# Effective lead selection for improved protein production in Aspergillus niger based on integrated genomics

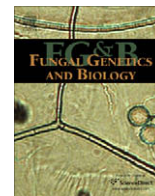Jacobs, D.I.; Olsthoorn, M.M.; Maillet, I.; Akeroyd, M.; Breestraat, S.; Donkers, S.; ... ; Sagt, C.M.

# Effective lead selection for improved protein production in *Aspergillus niger* based on integrated genomics

Denise I. Jacobs [a], Maurien M.A. Olsthoorn [a], Isabelle Maillet [b], Michiel Akeroyd [a], Stefaan Breestraat [a], Serge Donkers [a], Rob A.M. van der Hoeven [a], Cees A.M.J.J. van den Hondel [c], Rolf Kooistra [c], Thomas Lapointe [b], Hildegard Menke [a], Rogier Meulenberg [a], Marijke Misset [a], Wally H. Müller [d], Noël N.M.E. van Peij [a], Arthur Ram [c], Sabrina Rodriguez [b], Marc S. Roelofs [d], Johannes A. Roubos [a], Marcel W.E.M. van Tilborg [a], Arie J. Verkleij [d], Herman J. Pel [a], Hein Stam [a], Cees M.J. Sagt [a,*]

[a] *DSM Food Specialties, P.O. Box 1, NL-2600 MA Delft, The Netherlands*
[b] *DSM Nutritional Products, P.O. Box 3255, CH-4002 Basel, Switzerland*
[c] *Leiden University, Molecular Microbiology, Wassenaarseweg 64 2333 AL Leiden, The Netherlands*
[d] *Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*

## ARTICLE INFO

## ABSTRACT

The filamentous fungus *Aspergillus niger* is widely exploited for industrial production of enzymes and organic acids. An integrated genomics approach was developed to determine cellular responses of *A. niger* to protein production in well-controlled fermentations. Different protein extraction methods in combination with automated sample processing and protein identification allowed quantitative analysis of 898 proteins. Three different enzyme overproducing strains were compared to their isogenic fungal host strains. Clear differences in response to the amount and nature of the overproduced enzymes were observed. The corresponding genes of the differentially expressed proteins were studied using transcriptomics. Genes that were up-regulated both at the proteome and transcriptome level were selected as leads for generic strain improvement. Up-regulated proteins included proteins involved in carbon and nitrogen metabolism as well as (oxidative) stress response, and proteins involved in protein folding and endoplasmic reticulum-associated degradation (ERAD). Reduction of protein degradation through the removal of the ERAD factor *doaA* combined with overexpression of the oligosaccharyl transferase *sttC* in *A. niger* overproducing β-glucuronidase (GUS) strains indeed resulted in a small increase in GUS expression.

© 2008 Elsevier Inc. All rights reserved.

## 1. Introduction

The filamentous fungus *Aspergillus niger* is able to secrete large amounts of a wide variety of enzymes and metabolites. In nature, the enzymes are needed to release nutrients from complex biopolymers while metabolite excretion gives the fungus a competitive advantage. These natural characteristics are exploited by industry in both solid state and submerged fermentations for the production of enzymes and metabolites.

Protein production and secretion in *A. niger* is a complex and widely studied process and significant differences in titres between homologous and heterologous proteins are frequently observed. While homologous proteins are produced in a 10–50 g/L range, heterologous proteins often are produced 10- to 100-fold less (van den Hondel et al., 1991), indicating that *A. niger* encounters

severe problems in expression and/or secretion of foreign proteins. Homologous protein production strains show high protein fluxes and have a high demand for amino acids, while heterologous protein production results in secretion stress and a typical unfolded-protein response (UPR) response. Activation of the UPR pathway in filamentous fungi due to heterologous protein production has been frequently observed and is manifested by induction of chaperone and foldase genes (e.g., Wiebe et al., 2001; Saloheimo et al., 2003; Collén et al., 2005; Arvas et al., 2006).

Recently, genome sequences of industrial *A. niger* strains have been published covering both enzyme and metabolite producing strains (Pel et al., 2007; Baker, 2006; Sun et al., 2007, http://genome.jgi-psf.org/Aspni1/Aspni1.home.html). In addition, genomic tools such as transcriptomics and various proteomics technologies have been developed for fungi, which enable a new way of working in order to improve strains and processes for industrial fermentations.

In our studies, strains were grown under strictly controlled fermentation conditions and samples for transcriptomics and proteo-

mics were obtained using a standardised rapid procedure. Based on these reproducible fermentations, differential transcriptomics and proteomics studies were performed on protein production strains and their isogenic control strains. Genes which showed co-regulation at the transcriptome and proteome level were selected as potential targets for strain improvement. To date, only a few papers have been published investigating the responses due to protein production in filamentous fungi (e.g., MacKenzie et al., 2005; Sims et al., 2004, 2005; Levin et al., 2007; Arvas, 2007; Guillemette et al., 2007; Yuan et al., 2008).

Despite their importance, the number of published proteomics studies on filamentous fungi is also rather limited, as was reviewed by Kim et al. (2007, 2008) and Carberry and Doyle (2007). However, fungal proteomics is gathering pace due to the increasing availability of genome sequence information and advances in proteomics technologies. Preparation of protein extracts is a critical step in reliably and reproducibly determining the presence and the abundance of proteins. Since the fungal cell wall is exceptionally robust (Bowman and Free, 2006) and proteases are extremely resistant, cell lysis and protein extraction is challenging. Several protocols have already been published for fungal protein extraction (e.g., Nandakumar and Marten, 2002; Grinyer et al., 2004; Shimizu and Wariishi, 2005; Kniemeyer et al., 2006; Kim et al., 2008). To date, 139 intracellular proteins have been identified for various *Aspergillus* species (Kim et al., 2008). Using different extraction methods in parallel, a proteome which covers 898 proteins was obtained in this study.

As a next step, integration of transcriptomics and proteomics data is challenging since, e.g., turnover of mRNA, initiation of translation and stability of proteins are independent processes controlled at different levels. This might result in dissimilar or even invert relationships for mRNA and protein levels, a phenomenon which is studied among others in *Saccharomyces cerevisiae* (Daran-Lapujade et al., 2007; de Groot et al., 2007). Nevertheless, pathways and functionally related factors that responded similar on both levels are potential targets for strain improvement. Moreover, discrepancies between proteomics and transcriptomics levels can give indications for more complex regulatory (posttranslational) mechanisms which do not justify simple gene overexpression as de-bottlenecking strategy.

Here, we use a proteomics-based integrated genomics approach of controlled *A. niger* fermentations to determine the cellular responses of strains producing different homologous and heterologous enzymes. Our in house developed protein extraction protocols were used and almost 900 intracellular proteins were identified from 2D gels using automated spot excision, tryptic digestion and MALDI-TOF-MS peptide mass fingerprinting. Genes up-regulated both in the proteome and the corresponding transcriptome were selected as leads for generic strain improvement. Modulation of two of the identified lead factors resulted in an increased expression of the model protein β-glucuronidase (GUS). To our knowledge, this is one of the first applications of an integrated genomic approach based on controlled fermentations of *A. niger* to improve protein production in a generic way.

## 2. Materials and methods

### 2.1. Strains, plasmids and fermentation

#### 2.1.1. Strains

For the selection of leads by proteomics and transcriptomics, six different *A. niger* strains were fermented and subsequently analysed: three producing strains carrying several gene copies of a homologous hydrolase, a homologous protease, a heterologous lipase and their isogenic host strains.

For lead follow-up, construction of a *doaA* knock-out strain (GKUΔ*doa*A) in *A. niger* strain GKU1 (Δ*gla*A, Δ*ku*70) was performed. GKU1 is a derivative of strain GBA107 (Δ*gla*A in CBS513.88 background) and was obtained after deletion of the *kusA* gene in GBA107 using plasmid pGBKUS-5 (Meyer et al., 2007) and subsequent removal of the *amdS* marker via recombination on the flanking repeats using selection on 5′-fluoro-acetamine plates. The resulting plasmid (pΔdoaA) was linearised with NotI and AscI and transformed into strain GKU1. Transformant GKUΔ*doa*A (Δ*gla*A, Δ*ku*70, Δ*doa*A) was selected using Southern blotting (data not shown) and used for further analysis.

The β-glucuronidase (GUS) expression construct pGBGLA-GUS#14 was transformed to strain GKU1 (Δ*gla*A, Δ*ku*70) resulting in strain GKUGUS. The pGBGLAGUS construct was made in such a way that it integrates at the partially deleted *gla*A locus in the GBA107/GKU1 background via single homologous recombination on the 3′ and 3″ flanks of the *gla*A locus. The presence of a single copy of the constructs was confirmed by diagnostic PCR (data not shown). The construction of pGBsttC was similar to pGBGLAGUS construction. pGBGLAGUS#14 and pGBsttC were co-transformed into GKUΔ*doa*A (Δ*gla*A, Δ*ku*70, Δ*doa*A), resulting in strain GKU GUSSTTCΔDOAA. The tPA expression construct was transformed to strain GKU1, resulting in GKUTPA. As a control the pGBGLA514 expression plasmid was transformed into GKU1, resulting in GKUGLA514.

