free living eukaryotic genome known to data), its close phylogenetic relationship with yeast and the limited range of carbon sources it utilises, it was expected that *A. gossypii* wouldn't secrete a large variety of proteins to the medium. This was confirmed by our results, which also offered further insights into the putative functionality of the proteins predicted to be secreted by this fungus. For instance, nine putative proteases were predicted to be secreted, even though extracellular protease activity in *A. gossypii* culture supernatants has been undetectable, probably due to low protease concentration and non-optimal conditions for its activity. A putative lipase (AER454C) and two putative *β*-glucosidases (AGL354C and AGL343C) were also predicted to be secreted. Both these enzymatic activities have been previously detected in *A. gossypii* culture supernatants, but the genes encoding these enzymes have not been characterized thus far. A putative invertase (AFR529W) yet uncharacterized in *A. gossypii* was also predicted to be secreted by this fungus. This invertase was further molecularly and functionally characterized and confirmed to be targeted to the extracellular cell-wall and also secreted to the culture medium in *A. gossypii*. Its production was found to be regulated by the fermentable sugars present in the medium.

The awareness about the proteins putatively secreted by *A. gossypii* is envisaged to have a positive impact on the development of this fungus as a heterologous protein producer, not only at the quality control level, to help prevent contamination and/or degradation of the recombinant products, but also at the strain and/or vector design level. As an example, these proteins putatively secreted by *A. gossypii* (such as α-factor mating pheromones (α-MF)) are good candidates for signal peptide screening, thereby increasing the signal sequence repertoire used for protein secretion in *A. gossypii*. For instance, the *S. cerevisiae* α-MF leader sequence is widely used in many yeast expression vectors and in some cases it has been shown to lead the secretion of higher amounts of recombinant protein than if using the native signal peptide.

The study of the transcriptional responses of *A. gossypii* to heterologous protein secretion also provided useful information about possible strategies that could much improve the productivity of *A. gossypii*. Among these, the use of stronger promoters and/or better expression strategies (like genome integration) were found to be of most importance, as the expression levels of one of the first two heterologous proteins reported to be secreted by *A. gossypii* (EGI from *T. reesei* preceded with the *S. cerevisiae PGK1* constitutive promoter) were rather low. In this context, the strength of several native (*A. gossypii TEF* and *GPD*) and heterologous (*S. cerevisiae PGK1* and *ADH1*) constitutive promoters was compared when expressing a model heterologous protein of other origin, β-galactosidase from *A. niger*. Given the ease of detection of this secreted enzyme, the screening of transformants was facilitated. Moreover, as this protein had already been efficiently secreted by recombinant *S. cerevisiae* strains in high amounts, it provided a good model protein to further explore the potential of *A. gossypii* as a heterologous protein producer.

The extracellular β-galactosidase from *A. niger* was, therefore, for the first time expressed and secreted by *A. gossypii*, enlarging the pool of heterologous proteins being successfully secreted by this fungus to the culture medium. The strong *A. gossypii* constitutive *TEF* promoter was shown to drive the highest expression levels of this enzyme from 2-micron plasmids either in *A. gossypii* and *S. cerevisiae* laboratorial strains. For the first time, the levels of active heterologous enzyme secreted by *A. gossypii* were (up to 37 times) higher than those secreted by recombinant *S. cerevisiae* strains transformed with the same plasmids. In similar culture conditions, *A. gossypii* also produced 2.5 times more extracellular β-galactosidase than that previously reported for a β-galactosidase-highproducing *S. cerevisiae* strain transformed with a different plasmid,

where β-galactosidase was under the regulation of the *S. cerevisiae ADH1* promoter. In flask productions in complex medium containing glycerol instead of glucose as primary carbon source, the secretion of extracellular β-galactosidase by *A. gossypii* increased approximately 1.5-fold, which indicates that much can still be improved just by manipulating culture conditions.

Besides good expression capacity, *A. gossypii* also presented in this work other desirable features for a heterologous protein producer. The native signal peptide of the *A. niger* β-galactosidase and *S. cerevisiae* invertase (*SUC2* gene) were recognised by *A. gossypii* as secretion signals, directing the production of these proteins into the culture medium, which facilitates its downstream processing. *A. gossypii* had previously recognized the native signal sequences of *T. reesei* cellulases EGI and CBHI. It, therefore, seems to efficiently recognize a wide range of heterologous signal sequences, which is advantageous in a recombinant host. The rather small number and low concentration of proteins secreted by *A. gossypii* to the culture medium, together with the negligible extracellular protease activity also represent important advantages for heterologous protein production in this host, as secreted products are less likely to be contaminated or degraded by its native proteins. Furthermore, recombinant β-galactosidase secreted by *A. gossypii* was apparently glycosylated, as were the recombinant *T. reesei* EGI and CBHI previously secreted by *A. gossypii*. Despite its higher heterogeneity, the same tendency to a less extensive hyperglycosylation of the recombinant β-galactosidase produced by *A. gossypii* in comparison to that produced by *S. cerevisiae* was also apparent in this work.

To characterize the *N*-glycans assembled on the glycoproteins natively secreted by *A. gossypii*, we undertook a global analysis to them using MALDI-TOF mass spectrometric profiling and NMR spectroscopy. The obtained data revealed that the *A. gossypii* secreted *N*-glycome is characterized by high-mannose type structures in the range $Man_{4-18}GlcNAc_2$, mostly containing neutral core-type *N*-glycans with 8 to 10 mannoses. This is similar to what has been reported as the major *N*-glycan species usually assembled on the glycoproteins from yeast species ($Man_{8-13}GlcNAc_2$). Filamentous fungi mostly produce *N*-glycans in the range $Man_{5-9}GlcNAc_2$. Yeast species usually hyperglycosylate its *N*-glycans, which is not so usual in filamentous fungi. Here, the culture conditions have been shown to influence the extent of *N*-glycosylation and *N*-glycan phosphorylation in *A. gossypii*. Homologs for all of the *S. cerevisiae* genes known to be involved in the ER and Golgi *N*-glycan processing were found in the *A. gossypii* genome. However, processing of

N-glycans by *A. gossypii* differs considerably from that by *S. cerevisiae*, allowing much shorter *N*-glycans. Two genes coding for putative *N*-glycan processing enzymes that were identified in the *A. gossypii* genome and which have no homologs in *S. cerevisiae* may possibly contribute for the observed differences between these organisms. Therefore, we are initiating the molecular and functional characterization of these genes to assess their role in the *A. gossypii N*-glycan biosynthetic pathway.

Analysis of the transcriptional responses of *A. gossypii* to DTT-induced protein secretion stress surprisingly revealed that *A. gossypii* did not activate a conventional UPR, as generally observed in other fungi (both yeast and filamentous fungi) when exposed to similar stress conditions. However, there was evidence of secretion stress induced by DTT in *A. gossypii*, as the expression of several genes involved in protein unfolding, ERAD, proteasome degradation, proteolysis, vesicle trafficking, vacuolar protein sorting and secretion significantly increased after DTT addition. This suggests that alternative mechanisms exist in *A. gossypii* to cope with protein secretion stress. The nature of its regulation is, however, still unknown. Therefore, further studies should focus on the study of the regulatory mechanisms that modulate the *A. gossypii* responses to secretion stress. A first strategy could be to try to overexpress the active form of the *S. cerevisae HAC1* in *A. gossypii* (as the active form of the *A. gossypii HAC1* homolog hasn't yet been identified) and evaluate the expression of several UPR target genes, to assess if a conventional UPR could be activated in this way.

In summary, the results presented herein demonstrate that the potential of *A. gossypii* as a heterologous protein production host is still underexplored and much could be learned about possible strategies to improve its productivity, among which: (**1**) culture medium optimization; (**2**) screening for better promoters and secretion signal sequences; (**3**) development of better expression strategies, preferably through the integration of stable expression cassettes. Soon, the genome-scale metabolic model of *A. gossypii* will be available (a task under development in project AshByofactory) and *in silico* simulations will facilitate the optimization of production conditions, and will help design strains with improved productivities. The molecular tool developed in the scope of this thesis and which now allows the creation of stable *A. gossypii* mutant strains free of exogenous selective markers will be very useful to generate superior strains for safe use in industrial applications.

A considerable advance was given in the understanding of the biotechnological potential of *A. gossypii* as a heterologous protein producer. However, with the availability

of its genome-scale metabolic model, the production of other metabolites of interest may be rationally evaluated. Further physiological characterization of *A. gossypii*, particularly in oil and biodiesel waste-based media, would also offer new insights into other possible applications, like for example, in the bioremediation technology area.

