



Universiteit
Leiden
The Netherlands

Identification and characterization of starch and inulin modifying network of *Aspergillus niger* by functional genomics

Yuan, X.L.

Citation

Yuan, X. L. (2008, January 23). *Identification and characterization of starch and inulin modifying network of Aspergillus niger by functional genomics*. Institute of Biology Leiden (IBL), Group of Molecular Microbiology, Faculty of Science, Leiden University. Retrieved from <https://hdl.handle.net/1887/12572>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/12572>

Note: To cite this publication please use the final published version (if applicable).

A microscopic image of Aspergillus niger hyphae, showing several elongated, segmented cells with thick, dark outlines and granular internal structures. The cells are arranged in a branching pattern against a dark blue background.

Identification and Characterization of Starch and Inulin Modifying Network of *Aspergillus niger* by Functional Genomics

Starch and Inulin Modifying Network of *Aspergillus niger*

Xiao-Lian Yuan

Xiao-Lian Yuan

Identification and Characterization of Starch
and Inulin Modifying Network of *Aspergillus*
niger by Functional Genomics

Xiao-Lian Yuan

Identification and Characterization of Starch and Inulin Modifying Network of *Aspergillus niger* by Functional Genomics

Proefschrift

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus prof.mr. P.F. van der Heijden,

volgens besluit van het College voor Promoties

te verdedigen op 23 januari 2008

klokke 13:45 uur

door

Xiao-Lian Yuan

geboren te Feixi - Anhui, China

in 1974

Promotion Committee

Promotor:	Prof. Dr. C.A.M.J.J. van den Hondel
Co-promotor:	Dr. A.F.J. Ram
Referent:	Dr. P. J. Punt (TNO)
Other Members:	Prof. Dr. P.J.J. Hooykaas Prof. Dr. H.P. Spaink Prof. Dr. H.A.B. Wösten (Utrecht University) Prof. Dr. J.H. de Winde (Technical University of Delft) Prof. Dr. M.J.E.C. van der Maarel (Groningen University) Dr. A.A. van Dijk (DSM Food Specialties Delft)

Cover:	<i>Aspergillus niger</i> PinuE-racA(G12V) reporter strain grown in raffinose Designed by Xiao-Lian Yuan and Patrick Vermeij
Layout:	Patrick Vermeij
Printed By:	Printpartners Ipskamp B. V. Enschede
ISBN:	978-90-9022708-5

*For my parents,
Patrick and Jason*

献给我的父母
爱人和儿子

Contents

Chapter 1	General Introduction	9
Chapter 2	Construction of inulin and starch specific <i>Aspergillus niger</i> cDNA expression libraries	25
Chapter 3	<i>Aspergillus niger</i> genome wide analysis reveals a large number of novel alpha-glucan acting enzymes with unexpected expression profiles	39
Chapter 4	Characterization of two novel, putatively cell wall associated and GPI-anchored, α -glucanotransferase enzymes of <i>Aspergillus niger</i>	65
Chapter 5	Database mining and transcriptional analysis of genes encoding inulin modifying enzymes of <i>Aspergillus niger</i>	85
Chapter 6	Molecular and biochemical characterization of a novel intracellular invertase from <i>Aspergillus niger</i> with transfructosylating activity	105
Chapter 7	Identification of InuR, a new Zn(II)2Cys(6) transcriptional activator involved in the regulation of inulinolytic genes in <i>Aspergillus niger</i>	119
Chapter 8	A novel screening method for isolation of mutants involved in inulin signalling in <i>Aspergillus niger</i>	143
Summary		159
Samenvatting		163
摘要		167
References		171
Publications		185
Curriculum Vitae		187
Acknowledgements		189

Chapter 1

General Introduction

General introduction to filamentous fungi and *Aspergillus niger*

The kingdom of fungi contains an estimated 1.5 million species and only about 100,000 species have been described in detail (Hawksworth, 2001). The fungal kingdom includes both yeasts and filamentous (mycelial) fungi. Whereas yeasts are considered as unicellular microorganism that divide by budding (e.g. *Saccharomyces cerevisiae*) or fission (e.g. *Schizosaccharomyces pombe*), filamentous fungi are multicellular microorganisms that grow as long, thread-like cellular strands called hyphae that together form a fungal mycelium. Most fungi have a saprophytic lifestyle which means that they can utilize dead organic materials. Fungi are able to produce and secrete a wide variety of enzymes capable of breaking down complex organic molecules, liberating nutrient that are subsequently taken up by the fungus. *Aspergillus niger* is a saprophytic, ubiquitous filamentous fungus commonly isolated from soil, plant debris and indoor air environment (Raper and Fennell, 1955). *A. niger* is well known for its ability to grow on dead plant materials and its property to secrete high levels of enzymes that are able to break down plant cell walls (cellulose, hemicellulose and pectin) and plant storage polysaccharides (starch and inulin).

The genus *Aspergillus* includes over 185 species and was originally divided into 12 distinct groups (Raper and Fennell, 1965). *Aspergillus niger* was originally a group of black aspergilli according to the brown to black-shaded colour of the conidiospores (Raper and Fennell, 1965). In recent years, by using new molecular and biochemical techniques, the black aspergilli were reclassified into nine distinct black aspergilli species (*A. niger*, *A. carbonarius*, *A. japonicus*, *A. ellipticus*, *A. heteromorphus*, *A. tubingensis*, *A. foetidus*, *A. aculeatus* and *A. vadensis* (Megnegneau et al., 1993; Varga et al., 1994; Parenicová et al., 1997; de Vries et al., 2005a). The black aspergilli are one of the most important decomposers of dead plant matter in many ecosystems (Wainright, 1992). The production of valuable enzymes and bioproducts by aspergilli has brought up tremendous potential in food and feed industry as well as in basic science.

Aspergillus niger is an industrial important metabolite and enzyme producer

A. niger has become an industrially important organism since its use as a citric acid producer early in the 20th century (Schuster 2002). In addition to citric acid, *A. niger* is also a rich source of different hydrolytic and catalytic enzymes that are used in a variety of different (industrial) applications. For example, pectinases are used in the wine and fruit juice production (Grassin and Fanguenbergue 1999); proteases are used in food processing; amyloglucosidases are used in the starch industry (Frost and Moss 1987); hemicellulase are used in baking process; glucose oxidase and catalase are used in food processing and diagnosis (Berka et al., 1992). The good fermentation capabilities and high levels of protein secretion, together with the long history of safe use, make *A. niger* a very important organism in modern biotechnology (Archer and Peberdy 1997; Punt et al., 2002).

With the development of DNA-mediated transformation of aspergilli, *A. niger* has received an increasing interest as host not only for overproduction of homologous proteins (Van Gorcom et al., 1991; Van Hartingsveldt et al., 1993) but also for the overproduction of heterologous proteins (Davies, 1994; Verdoes et al., 1995; Archer and Peberdy 1997; Punt et al., 2002). Moreover, *A. niger* has gained the Generally Recognized As Safe (GRAS) status from the United States Food and Drug Administration (FDA) allowing its use in metabolite and enzyme production (Bigelis and Lasure, 1987; Schuster et al., 2002). In particular, the wide range of enzymes *A. niger* produces for the degradation of plant cell wall polysaccharides and plant storage polysaccharides are of major importance to the food and feed industry.

The *Aspergillus niger* genome

A. niger has become a industrially important host for the production of enzymes and metabolites. One of the Europe's leading bio-technology companies, DSM, has sequenced the *A. niger* genome. The *A. niger* genome is almost 3 times larger than the *S. cerevisiae* genome which was the first fungus with a sequenced genome. There are 14,165 ORFs encoded in approximately 33.9 Mb and 6,506 ORFs identified to have strong similarities to known function (Pel et al., 2007). The genome sequence provides a unique opportunity to further explore and exploit the industrially important organism for the discovery of new enzyme activities and functions and their interplay, with respect to carbohydrate modification and degradation. In addition to the genome sequence, Affymetrix microarrays are available via DSM that can be used to unravel the complex relationships between processes in living cells and their environment at a transcriptome level.

Plant storage carbohydrates

Carbohydrates are widely distributed on Earth, occur in many different forms, and are present in a wide variety of substances and materials. Most carbohydrates are synthesized by plants during the process of photosynthesis and can be divided into structural components and storage components. Structural carbohydrates are found as components of the plant cell wall. Non-structural carbohydrates are known as energy-rich compounds used for plant metabolism and energy storage compounds. The research described in this thesis focuses on two plant storage polysaccharides, starch and inulin.

Starch

Starch is one of the most abundant carbohydrates in nature and synthesized naturally in a variety of plants. Starch is synthesized in plastids found in leaves as a storage compound for respiration during dark periods. It is also synthesized in amyloplasts found in tubers, seeds and roots as a long-term storage compound. In these latter organelles, large amounts of starch accumulate as water-insoluble granules (Robyt, 1998; van der Maarel et al., 2002).

Some plant examples with high starch content are potato, corn, rice, sorghum, wheat, and cassava. There are two types of starch molecules in nature: one is amylose, an unbranched, single chain polymer of 200 to 6000 glucose units with only the α -1,4- glucosidic bonds (Fig. 1). The other type of polymer is amylopectin. Amylopectin also consists of a backbone of glucose units that are linked to each other by α -1,4- glucosidic bonds but also contains α -1,6- glucosidic branches, at once every 12-25 glucose residues along the backbone (Fig. 1). The degree of branching and the side chain length vary from source to source, but in general the more the chains are branched, the more water-soluble the starch becomes. A complete amylopectin molecule contains on average about 2.000.000 glucose units, thereby being one of the largest molecules in nature (Myers et al., 2000). The hydrolyzed products such as maltodextrin, maltose, or glucose and fructose syrups have various applications in starch processing industry.

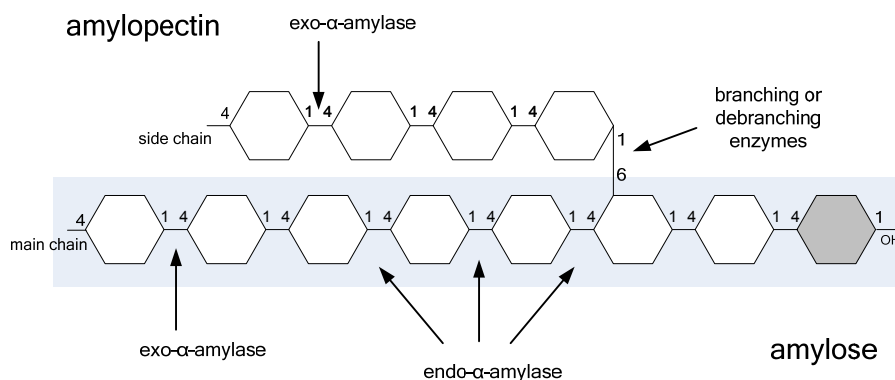


Fig. 1. Schematic presentation of the structure of amylopectin and amylose (shaded part only). Activities of fungal starch modifying enzymes are indicated. The filled box symbolizes the reducing end of starch.

Biodegradation and modification of starch

The storage polysaccharide starch is not only used by plants that have produced it, it is also an important nutrient source for many microorganisms (both bacteria and fungi) as well as for higher organisms, including humans. Both microorganisms and higher eukaryotes have a great diversity of enzymes able to hydrolyze starch (Steup 1988). Four groups of enzymes have been defined that are involved in starch degradation or modification and include endo-amylases, exo-amylases, debranching enzymes and transferases (van der Maarel et al., 2002). However, only three groups of enzymes, endo-amylases, exo-amylases and branching enzymes (transferases) have been characterized and described in *Aspergillus* species (Table 1).

Enzymes involved in starch degradation in *Aspergillus*

α -amylases (EC 3.2.1.1), belonging to endo-amylases, are able to cleave internal α -1,4-glycosidic bonds in amylose or amylopectin chain, but not the α -1,6-glycosidic linkages at the branching sides in amylopectin (Fig. 1). The end products of α -amylase action are oligosaccharides with varying length with an α -configuration and α -limit dextrins, which constitute branched oligosaccharides. The branched oligosaccharides can then be further hydrolyzed into glucose by exo-amylases such as α -glucosidases (EC 3.2.1.20) or α -glucoamylases (EC 3.2.1.3). Exo-amylases are able to cleave both α -1,4- and α -1,6-glycosidic bonds and act on the external glucose residues of amylose and amylopectin and thus produce only glucose (Fig. 1). Glucoamylase and α -glucosidase differ in their substrate preference: α -glucosidase acts best on short malto-oligosaccharides from non-reducing ends of the chains and liberates glucose with an α -configuration while glucoamylase hydrolyzes long-chain polysaccharides best from non-reducing ends of the chains with release of β -D-glucose (van der Maarel et al., 2002). Genes encoding α -amylases, glucoamylase and α -glucosidase from different *Aspergillus* species have been cloned and characterized (Table 1). Based on the derived amino acid sequence the gene products have been assigned to the glycosyl hydrolase family 13, 15, and 31 respectively (Henrissat 1991; Henrissat and Bairoch, 1993) as described at URL: <http://afmb.cnrs-rs.fr/CAZY/index.html> (Table 1).