#### 2.1.2. Plasmids

The bacterial β-glucuronidase (GUS) and human tissue plasminogen activator (tPA-for EM studies) were used as heterologous model proteins for lead validation. The GUS and tPA genes were fused to the first 512 amino acids of the glucoamylase (GLA) protein to be expressed in *A. niger* as described (Punt et al., 1994). The *GLAGUS* fusion gene was amplified from strain (AB4-1 dGlaA36[pBB19-3] #3 (Punt et al., 1994) using long template DNA polymerase (Roche Diagnostics, Germany). The 3-kb PCR fragment was digested with PacI and AscI and cloned into PacI and AscI digested pGBFIN-5 (EP1998/08577). The DNA sequence of the PCR amplified fragment was verified by sequence analysis. The resulting plasmid pGBGLAGUS#14 was used in subsequent experiments. The tPA construct is previously described (Guillemette et al., 2007).

The Stt3 homologue present in the *A. niger* genome was identified by BlastP searches using yeast Stt3 sequences. The identified gene was named *sttC* (An16g08570—GenBank Accession No. 4989248), SttC (previously described as Sstc (WO2008/053018)) is a putative transmembrane protein localised in the ER.

The Doa1 homologue present in the *A. niger* genome was identified by BlastP searches using yeast Doa1 protein sequences. The gene identified was named *doa*A (An03g04600—GenBank Accession No. 145235384). The *A. niger* DoaA protein consists of 776 amino acids and is similar (around 35% amino acid identity) to the Doa1/Ufd3 orthologue in *S. cerevisiae*. The deletion construct pΔdoaA consists of a 1.2-kb upstream (5′ flank) and a 1.0-kb downstream (3′ flank) sequences flanking the *doa*A predicted open reading frame. In addition, a 0.6-kb repeat of the upstream region (5′ flank) was inserted between the selection marker and the 3′ flank. This repeat allows the loop out of the *amdS* marker (Kelly and Hynes, 1985) via recombination during selection on 5-fluoroacetamide plates.

#### 2.1.3. Fermentation conditions

The *A. niger* strains were grown on defined medium with glucose as carbon source and ammonia as nitrogen source using 20 L submerged stirred (Rushton turbines) fermenters at 34 °C. Glucose feeding started after 24 h. pH was controlled at 4.5 or 5.5 with ammonia or phosphoric acid. Biomass samples for transcriptome and proteome analysis were taken at the same time

point during pseudo-steady state. Foaming was prevented by automatic addition of the antifoaming agent clerol. Fermenters were inoculated from pre-grown cultures in shake flask. On-line off gas MS analysis was performed in order to calculate oxygen consumption and carbon dioxide production.

## 2.2. Transcriptome analysis

### 2.2.1. GeneChip design

Based on the genome sequence of *A. niger* CBS513.88, a proprietary GeneChip®, dsmM_ANIGERa_coll511030F, was designed according to the GeneChip® CustomExpress™ program by Affymetrix (Affymetrix, Inc., Santa Clara, CA). The chips of the 49 format contain 11 μm features. In each probe set 12 pairs of 25-oligonucleotides represent each of approximately 14,259 annotated ORFs and genetic elements of *A. niger*. The probe sequences were designed based on maximally the 600 bp region in the (predicted) coding sequence stretching from the 3′-end. Typical Affymetrix control probe sets were included (*Bacillus subtilis*, *Escherichia coli*, *bacteriophage P1* and *A. thaliana* spike controls). For several constitutive *A. niger* genes, separate 3′-end, middle and 5′-end probe sets were designed that allow monitoring of the efficiency of the *in vitro* transcription reaction. Genes of one gene family can be represented by a unique and a shared probe set. In total 14,554 probe sets were measured using the dsmM_ANIGERa_coll511030F GeneChip®. The dsmM_ANIGERa_coll511030F library and platform information is deposited at GEO under number GPL6758.

### 2.2.2. Sample preparation

Biomass samples for transcriptomics were directly frozen into liquid nitrogen and stored at −80 °C. Mycelium was grinded using liquid nitrogen, pestle and mortar and subsequently total RNA was extracted with TRIzol Reagent (Invitrogen) followed by chloroform extraction according to the manufacturer's recommendations. Total RNA was further purified using the RNA easy kit (Qiagen). Concentration of total RNA was determined on a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). Quality and integrity of the RNA samples was confirmed by verification of the $A_{260nm}/A_{280nm}$ ratio and on an RNA 6000 Nano Assay Labchip on the Agilent 2100 Bioanalyser (Agilent Technologies).

### 2.2.3. GeneChip analysis

Fifteen microgram of total RNA isolated from mycelial samples was processed strictly following the Affymetrix protocols for eukaryotic target preparation. Probe synthesis of biotin-labeled cRNA was performed using the Bioarray High Yield RNA transcript labelling kit (Enzo Life Sciences). This probe synthesis was analysed using an Agilent 2100 Bioanalyzer (Agilent Technologies). Hybridisation was performed with 15 μg of fragmented cRNA, followed by washing, staining and scanning, all according to the Affymetrix protocols for eukaryotic target hybridization and staining (Affymetrix, Inc. "GeneChip Expression Analysis Technical Manual", August 2002). Micro Array Suite (MAS 5.0 Affymetrix) and GCOS (Affymetrix) software were used for data extraction of raw intensity values and present/absent calls. For data normalisation the arithmetic mean with a target value of 100 was used. Spotfire Decision Site for functional genomics 7.1 (Spotfire Inc.) and Genedata Expressionist analyst (GeneData AG) were used for further data analysis.

## 2.3. Proteome analysis

### 2.3.1. Sample preparation

Biomass samples for proteomics were rapidly washed three times with 10 volumes cold water in order to remove extracellular proteins. Subsequently, the cells were lyophilised to prevent any

protease activity. Two different protocols for protein extraction were used.

In the first protocol, the lyophilised cells were broken with a micro-dismembrator (Braun Biotech International) using teflon vials (3 mL) and six stainless steel balls (⌀ 5 mm). 1.5 mL cold 10% TCA in acetone containing 0.07% β-mercaptoethanol was added to 50 mg lyophilised cells. Proteins were precipitated overnight at 4 °C. After centrifugation (15,000 rpm) for 20 min at 4 °C, supernatant was discarded. To wash out TCA, 1.5 mL acetone containing 20 mM DTT was added and after vortexing and centrifugation (10 min at 4 °C, 12,000 rpm), the supernatant was discarded. This washing procedure was carried out three times and the resulting pellet was air dried. The pellet was dissolved in 1.5 mL solubilisation buffer (7 M urea, 2 M thiourea, 4% CHAPS, 1% DTT, 0.5% carrier ampholytes). After solubilisation, the protein extract was centrifuged for 10 min at 15,000 rpm at room temperature. An aliquot of this extract was used for protein quantification with a Bradford assay using BSA as standard.

In the second protocol, 80 mg of the lyophilised cells was resuspended and washed with TE buffer (10 mM Tris–HCl, pH 7.5, 1 mM EDTA, 1 mM PMSF). After two washing steps, cells were resuspended in 5 mL TE buffer in the presence of DNase (200 U), RNase (10 U), protease inhibitors (1 tablet of Complete—Roche Diagnostics, Mannheim/25 mL TE buffer) and HSP25 (10 μg—Sigma H7658). This latter protein was added in order to further prevent degradation, as described in Han et al. (2005). Mechanical breaking with the French press was performed as described in Maillet et al. (2007). The lysed cell suspension was centrifuged at 14,000 rpm for 30 min at 4 °C to sediment intact cells and debris. A small aliquot of supernatant was kept to quantify the protein amount with a Bradford assay. The supernatant was lyophilised and protein solubilisation was performed according to the procedure described in Maillet et al. (2007).

The proteins from the homologous protease producing strain and its host strain were extracted using both protocols, while those from the heterologous lipase producing strain and its host strain were obtained using the second (French press, FP) protocol and those from the homologous hydrolase producing strain and its host strain were extracted using the first (dismembrator, DM) protocol.

### 2.3.2. 2D gel electrophoresis and relative quantification

For analytical gels, 300 μg protein was loaded on 240 mm Immobiline DryStrips pH 4–7 (GE Health). Two-dimensional gels were prepared according to the protocol described in Maillet et al. (2007). The analytical 2D gels were stained with SyproRuby® (Bio-Rad) and used for relative spot quantification. For each strain, three analytical gels were run. Gel images were acquired using a Fuji FLA3000 scanner.

For preparative gels, 1 mg protein was loaded on 240 mm Immobiline DryStrips pH 4–7 (GE Health) and 2D gels were prepared as described before in Maillet et al. (2007). The preparative 2D gels were stained with colloidal Coomassie blue and used for spot identification using mass spectrometry. Gel images of Coomassie blue stained gels were acquired using an AGFA Scanner.

The analytical 2D gel images were analysed by the service-provider company Ludesi (Lund, Sweden): spots were detected in all the gels, matched between the 2D gels according to an all-to all matching and quantified. For each spot, the spot volume was determined and normalisation was based on a gel-to-gel pair match relation (Ludesi, personal communication): the algorithm computes an optimised normalisation factor for each gel and uses it to normalise the spot volumes in that gel. The algorithm assumes that some spots contain proteins whose expression level is constant during the whole experiment. The algorithm uses these proteins to find the optimal normalisation for the gels. The effects are that outliers and differentially expressed proteins are filtered out

and a more homogeneous spot set is used during the normalisation. The SyproRuby stained gels and the Coomassie blue stained gels have been normalised separately, since the Coomassie gels were only used for protein identification and not quantification. Quantitative data analysis is based on the comparison of the different producing strains with their corresponding host strain. For each strain, triplicate analytical gels have been used for statistical interpretation.