REFERENCES

Alberti-Segui C, Dietrich F, Altmann-Jöhl R, Hoepfner D, Phillipsen P. 2001. Cytoplasmic dynein is required to oppose the force that moves nuclei towards the hyphal tip in the filamentous ascomycete *Ashbya gossypii*. *J Cell Sci*. 114:975–986.

Alméciga-Díaz CJ, Gutierrez AM, Bahamon I, Rodríguez A, Rodríguez MA, Sánchez OF. 2011. Computational analysis of the fructosyltransferase enzymes in plants, fungi and bacteria. *Gene*. 484:26–34.

Al-Sheikh H, Watson AJ, Lacey GA, Punt PJ, MacKenzie DA, Jeenes DJ, Pakula T, Penttilä M, Alcocer MJC, Archer DB. 2004. Endoplasmic reticulum stress leads to the selective transcriptional downregulation of the glucoamylase gene in *Aspergillus niger*. *Mol Microbiol*. 53:1731–1742.

Altmann-Jöhl R, Philippsen P. 1996. *AgTHR4*, a new selection marker for transformation of the filamentous fungus *Ashbya gossypii,* maps in a four-gene cluster that is conserved between *A. gossypii* and *Saccharomyces cerevisiae*. *Mol Gen Genet*. 250:69–80.

Arvan P, Zhao X, Ramos-Castaneda J, Chang A. 2002. Secretory pathway quality control operating in golgi, plasmalemmal, and endosomal systems. *Traffic*. 3:771–780.

Arvas M, Pakula T, Lanthaler K, Saloheimo M, Valkonen M, Suortti T, Robson G, Penttilä M. 2006. Common features and interesting differences in transcriptional responses to secretion stress in the fungi *Trichoderma reesei* and *Saccharomyces cerevisiae*. *BMC Genomics*. 7:32.

Ashby SF, Nowell W. 1926. The Fungi of Stigmatomycosis. *Ann Bot*. 40:69–84.

Ayad-Durieux Y, Knechtle P, Goff S, Dietrich F, Philippsen P. 2000. A PAK-like kinase is required for maturation of young hyphae and septation in the filamentous ascomycete *Ashbya gossypii*. *J Cell Sci*. 113:4563–4575.

Ballou L, Hernandez LM, Alvarado E, Ballou CE. 1990. Revision of the oligosaccharide structures of yeast carboxypeptidase Y. *Proc Natl Acad Sci USA*. 87:3368–3372.

Batra LR. 1973. Nematosporaceae (Hemiascomycetidae): taxonomy, pathogenicity, distribution, and vector relations. *US Dept Agr Tech Bull*. 1469:1–71.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol*. 340:783–795.

Bigelis R. 1989. Industrial products of biotechnology: application of gene technology. In: *Biotechnology* vol. 7b. Rehm HJ, Reed G (eds). VCH, Weinheim. 230–269.

Blanchard V, Gadkari RA, Gerwig GJ, Leeflang BR, Dighe RR, Kamerling JP. 2007. Characterization of the *N*-linked oligosaccharides from human chorionic gonadotropin expressed in the methylotrophic yeast *Pichia pastoris*. *Glycoconj J*. 24:33–47.

Boeke JD, LaCroute F, Fink GR. 1984. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol Gen Genet*. 197:345–346.

Böer E, Wartmann T, Luther B, Manteuffel R, Bode R, Gellissen G, Kunze G. 2004. Characterization of the *AINV* gene and the encoded invertase from the dimorphic yeast *Arxula adeninivorans*. *Antonie van Leeuwenhoek*. 86:121–134.

Brachat S, Dietrich FS, Voegeli S, Zhang Z, Stuart L, Lerch A, Gates K, Gaffney T, Philippsen P. 2003. Reinvestigation of the *Saccharomyces cerevisiae* genome annotation by comparison to the genome of a related fungus: *Ashbya gossypii*. *Genome Biol*. 4:R45.

Brustolini OJB, Fietto LG, Cruz CD, Passos FML. 2009. Computational analysis of the interaction between transcription factors and the predicted secreted proteome of the yeast *Kluyveromyces lactis*. *BMC Bioinformatics*. 10:194.

Buser R, Lazar Z, Käser S, Künzler M, Aebi M. 2010. Identification, characterization, and biosynthesis of a novel *N*-glycan modification in the fruiting body of the basidiomycete *Coprinopsis cinerea*. *J Biol Chem*. 285:10715–10723.

Buston HW, Kasinathan S. 1933. The accessory factor necessary for the growth of *Nematospora gossypii*. III. The preparation of concentrates of the second accessory factor. *Biochem J*. 27:1859–1868.

Buston HW, Kasinathan S, Wylie SM. 1938. The nitrogen requirements of *Nematospora gossypii* in synthetic media. *Ann Bot*. 2:373–379.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res*. 37:D233–D238.

Caramelo JJ, Parodi AJ. 2008. Getting in and out from calnexin/calreticulin cycles. *J Biol Chem*. 283:10221–10225.

Carlson M, Botstein D. 1982. Two differentially regulated mRNAs with different 5′ ends encode secreted and intracellular forms of yeast invertase. *Cell*. 28:145–154.

Carvalho ND, Arentshorst M, Kooistra R, Stam H, Sagt CM, van den Hondel CA, Ram AF. 2011. Effects of a defective ERAD pathway on growth and heterologous protein production in *Aspergillus niger*. *Appl Microbiol Biotechnol*. 89:357–373.

Carvalho P, Goder V, Rapoport TA. 2006. Distinct ubiquitin-ligase complexes define convergent pathways for the degradation of ER proteins. *Cell*. 126:361–373.

Caskey C, Gallup WD. 1931. Changes in the sugar, oil, and gossypol content of the developing cotton boll. *J Agric Res*. 42:671–673.

Chen JS, Saxton J, Hemming FW, Peberdy JF. 1996. Purification and partial characterization of the high and low molecular weight form (S- and F-form) of invertase secreted by *Aspergillus nidulans*. *Biochim Biophys Acta*. 1296:207–218.

Chi Z, Liu Z. 1997. Inositol-mediated invertase secretion in *Saccharomyces* sp. W4. *Enzyme Microb Technol*. 21:463–467.

Choi BK, Bobrowicz P, Davidson RC, Hamilton SR, Kung DH, Li H, Miele RG, Nett JH, Wildt S, Gerngross TU. 2003. Use of combinatorial genetic libraries to humanize *N*-linked glycosylation in the yeast *Pichia pastoris*. *Proc Natl Acad Sci USA*. 100:5022–5027.

Conesa A, Punt PJ, van Luijk N, van den Hondel CA. 2001. The secretion pathway in filamentous fungi: a biotechnological view. *Fungal Genet Biol*. 33:155–171.

Cox JS, Shamu CE, Walter P. 1993. Transcriptional induction of genes encoding endoplasmatic reticulum resident proteins requires a transmembrane protein kinase. *Cell*. 73:1197–1206.

Cox JS, Walter P. 1996. A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response. *Cell*. 87:391–404.

Cubero B, Scazzocchio C. 1994. Two different, adjacent and divergent zinc finger binding sites are necessary for CREA-mediated carbon catabolite repression in the proline gene cluster of *Aspergillus nidulans*. *EMBO J*. 13:407–415.

Cudna RE, Dickson AJ. 2003. Endoplasmic reticulum signaling as a determinant of recombinant protein expression. *Biotechnol Bioeng*. 81:56–65.

Cullen PJ, Sprague GF Jr. 2012. The regulation of filamentous growth in yeast. *Genetics*. 190:23–49.

Cummings RD, Doering TL. 2009. Fungi. In: *Essentials of Glycobiology* 2nd Ed. Varki A, Cummings RD, Esko JD, et al. (eds). Cold Spring Harbor Laboratory Press, New York. Chapter 21.

Damasceno LM, Huang CJ, Batt CA. 2012. Protein secretion in *Pichia pastoris* and advances in protein production. *Appl Microbiol Biotechnol*. 93:31–39.

Dammer KH, Ravelo HG. 1990. Verseuchung von *Leptoglossus gonagra* (Fabr.) mit *Nematospora coryli* Peglion und *Ashbya gossypii* (Ashby et Nowell) Guilliermond in einer zitrusanlage der Republik Kuba. *Arch Phytopathol Pflanzenschutz Berlin*. 26: 71–78.

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

Demain AL. 1972. Riboflavin oversynthesis. *Annu Rev Microbiol*. 26:369–388.

Demain AL, Vaishnav P. 2009. Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv*. 27:297–306.

Denis CL, Young ET. 1983. Isolation and characterization of the positive regulatory gene *ADR1* from *Saccharomyces cerevisiae*. *Mol Cell Biol*. 3:360–370.

Deshpande N, Wilkins MR, Packer N, Nevalainen H. 2008. Protein glycosylation pathways in filamentous fungi. *Glycobiology*. 18:626–637.