Table 1. Description of starch degrading or modifying enzymes in aspergilli.

Enzyme (gene)	EC no.	GH Family ¹	Organism	Access no.	Reference
α -amylase (<i>amyl I</i>)	3.2.1.1	GH-13	<i>A. awamori</i>	BAD06002	Matsubara et al., 2004
α -amylase (<i>amyl III</i>)	3.2.1.1	GH-13	<i>A. awamori</i>	BAD06003	Matsubara et al., 2004
α -amylase (<i>amyl I</i>)	3.2.1.1	GH-13	<i>A. flavus</i>	AAF14264	Fakhoury and Woloshuk 1999
α -amylase	3.2.1.1	GH-13	<i>A. kawachi</i>	BAA22993	Kaneko et al., 1996
α -amylase (<i>amylA</i>)	3.2.1.1	GH-13	<i>A. niger</i>	CAA36966	Korman et al., 1990
α -amylase (<i>amylB</i>)	3.2.1.1	GH-13	<i>A. niger</i>	CAA36967	Korman et al., 1990
Acid α -amylase	3.2.1.1	GH-13	<i>A. niger</i>	P56271	Boel et al., 1990
α -amylase (<i>taa2</i>)	3.2.1.1	GH-13	<i>A. oryzae DSM63303</i>	BAA00336	Tada et al., 1989
α -amylase	3.2.1.1	GH-13	<i>A. shirousami</i>	BAA01255	Shibuya et al., 1992
glucoamylase (<i>gal</i>)	3.2.1.3	GH-15	<i>A. awamori</i>	BAD06004	Matsubara et al., 2004
glucoamylase	3.2.1.3	GH-15	<i>A. kawachi</i>	BAA00331	Hayashida et al., 1989
glucoamylase (<i>glaA</i>)	3.2.1.3	GH-15	<i>A. niger</i>	AAB59296	Nunberg et al., 1984
glucoamylase (<i>glaB</i>)	3.2.1.3	GH-15	<i>A. oryzae O-1013</i>	BAA25205	Hata et al., 1997
glucoamylase (<i>glaA</i>)	3.2.1.3	GH-15	<i>A. oryzae RIB 40</i>	AAB20818	Hata et al., 1991
glucoamylase (<i>glaB</i>)	3.2.1.3	GH-15	<i>A. oryzae RIB 40</i>	BAE57516	Machida et al., 2005
glucoamylase	3.2.1.3	GH-15	<i>A. shirousami</i>	BAA01254	Shibuya et al., 1990
Glucoamylase (<i>gla1</i>)	3.2.1.3	GH-15	<i>A. terreus</i>	L15383	Ventura et al., 1995
α -glucosidase (<i>aglA</i>)	3.2.1.20	GH-31	<i>A. niger</i>	AAB23581	Kimura et al., 1992
α -glucosidase (<i>agdA</i>)	3.2.1.20	GH-31	<i>A. oryzae RIB 40</i>	BAA08125	Minetoki et al., 1995
glycogen branching enzyme (<i>gbeA</i>)	2.4.1.18	GH-13	<i>A. oryzae RIB 40</i>	BAB69770	Sasangka et al., 2002
glycogen branching enzyme (<i>gbeA</i>)	2.4.1.18	GH-13	<i>A. nidulans</i>	BAA78714	Sasangka et al., 2002
Glycogen debranching enzyme (<i>gdb1</i>)	2.4.1.25 3.2.1.33	GH-13	<i>S. cerevisiae</i>	Q06625	Teste et al., 2000

¹ GH Family = Glycosyl hydrolase family

Enzymes involved in glycogen synthesis and degradation in *A. niger*

A. niger is able to store glucose as glycogen (Mattey and Allan, 1990). Glycogen is an analogue of starch, having a similar structure to amylopectin but with highly branched glucose polymer. The genes involved in the initial steps in glycogen synthesis have not been studied in *A. niger*. The final step of glycogen synthesis is catalysed by a glycogen branching enzyme (EC 2.4.1.18). The enzyme belongs to the class of transferases, which cleaves α -1,4-glycosidic bonds of amylose and amylopectin, and transfer the newly formed amylose chains to other α -1,4-linked chains, leading to formation of α -1-6 linked branches (Fig. 1). The genes encoding glycogen or starch branching enzymes have been identified in *A. oryzae* and *A. nidulans* (Table 1, Sasangka et al., 2002). Also in the *A. niger* genome, the gene encoding a putative glycogen branching enzyme (An14g04190) has been identified (Pel et al., 2007).

An enzyme supposed to be involved in the degradation of glycogen is a glycogen debranching enzyme (Fig. 1). This enzyme possesses two different catalytic activities (an oligo-1,4 $\rightarrow$ 1,4-glucanotransferase activity (EC 2.4.1.25) and an α -1,6-glucosidase activity (EC 3.2.1.33) on a single polypeptide chain. The enzyme eliminates the branch point in a two-step process that includes (i) the transfer of a maltotriosyl (or maltosyl) unit from the branch to an adjacent α -1,4-glucosyl chain by its oligo-1,4 $\rightarrow$ 1,4-glucanotransferase activity and (ii) the subsequent hydrolysis of the residual α -1,6-linked glucose by its α -1,6-glucosidase. The gene encoding the debranching enzyme has been identified and characterized in *S. cerevisiae* (Teste et al., 2000), but not yet in aspergilli. In *A. niger* a candidate glycogen debranching enzyme (An01g06120) has been identified (Pel et al., 2007). According to the classification of Henrissat, both starch or glycogen branching or debranching enzymes have been assigned to the glycosyl hydrolase family 13 (Table 1).

Inulin (fructan)

Most plants use starch as a storage polysaccharide. Fructans are the second most abundant storage carbohydrate that can be found in plants. About 15 % of flowering plant species store a proportion of their carbon as polymers of fructose (fructans) (Pollock and Chatterton, 1988; Hendry 1993).

There are three types of fructans present in nature. The first group are inulins, named after the genus (*Inula helenium*) from which inulin was first isolated by Rose in 1804. Inulin is composed of a linear fructose polymer linked by β -2,1-glycosidic bonds (Waterhouse and Chatterton, 1993) (see Fig. 2). A starting glucose moiety in the chain can be present but not necessary (Franck, 2002). These fructans are synthesized via fructosyltransferases using sucrose as the initial fructosyl acceptor and upon chain elongation as the fructosyl donor (see review Gupta and Kaur, 2000). All fructans found in the dicotyledons, as well as some monocotyledons are of this type.

The second type of fructans is the levan-type, which are also linear fructans, but in which the fructose units are (mostly) linked via a β -2, 6- glycosidic bonds (Suzuki and

Pollock, 1986). This type of fructan is found in a large part of the monocotyledons and in almost all bacterial fructans. The third type is the fructans of the mixed type, which are also referred to as the graminan type (Carpita et al., 1989). These fructans have both β -2, 1- and β -2, 6- glycosidic linkage bonds between the fructose units, and thus contain branches. These graminan type fructans are only found in grasses. Plant fructans are naturally present in the vacuole (Darwen and John, 1989). A single fructan molecule contains up to 200 fructose units in plants while polymers up to 100,000 fructose units are found in bacteria.

Inulin has been part of man's daily diet for many centuries. Inulin-containing plants that are used for human nutrition belong mainly to either the Liliacea, e.g., leek, onion, garlic and asparagus, or the Compositae, e.g., Jerusalem artichoke, dahlia (eaten in older days), chicory and yacon (Van Loo et al., 1995; Franck, 2002). In the food industry, inulin is obtained from chicory root and is used as a functional food ingredient.

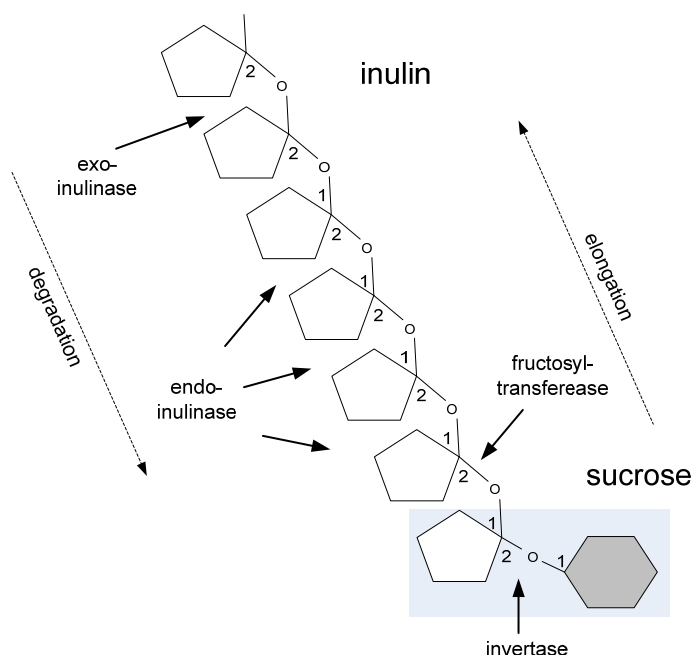


Fig. 2. Schematic presentation of the structure of inulin and sucrose (shaded part only). Activities of fungal inulin modifying enzymes are indicated. The filled box symbolizes the glucose residue attached to the reducing end of inulin.

Inulin and its partially hydrolyzed products (oligofructoses) may significantly improve organoleptic characteristics such as upgrading of both taste and mouthfeel in a wide range of food applications (Niness, 1999, Coussemont, 1999). In addition, inulin is used as fat and carbohydrate replacement and offers the advantage of not compromising on taste and texture, while delivering nutritionally enhanced products (Franck 2002). Furthermore, the complete hydrolyzed product, fructose, has twice the sweetness of normal table sugar.

Biodegradation and modification of inulin

Several fungal enzymes have been described that are involved in the degradation or modification of inulin and include exo-inulinases, endo-inulinases, invertases, and fructosyl-transferases (Fig. 2). Exo-inulinase (β -D-fructan fructohydrolase, EC 3.2.1.80), hydrolyzes the terminal β -2,1-fructofurosodic bonds in inulin to produce mainly fructose and partially glucose; endoinulinase (β -D-fructan fructanohydrolase, EC 3.2.1.7), which is specific for inulin and hydrolyzes the internal linkages in inulin to release inulotriose, -tetraose and -pentaose as the main products. A third enzyme, invertase (β -D-fructofuranoside fructohydrolase, EC 3.2.1.26) hydrolyzes β -2,1- fructosidic bond from sucrose to fructose and glucose. Because of the overlapping substrate specificity of invertases and exo-inulinases, the enzymes are further classified by the substrates hydrolytic ratio. When the sucrose to inulin hydrolytic ratio (S/I ratio) is 1.5-20, it is generally considered as exo-inulinase, otherwise the enzyme is considered as an invertase (Vandamme and Derycke, 1983). β -Fructosyltransferases (EC 2.4.1.99, 2.4.1.9) transfers fructose residues from the non-reducing terminal β -2,1- fructofurosodic bonds in sucrose or inulin to another sucrose or inulin molecule to form kestose or higher fructooligosaccharides. The difference between these two enzymes EC 2.4.1.9 and EC 2.4.1.99 is clarified by the products: when the predominant product is 1-kestose the enzyme is classified as EC 2.4.1.99, when the predominant products are higher fructooligosaccharides the enzyme is classified as EC 2.4.1.9. All enzymes with above activities are considered in this study as inulin modifying enzymes (IMEs).

Transcriptional regulation of genes encoding extracellular enzymes

The expression of extracellular enzymes in *A. niger* is highly regulated and controlled. The control of expression is mediated via a complex system of transcription factors that either activate or repress transcription in response to environmental clues such as the availability of certain carbon- or nitrogen-sources or extracellular pH. In this section, we will discuss some of the most important transcription factors involved in the regulation of extracellular enzymes in fungi.

General introduction to transcription factors

Transcription factors are proteins involved in the regulation of gene expression by binding to the promoter elements upstream of genes to either activate or repress transcription. Transcription factors contain two essential functional regions: a DNA-binding domain and an activator domain. The DNA-binding domain is characterized by a particular protein sequence motif that recognizes and binds to a specific DNA sequence near the start of transcription. Binding of the transcription factor to the promoter sequence either activates or represses RNA polymerase binding thereby induce or repress gene expression. In addition, transcription factors often contain a second domain, the activation domain. The

activation domain of transcription factors is used to regulate the activity of the transcription factor in response to internal or external signals (Stryer, 1999).