### 2.3.3. Spot identification and mapping

All visible spots on the Coomassie blue stained 2D gels were excised with an automated home-made spot picker, in-gel destained and digested with trypsin (Fountoulakis et al., 2004). The protein digests were analysed using MALDI-TOF-MS on a Bruker Ultraflex II TOF/TOF system (Bruker Daltonik, Bremen, Germany). For protein identification, the experimental peptide mass fingerprints (PMFs) were compared with the theoretical trypsin digests from the DSM *A. niger* protein sequences according to the procedure described in Maillet et al., 2007.

A script was developed to map quantification and identification data. Spot coordinates were used to link spot numbers from identification data to match identifiers from quantification data. This script was also developed to reduce the complexity of identification data: a spot present on a Coomassie blue gel can be matched to other spots on other Coomassie blue gels and then several identifications can be assigned to this spot. The criteria for this step are the number of times protein identification has been assigned to this spot and the probability of a false identification assigned to each MS measurement. As a final result, a consensus in protein identification was generated for each spot.

For data mining, the Expressionist module from Genedata (Basel, Switzerland) was used for statistical data analysis. For each strain comparison (overproducing strain compared to its host strain), statistically significant results were defined by filtering according to both a p-value of 0.05 and a cut-off ratio of 1.5.

### 2.4. Protein oxidation analysis

Protein extracts were derivatised using DNPH (2,4-dinitrophenylhydrazine) according to the OxyBlot™ Protein Oxidation Detection Kit (Chemicon). These samples and their controls without derivatization were used for SDS–PAGE using the XCELL Surelock system (Invitrogen). The gel (NuPAGE™ 4–12% BisTris) was run for 50 min at 200 V, with 600 mL 20× diluted MOPS–SDS buffer in the lower buffer chamber and 200 mL 20× diluted MOPS–SDS buffer (Invitrogen), containing 0.5 mL of antioxidant (Invitrogen) in the upper buffer chamber.

After 1D SDS–PAGE, the gel was blotted on a nitrocellulose membrane. The blot module was placed into Xcell II™ Blot Module (Invitrogen) and filled with transfer buffer (50 mL (20×) transfer buffer, Invitrogen) 200 mL methanol and 750 mL milliQ water). Blotting was performed for 90 min at a constant voltage of 25 V. After blotting, the membrane was incubated for 60 min with gently shaking in 20 mL blocking/dilution buffer (1% skim milk in PBS-T (phosphate-buffered saline, containing 0.05% Tween-20)). The 1st antibody stock (Rabbit Anti-DNP antibody) was 150 times diluted with blocking/dilution buffer just for use. The membrane was incubated in 15 mL of this solution during 60 min at room temperature with gently shaking. The membrane was rinsed and washed several times with PBS-T and then incubated with 15 mL of the diluted the 2nd antibody (Goat Anti-Rabbit IgG-Cy5 (GE Health)). The ECL Plex CyDye conjugated 2nd antibody was diluted 2500 times in PBS-T and the membrane was incubated in 25 mL of this solution at room temperature with gently shaking for 60 min (protected from light). After the membrane was rinsed and washed (4 × 5 min in PBS-T), the 2nd antibody signal was detected by scanning the membrane using the Typhoon 9200 (GE Health) 670 BP 30, Red (633 nm), PMT 500 V, 100 μm.

### 2.5. Morphological electron microscopy studies of protein-expressing strains

Cells were grown between PCTE membranes on complete medium (as described by Bennet and Lasure (1991), complemented with 0.1% (w/v) Casamino acids and 0.5% (w/v) yeast extract) agar plates containing 1% maltose for 5 days at 30 °C. Mycelium was fixed for 30 min in freshly prepared 1% $KMnO_4$ (Merck) on ice. Fixed mycelium was washed with 10% ethanol and pieces of approximately 2 by 4 mm were cut out of the colony periphery. The mycelium pieces were then dehydrated by a graded series of ethanol (30%, 50%, 70%, 80%, 90%, 96%, 100%) and further dehydrated with 1,2-propylene oxide (Merck; 50%, 75% and 100%). The mycelium pieces were infiltrated (25%, 50%, 75% and 100%) with and embedded in Spurr's resin (Spurr, 1969). The resin was polymerised for 48 h at 70 °C. Ultra-thin sections (90 nm) were cut with an Ultracut E ultramicrotome (Leica). Sections were post-stained for 10 min with aqueous 4% uranyl acetate in water (Merck) and 2 min with 0.4% lead citrate (Merck) (Venable and Coggeshall, 1965) and studied using a Tecnai 10 transmission electron microscope (FEI Company) at an acceleration voltage of 100 kV.

### 2.6. Glucuronidase assay

To analyse the intracellular levels of GUS protein, the GUS strains with and without *doaA* deletion were grown for 16 h in *Aspergillus* complete medium with 1% maltose as a carbon source. Total proteins were extracted after grinding frozen mycelium using pestle and mortar in liquid nitrogen. Proteins were extracted using a 10 mM sodium phosphate buffer, pH 6.0, containing 2% SDS, 10 mM EDTA and 1 mM PMSF. Protein concentrations of the different samples were determined with a Bradford assay using BSA as standard and equal amounts of protein were loaded on a SDS–PAGE gel and blotted to nitrocellulose-membranes. GUS protein was detected using a GUS-specific antibody followed by a goat-anti-rabbit-HRP secondary antibody and standard chemiluminescence technology (ECL-kit). In addition, total GUS activity was determined on plates as described (Damveld et al., 2005).

## 3. Results and discussions

### 3.1. Fermentation reproducibility

A first prerequisite for the development of a robust and sensitive quantitative functional genomics approach is a reproducible fermentation protocol. Reproducibility at 10 L fermentation scale was tested by fermenting an *A. niger* strain in triplicate. After prolonged fed batch cultivation the fermentations reached a pseudo-steady state. At this state the coefficient of variation of the fermentation parameters analysed appeared to be 2–5%. Transcriptome analysis was used to verify the reproducibility at the gene expression level. Approximately, 5700 genes of the 14,165 *A. niger* genes were expressed. Within the triplicates only 0.3% of the expressed genes was differentially expressed (>2-fold change). These results indicate that both the fermentations and the transcriptome analysis were highly reproducible.

### 3.2. Proteome and transcriptome analysis

#### 3.2.1. Protein extraction and identification

Cell lysis, protein extraction and separation determine the subset of proteins that can be analysed using proteomics. To compare the impact of the extraction method, washed biomass from the

homologous protease producing strain as well as that of the corresponding isogenic host strain were extracted using either the dismembrator (DM) or French press (FP) protocol (see Section 2.3.1). The visible spots from the preparative 2D gels of the different extracts were excised and analysed using MALDI-TOF-MS for protein identification. Independent of the extraction method approximately 1000 spots could be identified on a single 2D gel. On average, the number of isoforms identified per protein is two. On the 2D gel of the protein extract from the homologous protease producing strain 989 spots were identified in the DSM *A. niger* database, which correspond to 497 different proteins. In total, 898 different proteins were identified from the 2D gels. The number of proteins identified with the different extraction methods is similar (Table 1). However, the overlap between these sets is 53%, indicating that the extraction protocol has a significant effect on the subset of proteins.

A comparison was made between the proteomes of the protease producing strain and the corresponding isogenic host for each protocol. Twenty-nine proteins were present in significantly higher amounts in the overproducing strain when the French press method was applied (Table 2). Using the dismembrator protocol the same experiment resulted in 24 up-regulated proteins. Ten up-regulated proteins were observed in both data sets illustrating the value of using different extraction methods. Since in this experiment extracts of the producing strain and the isogenic host strain were prepared using the same protocol, the individual data sets were combined resulting in a total set of 43 up-regulated proteins in the protease producing strain compared to the isogenic host strain. This result confirms the importance to use different extraction methods in order to generate a broad proteome spectrum.

### 3.2.2. Virtual 2D gel

A virtual 2D gel for *A. niger* was generated by plotting for all 14,165 putative gene products the theoretical molecular weight (MW) against the calculated p*I* based on the amino acid composition (Fig. 1a). Interestingly, two clusters of theoretical spots are present on the virtual 2D gel, comparable to the virtual 2D gels of *S. cerevisiae* and different bacteria (Wildgruber et al., 2000; Hiller et al., 2003). The clear gap at p*I* 7–7.5 probably corresponds to the intracellular pH, which has experimentally been determined to be 7.2–7.5 (Hesse et al., 2002). A net zero charge is unfavourable for the solubilisation properties of proteins and would make them nonfunctional in the cell. Fig. 1b shows a virtual 2D gel of the set of proteins identified in this work. A large group (84%) are clustered in the expected p*I* range 4–7 and MW range 10–200 kDa of the 2D gel. However, proteins are also located in the p*I* range 7.5–10.5 and MW range 10–200 kDa, indicating a discrepancy between the theoretical and experimental p*I*.

**Table 1**
Total number of different proteins identified on preparative 2D gels from six *A. niger* strains

| Strain | Extraction protocol | Number of different proteins identified[a] |
| --- | --- | --- |
| Homologous hydrolase host | FP | 301 |
| Homologous hydrolase | FP | 320 |
| Heterologous lipase host | DM | 530 |
| Heterologous lipase | DM | 342 |
| Homologous protease host | DM | 497 |
| Homologous protease | DM | 232 |
| Homologous protease | FP | 233 |

Proteins were extracted either using the French press (FP) or dismembrator (DM) protocol.
[a] Isoforms are not included.