Dietrich FS, Voegeli S, Brachat S, Lerch A, Gates K, Steiner S, Mohr C, Pohlmann R, Luedi P, Choi S, Wing RA, Flavier A, Gaffney TD, Philippsen P. 2004. The

Ashbya gossypii genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. *Science*. 304:304–307.

Dietrich FS, Voegelli S, Kuo S, Philippsen P. 2013. Genomes of *Ashbya* fungi isolated from insects reveal four mating-type loci, numerous translocations, lack of transposons, and distinct gene duplications. *G3 (Bethesda)*. pii: g3.112.002881v1.

Domingues L, Onnela ML, Teixeira JA, Lima N, Penttilä M. 2000. Construction of a flocculent brewer's yeast strain secreting *Aspergillus niger* beta-galactosidase. *Appl Microbiol Biotechnol*. 54:97–103.

Domingues L, Teixeira JA, Penttilä M, Lima N. 2002. Construction of a flocculent *Saccharomyces cerevisiae* strain secreting high levels of *Aspergillus niger* beta-galactosidase. *Appl Microbiol Biotechnol*. 58:645–650.

Domingues L, Oliveira C, Castro I, Lima N, Teixeira JA. 2004 Production of *β*-galactosidase from recombinant *Saccharomyces cerevisiae* grown on lactose. *J Chem Technol Biotechnol*. 79:809–815.

Druzhinina I, Shelest E, Kubicek CP. 2012. Novel traits of *Trichoderma* predicted through the analysis of its secretome. *FEMS Microbiol Lett*. 337:1–9.

Dünkler A, Wendland J. 2007. Use of *MET3* promoters for regulated gene expression in *Ashbya gossypii*. *Curr Genet*. 52:1–10.

Efron B, Tibshirani R. 2007. On testing the significance of sets of genes. *Ann Appl Stat*. 1:107–129.

Eisenhaber B, Schneider G, Wildpaner M, Eisenhaber F. 2004. A sensitive predictor for potential GPI lipid modification sites in fungal protein sequences and its application to genome-wide studies for *Aspergillus nidulans*, *Candida albicans*, *Neurospora crassa*, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *J Mol Biol*. 337:243–253.

Elemento O, Slonim N, Tavazoie S. 2007. A universal framework for regulatory element discovery across all genomes and data types. *Mol Cell*. 28:337–350.

Emanuelsson O, Brunak S, von Heijne G, Nielsen H. 2007. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc*. 2:953–971.

Falcon S, Gentleman R. 2007. Using GOstats to test gene lists for GO term association. *Bioinformatics*. 23:257–258.

Farries EHM, Bell AF. 1930. On the metabolism of *Nematospora gossypii* and related fungi, with special reference to the source of nitrogen. *Ann Bot*. 44:423–455.

Fickers P, Le Dall MT, Gaillardin C, Thonart P, Nicaud JM. 2003. New disruption cassettes for rapid gene disruption and marker rescue in the yeast *Yarrowia lipolytica*. *J Microbiol Meth*. 55:727–737.

Fischer CR. 1969. Enzymology of the pigmented adenine-requiring mutants of *Saccharomyces* and *Schizosaccharomyces*. *Biochem Biophys Res Commun*. 34:306–310.

Fitzpatrick DA, Logue ME, Stajich JE, Butler G. 2006. A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis. *BMC Evol Biol*. 6:99.

Florea S, Andreeva K, Machado C, Mirabito PM, Schardl CL. 2009. Elimination of marker genes from transformed filamentous fungi by unselected transient transfection with a Cre-expressing plasmid. *Fungal Genet Biol*. 46:721–730.

Forment JV, Ramón D, MacCabe AP. 2006. Consecutive gene deletions in *Aspergillus nidulans:* application of the Cre/*lox*P system. *Curr Genet*. 50:217–224.

Förster C, Marienfeld S, Wendisch VF, Krämer R. 1998. Adaptation of the filamentous fungus *Ashbya gossypii* to hyperosmotic stress: different osmoresponse to NaCl and mannitol stress. *Appl Microbiol Biotechnol*. 50:219–226.

Föster C, Santos M, Ruffert S, Krämer R, Revuelta JL. 1999. Physiological consequence of disruption of the *VMA1* gene in the riboflavin overproducer *Ashbya gossypii*. *J Biol Chem*. 274:9442–9448.

Fragoso RG, Ciferri R. 1928. Hongos parasitos y saprifitos de la Republica Dominicana. *Boll R Soc Espan Hist Nat*. 28:377–388.

Fujita K, Kobayashi K, Iwamatsu A, Takeuchi M, Kumagai H, Yamamoto K. 2004. Molecular cloning of *Mucor hiemalis* endo-beta-*N*-acetylglucosaminidase and some properties of the recombinant enzyme. *Arch Biochem Biophys*. 432:41–49.

Futschik ME, Carlisle B. 2005. Noise-robust soft clustering of gene expression time-course data. *J Bioinform Comput Biol*. 3:965–988.

Gancedo, JM. 1998. Yeast carbon catabolite repression. *Microbiol Mol Biol Rev*. 60:334–361.

Gardner BM, Walter P. 2011. Unfolded proteins are Ire1-activating ligands that directly induce the unfolded protein response. *Science*. 333:1891–1894.

Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Moll Biol Cell*. 11:4241–4257.

Gasch AP, Moses AM, Chiang DY, Fraser HB, Berardini M, Eisen MB. 2004. Conservation and evolution of cis-regulatory systems in ascomycete fungi. *PLoS Biol*. 2:e398.

Gascón S, Neumann NP, Lampen JO. 1968. Comparative study of the properties of the purified internal and external invertases from yeast. *J Biol Chem*. 243:1573–1577.

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

Gatignol A, Baron M, Tiraby G. 1987. Phleomycin resistance encoded by the *ble* gene from transposon Tn*5* as a dominant selectable marker in *Saccharomyces cerevisiae*. *Mol Gen Genet*. 207:342–348.

Gattiker A, Rischatsch R, Demougin P, Voegeli S, Dietrich FS, Philippsen P, Primig M. 2007. Ashbya Genome Database 3.0: a cross-species genome and transcriptome browser for yeast biologists. *BMC Genomics*. 8:9.

Geiler-Samerotte KA, Dion MF, Budnik BA, Wang SM, Hartl DL, Drummond DA. 2011. Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. *Proc Natl Acad Sci USA*. 108:680–685.

Gemmill TR, Trimble RB. 1999. Overview of *N*- and *O*-linked oligosaccharide structures found in various yeast species. *Biochim Biophys Acta*.1426:227–237.

Gietz RD, Schiestl RH, Willems AR, Woods RA. 1995. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast*. 11:355–360.

Gonçalves C, Rodriguez-Jasso RM, Gomes N, Teixeira JA, Belo I. 2010. Adaptation of dinitrosalicylic acid method to microtiter plates. *Anal Methods*. 2:2046–2048.

Gouka RJ, Punt PJ, van den Hondel CA. 1997. Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. *Appl Microbiol Biotechnol*. 47:1–11.

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

Grass J, Pabst M, Kolarich D, Pöltl G, Léonard R, Brecker L, Altmann F. 2011. Discovery and structural characterization of fucosylated oligomannosidic *N*-glycans in mushrooms. *J Biol Chem*. 286:5977–5984.

Gueldener U, Heinisch J, Koehler GJ, Voss D, Hegemann JH. 2002. A second set of *lox*P marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res*. 30:e23.

Guillemette T, van Peij NN, Goosen T, Lanthaler K, Robson GD, van den Hondel CA, Stam H, Archer DB. 2007. Genomic analysis of the secretion stress response in the enzymeproducing cell factory *Aspergillus niger*. *BMC Genomics*. 8:158.

Guilliermond A. 1928. Recherches sur quelques Ascomycetes inferieurs isoles de la Stigmatomycose des graines de Cotonnier. Essai sur la phylogenie des Ascomycetes. *Rev Gen Bot*. 40:328–342; 397–414; 474–485; 555–574; 606–624; 690–704.

Guimarães LHS, Terenzi HF, Polizeli MLTM, Jorge JA. 2007. Production and characterization of thermostable extracellular β-D-fructofuranosidase produced by *Aspergillus ochraceus* with agroindustrial residues as carbon sources. *Enzyme Microb Technol*. 42:52–57.

Güldener U, Heck S, Fiedler T, Beinhauer JD, Hegemann JH. 1996. A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res*. 24:2519–2524.

Gupta R, Jung E, Brunak S. 2004. Prediction of *N*-glycosylation sites in human proteins. www.cbs.dtu.dk/services/NetNGlyc/

Harris SD, Read ND, Roberson RW, Shaw B, Seiler S, Plamann M, Momany M. 2005. Polarisome meets spitzenkörper: microscopy, genetics, and genomics converge. *Eukaryot Cell*. 4:225–229.