Transcription factors are typically classified according to the structure of its DNA binding domain. Transcription factors with various DNA binding motifs include the homeodomain, helix-turn-helix (HTH), leucine zipper, basic helix-loop-helix (bHLH), high mobility group (HMG) box, basic region-leucine zipper (bZIP), MADS box, ATTS domain and zinc fingers transcription factors (Pel et al., 2007 and reference herein).

DNA binding domains of a particular class of transcription factors generally bind similar to DNA target sequences (Suzuki and Yagi, 1994). The functional conservation of DNA binding motifs and their cognate DNA binding sites are often exploited to indicate functions of unknown DNA binding proteins identified from genome sequence data by comparisons with proteins of known function (Todd and Andrianopoulos, 1997).

In this thesis, we focus on Zinc(II) coordinating transcription factors of the C₂H₂, C₂X₁₇C₂ (GATA), and Zn(II)₂Cys₆ type. A zinc cluster is one of the most abundant DNA binding motifs in eukaryotes (Stryer, 1999). The DNA binding domain consists of two anti-parallel β sheets and one α helix to which a zinc ion is bound. The zinc ion is crucial for the stability of this domain type - in absence of the metal ion the domain unfolds. As we will see, representative members of these classes of transcription factors are important transcriptional regulators of genes encoding extracellular enzymes in fungi.

Cys₂His₂ zinc finger motif

The Cys₂His₂ zinc finger motif was originally found and described in the TFIIIA transcription factor, which activates the transcription of the 5S rRNA genes (Miller et al., 1985). The name of the motif is derived from their structure, in which a small group of conserved amino acids binds a zinc ion, giving a finger-like structure. The conserved two cysteine and two histidine residues were proposed to form a tetrahedral coordination complex with each zinc atom to generate peptide domains- zinc fingers- that interact with DNA (Fig. 3A; Vallee et al., 1991). The consensus sequence of a single finger is: Cys-X₂₋₄-Cys-X₃-Phe-X₃-Leu-X₂-His-X₃-His (Guasconi et al., 2002). Two examples of fungal Cys₂His₂ zinc finger transcription factors, CreA and PacC are discussed.

Carbon catabolite repressor (CreA)

The well studied C₂H₂-type zinc finger protein in aspergilli is CreA, encoding the carbon catabolite repressor. Carbon catabolite repression is a global regulatory mechanism in which the presence of glucose or other preferred carbon source represses expressions of genes involved in the utilization of less-favoured carbon sources. Filamentous fungi are characterized by their capacity to metabolise a variety of carbon sources, such as plant cell wall materials e.g. cellulose, xylan, pectins and plant storage polysaccharides e.g. starch and fructan, by secreting high amount of enzymes capable of braking down the complex substrates. Most enzyme systems involved in degradation of these polysaccharides are

repressed by glucose, a preferable carbon source. The major system responsible for carbon repression in aspergilli is mediated via the carbon catabolite repressor protein CreA (Drysdales et al., 1993; Ruijter and Visser, 1997). The CreA protein has two zinc finger structures of the C₂H₂ type, which binds to specific sites (5'-SYGGRG-3') in the promoters of its target genes (Kulmburg et al., 1993; Cubero and Scazzocchio, 1994). In the presence of preferred substrates, such as glucose, CreA represses the expression of target genes. Several *creA* mutants have been isolated from *A. nidulans* and *A. niger*, displaying (partially) derepressed phenotypes (Shroff et al., 1997; Ruijter and Visser, 1997). CreA is a global carbon catabolite repressor, not only repressing structural genes encoding extracellular enzymes, but also represses other genes encoding transcriptional activators e.g. AlcR (Mathieu et al., 2005, see below).

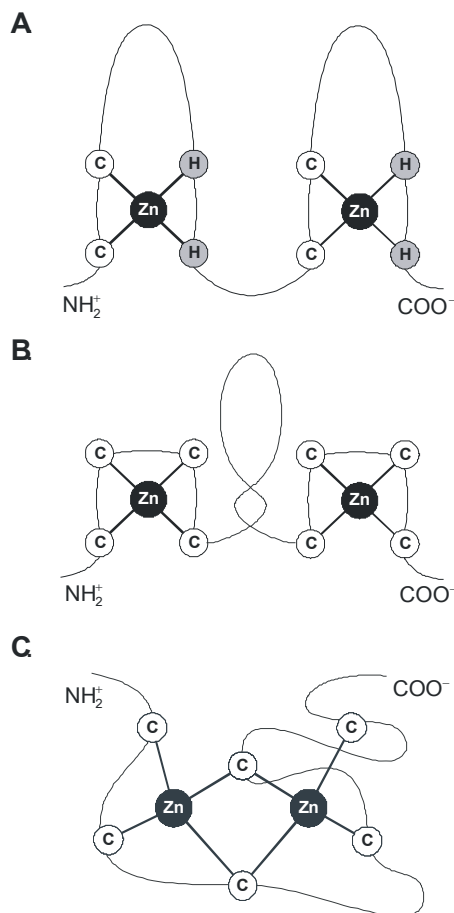


Fig. 3. Schematic presentation of DNA-binding domains of zinc coordinating transcription factors. A, Cys₂His₂ zinc finger motif; B, Cys₄ zinc finger motif; C, Zn(II)₂Cys₆ binuclear cluster motif. (This figure is modified from Vallee et al., 1991)

pH regulation(PacC)

Another well studied C₂H₂ zinc finger protein is *A. nidulans* PacC, a transcription factor mediating gene expression in relation to the extracellular pH regulation. PacC protein contains three C₂H₂ zinc fingers in the N-terminal region. The target sequence is 5'-GCCARG with a preference for T at the -1 position (Espeso et al., 1997; Penalva and Arst, 2002). The active form of PacC is a transcriptional repressor of acid-expressed genes and a transcriptional activator of alkaline expressed genes. The *pacC* gene itself is preferentially expressed at alkaline pH (Tilburn et al., 1995). The complex molecular mechanism of PacC activation has been reviewed extensively and the reader is referred to these reviews for further details (Penalva and Arst, 2002 and 2004).

Cys4 zinc finger motif (GATA type)

The GATA-type of transcription factor was first identified as a protein involved in erythroid-specific gene expression in vertebrates. This protein specifically binds to (A/T)GATA(A/G) sequence in the regulatory regions of target genes. The conserved core in the DNA binding motif gives the name to this class of transcription factors (GATA) (Yamamoto et al., 1990). The DNA-protein interaction occurs via highly conserved zinc finger domains in which the zinc ion is coordinated by four cysteine residues (Fig. 3B; Vallee et al., 1991; Omichinski et al., 1993). The consensus binding motif is Cys-X₂-Cys-X₁₇₋₁₈-Cys-X₂-Cys (Gronenborn, 2005). GATA factors in animals contain two zinc finger domains, while the majority of known fungal GATA factors contain a single Cys4 zinc finger which has the greatest similarity to the carboxyl (C) terminal finger of animal GATA factors. The consensus binding motif in fungi is in most cases Cys-X₂-Cys-X₁₈-Cys-X₂-Cys (Wilson and Arst, 1998; Gronenborn, 2005). The fungal GATA type transcription factors are involved in various cellular processes including metabolism, siderophore production, photoinduction and mating type switching (Scazzocchio, 2000). The well studied example, AreA protein involved in nitrogen metabolism is discussed here (Wilson and Arst, 1998).

Nitrogen catabolite activator (AreA)

AreA is a wide-domain transcription factor involved in nitrogen source utilization. This transcription activator is required for expression of the enzymes for utilization of secondary nitrogen sources e.g. nitrate, purines or amino acids, other than the preferred primary compounds e.g. ammonium, glutamate and glutamine. AreA contains only one single zinc finger DNA-binding domain in the C-terminal region and binds specifically to the sequence, HGATAR (Merika and Orkin 1993; Ravagnani et al., 1997) in target genes. The function of AreA is often through the interaction with other pathway specific regulatory proteins for the co-ordinated regulation of specific subsets of genes. For example, in nitrate assimilation pathway, AreA interacts with the pathway activator NirA and forms a transcriptional complex to regulate the nitrate pathway gene expression (see review Krappmann and Braus, 2005). The pathway inducer nitrate and AreA is required for NirA

mediated gene expression (Narendja et al., 2002). The AreA regulates localization and binding site occupancy of nitrate activator NirA (Berger et al., 2006).

Zn(II)₂Cys₆ binuclear cluster motif

The DNA binding motif Zn(II)₂Cys₆ type was first characterized in the *Saccharomyces cerevisiae* Gal4 protein, which is required for transcriptional activation of the genes coding for galactose metabolizing enzymes (Laughon and Gesteland, 1982; Johnston, 1987). Zn(II)₂Cys₆ binuclear cluster proteins are exclusively found in fungi (Schjerling and Holmberg, 1996; Borkovich et al., 2004) and therefore they are classified as typically fungal specific transcriptional factors (Todd and Andrianopoulos, 1997). The binding domain of this class contains six cysteine residues that are conserved. The consensus binding motif is Cys-X₂-Cys-X₆-Cys-X₍₅₋₁₆₎-Cys-X₂-Cys-X₍₆₋₈₎-Cys, which coordinate two zinc(II) ions to form zinc cluster, where the cysteine residues are known to be the only zinc ligands (Fig. 3C; Vallee et al., 1991; Todd and Andrianopoulos, 1997).

Zn(II)₂Cys₆ transcription factors contains three functional domains: a C₆ zinc cluster, usually at the N-terminal end of the protein, which is involved in DNA binding; a middle homology region (MHR), which is thought to be necessary for assisting the Cys₆ zinc cluster in DNA binding specificity; and a third, less well understood activation domain (Schjerling and Holmberg, 1996; Borkovich et al., 2004).

Several fungal transcriptional factors characterized as Zn(II)₂Cys₆ type are involved in carbon catabolite metabolism and the GAL4, AlcR and AmyR, transcription factors will be discussed in further detail.

The galactose utilization activator (Gal4)

The *S. cerevisiae* Gal4 protein is the first identified and the most extensively studied member of Zn(II)₂Cys₆ binuclear cluster DNA binding protein. Gal4p is required for the transcriptional activation of the genes coding for galactose metabolizing enzymes (Laughon and Gesteland, 1982; Johnston, 1987). Gal4p contains a DNA binding and dimerization domain and an N-terminal activating domain at N-terminus, and an activation domain at C-terminus (Ma and Ptashne, 1987). The Gal4 DNA binding domain localize at N-terminus and contains the 6 cysteine residues that coordinate binding of the two Zn(II) ions. This domain binds DNA as a dimer and each Gal4 monomer contacts one half of the dyad recognition site (Carey et al., 1989). The protein recognizes sites with two CGG triplets on opposite strands with spacing 11 nts in its cognate genes (CGG_NCCG) (Marmorstein et al., 1992).

In galactose metabolism pathway, a negative regulatory protein Gal80, and a galactose and ATP activated protein, Gal3, have been shown to associate with Gal4 to regulate Gal4 activity. In the absence of galactose, Gal80 interacts with the Gal4 protein transcriptional activation domain to prevent it from activating the galactose-inducible genes (Pilauri et al., 2005; Diep et al., 2006). In the presence of inducer (galactose or a related

metabolite), Gal4 is released from this repression by the activation of Gal3, which binds to Gal80 and blocks the repressive function of Gal80. The released Gal4 protein activates transcription of the structure genes through binding to upstream activated sequences (Pilauri et al., 2005).

Alcohol utilization activator (AlcR)

AlcR is the best studied Zn₂Cys₆ type transcription factor in filamentous fungi. The AlcR protein is a transcriptional activator involved in the ethanol oxidation pathway in *A. nidulans*. It contains an asymmetric DNA-binding domain with an extra helix inserted between the third and fourth cysteine, and lacks the conserved proline residue (Kulmburg et al., 1991; Ascone et al., 1997). Like Gal4, the AlcR protein was shown to interact with the CGG triplet in its target genes. However, the consensus site is extended to 5'-T/AGCGG-3' both in vitro and *in vivo* (Panozzo et al., 1997). AlcR, is so far, the sole activator among the zinc binuclear cluster Zn₂Cys₆ type proteins that binds as a monomer to single sites (Nikolaev et al., 1999; Cahuzac et al., 2001).