**Table 2**
Number of proteins significantly up-regulated between a producing strain and its host

| Production vs host strain | Extraction protocol | Number of identified proteins significantly up-regulated |
| --- | --- | --- |
| Homologous hydrolase | FP | 136 |
| Heterologous lipase | DM | 94 |
| Homologous protease | DM | 24 |
| Homologous protease | FP | 29 |

Proteins were extracted either using the French press (FP) or dismembrator (DM) protocol.

In general, the positions of the protein spots on the virtual 2D gels show a good similarity with the experimentally obtained gels, mostly in p*I* range ±0.5 U (data not shown). However, differences have been observed and can be explained either by (i) wrong predictions of the p*I* and MW, or (ii) due to posttranslational modifications of a protein. The first type of difference could originate from a wrongly predicted gene-model (splicing and correct start and stop identification), resulting in a different gene length and wrong p*I* and MW values for the protein. The second type of differences can be further split into posttranslational modifications such as phosphorylation, glycosylation, splicing, which can have a significant effect on the MW and/or p*I*.

Eventually, 898 proteins were identified from 2D gel, representing only 6% of the 14,165 proteins predicted for the *A. niger* CBS513.88 genome. These proteins represent 12% of all proteins that could theoretically be present, when expressed, on the 2D gels based on their calculated theoretical p*I* and MW between (4<p*I*<7 and 10 < MW < 200 kDa, 7578 proteins are predicted), (see Fig. 1b). It is evident that besides using different extraction methods, the usage of 2D gels with different pH and MW ranges ensures a further increase of the spectrum of identifiable proteins. Alternatively, an LC-based approach allows a different part of the proteome to be identified. Using off-line 2D LC–MS/MS to analyse a tryptic digest of the DM extract from the protease producing strain (data not shown), we were able to identify significantly more proteins than on the respective 2D gel. Only 35% of the proteins were identified using both approaches, showing again that the subset of proteins to be analysed depends on the sample preparation, separation and detection. As expected, many more proteins were identified in the p*I* 5–6 region using the 2D gel approach. Interestingly, 33% of the proteins identified using 2D LC–MS/MS have a theoretical molecular mass above 150 kDa and a theoretical p*I* larger that 8 and, therefore, fall out the range of the used 2D gels. Further increase in the number of protein identifications, especially low abundant ones, is possible using high mass accuracy mass spectrometry (e.g., Lu et al., 2008; Zimmer et al., 2006).

### 3.2.3. Transcriptome analysis

Samples from fermentations used for proteomics analysis were also analysed at the transcriptomics level using GeneChips. Thus, allowing a comparative analysis between the transcriptome and proteome. The transcriptome analysis resulted in numerous genes which were differentially expressed in the production strains vs their isogenic host strains. The protease producing strain showed 136 genes which were up-regulated and 175 genes which were down-regulated compared to the isogenic host strain. The lipase producing strain showed 312 genes which were up-regulated and 308 genes which were down-regulated compared to the isogeneic host strain. The hydrolase producing strain showed 190 up-regulated genes and 141 down-regulated genes compared the
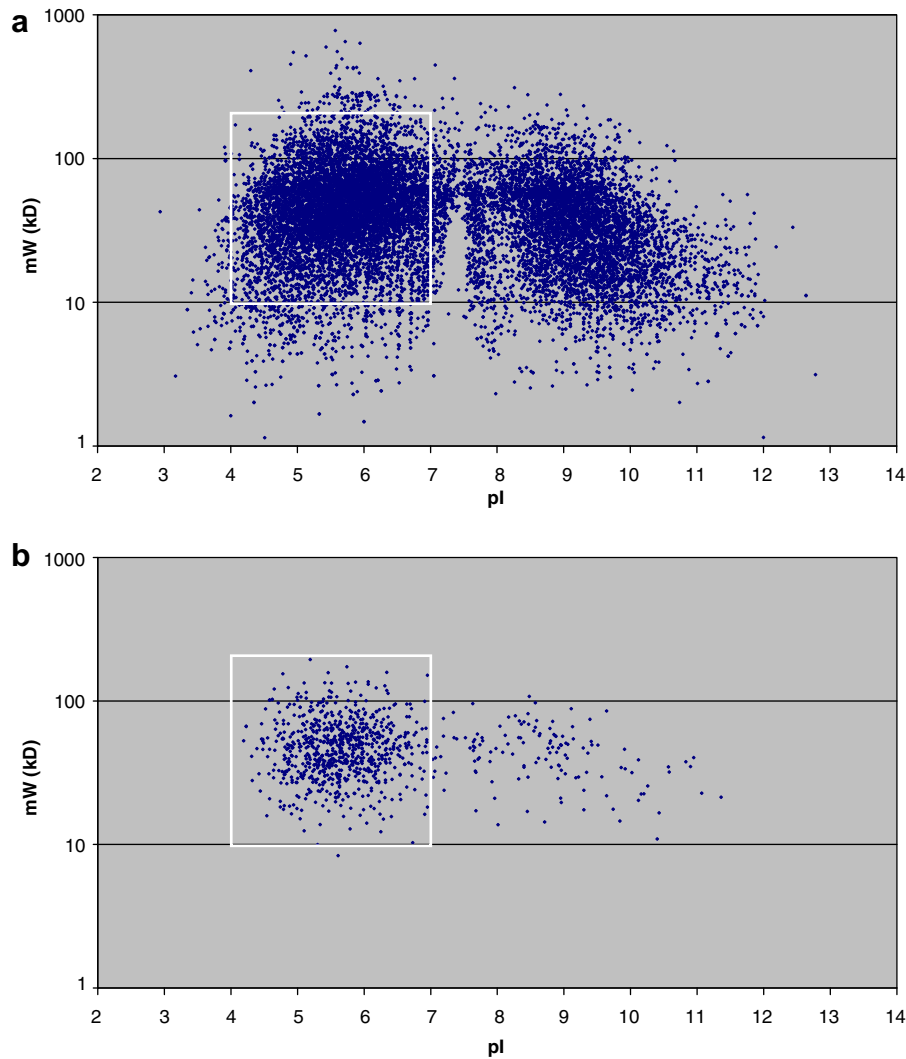
**Fig. 1.** Virtual 2D gel of (a) all 14,165 *A. niger* genes and (b) the 898 proteins identified in this study. The white rectangle indicates the ranges of the 2D gels used.

isogenic host strain. The aim of this analysis was to integrate the transcriptome data set with the proteome data set enabling selection of biologically relevant leads for strain improvement. The transcriptomics data set is used to overlay with the differential proteins resulting from the proteomics analysis in order to filter out genes which show co-regulation both at the proteome and transcriptome level.

*3.3. Integrative data comparison and lead selection*

Comparing the identified proteins with the mRNA expression levels of the encoding genes revealed that these proteins are mainly derived from the group of highly expressed genes (Fig. 2). Moreover, annotation shows that the majority of these proteins belong to the classes "known" or "strong similarity with known protein". Comparing the functional categories based on the known genes within the genome, the proteomics data demonstrate that well-known metabolic proteins are overrepresented, while the unclassified group is underrepresented (Fig. 3). It is clear that proteins which are abundantly present are well studied resulting in improved annotation. In conclusion, the proteomics-identified protein-subset is biased towards abundant proteins, that are stable, soluble and tend to have a strong similarity to well-studied proteins.

The strains used in this study produce different amounts of enzymes. The heterologous lipase is produced at low quantities, while the homologous protease and hydrolase are produced at high levels. The proteomes of the protein producing strains have been compared with their respective isogenic host strains. In Table 2, the number of up-regulated proteins between the producing and corresponding isogenic strains is presented. For example, from the 2D gels corresponding to the heterologous lipase producing strain and its isogenic host, 174 spots are significantly differentially expressed. Among these spots, 94 have been identified as proteins, with known function.