Hayakawa Y, Ishikawa E, Shoji JY, Nakano H, Kitamoto K. 2011. Septum-directed secretion in the filamentous fungus *Aspergillus oryzae*. *Mol Microbiol*. 81:40–55.

Hayes A, Zhang N, Wu J, Butler PR, Hauser NC, Hoheisel JD, Lim FL, Sharrocks AD, Oliver SG. 2002. Hybridization array technology coupled with chemostat culture: Tools to interrogate gene expression in *Saccharomyces cerevisiae*. *Methods*. 26:281–290.

Helenius A, Aebi M. 2004. Roles of *N*-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem*. 73:1019–1049.

Hemmoranta H, Satomaa T, Blomqvist M, Heiskanen A, Aitio O, Saarinen J, Natunen J, Partanen J, Laine J, Jaatinen T. 2007. *N*-glycan structures and associated gene expression reflect the characteristic *N*-glycosylation pattern of human hematopoietic stem and progenitor cells. *Exp Hematol*. 35:1279–1292.

Hernandez LM, Ballou L, Alvarado E, Tsai P, Ballou CE. 1989. Structure of the phosphorylated *N*-linked oligosaccharides from the *mnn9* and *mnn10* mutants of *Saccharomyces cerevisiae*. *J Biol Chem*. 264:13648–13659.

Heyer AG, Wendenburg R. 2001. Gene cloning and functional characterization by heterologous expression of the fructosyltransferase of *Aspergillus sydowi* IAM 2544. *Appl Environ Microbiol*. 67:363–370.

Hittinger CT, Rokas A, Carrol SB. 2004. Parallel inactivation of multiple *GAL* pathway genes and ecological diversification in yeasts. *Proc Natl Acad Sci USA*. 101:14144–14149.

Hoffman CS. 2005. Glucose sensing *via* the protein kinase A pathway in *Schizosacharomyces pombe*. *Biochem Soc Trans*. 33:257–260.

Hollien J, Weissman JS. 2006. Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response. *Science*. 313:104–107.

Horton P, Park K, Obayashi T, Fujita N, Harada H, Adams-Collier C, Nakai K. 2007. WoLF PSORT: protein localization predictor. *Nucleic Acids Res*. 35:W585-W587.

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. 2003. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res*. 31:e15.

Iwaki T, Takegawa K. 2004. A set of *lox*P marker cassettes for Cre-mediated multiple gene disruption in *Schizosaccharomyces pombe*. *Biosci Biotechnol Biochem*. 68:545–550.

Jigami Y, Odani T. 1999. Mannosylphosphate transfer to yeast mannan. *Biochim Biophys Acta*. 1426:335–345.

Jiménez A, Santos MA, Pompejus M, Revuelta JL. 2005. Metabolic engineering of the purine pathway for riboflavin production in Ashbya gossypii. *Appl Environ Microbiol*. 71:5743–5751.

Jiménez A, Santos MA, Revuelta JL. 2008. Phosphoribosyl pyrophosphate synthetase activity affects growth and riboflavin production in *Ashbya gossypii*. *BMC Biotechnol*. 8:67–78.

Jin C. 2012. Protein glycosylation in *Aspergillus fumigatus* is essential for cell wall synthesis and serves as a promising model of multicellular eukaryotic development. *Int J Microbiol*. 2012:654251.

Jongen SP, Gerwig GJ, Leeflang BR, Koles K, Mannesse ML, van Berkel PH, Pieper FR, Kroos MA, Reuser AJ, Zhou Q, Jin X, Zhang K, Edmunds T, Kamerling JP. 2007. *N*-glycans of recombinant human acid α-glucosidase expressed in the milk of transgenic rabbits. *Glycobiology*. 17:600–619.

Julenius K, Mølgaard A, Gupta R, Brunak S. 2005. Prediction, conservation analysis and structural characterization of mammalian mucin-type *O*-glycosylation sites. *Glycobiology*. 15:153–164.

Kainz E, Gallmetzer A, Hatzl C, Nett JH, Li H, Schinko T, Pachlinger R, Berger H, Reyes-Dominguez Y, Bernreiter A, Gerngross T, Wildt S, Strauss J. 2008. *N*-Glycan modification in *Aspergillus* species. *Appl Environ Microbiol*. 74:1076–1086.

Käll L, Krogh A, Sonnhammer EL. 2007. Advantages of combined transmembrane topology and signal peptide prediction - the Phobius web server. *Nucleic Acids Res*. 35:W429–432.

Karos M, Vilariño C, Bollschweiler C, Revuelta JL. 2004. A genome-wide transcription analysis of a fungal riboflavin overproducer. *J Biotechnol*. 113:69–76.

Kato T, Park EY. 2006. Expression of alanine:glyoxylate aminotransferase gene from *Saccharomyces cerevisiae* in *Ashbya gossypii*. *Appl Microbiol Biotechnol*. 71:46–52.

Kato T, Park EY. 2012. Riboflavin production by *Ashbya gossypii*. *Biotechnol Lett*. 34:611–618.

Kauffman KJ, Pridgen EM, Doyle FJ 3rd, Dhurjati PS, Robinson AS. 2002. Decreased protein expression and intermittent recoveries in BiP levels result from cellular stress during heterologous protein expression in *Saccharomyces cerevisiae*. *Biotechnol Prog*. 18:942–950.

Kauffmann A, Gentleman R, Huber W. 2009. arrayQualityMetrics - a bioconductor package for quality assessment of microarray data. *Bioinformatics*. 25:415–416.

Kaufman RJ. 1999. Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls. *Genes Dev*. 13:1211–1233.

Kaufmann A. 2009. A plasmid collection for PCR-based gene targeting in the filamentous ascomycete *Ashbya gossypii*. *Fungal Genet Biol*. 46:595–603.

Kavitha S, Chandra TS. 2009. Effect of vitamin E and menadione supplementation on riboflavin production and stress parameters in *Ashbya gossypii*. *Proc Bio*. 44:934–938.

Kim MW, Rhee SK, Kim J, Shimma Y, Chiba Y, Jigami Y, Kang HA. 2004. Characterization of *N*-linked oligosaccharides assembled on secretory recombinant glucose oxidase and cell wall mannoproteins from the methylotrophic yeast *Hansenula polymorpha*. *Glycobiology*. 14:243–251.

Kim W, Spear ED, Ng DT. 2005. Yos9p detects and targets misfolded glycoproteins for ER-associated degradation. *Mol Cell*. 19:753–764.

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 Cells*. 11:59–69.

Kimata Y, Kohono K. 2011. Endoplasmatic reticulum stress-sensing mechanisms in yeast and mammalian cells. *Curr Opin Cell Biol*. 23:135–142.

Knechtle P, Dietrich F, Philippsen P. 2003. Maximal polar growth potential depends on the polarisome component AgSpa2 in the filamentous fungus *Ashbya gossypii. Mol Biol Cell*. 14:4140–4154.

Knechtle P, Dietrich F, Philippsen P. 2004. Maximal polar growth depends on the polarisome component AgSpa2 in the filamentous fungus *A. gossypii*. *Mol Biol Cell*. 14:4140–4154.

Kögl F, Fries N. 1937. Uber den einfluss von biotin, aneurin, und meso-inosit auf das Wachstum verschiedenen Pilzarten. 26. Mitteil-ung iiber pflanzliche wachstumstoffe. *Ztschr Physiol Chem*. 249:93–110.

Köhli M, Galati V, Boudier K, Roberson RW, Philippsen P. 2008. Growth-speed-correlated localization of exocyst and polarisome components in growth zones of *Ashbya gossypii* hyphal tips. *J Cell Sci*. 121:3878–3889.

Kotz A, Wagener J, Engel J, Routier FH, Echtenacher B, Jacobsen I, Heesemann J, Ebel, F. 2010. Approaching the secrets of *N*-glycosylation in *Aspergillus fumigatus*: characterization of the AfOch1 protein. *PLoS One*. 5:e15729.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*. 305:567–580.

Kruger NJ, von Schaewen A. 2003. The oxidative pentose phosphate pathway: structure and organisation. *Curr Opin Plant Biol*. 6:236–246.

Kulmburg P, Mathieu M, Dowzer C, Kelly J, Felenbok B. 1993. Specific binding sites in the *alcR* and *alcA* promoters of the ethanol regulon for the CREA repressor mediating carbon catabolite repression in *Aspergillus nidulans*. *Mol Microbiol*. 7:847–857.

Kumar V, Ramakrishnan S, Teeri TT, Knowles JKC, Hartley BS. 1992. *Saccharomyces cerevisiae* cells secreting an *Aspergillus niger* β-galactosidase grow on whey permeate. *Biotechnology (NY)*. 10:82–85.