Regulation of ethanol utilization pathway has been studied in detail in *A. nidulans*. It is clear that the regulation is controlled by two regulatory mechanisms and includes a specific activation of AlcR by inducer molecules and repression by the carbon catabolite repression mechanism, mediated by CreA repressor (Mathieu et al., 2000; Felenbok et al., 2001). Activation of AlcR requires the presence of an inducing compound in the cell and induces the transcription of its structural genes, *alcA* and *aldA* encoding enzymes involved in the utilization of ethanol (Lockington et al., 1987; Fillinger and Felenbok, 1996; Felenbok et al., 2001). Acetaldehyde is the sole physiological inducer, many compounds carrying a carbonyl function act as direct inducers of the alc system mediated by their break down product, aldehyde (Flipphi et al., 2002).

In addition, the *alcR* gene itself is subject to the autoregulatory circuits. AlcR consensus binding elements present in *alcR* promoter has been proved to be functional in vivo (Kulmburg et al., 1992; Mathieu et al., 2000). It has been shown that there is a close correlation between the level of expression of these *alc* genes and that of *alcR* (Fillinger and Felenbok, 1996; Panozzo et al., 1997). Therefore *alcR* positive autoregulation is an important mechanism increasing the expression of the responsive genes.

The CreA mediated repression of AlcR target genes is mediated at two levels: (i) by repressing directly *alcR* expression and thereby that of AlcR-target genes and by (ii) repressing directly the AlcR target genes (Mathieu et al., 2005).

The starch utilization activator (AmyR)

Aspergilli produce various starch-degrading enzymes (amylolytic enzymes), such as α -amylase, glucoamylase and α -glucosidase, and hydrolyze starch synergistically to produce maltose and/or glucose which can be taken up by the fungus using sugar transporters. A number of genes encoding these enzymes have been cloned and sequenced (Boel et al., 1984; Wirsal et al., 1989; Hata et al., 1992; Minetoki et al., 1995).

Regulation of amylolytic system in aspergilli has been studied extensively. In general, the expression of amylolytic genes is induced by starch and maltose (Fowler et al., 1990; Hata et al., 1992; 1993; Morkeberg et al., 1995). Moreover, α -linked glucosides such as maltooligosaccharides, kojibiose, nigerose, isomaltose, panose, and artificially modified maltosides such as α -phenyl maltoside and α -isoamyl maltoside also function as effective inducers for α -amylase synthesis in *A. oryzae* as well as *A. nidulans* (Tsukagoshi et al., 2001). Among them, isomaltose was found to be the most effective inducer for amylase synthesis in *A. nidulans* at an extremely low concentration of 3 μ M (Kato et al., 2002a). Isomaltose is known to be produced by transglycosylation activity of α -glucosidases, such as *A. nidulans* α -glucosidase B (AgdB; Kato et al., 2002b). Taken together, isomaltose may be the most probable candidate for the physiological inducer of the amylase synthesis in aspergilli (Tsukagoshi et al., 2001; Kato et al., 2002a).

Synthesis of the amylolytic enzymes is repressed by an excess of a preferred carbon sources such as glucose and xylose (Fowler et al., 1990, Hata et al., 1992; Morkeberg et al., 1995) and mediated by the CreA carbon catabolite repressor (Ruijter and Visser, 1997). Potential CreA binding site are found in promoters of a number of amylolytic genes. Thus, it is likely that repression of amylolytic genes in the presence of an excess of good carbon source is due to the action of CreA. Kato et al., (1996) have shown that the DNA binding domain of CreA can bind to the *A. oryzae* α -amylase promoter.

The gene encoding the transcriptional activator related to starch degradation, *amyR*, was first cloned independently in *A. oryzae* by Petersen et al. (1999) and Gomi et al. (2000) using two different strategies. Petersen et al. used an *amyB* promoter driving reporter gene to isolate negative regulatory mutants which fail to produce the reporter protein. The mutant was complemented by *A. oryzae* cosmid library to obtain the gene encoding the AmyR transcription factor. Gomi et al. cloned the *amyR* gene by screening for clones that suppress the titration phenomenon observed when multiple copies of an *amyB*-LacZ reporter protein were introduced in *A. nidulans*. Transformants were isolated in which the reduced expression of the LacZ gene was suppressed leading to the identification of *amyR*.

amyR genes have also been cloned in *A. nidulans* and *A. niger* by heterologous hybridization of the *A. oryzae amyR* gene (Tani et al., 2001). The AmyR protein showed very high similarity (more than 70% overall amino acid identity) among aspergilli (Tani et al., 2001). The AmyR proteins from aspergilli all contain the Zn(II)₂Cys₆ binuclear cluster binding motif at the N-terminus and the motif is completely conserved among the three AmyR proteins, which suggest that these AmyR proteins might recognize and bind the same DNA sequences in the promoter regions of target genes. Petersen et al (1999) reported that the AmyR recognizes two types of sequences in its cognate genes: one is two CGG direct repeat triplets separated by 8 nucleotides (CGG_NCGG); the other is CGG single triplets followed by AAATTTAA. The latter binding elements was later improved and corrected into CGG_NAGG (Ito et al., 2004). AmyR binding site in the *A. nidulans agdA* promoter indicates that AmyR can bind to a single CGG triplet as a monomer but requires the CGG direct repeat for high affinity binding and transcriptional activation (Tani et al., 2001). Furthermore, analysis of AmyR binding site in *A. oryzae* *taaG2* promoter indicate that

two AmyR molecules bind cooperatively to the promoter by recognizing the CGG triplet at the 5'-end and the AGG triplet at the 3'-end. The AGG can function as second binding site only in the presence of CGG (Ito et al., 2004). Mutation analysis of AGG to CGG indicated that CGGN₈CGG is preferable to CGGN₈AGG for AmyR binding (Ito et al., 2004).

Aim and outline of the thesis

A. niger is well known for secreting a wide variety of enzymes to break down plant cell wall materials and plant storage polysaccharides to serve as its energy and carbon sources. The break-down products have many applications in the food processing industry and are also major constituents of human and animal diets. The expectation before the *A. niger* genome sequencing project was that only a small fraction of potential enzymes produced by *A. niger* was exploited. Besides the full genome sequence of *A. niger*, we also made use of Affymetrix GeneChip arrays which allowed us to perform genome-wide expression analysis after growth on different carbon sources. In this thesis we focussed on 1) the identification and characterization of the structural genes involved in the degradation and modification of starch and inulin; 2) examining the genome-wide transcriptional responses in relation to growth on starch, inulin and related carbon sources using Affymetrix GeneChip arrays; 3) unravelling the corresponding pathway specific transcription factors and inducing or signalling molecules.

Chapter 1 reviews current knowledge on starch and inulin degradation or modification in aspergilli as well as their transcriptional regulation. The most important transcription factors involved in regulating the expression of the enzymes for degradation of sugar polymers are discussed in details.

To identify genes encoding enzymes with starch or inulin modifying activities in *A. niger*, one approach was to construct cDNA expression libraries. *A. niger* starch or inulin specific cDNA expression libraries were generated using Gateway cloning technology and characterized and evaluated in **Chapter 2**.

In **Chapter 3**, we described alternative way to identify starch modifying enzymes in *A. niger* by mining the genome sequence for the presence of GH 13, 15, 31 family members, to which potential starch degradation activities belong. Combining the analysis of the genome sequence with transcriptional expression profiles, we identified a large number of novel enzymes with unexpected expression profiles. Only two enzymes (AgdB and AmyC) were predicted to play a role in starch degradation. The potential roles of the other novel enzymes might be related to cell wall alpha-glucan synthesis and modification and are discussed.

Among them, three GH13 family enzymes (AgtA, AgtB and AgtC), sharing high similarity to the α -amylases, were predicted to be glycosylphosphatidylinosol-anchored and lacked some highly conserved amino acids of the α -amylase family. The overexpressed protein analysis and deletion analysis of *agtA*, *agtB* and *agtC* were studied in **Chapter 4**. We showed that AgtA and AgtB possess activity of 4- α -glucanotransferases (EC 2.4.1.25). The

knock-out of *agtA* has characteristics indicating a cell wall integrity defect. Taken together, we concluded that AgtA has a possible role in the synthesis and / or maintenance of the alpha-glucan in the fungal cell wall.

In **Chapter 5**, we surveyed the *A. niger* genome sequence for the presence of gene/enzyme network involved in inulin and sucrose metabolism. Two new intracellular proteins SucB and SucC were identified in addition to the three known extracellular enzymes InuE, SucA, InuA. Transcription analysis revealed that the extracellular inulinolytic genes were co-ordinately regulated and induced by inulin and sucrose. Further analysis identified sucrose, but not previously described fructose, as an efficient inducing molecule, involved in inulin degradation pathway.

To examine if SucB, one of the two novel intracellular invertases, play an essential role in generating pathway inducing molecules in inulin or sucrose catabolism, the phylogenetic, molecular and biochemical characteristics of SucB were studied and described in **Chapter 6**.

The expression of extracellular inulinolytic genes in *A. niger* is co-regulated and induced by sucrose and inulin. In **Chapter 7**, we described the identification and characterization of the pathway transcription activator, InuR.

In **Chapter 8**, we developed a positive screening method for the isolation of mutants involved in inulin signalling pathway, using pathway tightly regulated *inuE* promoter. The research described in the last four Chapters, allowed us to formulate working hypothesis of the inulin signalling pathway in *Aspergillus niger*, presented at the end of Chapter 8.

Chapter 2

Construction of inulin and starch specific *Aspergillus niger* cDNA expression libraries

Xiao-Lian Yuan, Mark Arentshorst, Cees A.M.J.J. van den Hondel and Arthur F.J. Ram

Abstract

To screen for genes encoding enzymes with starch or inulin modifying activities from *Aspergillus niger*, cDNA expression libraries were generated using Gateway cloning technology. RNA was isolated from different time points during submerged growth in cultures containing starch or inulin as carbon sources. RNA from the different time points (16, 32, 48 and 64 h after inoculation) of inulin or starch grown cultures were isolated and equally pooled for cDNA synthesis. The synthesized cDNAs were cloned into both *Escherichia coli* expression vector (pSPORT1) and *Saccharomyces cerevisiae* expression vector (pDEST-YES52). Analysis of 24 randomly picked clones from the inulin and starch specific *E. coli* cDNA expression libraries showed that 95% of clones harbor cDNA inserts with average size of 1.4 kb for inulin libraries and 100% of clones contain cDNA inserts with average size of 1.8 kb for starch libraries, respectively. For construction of the *S. cerevisiae* expression libraries the construction procedure was more complex. First, cDNA fragments were cloned into a mammalian expression vector (pCMV-SPORT6) and the resulting inserts from the primary libraries were transferred to pDONR201 to give so-called entry libraries. The inserts from the starch or inulin entry libraries were transferred to the destination vector (pDEST-YES52), to give the *S. cerevisiae* expression libraries. Analysis of the yeast expression libraries showed that 33% of the ampicillin resistant cDNA clones from the starch library and 54% of clones from the inulin library contained a cDNA in the yeast expression vector. The remaining clones were from previous libraries, either primary vector pCMV-SPORT6 or entry vector pDONR201 or a combination of the two. To correct for the relatively low percentage of *E. coli* clones harboring the right expression vector, a relative large number of primary transformants were produced to have enough correct yeast expression clones for screening. As only cDNAs clones inserted in pYES-DEST52 will be transformed to *S. cerevisiae* the presence of other plasmids in the pool of clones will not hamper the screening procedure. Analysis of size of the inserts showed, in general, a reduction of the average length of the insert in subsequent libraries. The average length of the inserts in the yeast expression library for inulin was 1.2 kb and 1.6 kb for starch. This suggested that the construction of cDNA libraries can be improved by reducing cDNA transfer steps, e.g. by insertion of cDNAs directly into an entry vector or by insertion cDNAs directly into a *S. cerevisiae* expression vector.

Introduction

Aspergillus niger are distributed worldwide and commonly present on decaying plant debris. As a saprophytic fungus, *A. niger* produces a variety of hydrolytic enzymes that are able to break down the plant polysaccharides into smaller molecules, which can be served as their nutrient source. Many enzymes secreted by *A. niger* have already found their applications in the baking, starch, textile and food and feed industries (Schuster et al., 2002; De Vries et al., 2001; Semova et al., 2006).