In order to identify proteins that are differentially expressed in all protein producing strains, a Venn diagram was made (Fig. 4) in which overlapping responses are depicted as overlapping circles. Up-regulated proteins and genes identified in the protein production strains compared to their isogenic host strains from the overlapping subsets 1, 2, 3 and 4 (Fig. 4) are considered to represent general cell responses on protein production and thereby to contain the most relevant generic leads for protein production improvement. Therefore, these proteins were functionally categorised (Fig. 5). The main functional categories of up-regulated proteins in the production strains are protein folding (12 proteins), protein degradation (10 proteins) and C/N metabolism (19 proteins), indicating a clear link between protein flux, folding capacity
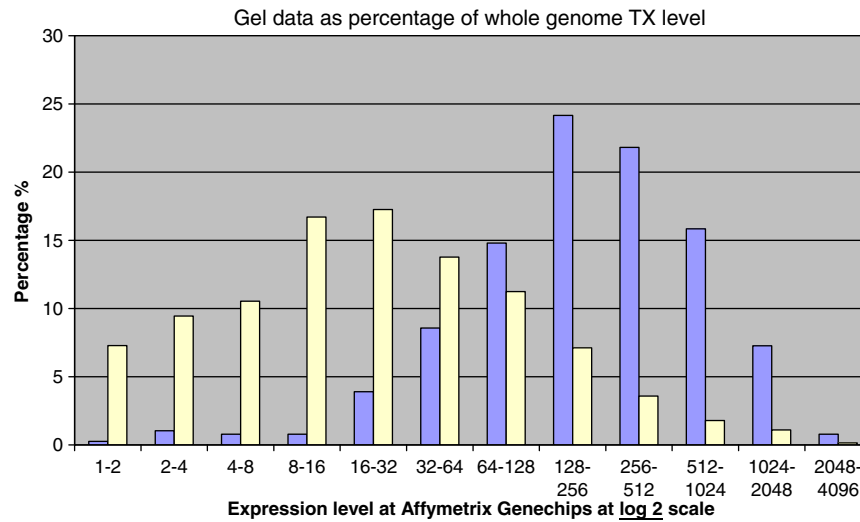
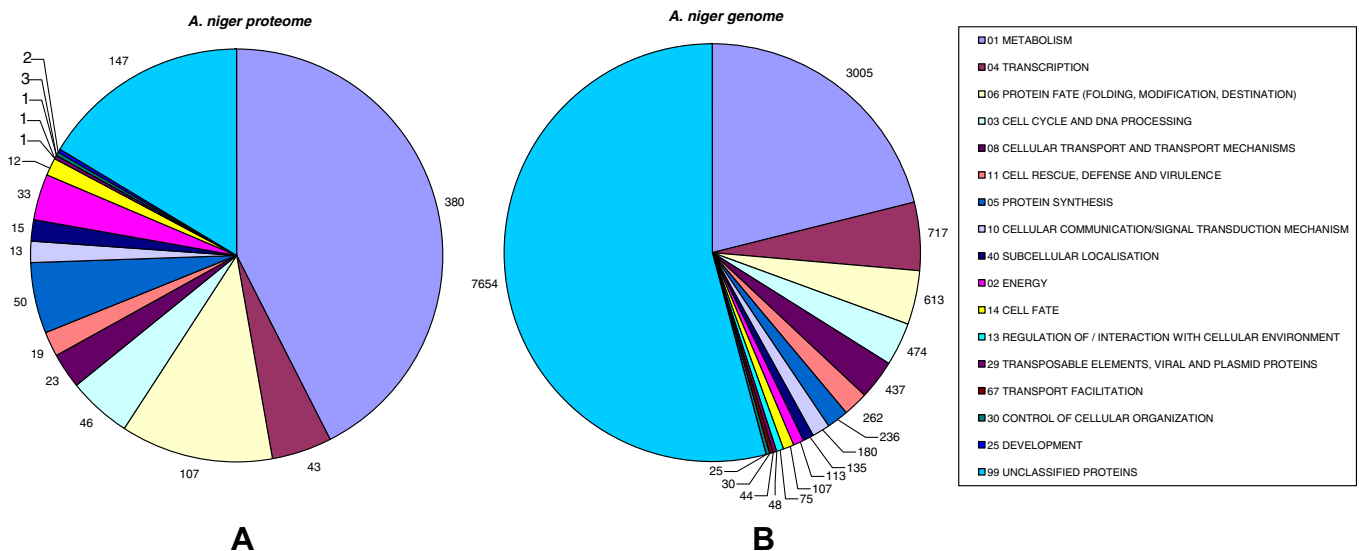**Fig. 2.** Expression levels of transcriptome (yellow) and proteome (blue) data.



**Fig. 3.** Functional classification according to Funtional Catatalogue (Fun Cat), as described by Ruepp et al. (2004), of (A) all 14,165 *A. niger* genes and (B) the 898 proteins identified in this study.

and a high demand for building blocks and energy. These findings are summarised in the pyramid in Fig. 6.

The up-regulated proteins for C/N metabolism could reflect the high demands for nitrogen (protein production) and carbon (energy consumption) due to the large amounts of protein being produced. A relatively large part of the proteins which are up-regulated in the homologous hydrolase and protease producing strains, belong to the central (C/N) metabolism, indicating a bottleneck at the first level of the pyramid (Fig. 6). This is in line with the high protein fluxes in these strains.

The synthesised proteins have to be folded correctly which is indicated by the second level of the pyramid. Our data suggest that at sufficient C/N supply, protein folding becomes a rate limiting step in protein production. All three protein producing strains clearly show an up-regulation of proteins involved in protein folding. The limitations in folding capacity can either be caused by the flux being too high (quantitative) or by the product containing intrinsic folding problems (qualitative), for example, in case of the heterologous

lipase due to its foreign nature, or by a combination of both. Remarkably, all proteins involved in protein folding are up-regulated upon production of a heterologous lipase, whereas no additional proteins involved in protein folding are up-regulated in the homologous (hydrolase and protease) producing strains. This observation indicates that heterologous lipase production causes problems at the level of protein folding whereas homologous protein production shows no specific protein folding response.

The third level of the pyramid is composed of proteins involved in protein degradation. Both cytoplasmic and vacuolar proteins involved in protein degradation are up-regulated. Apparently, when the cell is not able to fold proteins at high fluxes it reacts with protein degradation. Misfolded proteins are disposed by the fungal cell which contains an effective machinery for intracellular protein degradation (Pel et al., 2007).

The last functional category as shown at the top of the pyramid comprises proteins involved in ROS (reactive oxygen species)/ stress. These proteins are only up-regulated in the homologous
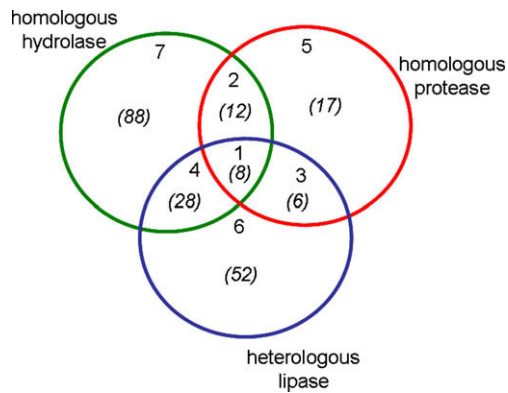
**Fig. 4.** Venn Diagram for proteome analysis of *A. niger* strains. The (sub)sets 1–7 contain proteins that are up-regulated in the producing strain compared to its host strain. Numbers indicate different subsets as indicated below, numbers in italics and between brackets represent the number of up-regulated proteins in each subset. Set 1: up-regulated proteins in all combinations. Set 2: up-regulated proteins in the homologous hydrolase and protease strains. Set 3: up-regulated proteins in the heterologous lipase and homologous protease strains. Set 4: up-regulated proteins in the homologous hydrolase and heterologous lipase strains. Set 5: up-regulated proteins in the homologous protease strain. Set 6: up-regulated proteins in the heterologous lipase strain. Set 7: up-regulated proteins in the homologous hydrolase strain.

hydrolase and heterologous lipase strains (Fig. 4, subset 4). A link between the number of S–S bridges and ROS has been reported (Schröder and Kaufmann, 2005; Tu and Weissman, 2002). In addition, the correlation between ROS and impaired protein secretion has been shown before in *S. cerevisiae* (Sagt et al., 2002) and, the other way around, an increase in misfolded proteins results in oxidative stress (Dukan et al., 2000). The heterologous lipase contains 14 cysteine residues, while the homologous hydrolase and the homologous protease have 10 and 7 cysteine molecules, respectively. The large amount of homologous hydrolase produced and the large amount of cysteines in the heterologous lipase, give a high pressure on formation of disulphide bridges in the production strains, which is probably reflected in the arise of ROS and oxidative stress response. The homologous protease producing strain apparently is still able to handle the amount of disulphide bridges to be formed and ROS/oxidative stress responses are not induced.

The three differential data sets were compared to determine the co-regulation of proteome and transcriptome level. Co-regulation of protein and gene level was observed in the functional categories C/N metabolism (26%), protein folding (30%), ROS/oxidative stress

(20%) and protein degradation (58%). It is clear that the protein degradation pathway, which was found to be up-regulated in all protein producing strains, showed the highest level of co-regulation on mRNA and protein level. This makes this pathway the most promising lead for generic strain improvement, since the high level of co-regulation indicates that straightforward overexpression or knock-out strategies are likely to result in altered protein levels and possibly altered phenotypes. Other categories showed much less co-regulation of gene and protein levels which makes strain modification using these categories less straightforward. In this integrative approach the differential expressed functional pathways which showed correlation on protein and gene level were ranked dominant over the fold changes of individual genes and proteins.

### 3.4. Follow-up and validation of leads

#### 3.4.1. Oxidative stress

Up-regulation of several proteins involved in the ROS/stress responses in the heterologous lipase and homologous hydrolase producing strains suggests that these strains suffer from oxidative stress, caused by an excess of ROS. Oxidative stress results amongst others in protein oxidation, which is an irreversible process leading to loss of protein activity. One of the first target proteins for oxidation are ER proteins involved in protein folding (van der Vlies et al., 2002). To verify protein oxidation in the production strains, the OxyBlot™ Protein Oxidation Detection Kit (Chemicon) that detects carbonyl groups on a Western blot, was used (Fig. 7). As shown in Fig. 7B the protein load of production strains and their isogenic hosts is equal. In Fig. 7A it is shown that the lipase overproducing strain contains carbonylated proteins which are indicative of ROS. The hydrolase overexpression strain shows no increase in carbonylated proteins compared to the isogenic host strain. These results are in alignment with the proteomic data obtained for the lipase producing strain.