Kurtzman CP. 1995. Relationships among the genera *Ashbya*, *Eremothecium*, *Holleya*, and *Nematospora* determined from rDNA sequence divergence. *J Industr Microbiol*. 14:523–530.

Kurtzman CP, Robnett CJ. 2003. Phylogenetic relationships among yeasts of the '*Saccharomyces* complex' determined from multigene sequence analyses. *FEMS Yeast Res*. 3:417–432.

Kutsal T, Özbas MT. 1989. Microbial production of vitamin B_2 (riboflavin). In: *Biotechnology of Vitamins, Pigments and Growth Factors*. Vandamme EJ (ed). Elsevier, London. 149–166.

Leber JH, Bernales S, Walter P. 2004. *IRE1*-independent gain control of the unfolded protein response. *PLoS Biol*. 2:e235.

Lee LM, DaSilva NA. 2005. Evaluation of the *Saccharomyces cerevisiae ADH2* promoter for protein synthesis. *Yeast*. 22:431–440.

Lee S, Wormsley S, Kamoun S, Lee A, Joiner K, Wong B. 2003. An analysis of the *Candida albicans* genome database for soluble secreted proteins using computer-based prediction algorithms. *Yeast*. 20:595–610.

Lennarz WJ. 2007. Studies on oligosaccharyl transferase in yeast. *Acta Biochim Pol*. 54:673–677.

Lieb JD, Liu X, Botstein D, Brown PO. 2001. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat Genet*. 28:327–334.

Liko D, Slattery MG, Heideman W. 2007. Stb3 binds to ribosomal RNA processing element motifs that control transcriptional responses to growth in *S. cerevisiae*. *J Biol Chem*. 282:26623–26628.

Lim, FL, Hayes A, West AG, Pic-Taylor A, Darieva Z, Morgan BA, Oliver SG, Sharrocks AD. 2003. Mcm1p-induced DNA bending regulates the formation of ternary transcription factor complexes. *Mol Cell Biol*. 23:450–461.

Linde D, Macias I, Fernández-Arrojo L, Plou FJ, Jiménez A, Fernández-Lobato M. 2009. Molecular and biochemical characterization of a β-fructofuranosidase from *Xanthophyllomyces dendrorhous*. *Appl Environ Microbiol*. 75:1065–1073.

Lum G, Min XJ. 2011. FunSecKB: the Fungal Secretome KnowledgeBase. *Database (Oxford)*. 2011:bar001.

Lundin M, Nehlin JO, Ronne H. 1994. Importance of a flanking AT-rich region in target site recognition by the GC box-binding zinc finger protein MIG1. *Mol Cell Biol*. 14:1979–1985.

Lutfiyya LL, Johnston M. 1996. Two zinc-finger-containing repressors are responsible for glucose repression of *SUC2* expression. *Mol Cell Biol*. 16:4790–4797.

Mach RL, Strauss J, Zeilinger S, Schindler M, Kubicek CP. 1996. Carbon catabolite repression of xylanase I (*xynl*) gene expression in *Trichoderma reesei*. *Mol Microbiol*. 21:1273–1281.

Maeting I, Schmidt G, Sahma H, Revuelta JL, Stierhofd YD, Stahmanna KP. 1999. Isocitrate lyase of *Ashbya gossypii* – transcriptional regulation and peroxisomal localization. *FEBS Lett*. 444:15–21.

Manzi AE, Norgard-Sumnicht K, Argade S, Marth JD, van Halbeek H, Varki A. 2000. Exploring the glycan repertoire of genetically modified mice by isolation and profiling of the major glycan classes and nano-NMR analysis of glycan mixtures. *Glycobiology*. 10:669–689.

Maras M, De Bruyn A, Schraml J, Herdewijn P, Claeyssens M, Fiers W, Contreras R. 1997. Structural characterization of *N*-linked oligosaccharides from cellobiohydrolase I secreted by the filamentous fungus *Trichoderma reesei* RUTC 30. *Eur J Biochem*. 245:617–625.

Maras M, van Die I, Contreras R, van den Hondel CA. 1999. Filamentous fungi as production organisms for glycoproteins of bio-medical interest. *Glycoconj J*. 16:99–107.

Martoglio B, Dobberstein B. 1998. Signal sequences: more than just greasy peptides. *Trends Cell Biol*. 8:410–415.

Mateos L, Jiménez A, Revuelta J, Santos M. 2006. Purine biosynthesis, riboflavin production, and trophic phase span are controlled by a Myb-related transcription factor in the fungus *Ashbya gossypii*. *Appl Environ Microbiol*. 72:5052–5060.

Matsumoto R, Akama K, Rakwal R, Iwahashi H. 2005. The stress response against denatured proteins in the deletion of cytosolic chaperones *SSA1/2* is different from heat-shock response in *Saccharomyces cerevisiae*. *BMC Genomics*. 6:141.

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

Metzger MB, Michaelis S. 2009. Analysis of quality control substrates in distinct cellular compartments reveals a unique role for Rpn4p in tolerating misfolded membrane proteins. *Mol Biol Cell*. 20:1006–1019.

Meyer zu Berstenhorst S, Hohmann HP, Stahmann KP. 2009. Vitamins and vitamin-like compounds: microbial production. In: *Encyclopedia of Microbiology*. Schaechter M (ed). Elsevier, Oxford. 549–561.

Mickelson MN. 1950. The metabolism of glucose by *Ashbya gossypii*. *J Bacteriol*. 59:659–566.

Mickelson MN, Schuler MN. 1953. Oxiadtion of acetate by *Ashbya gossypii*. *J Bacteriol*. 65:297–304.

Miller GL. 1959. Use of dinitrosalicylic reagent for determination of reducing sugars. *Anal Chem*. 31:426-428.

Miyazaki T, Nakayama H, Nagayoshi Y, Kakeya H, Kohno S. 2013. Dissection of Ire1 functions reveals stress response mechanisms uniquely ivolved in *Candida glabrata*. *PLoS Pathog*. 9:e1003160.

Mohr C. 1997. Genetic engineering of the filamentous fungus *Ashbya gossypii*: construction of a genomic library, isolation of genes for beta-isopropylmalatedehydrogenase (*LEU2*) and a protein kinase (*APK1*) by heterologous complementation, and characterization of non-reverting mutants. PhD thesis, University of Basel, Switzerland.

Monschau N, Sahm H, Stahmann KP. 1998. Threonine aldolase overexpression plus threonine supplementation enhanced riboflavin production in *Ashbya gossypii*. *Appl Environ Microbiol*. 64:4283–4290.

Montesino R, Nimtz M, Quintero O, García R, Falcón V, Cremata JA. 1999. Characterization of the oligosaccharides assembled on the *Pichia pastoris*-expressed recombinant aspartic protease. *Glycobiology*. 9:1037–1043.

Moreno S, Sánchez Y, Rodríguez L. 1990. Purification and characterization of the invertase from *Schizosaccharomyces pombe*. A comparative analysis with the invertase from *Saccharomyces cerevisiae*. *Biochem J*. 267:697–702.

Mori K, Kawahara T, Yoshida H, Yanagi H, Yura T. 1996. Signalling from endoplasmatic reticulum to nucleous: transcription factor with a basic-leucine zipper motif is required for the unfolded protein-response pathway. *Genes Cells*. 1:803–817.

Mori K. 2000. Tripartite management of unfolded proteins in the endoplasmatic reticulum. *Cell*. 101:451–454.

Mori K. 2009. Signalling pathways in the unfolded protein response: development from yeast to mammals. *J Biochem*. 146:743–750.

Motteram J, Lovegrove A, Pirie E, Marsh J, Devonshire J, van de Meene A, Hammond-Kosack K, Rudd JJ. 2011. Aberrant protein *N*-glycosylation impacts upon infection-related growth transitions of the haploid plant-pathogenic fungus *Mycosphaerella graminicola*. *Mol Microbiol*. 81:415–433.

Nakatsukasa K, Brodsky JL. 2008. The recognition and retrotranslocation of misfolded proteins from the endoplasmic reticulum. *Traffic*. 9:861–870.

Nevalainen KMH, Teo VSJ, Bergquist PL. 2005. Heterologous protein expression in filamentous fungi. *Trends Biotechnol*. 23:468–474.

O'Connell S, Walsh G. 2010. A novel acid-stable, acid-active *β*-galactosidase potentially suited to the alleviation of lactose intolerance. *Appl Microbiol Biotechnol*. 86:517–524.

Odani T, Shimma Y, Wang XH, Jigami Y. 1997. Mannosylphosphate transfer to cell wall mannan is regulated by the transcriptional level of the *MNN4* gene in *Saccharomyces cerevisiae*. *FEBS Lett*. 420:186–190.