Enzymes secreted by *A. niger* that are involved in starch or inulin catabolism have been previously described. The known starch degrading enzymes include glucoamylase

(*glaA*) (Nunberg et al., 1984), acid alpha-amylase (*aamA*) (Boel et al., 1990; Brady et al., 1991), alpha-amylases (*amyA* and *amyB*) (Korman et al., 1990) and alpha-glucosidase (*aglA*) (Kimura et al., 1992). Well described inulin degrading enzymes consist of invertase (*sucA*) (Bergès et al 1993, Boddy et al., 1993; L'Hocine et al., 2000), endo-inulinase (*inuA* and *inuB*) (Ohta et al., 1998; Aikimoto et al., 1999), and exo-inulinase (*inuE*) (Arand et al., 2002; Kulminskaya et al., 2003; Moriyama et al., 2003). Since *A. niger* grows well on starch and inulin as carbon and energy sources, it is predicted that a full set of enzymes converting starch into oligosaccharides, maltose, glucose or enzymes converting inulin into oligofructose, sucrose, glucose and fructose are present in this organism.

To identify additional genes encoding enzymes with starch or inulin modifying activities, one approach is to use the deduced amino acids sequences of known starch or inulin degrading enzymes and to perform Blast or HMM searches against the *A. niger* genome to identify related proteins. This approach has been successfully used to predict or identify additional enzymes involved in starch or inulin catabolism (chapters 3 and chapter 7). The disadvantage of this approach is that completely new starch or inulin modifying enzymes can be missed if these enzymes may have amino acid sequences which differ from the known enzymes. Consequently, those proteins are not recognized in Blast or HMM searches. In addition, site activities of enzymes from other Glycosyl Hydrolyases (GH) families might be missed by performing only the bioinformatics based genome mining approach. An alternative approach to identify starch or inulin modifying enzymes is to construct cDNA expression libraries in combination with a High Throughput Screening (HTS) using appropriate substrates. The screening of cDNA libraries for interesting enzyme activities has been proven to be a powerful and efficient method for investigation of the function of gene products (van der Vlugt-Bergmans and van Ooyen, 1999; Meeuwsen et al., 2000).

Construction of high-quality (full length) cDNA libraries in proper expression vectors is essential for successful screening and further facilitates the cDNA product characterization. Genetic manipulation of *E. coli* colonies are easy to handle in HTS and large amount of recombinant proteins can be expressed in a short time. However, the expression of eukaryotic proteins in *E. coli* can be problematic, due to aggregation, formation of insoluble inclusion bodies, and/or degradation of the expression product (Hannig and Makrides, 1998; Baneyx, 1999). Eukaryotic hosts e.g. yeast species (*Pichia pastoris* and *S. cerevisiae*) or filamentous fungi (e.g. *A. niger*) as an alternative expression system, provide the specific cellular environment for the expression of eukaryotic proteins. However, possible bottlenecks in the use of *P. pastoris* or *S. cerevisiae* as host for the expression of *Aspergillus* cDNAs include that expression could suffer from lower yields of heterologous proteins (Buckholz and Gleeson, 1991; Punt et al., 2002; Holz et al., 2003) and that these organisms are more laborious in HTS compared to *E. coli*.

Gateway Cloning Technology is a universal system for cloning and subcloning DNA/cDNA fragments into many expression vectors (Ohara and Temple, 2001). This technology uses the λ -recombination system transfer cDNA fragments between vectors that

contain recombination sites for the λ -recombinase machinery while maintaining reading frame and orientation. This approach is characterized by several advantages over conventional restriction-assisted cloning methods which include the facts that, (i) it eliminates restriction digestion for directional cloning, ii) it generates lower level of chimeric clones, iii) produces higher amount of full-length cDNA clones and iv) gives higher cloning efficiency (Hartley et al., 2000; Ohara and Temple, 2001). Moreover, this approach facilitates the construction of same cDNAs input into different expression vectors, especially when it is not clear which system or host background will provide sufficient expression levels to allow interesting protein purification. Thus, this approach reduces enormous amount of work and further prevents the significant barrier to progress.

In this paper, we have used the Gateway cloning technology to construct *A. niger* cDNA expression libraries to screen for starch and inulin degrading activities. Molecular characterization of the *E. coli* and *S. cerevisiae* libraries indicated that useful expression libraries have been obtained. Approaches that might further improve the construction of cDNA expression libraries as well as the screening are discussed.

Materials and Methods

Strains and culture conditions

A. niger strain N402 used in this study was derived from the wild-type strain *A. niger* van Tieghem (CBS 120.49, ATCC 9029) (Bos et al., 1988). Growth curves were determined by inoculating 2×10^6 spores/ml in 50 ml medium in 150-ml of Erlenmeyer flasks. Cultures were incubated at 30 °C in a rotary shaker at 300 rpm for the time indicated. Each flask contained 50 ml of Minimal Medium (pH 6.5) (Bennett & Lasure, 1991) supplemented with 0.1% (w/v) casamino acids and 2% (w/v) carbon source. Inulin (Sensus, Frutafit® TEX, Cosun, The Netherlands) and starch (Windmill starch, Avebe, The Netherlands) were used as carbon sources. Mycelia were harvested over a Myracloth filter followed by freezing in liquid nitrogen. The samples were stored at -80 °C prior to the isolation of total RNA or to the determination of biomass. For determination of biomass, the frozen mycelia were freeze dried. Conidiospores were obtained by harvesting spores from a complete medium-plate (minimal medium with 0.5% (w/v) yeast extract and 0.1% (w/v) casamino acids) containing 1% (w/v) glucose, after 4-6 days of growth at 30 °C, using a 0.9% (w/v) NaCl solution.

The bacterial strain used for transformation and amplification of cDNA libraries was *Escherichia coli* ElectroMAX DH10B (Invitrogen). Transformation of ElectroMAX DH10B cells was performed according to the electroporation protocol as described by suppliers. The electroporation voltage in the gene pulser is 2.5 kV in a 0.1 cm gap chamber at settings of 100 Ω and 25 μ F.

Northern analysis

Total RNA was isolated by grinding frozen (-80 °C) mycelium in liquid nitrogen with a pestle and mortar. Powdered mycelium (200 mg) was extracted with 1 ml TRIzol Reagent (Invitrogen, U.S.A) in accordance with the supplier's instructions. Northern analysis and generation of probes for the *A. niger* invertase (*sucA*) and the exo-inulinase (*inuE*) was performed as described in Yuan et al. (2006). The probe for glucoamylase (*glaA*) was a 0.5-kb *KpnI* fragment from pAN56-2 and the probe for glyceraldehyde-3-phosphate dehydrogenase (*gpdA*) was a 1.5-kb of *HindIII* fragment from pAB5-2. Both pAB56-2 and pAB5-2 were obtained from Dr. Peter Punt (TNO, the Netherlands).

Construction of cDNA library in *E. coli* expression vector

Total RNA was isolated using Trizol reagent (Invitrogen) according to the supplier's instructions. mRNA was isolated using FastTrack kit (Invitrogen, K1593). cDNA libraries were constructed using Superscript plasmid system with Gateway technology for cDNA synthesis and plasmid cloning (Invitrogen, 18248-013). Briefly, the first strand cDNA was primed with an oligo(dT)-primer/*NotI*-linker in the presence of SuperScript II RT and dNTPs. The second strand cDNA was synthesized by the combined actions of RNase H and DNA polymerase I. The ends of the cDNA copies were first filled in with DNA polymerase I in the presence of dNTPs, ligated with *SalI* adaptors using T4 DNA ligase, then digested with *NotI*, to obtain *NotI*-*SalI* termini cDNA. The obtained ds cDNAs with *NotI*-*SalI* termini were further directionally cloned into pSPORT1. Plasmid pSPORT1 contains the ampicillin resistant gene for selection of transformants. The cloned products were transformed into *E. coli* strain ElectroMAX DH10B (MAX Efficiency®, Invitrogen) by electroporation. Independent transformants were grown into colonies on LB plates with ampicillin at a density of $(1-3) \times 10^4$ colonies per plate. Library cells were washed off from the plates with 0.9% NaCl and frozen down with 40% glycerol in 1-ml aliquots and stored at -80°C. Library DNA was isolated from the *E. coli* transformants using Plasmid Maxiprep Kit (QIAGEN) and 100-µl aliquots were stored at -20°C.

Construction of cDNA libraries in the *S. cerevisiae* expression vector

Gateway Cloning Technology provides the possibility to transfer DNA inserts between vectors using λ -att site-specific recombination and the technology was used to construct *A. niger* cDNA libraries in yeast expression vector. cDNAs with *NotI*-*SalI* termini (see above) were first directionally cloned into attB containing vector pCMV-SPORT6 which resulted in the pCMV-SPORT6-cDNAs libraries (primary libraries). Plasmid DNA from the pCMV-SPORT6-cDNAs libraries was isolated and subjected for BP reaction with pDONR201. In short, 500 ng of the plasmid library was incubated with 300 ng of attP containing entry vector pDONR201, 4 µl of BP reaction buffer and 4 µl of BP clonase enzyme mix at 25°C for 1 hour. The reaction was stopped by adding 2 µl of proteinase K solution for 10 min at 37°C. 1 µl of the BP reaction were electroporated into 50 µl of library efficiency competent cells *E. coli* strain ElectroMAX DH10B and plated out on of Kanamycin (50 µg/ml)

containing LB plates. Transformants washed from the plates and plasmid DNA was isolated which resulted in the entry library (pDONR201-cDNA). To construct final expression libraries, the entry plasmid pDONR201-cDNAs were used for LR reaction. Briefly, 300ng of entry plasmid was incubated with 300 ng of attR containing destination vector pYES-DEST52, 4 µl of LR reaction buffer and 4 µl of LR clonase enzyme mix at 25°C for 1 hour. 2 µl of proteinase K solution was added and incubated at 37°C for 10 min to stop the reaction. 1 µl of LR reaction mix were then transformed into 50 µl of strain ElectroMAX DH10B competent cells and plated out on Ampicillin (100 µg/ml) containing LB plates and resulted in the destination expression library harboring plasmid pYES-DEST52-cDNAs (see Fig. 2 for the construction procedure).

Digestion analysis

For the analysis of cDNA inserts in each library, 24 clones were randomly selected and cultured in 3 ml of LB medium containing 50 µg/ml of ampicillin (*E. coli* expression libraries, Gateway Primary libraries and Destination libraries) or 25 µg/ml of kanamycin (Gateway Entry libraries) individually. The harboring plasmids were isolated and used for digestion analysis. For each enzyme digestion reaction, 500 ng of plasmid was incubated with 10 U of the indicated enzymes, 2 µl of 10X corresponding enzyme buffer (Invitrogen), TE buffer to a total volume of 20 µl at 37°C for 2h. The digested reaction was then running in the agarose gel by electrophoresis at 80 volts for 2h.

Results and discussion

Set up growth conditions for generation of cDNA library from inulin and starch grown cultures

To construct inulin and starch specific cDNA libraries, the growth of *A. niger* in liquid medium containing inulin and starch as sole carbon source was examined. *A. niger* wild type strain N402 was cultivated in 2% (w/v) inulin or 2% (w/v) starch and the biomass during the growth period was determined at each time points (16, 32, 48, 64 and 80 h after inoculation) (Fig. 1A). In both the inulin and starch grown cultures, *A. niger* showed a typical growth curve with an exponential growth phase (between 16 and 48 h), a stationary growth phase (48-64 h) and a death or autolytic phase (64-80 h). During the stationary phase the cultures reached the highest biomass values with a density of 11.0 g/L for inulin, 9.0g/L for starch. During the autolysis phase the biomass dropped to a density of 9.0 g/L for inulin and 8.0 g/L for starch at 80 h respectively (Fig. 1A).

To examine the expression of genes encoding representative inulin or starch degrading enzymes, Northern blot analysis was performed using RNA isolated at the different time points. For the inulin culture, the genes encoding invertase (*sucA*) and exo-inulinase (*inuE*) were used as probes. *sucA* was highly expressed at very early stage (16 h) and the expression level was soon strongly reduced with the later time points. Almost no expression could be detected at the latest time point (80 h). *inuE* was expressed specifically

at 16 h and 48 h and not detectable at other time points (Fig. 1B). The lack of expression of *sucA* and *inuE* at 80 h might be due to the total consumption of inulin at this stage, possibly in combination with the degradation of mRNA during the autolytic phase of the growth.

A similar Northern blot analysis was also performed for the starch culture. The gene encoding extracellular starch degrading enzyme, glucoamylase (*glaA*) and an additional gene encoding glyceraldehyde-3-phosphate dehydrogenase (*gpdA*) were chosen to detect mRNA levels. *glaA* was expressed at exponential phase and stationary phase, 32, 48 and 64 h and no or very weak expression was detected at 16h and 80h. *gpdA* is usually a constitutively expressed gene, and often used as marker for RNA loading control. This gene was only highly expressed at early stages 16, 32 and 48h and its expression was strongly reduced and hardly detected at later stages (64 h and 80 h) (Fig. 1B). Thus, the expression of *gpdA* is only observed when the culture is actively growing. Reaching the stationary or autolysis phase of the culture, results in a strong reduction in *gpdA* mRNA levels, indicating that the constitutive expression of *gpdA* seems to be true for exponentially growing cultures.