#### 3.4.2. The effect of protein production on the subcellular morphology

Heterologous protein production is often accompanied by reactive oxygen species (ROS) stress and enlargement of the ER membrane (Sagt et al., 2002). We studied enlargement of ER membranes in *A. niger* due to heterologous protein production using GUS and tPA as a model protein. It has been shown before that GUS and tPA expression in *A. niger* evokes a strong UPR response, similar to the expression of lipase in *A. niger* (Rolf Kooistra personal communication). Subcellular morphology of the two heterologous pro-
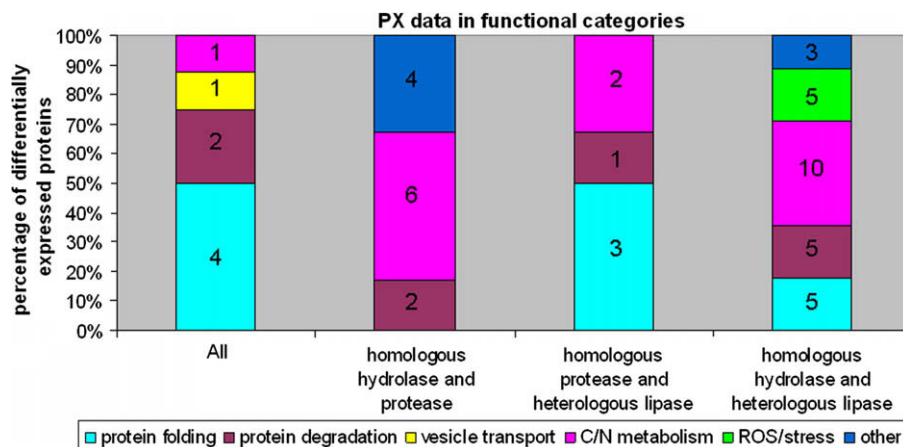


**Fig. 5.** Graphical representation of up-regulated proteins in the production strains classified in different functional categories of sets 1–4 of the Venn diagram (Fig. 4). Proteins are classified according to their annotation, the number of proteins is given and their relative contributions are calculated.
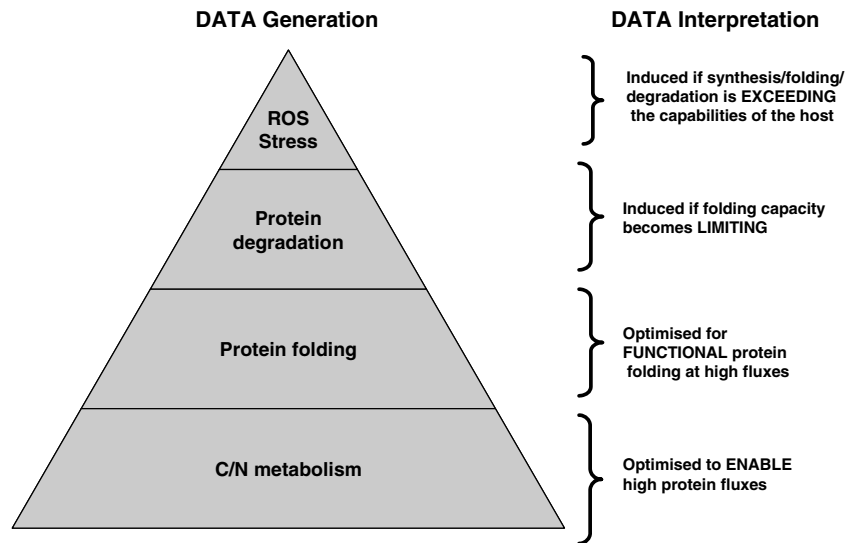
**DATA Generation**



**DATA Interpretation**

Induced if synthesis/folding/degradation is EXCEEDING the capabilities of the host

Induced if folding capacity becomes LIMITING

Optimised for FUNCTIONAL protein folding at high fluxes

Optimised to ENABLE high protein fluxes

**Fig. 6.** Summary of the data interpretation related to protein production. Different levels of cellular responses are depicted as a pyramid. Basic responses enabling other responses on a higher level are indicated.
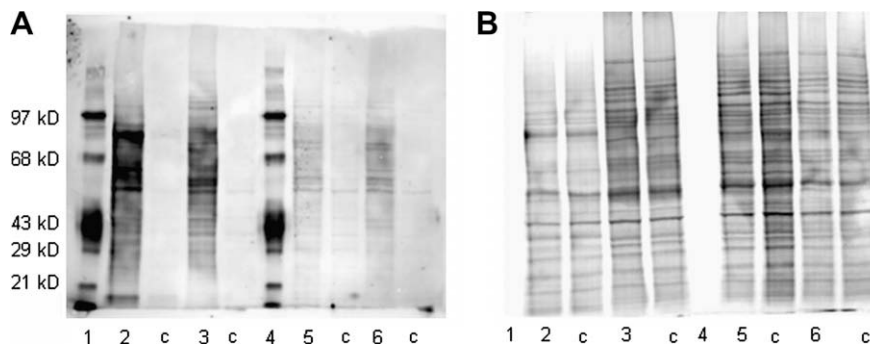


**Fig. 7.** (A) Western blot of intracellular protein extracts immunostained for carbonylated proteins using the OxyBlot™ Protein Oxidation Detection Kit (Chemicon). Lanes 1 and 4, DNP-marker proteins; lane 2, heterologous lipase strain, and its control without DNPH derivatization; lane 3, host strain of the heterologous lipase strain, and its control; lane 5, homologous hydrolase, and its control; lane 6, host strain of the homologous hydrolase, and its control. (B) SDS–PAGE gel after Western blotting stained with Sypro Ruby (general protein stain). Lanes as described for Fig. 3A.

tein-expressing strains GKUGUS and GKUTPA were compared to their parental strain GKUGLA514 by electron microscopy. A striking difference in amount and organisation of the ER membranes could be observed. GKUGUS and GKUTPA cells contain more ER membranes that are occasionally organised in stacks of several membranes (Fig. 8). Some of the ER membranes also seemed thicker than the wild-type ER, although the methods used for fixation and staining do not allow for very accurate measurements. The heterologous protein expressing strains also exhibit an increase in vacuole number. Only cells from the colony periphery of an agar plate culture were studied. These cells normally do not contain many vacuoles as can be seen in the glucoamylase producing reference strain (Fig. 8a). In yeast, ER membranes have previously been shown to proliferate upon induction of the unfolded-protein response by DTT stress (Bernales et al., 2007). Both the amount of membranes and the volume of the ER lumen increased. A similar effect was shown in the heterologous protein producing strains GKUGUS and GKUTPA. Since expression of many heterologous proteins is known to induce UPR, likely also genes involved in lipid biosynthesis are up-regulated. This may have resulted in ER proliferation. In addition, accumulation of unfolded or misfolded proteins in the ER may also directly contribute to enlargement of the ER lumen. Furthermore, when the yeast ER is overloaded with

unfolded proteins, some of these proteins can escape degradation by ERAD and exit the ER to be transported to the Golgi. In the Golgi, however, the unfolded proteins are subjected to an additional control system that sorts them to the vacuole for degradation (Kincaid and Cooper, 2007). A similar mechanism might be present in A. niger, which could explain the increased vacuolization of GKUGUS and GKUTPA in peripheral hyphal cells.

### 3.4.3. Increasing GUS production by disruption of DoaA and overexpression of SttC

The results of the integrated genomics approach (Fig. 6) indicated that protein degradation and protein folding play an important role in determining protein titres for production strains. The ultrastructural studies described above also indicate that protein folding and ERAD might be an important factor determining productivity of A. niger. An essential gene in this ERAD route might be doaA, a factor required for ubiquitin-mediated proteolysis (Rabinovich et al., 2002). The sttC gene is involved in glycosylation of secretory proteins which is important for protein folding (Yan and Lennarz, 2002). In order to decrease the ERAD activity and increase the glycosylation and folding potential of the cellular protein factory we knocked out doaA and SttC levels were increased by overexpression of this oligosaccharyl transferase gene.

**GKUGLA514**



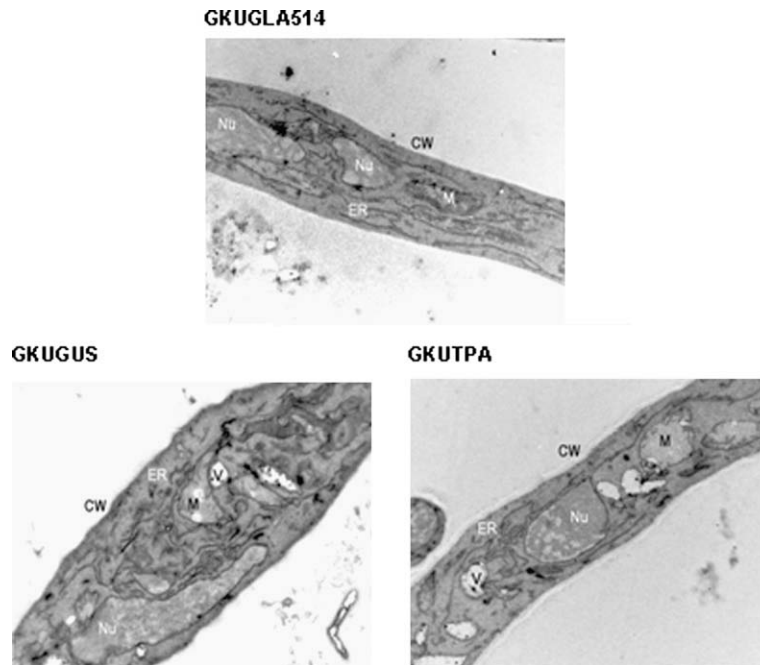**GKUGUS**                                    **GKUTPA**



**Fig. 8.** GKUGLA514 morphology compared to GKUGUS and GKUTPA. Electron micrographs of parts of hyphal cells of *Aspergillus niger* strain GKUGLA514 (A) and two heterologous protein producing strains GKUGUS (B) and GKUTPA (C). GKUGLA514 shows a normal endoplasmic reticulum (ER) morphology, several strands of ER membrane can be observed around the nucleus and throughout the hypha. In both GKUGUS and GKUTPA much more ER can be observed. The ER in these two strains is sometimes organised in multiple stacks of membrane. Vacuoles are also more common in the heterologous protein producing strains. ER, endoplasmic reticulum; M, mitochondrion; Nu, nucleus; V, vacuole.