Oliveira C, Teixeira JA, Lima N, Da Silva NA, Domingues L. 2007. Development of stable flocculent *Saccharomyces cerevisiae* strain for continuous *Aspergillus niger* beta-galactosidase production. *J Biosci Bioeng*. 103:318–324.

Oliveira C, Guimarães PMR, Domingues L. 2011. Recombinant microbial systems for improved *β*-galactosidase production and biotechnological applications. *Biothecnol Adv*. 29:600–609.

Orban PC, Chui D, Marth JD. 1992. Tissue- and site-specific DNA recombination in transgenic mice. *Proc Natl Acad Sci USA*. 9:6861–6865.

Orikasa Y, Oda Y. 2012. Molecular characterization of β-fructofuranosidases from *Rhizopus delemar* and *Amylomyces rouxii*. *Folia Microbiol*. 10.1007/s12223-012-0211-9.

Özbas MT, Kutsal T. 1986. Comparative study of riboflavin production from two microorganisms: *Eremothecium ashbyi* and *Ashbya gossypii*. *Enzyme Microb Technol*. 8:593–596.

Pakula TM, Laxell M, Huuskonen A, Uusitalo J, Saloheimo M, Penttilä 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*. 278:45011–45020.

Park EY, Zhang JH, Tajima S, Dwiarti L. 2007. Isolation of *Ashbya gossypii* mutant for an improved riboflavin production targeting for biorefinery technology. *J Appl Microbiol*. 103:468–476.

Park JN, Lee DJ, Kwon O, Oh DB, Bahn YS, Kang HA. 2012. Unraveling unique structure and biosynthesis pathway of *N*-linked glycans in human fungal pathogen *Cryptococcus neoformans* by glycomics analysis. *J Biol Chem*. 287:19501–19515.

Patel RD, Lodge JK, Baker LG. 2010. Going green in *Cryptococcus neoformans*: The recycling of a selectable drug marker. *Fungal Genet Biol*. 47:191–198.

Patil CK, Li H, Walter P. 2004. Gcn4p and novel upstream activating sequences regulate targets of the unfolded protein response. *PLoS Biol*. 2:e246.

Perlman D, Raney P, Halvorson HO. 1984. Cytoplasmic and secreted *Saccharomyces cerevisiae* invertase mRNAs encoded by one gene can be differentially or co-ordinately regulated. *Mol Cell Biol*. 4:1682–1688.

Pfeifer VF, Tanner FW, Vojnovich C, Traufler DH. 1950. Riboflavin by fermentation with *Ashbya gossypii*. *Ind Eng Chem*. 42:1776–1781.

Pittet M, Conzelmann A. 2007. Biosynthesis and function of GPI proteins in the yeast *Saccharomyces cerevisiae*. *Biochim Biophys Acta*. 1771:405–420.

Pompejus M, Revuelta JL, Santos MA. 1999. Orotidine-5'-phosphate decarboxylase-gene, gene construct containing said gene and the utilization thereof. *Patent WO/1999/036432*.

Porro D, Sauer M, Branduardi P, Mattanovich D. 2005. Recombinant protein production in yeasts. *Mol Biotechnol*. 31:245–259.

Potter BA, Hughey RP, Weisz OA. 2006. Role of *N*- and *O*-glycans in polarized biosynthetic sorting. *Am J Physiol Cell Physiol*. 290:C1–C10.

Pridham TG, Raper KB. 1950. *Ashbya gossypii* – Its significance in nature and in the laboratory. *Mycologia*. 42:603–623.

Prillinger H, Schweigkofler W, Breitenbach M, Briza P, Staudacher E, Lopandic K, Molnár O, Weigang F, Ibl M, Ellinger A. 1997. Phytopathogenic filamentous (*Ashbya*, *Eremothecium*) and dimorphic fungi (*Holleya*, *Nematospora*) with needle-shaped ascospores as new members within the Saccharomycetaceae. *Yeast*. 13:945–960.

Qian W, Song H, Liu Y, Zhang C, Niu Z, Wang H, Qiu B. 2009. Improved gene disruption method and Cre-*lox*P mutant system for multiple gene disruptions in *Hansenula polymorpha*. *J Microbiol Meth*. 79:253–259.

Ramakrishnan S, Hartley BS. 1993. Fermentation of lactose by yeast cells secreting recombinant fungal lactase. *Appl Environ Microbiol*. 59:4230–4235.

Rawlings ND, Barrett AJ, Bateman A. 2012. *MEROPS*: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 40:D343–D350.

Read ND. 2011. Exocytosis and growth do not occur only at hyphal tips. *Mol Microbiol*. 81:4–7.

Reddy A, Maley F. 1996. Studies on identifying the catalytic role of Glu-204 in the active site of yeast invertase. *J Biol Chem*. 271:13953–13957.

Resina D, Maurer M, Cos O, Arnau C, Carnicer M, Marx H, Gasser B, Valero F, Mattanovich D, Ferrer P. 2009. Engineering of bottlenecks in *Rhizopus oryzae*

lipase production in *Pichia pastoris* using the nitrogen source-regulated *FLD1* promoter. *Nat Biotechnol*. 25:396–403.

Revuelta Doval JL, Santos Garcia MA, Pompeus M, Seulberger H. 1999. Promoter from *Ashbya gossypii*. *Patent WO/1999/033933*.

Ribeiro O, Gombert AK, Teixeira JA, Domingues L. 2007. Application of the Cre-*lox*P system for multiple gene disruption in the yeast *Kluyveromyces marxianus*. *J Biotechnol*. 131:20–26.

Ribeiro O, Wiebe M, Ilmén M, Domingues L, Penttilä M. 2010. Expression of *Trichoderma reesei* cellulases CBHI and EGI in *Ashbya gossypii*. *Appl Microbiol Biotechnol*. 87:1437–1446.

Ribeiro O, Domingues L, Penttilä M, Wiebe MG. 2011. Nutritional requirements and strain heterogeneity in *Ashbya gossypii*. *J Basic Microbiol*. 51:1–8.

Ribeiro O. 2012. Physiological characterization of *Ashbya gossypii* and strain development for recombinant protein production. PhD thesis, University of Minho, Portugal.

Ribeiro O, Magalhães F, Aguiar TQ, Wiebe MG, Penttilä M, Domingues L. 2013. Random and direct mutagenesis to enhance protein secretion in *Ashbya gossypii*. *Bioengineered*. 4:5.

Robb A, Brown JD. 2001. Protein transport: two translocons are better than one. *Mol Cell*. 8:484–486.

Robinson AS, Hines V, Wittrup KD. 1994. Protein disulfide isomerise overexpression increases secretion of foreign proteins in *Saccharomyces cerevisiae*. *Bio/Technology*. 12:381–384.

Rolfes RJ. 2006. Regulation of purine nucleotide biosynthesis: in yeast and beyond. *Biochem Soc Trans*. 34:786–790.

Romero PA, Lussier M, Véronneau S, Sdicu AM, Herscovics A, Bussey H. 1999. Mnt2p and Mnt3p of *Saccharomyces cerevisiae* are members of the Mnn1p family of α1,3-mannosyltransferases responsible for adding the terminal mannose residues of *O*-linked oligosaccharides. *Glycobiology*. 9:1045–1051.

Rubio CM, Maldonado CM. 1995. Purification and characterisation of invertase from *Aspergillus niger*. *Curr Microbiol*. 31:80–83.

Rubio MC, Runco R, Navarro AR. 2002. Invertase from a strain of *Rhodotorula glutinis*. *Phytochemistry*. 61:605–609.

Rubio MC, Navarro AR. 2006. Regulation of invertase synthesis in *Aspergillus niger*. *Enzyme Microb Technol*. 39:601–606.

Ruohonen L, Aalto MK, Keränen S. 1995. Modifications of the *ADH1* promoter of *Saccharomyces cerevisiae* for efficient production of heterologous proteins. *J Biotechnol*. 39:193–203.

Rutkowski DT, Kaufman RJ. 2004. A trip to the ER: coping with stress. *Trends Cell Biol*. 14:20–28.

Saloheimo M, Lund M, Penttilä ME. 1999. The protein disulphide isomerase gene of the fungus *Trichoderma reesei* is induced by endoplasmic reticulum stress and regulated by the carbon source. *Mol Gen Genet*. 262:35–45.

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

Saloheimo M, Pakula TM. 2012. The cargo and the transport system: secreted proteins and protein secretion in *Trichoderma reesei* (*Hypocrea jecorina*). *Microbiology*. 158:46–57.

Sambrook J, Russell DW. 2001. Molecular Cloning: A Laboratory Manual, third ed. Cold Spring Harbor Laboratory Press, New York.