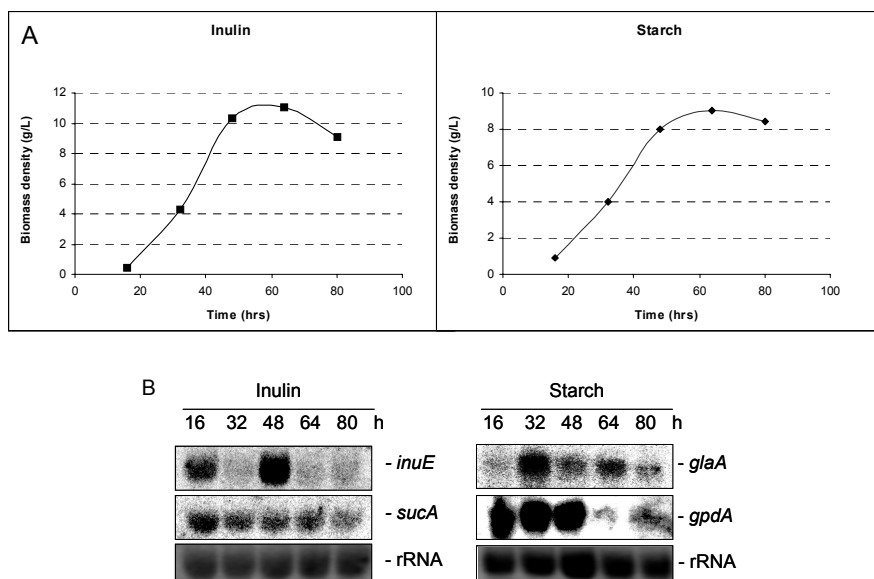


Fig. 1. Examination of growth conditions regarding to inulin or starch degrading activities. A, Growth curve of *A. niger* strain N402 on MM containing 2% (w/v) inulin or starch as sole carbon source. B, Expression analysis of *A. niger* genes on inulin or starch cultures at different time points.

Although these various genes showed different expression patterns at different time points, it was obvious that no or very weak expression of any of these genes was detected at latest time point (80 h) for both the starch and inulin cultures. The concern was that the RNA fractions of the 80 h time point contained high RNase activity, which might cause RNA degradation of the other RNA samples upon pooling. Therefore, equal amounts of total RNA from the first four time points 16, 32, 48 and 64 h, excluding the 80 h time point, were pooled and used for generating cDNA libraries from both inulin and starch grown cultures.

Construction and characterization of inulin and starch cDNA libraries in the *E. coli* expression vector

An overview of the experimental procedure to construct the cDNA libraries is given in Fig. 2. Equal amount of total RNA from each time point (16, 32, 48, 64 h) of starch and inulin samples was pooled and mRNA was isolated using TRIzol Reagent (Invitrogen, U.S.A). The purified mRNA was used for cDNA synthesis and cloning using SuperScript plasmid system with Gateway Technology (Invitrogen). The synthesized cDNA was adapted with *Sal*I at its 5' end and *Not*I at its 3' end which allowed directional cloning in pSPORT1. Expression of cDNA in *E. coli* expression vector pSPORT1 is driven by the *lac*P promoter. The cDNA clones were transformed to *E. coli* by electroporation. Due to limitation of posttranslational modification of eukaryotic proteins in *E. coli*, we have only generated a limited number of cDNA clones in *E. coli* expression vector. In total, 0.24×10^5 primary transformants for inulin cDNA library and 1.92×10^5 primary transformants for starch cDNA library were created, respectively (Table 1).

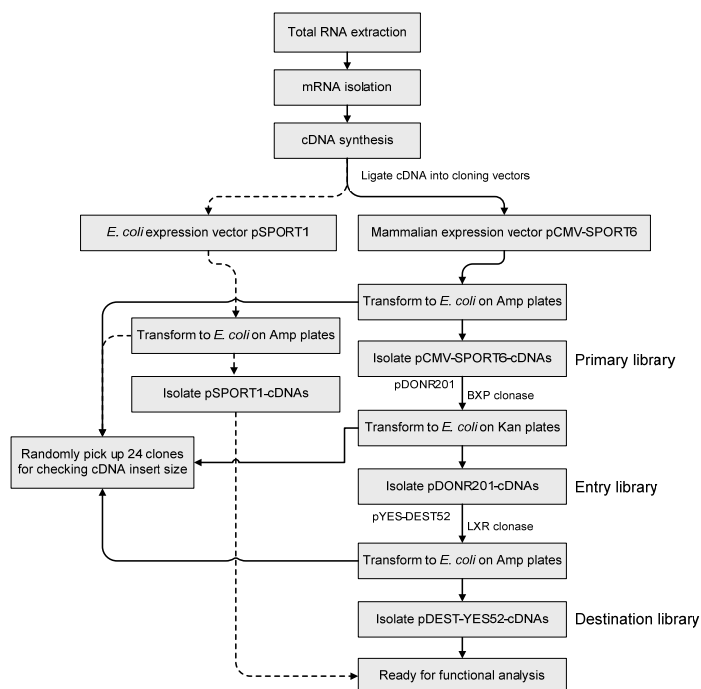


Fig. 2. Flow chart of construction of *A. niger* cDNA expression libraries using Gateway cloning technology. Solid line indicates yeast expression system; broken line indicates *E. coli* expression system.

Table 1. Molecular characterization of *A. niger* cDNA libraries in *E. coli* expression vector.

Cloning vector	cDNA source	No. of clones	No. of clones with insert	average insert size
pSPORT-1	Inulin	0.24×10^5	95%	1.5 kb
	Starch	1.92×10^5	100%	1.7 kb

To evaluate the cDNA insert size in the libraries, 24 single clones was randomly selected for each library and the plasmid was extracted and digested with *EcoRV* (next to cDNA 5' termini adapter *SalI*) and *HindIII* (next to cDNA 3' termini adapter *NotI*), thereby releasing the cDNA insert. A variety of insert sizes range from 200 bp, the size expected for a clone with no insert, to 3 kb was observed for each library (data not shown). Empty clones, based on this digestion analysis represented 5% for inulin cDNA library while 0 for starch cDNA library. When the number of empty clones in each library was subtracted, the average insert size was 1.5 kb and 1.7 kb for inulin and starch cDNA libraries respectively (Table 1). Among them, 54% of inulin clones and 61% of starch clones harbor cDNA inserts with size between 1.0-2.0 kb, while 21% of inulin clones and 26% of starch clones contain cDNA fragments with size of over 2.0 kb (data not shown). This data indicated that both inulin and starch cDNA expression libraries in pSPORT1 are of good quality.

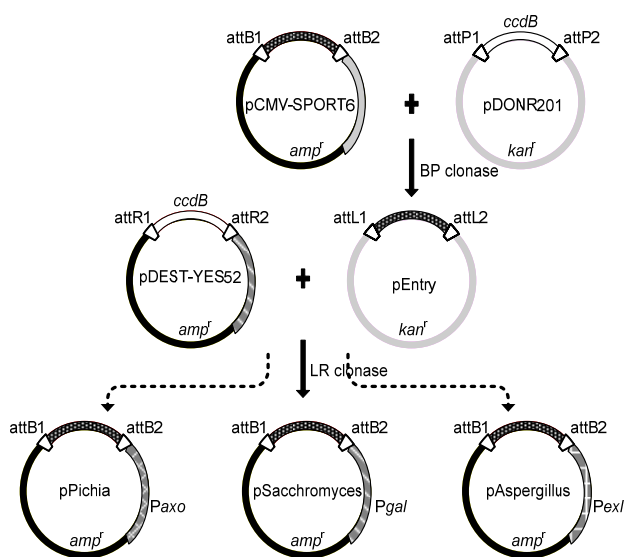


Fig. 3. Schematic representation of Gateway compatible cloning technology. BP reaction resulted in entry clones, LR reaction resulted in destination expression clones. Dot filled frame means cDNA fragment. Broken arrows mean possible destination expression clones generated from same entry clones. Possible promoters used to drive cDNA expression and antibiotic resistant genes for selection of clones are indicated.

Construction and characterization of inulin and starch cDNA libraries in the *S. cerevisiae* expression vector

To construct cDNA libraries in the *S. cerevisiae* expression vector, the same method was used as described above for the synthesis of cDNA insert flanked with *SalI* and *NotI* restriction sites. Since no suitable entry vector was available for direct cDNA cloning in the yeast expression vector at that time of library construction (first half 2002), cDNA libraries in yeast expression vector were generated in three subsequent steps using Gateway cloning

technology (see materials and methods and Fig. 2). As shown in Fig. 3, this technology uses the λ -phage att recombination sites in combination with λ -recombinase enzyme (Clonase) to transfer cDNA segments from one vector to another. For example, cDNA fragments that are flanked by attB sequences in Gateway primary vector (e.g. pCMV-SPORT6) are efficiently transferred by recombination into entry vector (e.g. pDONR201) containing attP sequences with the use of BP Clonase. The obtained cDNA-entry clones after the BP reaction are properly oriented and now contain attL sequences. These attL flanked cDNA inserts can be efficiently transferred to other destination vectors (e.g. pDEST-YES52) containing attR sequences using LR clonase. Recombination of attL and attR sites results in the formation of attB sites, allowing subsequent transfer of the inserts to another BP reaction, if required. In this cloning system, two types of antibiotic markers are used to select for transformants. The attB-containing primary plasmid (pCMV-SPORT6) and the attR-containing destination plasmid (pDEST-YES52) contain the ampicillin resistant gene while the attP-containing donor plasmid contains the kanamycin resistant gene. Moreover, the donor vector contains the *ccdB* gene flanked by the attP sites for negative selection to prevent growth of *E. coli* harboring *ccdB* containing λ -phages, by products of BP and LR reaction, in medium, thus increasing recombination efficiency.

For the construction of the *S. cerevisiae* expression libraries, ds-cDNAs within *SalI*-*NotI* fragments were first cloned into Gateway primary cloning vector pCMV-SPORT6. In the primary libraries, 2.80×10^5 and 1.75×10^5 independent clones were created for inulin and starch cDNA libraries, respectively. Based on the *EcoRV*-*HindIII* digestion analysis of 24 randomly selected clones, 95% of clones from inulin library and 100% of clones from starch library contain cDNA inserts. The average insert size excluding empty clones was 1.4 kb for inulin library and 2.2 kb for starch library, respectively (Table 2). For the inulin library, 71% of the clones contained an insert of 1 kb or larger and 83% of the inserts of the starch library was over 1 kb in length (data not shown). This preliminary data showed the primary cDNA libraries were satisfying and further construction of yeast expression libraries was preceded.

Table 2. Molecular characterization of *A. niger* cDNAs libraries in Gateway compatible vectors.

Gateway vector	cDNA source	No. of clones	No. of clones with insert	average insert size
pCMV-SPORT6	Inulin	2.80×10^5	95%	1.4 kb
	Starch	1.75×10^5	100%	2.2 kb
pDONOR201	Inulin	2.87×10^5	100%	1.7 kb
	Starch	1.88×10^5	100%	1.8 kb

Using the BP reaction, cDNA inserts were transferred from Gateway primary vector pCMV-SPORT6 to entry vector pDONR201 resulting in the so-called entry libraries. In the entry libraries, 2.87×10^5 and 1.88×10^5 recombinants grown on kanamycin plates were harvested for inulin and starch library, respectively. All the clones analyzed from both inulin and starch entry libraries harbor cDNA fragments. The average size was 1.7 kb for inulin library

and 1.8 kb for starch library (Table 2). The data indicated that these entry libraries were of good quality and used to produce final yeast expression libraries.

Table 3. Molecular characterization of *A. niger* cDNA libraries in yeast expression vector.