Strains with increased SttC levels and disruption of *doaA* (GKUGUSSTTCΔDOAA,) were analysed with respect to (heterologous) protein production using GUS as a model protein. To analyse and compare the (intracellular) levels of GUS protein in GKUGUS and GKUGUSSTTCΔDOAA, both strains were grown simultaneously. Higher GUS activity (Fig. 9) was detected in the GKUGUSSTTCΔDOAA strain compared to its reference strain GKUGUS. This indicates a combined positive effect of the *doa*A disruption and overexpression of *sttC* on the expression of the heterologous protein β-glucuronidase in *A. niger*.

## 4. Conclusions

In this era of genomic research the main challenge is how to use the avalanche of data which is generated. Data handling is being automated and PCA (principle component analysis) is being used to prioritize leads. However, even an integrative genomics approach as described here, results in a wide spectrum of possible leads for strain improvement. Since the throughput of strain construction is by far not on the same level as data generation, an effective selection is the key for successful strain improvement. Therefore the selection of biological relevant data remains a com-

bination of profound knowledge on cellular architecture and classical disciplines like biochemistry and microscopy.

Most straightforward candidates for strain improvement are leads, which show correlation in their response on proteome and transcriptome level and belong to the same functional pathway. In this comparative integrative study, genes linked to protein folding and degradation (ER folding capacity) and amino acid metabolism show a high level of correlation and show repeatedly differential expression.

Interestingly, correlation between transcriptomics and proteomics data is not always found and can sometimes even be opposite. When transcriptomics and proteomics signals are not correlating well, this might be an indication for additional layers of control of regulation, complicating selection of leads for strain development. These discrepancies have been observed before (e.g., de Groot et al., 2007) and can be caused by a variety of reasons. Some proteins are difficult to visualise on 2D gels or show severe instability. Other proteins might be difficult to extract or show migration problems in the gel. Therefore, we recommend to use different extraction procedures to increase the number of proteins which can be identified. The generation of a broad spectrum proteome is a prerequisite towards effective integration of proteome and transcriptome.

Biological regulation is much more complex than the central dogma of DNA–RNA–protein. It is this variety of regulatory mechanisms which calls for a multidisciplinary approach for effective improvement of protein production strains. The different layers of regulation, on protein level and mRNA level, make straightforward genome modification not always successful in strain optimization. This is the most interesting feature of integrated genomics since posttranslational and posttranscriptional modifications might give more focus on the approach which should be followed to relieve bottlenecks. Proteomics will mature in different specialised expertises like phospho-proteomics, subcellular proteomics or combined with pulse chase experi-



**Fig. 9.** Detection of total GUS activity in colonies on plates. Indicated strains were measured for GUS activity as previously described by Damveld et al. (2005).

ments. This will enable the overexpression of genes encoding modified proteins like constitutive phosphorylated proteins, proteins modified in localisation signal or proteins with altered degradation signals. This approach is very promising for the future and depicts the true power of strain improvement based on integrated genomics.

This study shows the importance of an integrated approach for effective strain improvement. Based on proteomic data the corresponding genes were selected from transcriptomics data and differential expressed proteins which showed regulation on mRNA level as well were selected for strain improvement. Elements of the protein folding pathway, showed the highest co-regulation on proteomics and transcriptomics level. The production systems for the lipase and hydrolase showed co-regulation of ROS involved genes and proteins as well. To validate these leads, carbonylated proteins were determined and ultrastructural studies were performed on tPA and GUS expressing model systems in *A. niger*. It is this combination of integrative genomics and biochemical and microscopical studies which eventually showed that *doaA* and *sttC* could be crucial factors for generic strain optimization. In yeast it is known that DOA1 is involved in degradation of ubiquitinated proteins by interaction with CDC48 (Mullally et al., 2006). SttC has been shown to be involved in oligosaccharyl transferase in yeast (Yan and Lennarz, 2002). The GUS model system showed that *doaA* removal in combination with sttC overexpression is increasing GUS production in *A. niger*. The generic improvement of the *doaA* knockout, *sttC* overexpression strain is currently being studied in our laboratory.

Advanced proteomic approaches as described above will enhance the success rate of strain improvement based on genomics results. These developments will form the basis of a new way of working in molecular biology research. Broad hypothesis generation based on integrative genomics will be followed by fine tuning using classical approaches like physiology, biochemistry and microscopy followed by strain modification and validation. Again the frontiers of science will be at the interface of different disciplines which provides a challenge for the future.

## Acknowledgments

## References

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 22, 7–32.

Arvas, M., 2007. Comparative and functional genome analysis of fungi for development of the protein production host *Trichoderma reesei*. VTT publications 654.

Baker, S.E., 2006. *Aspergillus niger* genomics: past, present and into the future. Med. Mycol. 44 (Suppl. 1), S17–S21.

Bennett, J.W., Lasure, L.L., 1991. More Gene Manipulations in Fungi. Academic Press, San Diego.

Bernales, S., Schuck, S., Walter, P., 2007. ER-phagy: selective autophagy of the endoplasmic reticulum. Autophagy 3, 285–287.

Bowman, S.M., Free, S.J., 2006. The structure and synthesis of the fungal cell wall. Bioessays 28, 799–808.

Carberry, S., Doyle, S., 2007. Proteomic studies in biomedically and industrially relevant fungi. Cytotechnology 53, 95–100.

Collén, A., Saloheimo, M., Bailey, M., Penttilä, M., Pakula, T.M., 2005. Protein production and induction of the unfolded protein response in *Trichoderma reesei* strain Rut-C30 and its transformant expressing endoglucanase I with a hydrophobic tag. Biotechnol. Bioeng. 89, 335–344.

Damveld, R.A., Arentshorst, M., Franken, A., van Kuyk, P.A., Klis, F.M., van den Hondel, C.A., Ram, A.F., 2005. The *Aspergillus niger* MADS-box transcription factor RlmA is required for cell wall reinforcement in response to cell wall stress. Mol. Microbiol. 58, 305–319.

Daran-Lapujade, P., Rossell, S., van Gulik, W.M., Luttik, M.A., de Groot, M.J., Slijper, M., Heck, A.J., Daran, J.M., de Winde, J.H., Westerhoff, H.V., Pronk, J.T., Bakker, B.M., 2007. The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. Proc. Natl. Acad. Sci. USA 104, 15753–15758.

De Groot, M.J., Daran-Lapujade, P., van Breukelen, B., Knijnenburg, T.A., de Hulster, E.A., Reinders, M.J., Pronk, J.T., Heck, A.J.R., Slijper, M., 2007. Quantitative proteomics and transcriptomics of anaerobic and aerobic yeast cultures reveals post-transcriptional regulation of key cellular processes. Microbiology 153, 3864–3878.

Dukan, S., Farewell, A., Ballesteros, M., Taddei, F., Radman, M., Nyström, T., 2000. Protein oxidation in response to increased transcriptional or translational errors. PNAS 97, 5746–5749.

EP1998/08577—Expression Cloning in Filamentous Fungi (WO1999/32617).

Fountoulakis, M., Juranville, J.F., Jiang, L., Avila, D., Röder, D., Jakob, P., Berndt, P., Evers, S., Langen, H., 2004. Depletion of the high-abundance plasma proteins. Amino Acids 27, 249–259.

Grinyer, J., McKay, M., Nevalainen, H., Herbert, B.R., 2004. Fungal proteomics: initial mapping of biological control strain *Trichoderma harzianum*. Curr. Genet. 45, 163–169.

Guillemette, T., Peij, N.N.M.E., Goosen, T., Lanthaler, K., Robson, G.D., van den Hondel, C.A.M.J.J., Stam, H., Archer, D., 2007. Genomic analysis of the secretion stress response in the enzyme-producing cell factory *Aspergillus niger*. BMC Genomics 8, 158.

Han, M.J., Lee, J.W., Lee, S.Y., 2005. Enhanced proteome profiling by inhibiting proteolysis with small heat shock proteins. J. Proteome Res. 4, 2429–2434.

Hesse, S.J., Ruijter, G.J., Dijkema, C., Visser, J., 2002. Intracellular pH homeostasis in the filamentous fungus *Aspergillus niger*. Eur. J. Biochem. 269, 3485–3494.

Hiller, K., Schobert, M., Hundertmark, C., Jahn, D., Münch, R., 2003. JVirGel: calculation of virtual two-dimensional protein gels. Nucleic Acids Res. 31, 3862–3865.