Sauer B, Henderson N. 1988. Site-specific DNA recombination in mammalian cells by the Cre recombination of bacteriophage P1. *Proc Natl Acad Sci USA*. 85, 5166–5170.

Sauer B. 1987. Functional expression of the Cre-*lox* site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol*. 7:2087–2096.

Schade D, Walther A, Wendland J. 2003. The development of a transformation system for the dimorphic plant pathogen *Holleya sinecauda* based on *Ashbya gossypii* DNA elements. *Fungal Genet Biol*. 40;65–71.

Schenk J, Balazs K, Jungo C, Urfer J, Wegmann C, Zocchi A, Marison IW, von Stockar U. 2008. Influence of specific growth rate on specific productivity and glycosylation

of a recombinant avidin produced by a *Pichia pastoris* Mut$^+$ strain. Biotechnol Bioeng. 99:368–377.

Schlösser T, Schmidt G, Stahmann KP. 2001. Transcriptional regulation of 3,4-dihydroxy-2-butanone 4-phosphate synthase. *Microbiology*. 147:3377–3386.

Schlösser T, Wiesenburg A, Gätgens C, Funke A, Viets U, Vijayalakshmi S, Nieland S, Stahmann KP. 2007. Transcriptional regulation of 3,4-dihydroxy-2-butanone 4-phosphate synthase. *Microbiology*. 147:3377–3386.

Schlüpen C, Santos MA, Weber U, De Graaf A, Revuelta JL, Stahmann KP. 2003. Disruption of the *SHM2* gene, encoding one of two serine hydroxymethyltransferase isozymes, reduces the flux from glycine to serine in *Ashbya gossypii*. *Biochem J*. 369:263–273.

Schlüpen C. 2003. Untersuchungen zur bedeutung verschiedener enzyme des glycin-stoffwechsels für die riboflavin-bildung in *Ashbya gossypii*. PhD thesis, Heinrich Heine University of Düsseldorf, Germany.

Schmidt G, Stahmann KP, Kaesler B, Sahm H. 1996a. Correlation of isocitrate lyase activity and riboflavin formation in the riboflavin overproducer *Ashbya gossypii*. *Microbiology*. 142:419–426.

Schmidt G, Stahmann KP, Sahm H. 1996b. Inhibition of purified isocitrate lyase identified itaconate and oxalate as potential antimetabolites for the riboflavin overproducer *Ashbya gossypii*. *Microbiology*. 142:411–417.

Schmitz HP, Philippsen P. 2011. Evolution of multinucleated *Ashbya gossypii* hyphae from a budding yeast-like ancestor. *Fungal Biol*. 115:557–568.

Schroder M, Clark R, Kaufman RJ. 2003. *IRE1*- and *HAC1*-independent transcriptional regulation in the unfolded protein response of yeast. *Mol Microbiol*. 49:591–606.

Shaka AJ, Barker PB, Freeman R. 1985. Computer-optimized decoupling scheme for wideband applications and low-level operation. *J Magn Reson*. 64:547–552.

Sherman F, Fink GR, Hicks JB. 1986. Methods in yeast genetics: A laboratory course manual. Cold Spring Harbor Laboratory Press, New York.

Shoji JY, Arioka M, Kitamoto K. 2008. Dissecting cellular components of the secretory pathway in filamentous fungi: insights into their application for protein production. *Biotechnol Lett*. 30:7–14.

Shusta EV, Raines RT, Pluckthum A, Wittrup KD. 1998. Increasing the secretory capacity of *Saccharomyces cerevisiae* for production of single-chain antibody fragments. *Nat. Biotechnol*. 16:773–777.

Sievers F, Wilm A, Dineen, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, R, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 7:539.

Sims AH, Gent ME, Lanthaler K, Dunn-Coleman NS, Oliver SG, Robson GD. 2005. Transcriptome analysis of recombinant protein secretion by *Aspergillus nidulan*s and the unfolded-protein response *in vivo*. *Appl Environ Microbiol*. 71:2737–2747.

Slot JC, Rokas A. 2010. Multiple *GAL* pathway gene clusters evolved independently and by different mechanisms in fungi. *Proc Natl Acad Sci USA*. 107:10136–10141.

Smiley KL, Sobolov M, Austin FL, Rasmussen RA, Smith MB, van Lanen JM, Stone L, Boruff CS. 1951. Biosynthesis of riboflavin, *L. bulgaricus* factor, and other growth factors. Laboratory and pilot plant studies of biosynthesis by *A. gossypii* cultivated on grain stillage media. *Ind Eng Chem*. 43:1380–1384.

Smyth GK. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. S*tat Appl Genet Mol Biol*. 3:Article3.

Song Y, Choi MH, Park J-N, Kim MW, Kim EJ, Kang HA, Kim J-Y. 2007. Engineering of the yeast *Yarrowia lipolytica* for the production of glycoproteins lacking the outer-chain mannose residues of *N*-glycans. *Appl Environ Microbiol*. 73:4446–4454.

Stahmann KP, Böddecker T, Sahm H. 1997. Regulation and properties of a fungal lipase showing interfacial inactivation by gas bubbles, or droplets of lipid or fatty acid. *Eur J Biochem*. 244:220–225.

Stahmann KP, Revuelta JL, Suelberger H. 2000. Three biotechnical process using *Ashbya gossypii*, *Candida famata*, or *Bacillus subtilis* compete with chemical riboflavin production. *Appl Microbiol Biotechnol*. 53:509–516.

Stahmann KP, Arst Jr HN, Althöfer H, Revuelta JL, Monschau N, Schlüpen C, Gätgens C, Wiesenburg A, Schlösser T. 2001. Riboflavin, overproduced during sporulation of *Ashbya gossypii*, protects its hyaline spores against ultraviolet light. *Environ Microbiol*. 3:545–550

Stals I, Samyn B, Sergeant K, White T, Hoorelbeke K, Coorevits A, Devreese B, Claeyssens M, Piens K. 2010. Identification of a gene coding for a deglycosylating enzyme in *Hypocrea jecorina*. *FEMS Microbiol Lett*. 303:9–17.

Stals I, Sandra K, Geysens S, Contreras R, Van Beeumen J, Claeyssens M. 2004. Factors influencing glycosylation of *Trichoderma reesei* cellulases. I: Postsecretorial changes of the *N*- and *O*-glycosylation pattern of Cel7A. *Glycobiology*. 14:713–724.

Steensma HY, Ter Linde JJ. 2001. Plasmids with the Cre-recombinase and the dominant nat marker, suitable for use in prototrophic strains of *Saccharomyces cerevisiae* and *Kluyveromyces lactis*. *Yeast*. 18:469–472.

Steiger MG, Vitikainen M, Uskonen P, Brunner K, Adam G, Pakula T, Penttilä M, Saloheimo M, Mach RL, Mach-Aigner AR. 2011. Transformation system for *Hypocrea jecorina* (*Trichoderma reesei*) that favors homologous integration and employs reusable bidirectionally selectable markers. *Appl Environ Microbiol*. 77:114–121.

Steinberg G. 2007. Hyphal growth: a tale of motors, lipids, and the Spitzenkörper. *Eukaryot Cell*. 6:351–360.

Steiner S, Philippsen P. 1994. Sequence and promoter analysis of the highly expressed *TEF* gene of the filamentous fungus *Ashbya gossypii*. *Mol Gen Genet*. 242:263–271.

Steiner S, Wendland J, Wright MC, Philippsen P. 1995. Homologous recombination as the main mechanism for DNA integration and cause of rearrangements in the filamentous ascomycete *Ashbya gossypii*. *Genetics*. 140:973–987.

Sternberg N, Hamilton D, Hoess R. 1981. Bacteriophage P1 site-specific recombination: I. Recombination between *lox*P sites. *J Mol Biol*. 150:467–486.

Stolz A, Wolf DH. 2010. Endoplasmic reticulum associated protein degradation: a chaperone assisted journey to hell. *Biochim Biophys Acta*. 1803:694–705.

Stolz J, Munro S. 2002. The components of the *Saccharomyces cerevisiae* mannosyltransferase complex M-Pol I have distinct functions in mannan synthesis. *J Biol Chem*. 277:44801–44808.

Storey J, Taylor JE, Siegmund D. 2004. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *J R Statist Soc B*. 66:187–205.

Sugimoto T, Morimoto A, Nariyama M, Kato T, Park EY. 2010. Isolation of an oxalate-resistant *Ashbya gossypii* strain and its improved riboflavin production. *J Ind Microbiol Biotechnol*. 37:57–64.

Supek F, Bošnjak M., Škunca N, Šmuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*. 6:e21800.

Swaim C, Anton B, Sharma S, Taron C, Benner J. 2008. Physical and computational analysis of the yeast *Kluyveromyces lactis* secreted proteome. *Proteomics*. 8:2714–2723.