Gateway vector	cDNA source	No. of clones	% of pYES clones	No. of pYES clones	average insert size
pDEST-YES52	Inulin	2.80×10^5	54%	1.50×10^5	1.2 kb
	Starch	5.00×10^5	33%	1.50×10^5	1.6 kb

Using the LR reaction, cDNA inserts from the entry library were then transferred to destination yeast expression vector pYES-DEST52. In destination libraries, clones were selected on ampicillin LB plates. 2.80×10^5 clones for inulin library and 5.00×10^5 for starch library were generated. Surprisingly, *EcoRV/HindIII* digestion analysis revealed that only 54% and 33% of analyzed clones for inulin and starch libraries, respectively, consist of pYES-DEST52 backbone containing the URA3 selection marker to select for transformants in *S. cerevisiae* (Table 3). This was further confirmed by *PstI* digestion of the clones. Proper destination plasmids should give an anticipated 500 bp fragment and only the plasmids that were identified as pYES-DEST52 vectors using the *EcoRV/HindIII* digestion contained the 500 bp fragment. Further digestion analysis of 11 unclear clones showed that one of them was from primary vector pCMV-SPORT6. The pattern of digestion of the remaining clones gave inconclusive results and the exact nature of these plasmids is not clear. Testing of the antibiotic resistance of the clones on ampicillin or kanamycin containing LB-plates revealed that about 30% of destination clones harbor both ampicillin and kanamycin resistance genes (data not shown). This indicates that contamination of previous vectors during recombinant reaction occurred; however the exact mechanism that is either from single cross-over during transformation or from co-transformation with previous vectors is still unknown. Importantly, the plasmids in this library whose exact nature is unclear do not contain the URA3 selection marker. Thus, upon transformation to *S. cerevisiae* only pDEST-YES52 expression vectors will result in yeast transformants. Therefore no more effort was taken for clarification of the unexpected clones. Excluding clones with the contaminated vectors, the average cDNA insert size in pDEST-YES52 vector was 1.2 kb for inulin library and 1.6 kb for starch library (Table 3). These sizes are reasonable for production of cDNA product. Because of the presence of addition plasmids other than with the pDEST-YES52 backbone, the number of primary transformants in the libraries to reach certain amounts (150.000) of clones with a pDEST-YES52 backbone for functional screening was adjusted (Table 3).

Future considerations for construction and screening of cDNA expression libraries

cDNA expression cloning is a powerful and efficient tool for investigation of secreted hydrolytic enzymes of fungal origin (Dalboge and Heldt-Hansen, 1994; Christgau et al., 1995; van der Vlugt-Bergmans and van Ooyen, 1999). However, selection of optimal protein expression system is very important for screening and functional analysis of the interesting cDNA clones.

Both bacteria expression system and eukaryotic system have advantages and disadvantages for expressing eukaryotic genes. For example, bacteria strain *E. coli* is a valuable host for the efficient, cost-effective and high level production of heterologous proteins, however, the problems of protein accumulation and protein processing, protein folding and posttranslational modification for eukaryotic genes is still an issue (Hannig and Makrides, 1998). On the other hand, yeast strains *Pichia pastoris* and *Saccharomyces cerevisiae* as well as *Aspergillus* strains have benefit for protein processing and protein posttranslation modification, but have limitation for lower yield and host strain instability (Buckholz and Gleeson, 1991; Da Silva and Bailey, 1991; Punt et al., 2002; Holz et al., 2003). Nevertheless, modification on both eukaryotic system and prokaryotic system has been improved for expression of eukaryotic genes (<http://www.invitrogen.com>). In the original experimental set-up, bacteria strain *E. coli*, yeast strains *P. pastoris*, *S.s cerevisiae* and *A. niger* were all considered as good candidates for expressing *A. niger* cDNA clones. Therefore we have used Gateway cloning technology which allows efficient constructing the cDNA libraries in variant expression vectors (Ohara and Temple, 2001) to construct the cDNA expression libraries (see Fig. 2 and Fig. 3).

Construction of *E. coli* expression libraries was easy by one single ligation step. The synthesized cDNA was cloned into *E. coli* expression vector pSPORT1 by single ligation step with directional cloning (see Fig. 2). Both the percentage and sizes of cDNA inserts matched to the requirements for cDNA screening (see Table 1). Due to limitation of posttranslational modification of eukaryotic proteins in *E. coli*, we have only generated a limited number of cDNA clones in *E. coli* expression vector (Table 1).

Construction of yeast expression libraries was complex using Gateway cloning technology (Fig. 2 and Fig. 3). The results from the yeast expression libraries may not show very efficient due to three steps cDNA transfer (Table 3). However, it is worth noting that these libraries cloned in Gateway system are easy to modify. One option is to modify the destination expression vector by replacing attR recombinant sequence with the attP sequence, by one step BP reaction, cDNA inserts could be transferred into destination expression vector. On the other hand, Gateway system has been improved in such way that the synthesized cDNA fragments are flanked by attB recombinant sequences on each end and can be directly transferred to attP-containing Gateway entry vector by BP reaction, skipping the step to the primary vector pCMV-SPORT6. By one step LR reaction, the cDNAs can be cloned into different expression vectors (Karnaoukhova et al., 2003; M. Arentshorst, personal communications). These modifications would reduce the recombination step, and thus lower the cDNA size reduction and minimize the contamination efficiency.

Our cDNA libraries have been used in preliminary transformation and screening experiments in both *E. coli* and *S. cerevisiae* to identify inulinolytic and amylolytic activities (C. Goosen, R. van der Kaaij, personal communication), but have not yielded clones with clear hydrolytic activities. PCR analysis, however, revealed the presence of known inulinolytic (*inuA*, *inuE* and *sucA*) and amylolytic (*glaA*) clones in these cDNA libraries,

indicating the problem occurs in protein expression and screening. Targeted overexpression of *sucB* using the same yeast expression system resulted in the expression of SucB protein in cell free lysates (Goosen et al., 2007, Chapter 6) indicating that the expression system is functional. Firm conclusions about the libraries can not be drawn at the moment since only a limited number of transformants (about 3,000 clones) have been analyzed. In addition, problems were encountered when expressing cDNA from the *gal1* promoter. The expression level of induction from the *gal1* promoter in *S. cerevisiae* was lower as expected in the strain used (BY4347 Δ *suc2*). A possible improvement of the screening system is the use of a yeast strain in which the expression from the *gal1* promoter is stronger.

Chapter 3

Aspergillus niger genome wide analysis reveals a large number of novel α -glucan acting enzymes with unexpected expression profiles

*Xiao-Lian. Yuan**, *Rachel M. van der Kaaij**, *Cees A.M.J.J. van den Hondel*, *Peter J. Punt*, *Marc J.E.C van der Maarel*, *Lubbert Dijkhuizen*, *Arthur F.J. Ram*

** These authors contributed equally to this study*

Submitted

Abstract

The filamentous ascomycete *Aspergillus niger* is well known for its ability to produce a large variety of enzymes for the degradation of plant polysaccharide material. A major carbon and energy source for this soil fungus is starch, which can be degraded by the concerted action of α -amylase, glucoamylase and α -glucosidase enzymes, members of the glycoside hydrolase (GH) families 13, 15 and 31, respectively. In this study we have combined analysis of the genome sequence of *A. niger* CBS 513.88 with microarray experiments to identify novel enzymes from these families and to predict their physiological functions. We have identified 17 previously unknown family GH13, 15 and 31 enzymes in the *A. niger* genome, all of which have orthologues in other aspergilli. Only two of the newly identified enzymes, a putative α -glucosidase (AgdB) and an α -amylase (AmyC), were predicted to play a role in starch degradation. The expression of the majority of the genes identified was not induced by maltose as carbon source, and not dependent on the presence of AmyR, the transcriptional regulator for starch degrading enzymes. The possible physiological functions of the other predicted family GH13, GH15 and GH31 enzymes, including intracellular enzymes and cell wall associated proteins, in alternative α -glucan modifying processes are discussed.

Introduction

Aspergillus niger is a saprophytic fungus well known for its production and secretion of a variety of hydrolytic enzymes contributing to its ability to degrade plant polysaccharides such as cellulose, hemicellulose, pectin, starch and inulin (De Vries and Visser 2001; Tsukagoshi et al. 2001; Yuan et al. 2006). Starch is the most abundant storage carbohydrate in the plant kingdom and is present in tubers, seeds and roots of a variety of crop plants including cereals, potatoes and manioc (Peters 2006). Starch is composed of two different molecules: (i) amylose, an unbranched, single chain of α -(1,4)-linked glucose residues and (ii) amylopectin, consisting of a α -(1,4)-linked glucose chain with α -(1,6)-branches on every 12-25 glucose residues along the α -(1,4)-linked backbone (Robyt 1998). The degradation of starch is performed by a variety of enzymes, which are divided over three Glycoside Hydrolase (GH) families based on their sequence similarity (<http://www.cazy.org>) (Coutinho and Henrissat 1999). The first step in starch degradation is the endo-hydrolysis of the long polysaccharide chains into shorter maltooligosaccharides and α -limit dextrans by α -amylases (EC 3.2.1.1). α -Amylases belong to family GH13, a large family containing various hydrolysing and transglycosylating enzymes, mostly acting on α -(1,4) or α -(1,6)-glycosidic bonds. Members of family GH13 have a $(\beta/\alpha)_8$ barrel structure and can be recognized by four highly conserved amino acid regions containing three catalytic residues (MacGregor et al. 2001; Nakajima et al. 1986). After endo-hydrolysis, subsequent steps in starch degradation involve exo-acting enzymes releasing glucose. This reaction is performed by glucoamylase type enzymes of family GH15 (EC 3.2.1.3), a relatively confined family with regard to enzyme specificity, as all its studied members hydrolyse either α -(1,4)

or α -(1,6)-bonds to release β -glucose from the non-reducing end of maltooligosaccharides. Most GH15 enzymes described thus far possess a starch binding domain (SBD) (Sauer et al. 2000), a discrete C-terminal region of the protein that binds to starch and facilitates hydrolysis (Southall et al. 1999). Additionally, α -(1,4)-glucosidases of family GH31 may release α -glucose from the non-reducing end of starch (EC 3.2.1.20). This family also harbours other enzyme specificities such as α -xylosidase activity.

Several *A. niger* enzymes involved in starch degradation, and their corresponding genes, have been characterized and isolated. *A. niger* glucoamylase GlaA (family GH15) is an important enzyme for the modification of starch in the food industry (Boel et al. 1984; van Dijck et al. 2003). Additionally, one GH31 α -glucosidase (AglA, renamed AgdA) (Nakamura et al. 1997) has been characterized previously, as well as three family GH13 α -amylases: acid amylase AamA, and the almost identical AmyA and AmyB (Boel et al. 1990; Korman et al. 1990). The transcriptional regulation of the genes encoding starch degrading enzymes has been studied in several aspergilli (Nakamura et al. 2006). In general, their expression is high on starch and induced by the presence of (iso)maltose (Tsukagoshi et al. 2001; Kato et al. 2002a). The presence of the inducer activates the Zn(II)₂Cys₆ transcription factor AmyR which binds to CGGN8(C/A)GG sequences in the promoter regions of AmyR target genes thereby activating their transcription (Petersen et al. 1999; Gomi et al. 2000; Tani et al. 2001; Ito et al. 2004).

Recent studies have indicated that some GH13 enzymes in fungi may be involved in the synthesis or modification of α -glucan in the fungal cell wall, rather than in starch degradation. The cell wall of aspergilli contains four major classes of polysaccharides: chitin, α -glucan, β -(1,3)-glucan and galactomannan (Fontaine et al. 2000; Beauvais and Latgé 2001). The α -glucan fraction identified in *A. niger* consists of two types of molecules: a linear polymer with alternating α -(1,3)/(1,4)-glycosidic bonds called nigeran (Barker and Carrington 1953) and pseudonigeran, a linear α -(1,3)-glucan molecule with some (3-10%) α -(1,4)-linkages (Johnston 1965; Horisberger et al. 1972). Synthesis of α -glucan is thought to be carried out by α -glucan synthase enzymes encoded by *ags* genes. The first putative α -D-glucan synthase encoding gene (*ags1*) was identified in *Schizosaccharomyces pombe* (Hochstenbach et al. 1998). The *ags1* gene encodes a large, three-domain protein. In addition to the multi-pass transmembrane domain in the C-terminal part of the protein, two predicted catalytic domains are present. The middle domain shows strong similarity to glycogen and starch synthases in Glycosyltransferase family (GT) 5 and is predicted to be involved in the synthesis of α -glucan. The N-terminal part of the protein is similar to α -amylases and belongs to family GH13. This part of the protein is predicted to be localized extracellularly, and might be involved in connecting two α -(1,3)-glucan chains (Grün et al. 2005). Apart from the α -glucan synthases, two types of family GH13 enzymes were recently identified in fungi to play a role in fungal cell wall biosynthesis. Marion *et al.* (Marion et al. 2006) provided evidence for the involvement of a putative α -amylase (Amy1p) in the formation of α -(1,3)-glucan in the cell wall of *Histoplasma capsulatum*. An AMY1 knockout was unable to produce α -(1,3) glucan and showed reduced virulence. In dimorphic fungi like *H. capsulatum*, cell wall α -glucan is a known virulence factor (Rappleye et al. 2004;

Rapplee and Goldman 2006). The second α -amylase-like enzyme Aah3p was first studied in *S. pombe* (Morita et al. 2006). Disruption of *aah3* encoding a GPI-anchored protein resulted in a hypersensitivity towards cell wall degrading enzymes and an aberrant cell shape, indicating that normal cell wall biosynthesis was affected. Disruption of a homologous gene (*agtA*) in *A. niger* also affected cell wall stability (van der Kaaij et al. 2007a).