Kelly, J.M., Hynes, M.J., 1985. Transformation of *Aspergillus niger* by the *amdS* gene of *Aspergillus nidulans*. EMBO J. 4, 475–479.

Kim, Y., Nadakumar, M.P., Marten, M.R., 2007. Proteomics of filamentous fungi. Trends Microbiol. 25, 395–400.

Kim, Y., Nadakumar, M.P., Marten, M.R., 2008. The state of proteome profiling in the fungal genus *Aspergillus*. Brief. Funct. Genomics Proteomics 7, 87–94.

Kincaid, M.M., Cooper, A.A., 2007. Misfolded proteins traffic from the endoplasmic reticulum (ER) due to ER export signals. Mol. Biol. Cell. 18, 455–463.

Kniemeyer, O., Lessing, F., Scheibner, O., Wertweck, C., Brakhage, A.A., 2006. Optimisation of a 2-D gel electrophoresis protocol for the human-pathogenic fungus *Aspergillus fumigatus*. Curr. Genet. 49, 178–189.

Levin, A.M., de Vries, R.P., Conesa, A., de Bekker, C., Talon, M., Menke, H.H., van Peij, N.N.M.E., Wösten, H.A., 2007. Spatial differentiation in the vegetative mycelium of *Aspergillus niger*. Eukaryot. Cell 6, 2311–2322.

Lu, B., Motoyama, A., Ruse, C., Venable, J., Yates, J.R., 2008. Improving protein identification sensitivity by combining MS and MS/MS information for shotgun proteomics using LTQ-Orbitrap high mass accuracy data. Anal. Chem. 80, 2018–2025.

MacKenzie, D.A., Guillemette, T., Al-Sheikh, H., Watson, A.J., Jeenes, D.J., Wongwathanarat, P., Dunn-Coleman, N.S., van Peij, N.N.M.E., Archer, D.B., 2005. UPR-independent dithiothreitol stress-induced genes in *Aspergillus niger*. Mol. Genet. Genomics 274, 410–418.

Maillet, I., Berndt, P., Malo, C., Rodriguez, S., Brunisholz, R.A., Pragai, Z., Arnold, S., Langen, H., Wyss, M., 2007. From the genome sequence to the proteome and back: evaluation of *E. coli* genome annotation with a 2-D gel-based proteomics approach. Proteomics 7, 1097–1106.

Meyer, V., Arentshorst, M., El-Ghezal, A., Drews, A.-C., Kooistra, R., van den Hondel, C.A.M.J.J., Ram, A.F.J., 2007. Highly efficient gene targeting in the *Aspergillus niger kusA* mutant. J. Biotechnol. 128, 770–775.

Mullally, J.E., Chernova, T., Wilkinson, K.D., 2006. Doa1 is a Cdc48 adapter that possesses a novel ubiquitin binding domain. Mol. Cell. Biol. 26, 822–830.

Nandakumar, M.P., Marten, M.R., 2002. Comparison of lysis methods and preparation protocols for one- and two-dimensional electrophoresis of *Aspergillus oryzae* intracellular proteins. Electrophoresis 23, 2216–2222.

Pel, H.J., de Winde, J.H., Archer, D.B., Dyer, P.S., Hofmann, G., Schaap, P.J., Turner, G., de Vries, R.P., Albang, R., Albermann, K., Andersen, M.R., Bendtsen, J.D., Benen, J.A., van den Berg, M., Breestraat, S., Caddick, M.X., Contreras, R., Cornell, M., Coutinho, P.M., Danchin, E.G., Debets, A.J., Dekker, P., van Dijck, P.W., van Dijk, A., Dijkhuizen, L., Driessen, A.J., d'Enfert, C., Geysens, S., Goosen, C., Groot, G.S., de Groot, P.W., Guillemette, T., Henrissat, B., Herweijer, M., van den Hombergh, J.P., van den Hondel, C.A., van der Heijden, R.T., van der Kaaij, R.M., Klis, F.M., Kools, H.J., Kubicek, C.P., van Kuyk, P.A., Lauber, J., Lu, X., van der Maarel, M.J., Meulenberg, R., Menke, H., Mortimer, M.A., Nielsen, J., Oliver, S., Olsthoorn, M.M.A., Pal, K., van Peij, N.N.M.E., Ram, A.F., Rinas, U., Roubos, J.A., Sagt, C.M., Schmoll, M., Sun, J., Ussery, D., Varga, J., Vervecken, W., van de Vondervoort, P.J., Wedler, H., Wosten, H.A., Zeng, A.P., van Ooyen, A.J., Visser, J., Stam, H., 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat. Biotechnol. 25, 221–231.

Punt, P.J., Veldhuisen, G., van den Hondel, C.A., 1994. Protein targeting and secretion in filamentous fungi. A progress report. Antonie Van Leeuwenhoek 65, 211–216.

Rabinovich, E., Kerem, A., Frohlich, K.U., Diamant, N., Bar-Nun, S., 2002. AAA-ATPase p97/Cdc48p, a cytosolic chaperone required for endoplasmic reticulum-associated protein degradation. Mol. Cell. Biol. 22, 626–634.

Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Güldener, U., Mannhaupt, G., Münsterkötter, M., Mewes, H.W., 2004. The

FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Res. 18, 5539–5545.

Sagt, C.M., Müller, W.H., van der Heide, L., Boonstra, J., Verkleij, A.J., Verrips, C.T., 2002. Impaired cutinase secretion in *Saccharomyces cerevisiae* induces irregular endoplasmic reticulum (ER) membrane proliferation, oxidative stress, and ER-associated degradation. Appl. Environ. Microbiol. 68, 2155–2160.

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

Schröder, M., Kaufman, R.J., 2005. ER stress and the unfolded protein response. Mutat. Res. 569, 29–63.

Shimizu, M., Wariishi, H., 2005. Development of a sample preparation method for fungal proteomics. FEMS Microbiol. Lett. 247, 17–22.

Sims, A.H., Gent, M.E., Lanthaler, K., Dunn-Coleman, N.S., Oliver, S.G., Robson, G.D., 2005. Transcriptome analysis of recombinant protein secretion by *Aspergillus nidulans* and the unfolded-protein response in vivo. Appl. Environ. Microbiol. 71, 2737–2747.

Sims, A.H., Robson, G.D., Hoyle, D.C., Oliver, S.G., Turner, G., Prade, R.A., Russell, H.H., Dunn-Coleman, N.S., Gent, M.E., 2004. Use of expressed sequence tag analysis and cDNA microarrays of the filamentous fungus *Aspergillus nidulans*. Fungal Genet. Biol. 41, 199–212.

Spurr, A.R., 1969. A low-viscosity epoxy resin embedding medium for electron microscopy. J. Ultrastruct. Res. 26, 31–43.

Sun, J., Lu, X., Rinas, U., Zeng, A.-P., 2007. Metabolic peculiarities of *Aspergillus niger* disclosed by comparative metabolic genomics. Genome Biol. 8, R182.

Tu, B.P., Weissman, J.S., 2002. The FAD- and $O_2$-dependent reaction cycle of Ero1-mediated oxidative protein folding in the endoplasmic reticulum. Mol. Cell 10, 983–994.

Venable, J.H., Coggeshall, R., 1965. A simplified lead citrate stain for use in electron microscopy. J. Cell Biol. 25, 407–408.

Van den Hondel, C.A.M.J.J., Punt, P.J., van Gorcom, R.F.M., 1991. Heterologous gene expression in fungi. In: Bennet, J.W., Lause, L.L. (Eds.), More Genetic Manipulations of Filamentous Fungi. Academic Press, Orlando, pp. 396–428.

Van der Vlies, D., Pap, E.H., Post, J.A., Celis, J.E., Wirtz, K.W., 2002. Endoplasmic reticulum resident proteins of normal human dermal fibroblasts are the major targets for oxidative stress induced by hydrogen peroxide. Biochem. J. 366, 825–830.

Wiebe, M.G., Karandikar, A., Robson, G.D., Trinci, A.P., Candia, J.L., Trappe, S., Wallis, G., Rinas, U., Derkx, P.M., Madrid, S.M., Sisiega, H., Faus, I., Montijn, R., van den Hondel, C.A., Punt, P.J., 2001. Production of tissue plasminogen activator (t-PA) in *Aspergillus niger*. Biotechn. Bioeng. 76, 164–174.

Wildgruber, R., Harder, A., Obermaier, C., Boguth, G., Weiss, W., Fey, S.J., Larsen, P.M., Görg, A., 2000. Toward higher resolution: two-dimensional electrophoresis of *Saccharomyces cerevisiae* proteins using overlapping narrow immobilized pH gradients. Electrophoresis 21, 2610–2616.

WO2008053018—Improved Production Of Secreted Proteins By Filamentous Fungi.

Yan, Q., Lennarz, W.J., 2002. Studies on the function of oligosaccharyl transferase subunits Stt3p is directly involved in the glycosylation process. J. Biol. Chem. 6 (277), 47692–47700.

Yuan, X.L., Roubos, J.A., van den Hondel, C.A., Ram, A.F., 2008. Identification of InuR, a new Zn(II)2Cys6 transcriptional activator involved in the regulation of inulinolytic genes in *Aspergillus niger*. Mol. Genet. Genomics 279, 11–26.

Zimmer, J.S., Monroe, M.E., Qian, W.J., Smith, R.D., 2006. Advances in proteomics data analysis and display using an accurate mass and time tag approach. Mass Spectrom. Rev. 25, 450482.