Tajima S, Itoh Y, Sugimoto T, Kato T, Park EY. 2009. Increased riboflavin production from activated bleaching earth by a mutant strain of *Ashbya gossypii*. *J Biosci Bioeng*. 108:325–329.

Takashima S, Iikura H, Nakamura A, Masaki H, Uozumi T. 1996. Analysis of Crel binding sites in the *Trichoderma reesei cbhl* upstream region. *FEMS Microbiol Lett*. 145:361–366.

Tanner FW Jr, Vojnovich C, Van Lanen JM. 1949. Factors affecting riboflavin production by *Ashbya gossypii*. *J Bacteriol*. 58:737–745.

Teixeira MC, Monteiro P, Jain P, Tenreiro S, Fernandes AR, Mira NP, Alenquer M, Freitas AT, Oliveira AL, Sá-Correia I. 2006. The YEASTRACT database: a tool for the analysis of transcription regulatory associations in *Saccharomyces cerevisiae*. *Nucleic Acids Res*. 34:D446–D451.

Toivari MH, Ruohonen L, Richard P, Penttilä M, Wiebe MG. 2010. *Saccharomyces cerevisiae* engineered to produce D-xylonate. *Appl Microbiol Biotechnol*. 88:751–760.

Travers KJ, Patil CK, Wodicka L, Lockhart DJ, Weissman JS, Walter P. 2000. Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell*. 101:249–258.

Trimble RB, Atkinson PH. 1986. Structure of yeast external invertase $Man_{8-14}GlcNAc$ processing intermediates by 500-megahertz ^{1}H NMR spectroscopy. *J Biol Chem*. 261:9815–9824.

Trimble RB, Atkinson PH, Tschopp JF, Townsend RR, Maley F. 1991. Structure of oligosaccharides on *Saccharomyces SUC2* invertase secreted by the methylotrophic yeast *Pichia pastoris*. *J Biol Chem*. 266:22807–22817.

Trumbly RJ. 1992. Glucose repression in the yeast *Saccharomyces cerevisiae*. *Mol Microbiol*. 6:15–21.

Tsang A, Butler G, Powlowski J, Panisko EA, Baker SE. 2009. Analytical and computational approaches to define the *Aspergillus niger* secretome. *Fungal Genet Biol*. 46:S153–S160.

UniProt Consortium. 2013. Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res*. 41:D43–D47.

Vainstein M, Peberdy J. 1991. Regulation of invertase in *Aspergillus nidulans*: effect of different carbon sources. *J Gen Microbiol*. 137:315–321.

Valkonen M, Penttilä M, Saloheimo M. 2003a. Effects of inactivation and constitutive expression of the unfolded-protein response pathway on protein production in the yeast *Saccharomyces cerevisiae*. *Appl Environ Microbiol*. 69:2065–2072.

Valkonen M, Ward M, Wang H, Penttilä M, Saloheimo M. 2003b. Improvement of foreign-protein production in *Aspergillus niger* var. *awamori* by constitutive induction of the unfolded-protein response. *Appl Environ Microbiol*. 69:6979–6986.

Van de Velde S, Thevelein JM. 2008. Cyclic AMP-protein kinase A and Snf1 signaling mechanisms underlie the superior potency of sucrose for induction of filamentation in *Saccharomyces cerevisiae*. *Eukaryot Cell*. 7:286–293.

van Duyne GD. 2001. A structural view of Cre-*lox*P site-specific recombination. *Annu Rev Biophys Biomol Struct*. 30:87–104.

Verduyn C, Postma E, Scheffers WA, Van Dijken JP. 1992. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast*. 8:501–517.

Verostek MF, Lubowski C, Trimble RB. 2000. Selective organic precipitation/extraction of released *N*-glycans following large-scale enzymatic deglycosylation of glycoproteins. *Anal Biochem*. 278:111–122.

Vitolo M, Duranti M, Pellegrini M. 1995. Effect of pH, aeration and sucrose feeding on the invertase activity of intact *Saccharomyces cerevisiae* cells grown in sugar cane blackstrap molasses. *J Ind Microbiol*. 15:75–79.

Vliegenthart JFG, Dorland L, van Halbeek H. 1983. High-resolution, ^{1}H-nuclear magnetic resonance spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. *Adv Carbohydr Chem Biochem*. 41:209–374.

von Arx JA, van der Walt JP. 1987. Ophiostomatales and Endomycetales. *Stud Mycol*. 30:167–176.

Wallis GLF, Easton RL, Jolly K, Hemming FW, Peberdy JF. 2001. Galactofuranoic-oligomannose *N*-linked glycans of α-galactosidase A from *Aspergillus niger*. *Eur J Biochem*. 268:4134–4143.

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

Walther A, Wendland J. 2004. Apical localization of actin patches and vacuolar dynamics in *Ashbya gossypii* depend on the *WASP* homolog Wal1p. *J Cell Sci*. 117:4947–4758.

Walther A, Wendland J. 2012. Yap1-dependent oxidative stress response provides a link to riboflavin production in *Ashbya gossypii*. *Fungal Genet Biol*. 49:697–707.

Wang H, Xu Z, Gao L, Hao B. 2009. A fungal phylogeny based on 82 complete genomes using the composition vector method. *BMC Evol Biol*. 9:195.

Ward OP. 2012. Production of recombinant proteins by filamentous fungi. *Biotechnol Adv*. 30:1119–1139.

Weerapana E, Imperiali B. 2006. Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology*. 16:91R–101R.

Wendland J, Ayad-Durieux Y, Knechtle P, Rebischung C, Philippsen P. 2000. PCR-based gene targeting in the filamentous fungus *Ashbya gossypii*. *Gene*. 242:381–391.

Wendland J, Philippsen P. 2000. Determination of cell polarity in germinated spores and hyphal tips of the filamentous ascomycete *Ashbya gossypii* requires a rhoGAP homolog. *J Cell Sci*. 113:1611–1621.

Wendland J, Philippsen P. 2001. Cell polarity and hyphal morphogenesis are controlled by multiple Rho-protein modules in the filamentous ascomycete *Ashbya gossypii*. *Genetics*. 157:601–610.

Wendland J, Walther A. 2005. *Ashbya gossypii*: a model for fungal developmental biology. *Nat Rev Microbiol*. 3:421–429.

Wendland J, Dünkler A, Walther A. 2011. Characterization of α-factor pheromone and pheromonereceptor genes of *Ashbya gossypii*. *FEMS Yeast Res*. 11:418–429.

Wickerham LS, Flickinger MH, Johnson RM. 1946. Production of riboflavin by *Ashbya gossypii*. *Arch Biochem*. 9:95–98.

Wiebe MG, Robson GD, Shuster J, Trinci AP. 2001. Evolution of a recombinant (gucoamylase-producing) strain of *Fusarium venenatum* A3/5 in chemostat culture. *Biotechnol Bioeng*. 73:146–156.

Wimalasena TT, Enjalbert B, Guillemette T, Plumridge A, Budge S, Yin Z, Brown AJP, Archer DB. 2008. Impact of the unfolded protein response upon genome-wide expression patterns, and the role of Hac1 in the polarized growth, of *Candida albicans*. *Fungal Genet Biol*. 45:1235–1247.

Wolska-Mitaszko B, Jaroszuk-Scisei J, Pszeniczna K. 2007. Isoforms of trehalase and invertase of *Fusarium oxysporum*. *Mycol Res*. (III):456–465.

Wright MC, Philippsen P. 1991. Replicative transformation of the filamentous fungus *Ashbya gossypii* with plasmids containing *Saccharomyces cerevisiae ARS* elements. *Gene*. 109:99–105.

Xu X, Kanbara K, Azakami H, Kato A. 2004. Expression and characterization of *Saccharomyces cerevisiae* Cne1p, a calnexin homologue. *J Biochem*. 135:615–618.

Yoon J, Aishan T, Maruyama J, KitamotovK. 2010. Enhanced production and secretion of heterologous proteins by the filamentous fungus *Aspergillus* oryzae via disruption

of vacuolar protein sorting receptor gene *Aovps10*. *Appl Environ Microbiol*. 76:5718–5727.

Yorimitsu T, Nair U, Yang Z, Klionsky DJ. 2006. Endoplasmic reticulum stress triggers autophagy. *J Biol Chem*. 281:30299–30304.

Zhang H, Chi Z. 2004. Inositol and phosphatidylinositol-mediated invertase secretion in *Schizosaccharomyces pombe*. *Enzyme Microb Technol*. 34:213–218.

Zhang Y, Liu H, Li B, Zhang JT, Li Y, Zhang H. 2009. Generation of selectable marker-free transgenic tomato resistant to drought, cold and oxidative stress using the Cre/*lox*P DNA excision system. *Transgenic Res*. 18:607–619.