In this study, we surveyed the *A. niger* genome sequence (Pel et al., 2007) to identify for the first time all GH13, GH15 and GH31 family members present in this important industrial source for amylolytic enzymes. This resulted in identification of a surprisingly large number of previously unknown enzymes. By studying their phylogeny, the presence of specific protein features and synteny with other *Aspergillus* species, members of each GH family could be divided into several groups. Additionally, we studied the transcriptional regulation of the genes encoding these proteins in a wild type *A. niger* strain, and in a derived *amyR* deletion strain, during their growth on xylose and maltose. Only few of the identified proteins were induced by maltose. Expression of many of the identified groups of enzymes, including the homologues of both *S. pombe* Aah3p and *H. capsulatum* Amy1p, was not induced by maltose and was not dependent on the presence of AmyR. The possible involvement of these enzymes in cell wall α -glucan synthesis and modification is discussed.

Material and methods

Database mining of *A. niger* genome and analysis of predicted proteins

The full genome sequence of *A. niger* strain CBS 513.88, has been deposited at the EMBL database with accession numbers AM270980-AM270998 (Pel et al. 2007) and was used for database mining. The nucleotide accession numbers of *A. niger* genes as listed in Table 1 refer to this database. Hidden Markov Model (HMM) profiles were built with the HMMER package (Durbin and Eddy 1998) (<http://hmmer.wustl.edu/>) based on the amino acid sequences of known members of GH13, GH15 and GH31. Proteins belonging to these families, originating from the different kingdoms of life, were retrieved from the CAZy website at <http://www.cazy.org> (Coutinho and Henrissat 1999), and the protein sequences were extracted from the GenBank/GenPept database at <http://www.ncbi.nlm.nih.gov/entrez/> and Swiss-Prot database at <http://www.expasy.org/sprot/>. The *A. niger* genome was searched with the HMM profiles using the WISE2 package (Birney et al. 2004) (<http://www.ebi.ac.uk/Wise2/>).

The presence of signal peptidase cleavage sites, glycosylphosphatidylinositol (GPI) attachment sites and starch binding domains (SBD) in the obtained sequences were predicted by web-based tools at URL: <http://www.cbs.dtu.dk/services/SignalP/> (Bendtsen et al. 2004), URL: http://mendel.imp.ac.at/sat/gpi/gpi_server (Eisenhaber et al. 2004), and URL: <http://www.ncbi.nlm.nih.gov/BLAST/> (Marchler-Bauer and Bryant 2004), respectively.

Multiple sequence alignments of GH13, 15 and 31 family members were performed using DNAMAN version 4.0 (Lynnon BioSoft, Canada). The alignments were

based on the full length of the predicted proteins, except in case of predicted α -glucan synthases for which only the N-terminal part, encoding the family GH13 domain, was used for the alignment. The phylogenetic relationship was calculated by using Optimal Alignment (Thompson et al. 1994) with gap opening penalty and gap extension penalty of 10 and 0.05, respectively. A bootstrapped test of phylogeny was performed by the Neighbor-Joining method using 1000 replicates. Wherever possible, one protein with described activity was included for each of the groups identified based on phylogenetic analysis.

Strains and transformations

The wild type *A. niger* strain used in this study is N402, a *cpsA1* derivative of *A. niger* van Tieghem (CBS 120.49, ATCC 9029) (Bos et al. 1988). Strain AB4.1 is a *pyrG* negative derivative of N402 (van Hartingsveldt et al. 1987) and was used to construct the *amyR* disruption strain. *A. niger* strains were grown in *Aspergillus* minimal medium (MM) (Bennett and Lasure 1991), or *Aspergillus* complete medium (CM) consisting of MM with the addition of 0.5% (w/v) yeast extract and 0.1% (w/v) casamino acids. Growth medium was supplemented with 10 mM uridine (Serva, Germany) when required. Transformation of *A. niger* AB4.1 was performed as described earlier (Punt and van den Hondel 1992) using lysing enzymes (L1412, Sigma, U.S.A.) for protoplastation. The bacterial strain used for transformation and amplification of recombinant DNA was *Escherichia coli* XL1-Blue (Stratagene, U.S.A.). Transformation of XL1-Blue was performed according to the heat shock protocol (Inoue et al. 1990).

Disruption of the maltose utilization activator gene *amyR* in *A. niger*

Plasmid pJG01 containing the *A. niger amyR* gene as a 4.3 kb *NsiI* fragment in pGEM11 was kindly provided by P. van Kuyk (Wageningen University, the Netherlands) and used to disrupt the *amyR* gene. The construction of the *amyR* deletion cassette was performed as follows. The *Bam*HI-*Eco*RI fragment and *NsiI*-*Sal*I fragment flanking the *amyR* ORF at the 5' and 3' region respectively were isolated from pJG01. The isolated *NsiI*-*Sal*I fragment was cloned into pUC19 to obtain plasmid pAmyRF3. Subsequently, a *Bam*HI-*Sal*I fragment carrying the *A. oryzae pyrG* gene, obtained from plasmid pAO4-13 (de Ruiter-Jacobs et al. 1989) was inserted into pAmyRF3 which resulted in plasmid pAmyRF3-*pyrG*. The *Bam*HI-*Eco*RI fragment isolated from pJG10 was ligated into pAmyRF3-*pyrG* resulting in the *amyR* deletion plasmid (p Δ *amyR*). Prior to transformation to AB4.1, p Δ *amyR* was linearized with *Eco*RI. Uridine prototrophic transformants were selected by their ability to grow on MM without uridine. After two rounds of purification, transformants were tested for their ability to grow on starch. Approximately 10% of the *pyrG*⁺ transformants showed defective growth on MM agar plates containing starch as sole carbon source. Six independent putative *amyR* deletion strains (YvdM1.1-1.6) with identical phenotypes were obtained. Southern blot analysis confirmed proper deletion and a single integration of the *amyR* disruption cassette

at the *amyR* locus. Strain YvdM1.1 was used for further analysis and we will refer to this strain as the $\Delta amyR$ strain in the remaining of this paper.

Culture conditions, RNA preparation, microarray analysis and data analysis

RNA extracted from the *A. niger* $\Delta amyR$ strain and its parental strain (N402) grown on different carbon sources were used for microarray experiments using custom made 'dsmM_ANIGERa_coll' Affymetrix GeneChip® Microarrays kindly provided by DSM Food Specialties (Delft, The Netherlands). All experiments for each growth condition (culturing the mycelia, RNA extractions and microarray hybridizations) were performed twice as independent biological experiments.

A. niger spores (2×10^6 spores ml⁻¹) were inoculated in 250 ml MM supplemented with 2% (w/v) xylose (Sigma) and 0.1% (w/v) casamino acids and grown for 18 h at 30 °C on a rotary shaker at 300 rpm. The mycelium was harvested by suction over a nylon membrane and washed with MM without carbon source. Aliquots of 1.6 g wet weight of mycelium were transferred to 300 ml Erlenmeyer flasks containing 70 ml MM supplemented with 1% (w/v) carbon source (maltose (Sigma) or xylose) and incubated at 30 °C for a further 2 or 8 h. The pH of all cultures grown for 2 h was equal to the pH at the time of transfer (pH 6.2). Cultures grown for 8 h were buffered at pH 4 by the addition of 100 mM of citric acid/sodium citrate to allow comparison between the N402 and the $\Delta amyR$ strain at this time point as medium of the $\Delta amyR$ strain did not acidify as quickly as the medium of the N402 strain. The mycelium was harvested over Miracloth filter, frozen in liquid nitrogen and stored at -80 °C. Total RNA was isolated from mycelia using TRIzol reagent (Invitrogen) and RNA quality was verified by analyzing aliquots with glyoxal/DMSO gel electrophoresis and Agilent Bioanalyzer "Lab on chip" system (Agilent Technologies, U.S.A.). Processing, labeling and hybridization of cRNA to *A. niger* Affymetrix GeneChips were performed according to the corresponding Affymetrix protocols for "Eukaryotic Target Preparation" and "Eukaryotic Target hybridization". For probe array washing and staining, the protocol "Antibody Amplification for Eukaryotic Targets" was followed. Hybridized probe array slides were scanned with Agilent technologies G2500A Gene Array Scanner at a 3 μ m resolution and a wavelength of 570 nm. Affymetrix Microarray Suite software MAS5.0 was used to calculate the signal and p-values and to set the algorithm's absolute call flag, which indicates the reliability of the data points according to P (present), M (marginal) and A (absent). The data on each chip were globally scaled to an arbitrary target gene intensity of 500. The complete microarray data were deposited into ArrayExpress with an accession E-TABM-324 at <http://www.ebi.ac.uk/miameexpress>.

The pre-scaled data from each hybridization experiment was then normalized for statistical analysis using Genespring 7.0 software (Silicon Genetics, U.S.A.). The per chip normalization was performed to ensure that the overall characteristic of the expression distribution such as median should be the same for all the chips. For the genome wide analysis, we focused on maltose induced genes and therefore a pre-filtering of data was performed to select for genes whose detection calls are present in both maltose duplicate samples in the wild type strain (N402). The selected dataset was used to perform a 1-way

ANOVA analysis under the test type of “parametric test, don’t assume variances equal”. Fold changes in expression between two different conditions were then computed for genes with $p < 0.08$ based on 1- way ANOVA analysis.

Results

Identification of glycoside hydrolase family 13, 15 and 31 genes in the *A. niger* CBS 513.88 genome sequence

α -Amylases, glucoamylases and α -glucosidases, members of families GH13, 15 and 31, respectively, are the three main types of enzymes involved in breakdown of starch by aspergilli (Tsukagoshi et al. 2001). To identify all genes encoding enzymes that might play a role in starch utilization, or other α -glucan modifying processes in *A. niger*, the genome of *A. niger* CBS 513.88 was searched with HMM profiles based on known enzymes from families GH13, 15 and 31. This resulted in the retrieval of a total of 27 protein sequences including 17 previously unknown proteins, listed in Table 1A.

Two approaches were combined to predict putative functions in cellular processes for this surprisingly large number of newly identified proteins. First, phylogenetic trees were constructed using the GH13, GH15 and GH31 family members identified in the *A. niger* genome, as well as functionally characterized proteins from other organisms with similarity to the identified *A. niger* proteins (Fig. 1). Second, using DNA microarrays, the expression of all the *A. niger* genes encoding GH13, GH15 and GH31 enzymes was examined in both the *A. niger* wild type strain N402 and the $\Delta amyR$ strain derived, after growth on different carbon sources (Fig. 2 and Table S1). Both the N402 and the $\Delta amyR$ strains were pregrown in xylose for 18 h, and mycelia were transferred to either xylose or maltose media and grown further for 2 h or 8 h. Expression levels were determined based on geometric mean data of biological duplicate samples. We will discuss each enzyme family in detail and combine the findings in *A. niger* with the predicted proteins present in the genomes of *A. fumigatus* (Nierman et al. 2005), *A. nidulans* (Galagan et al. 2005) and *A. oryzae* (Machida et al. 2005).

Identification and transcriptional regulation of starch and α -glucan acting enzymes

Table 1A. All members of family GH13, GH15 and GH31 identified in the genome sequence of *A. niger* CBS 513.88 using HMM profiles. The newly identified proteins are indicated in bold.

Accession no.	Gene	Family	Enzyme activity	Features ^a	AmyR binding sites ^b	Proposed biological function	Ref ^c
An11g03340	<i>aamA</i>	GH13	acid α -amylase	SS		Starch degradation	1
An12g06930	<i>amyA</i>	GH13	α -amylase	SS	+970; +252	Starch degradation	2
An05g02100	<i>amyB</i>	GH13	α -amylase	SS	+252	Starch degradation	2
An04g06930	<i>amyC</i>	GH13	α -amylase	SS	+787; +664; -531	Starch degradation	3
An09g03100	<i>agtA</i>	GH13	α -glucanotransferase	SS, GPI		Cell wall α -glucan synthesis	4
An12g02460	<i>agtB</i>	GH13	α -glucanotransferase	SS, GPI	+810	Cell wall α -glucan synthesis	4
An15g07800	<i>agtC</i>	GH13	putative α -glucanotransferase	SS, GPI		Cell wall α -glucan synthesis	4
An02g13240	<i>agdC</i>	GH13	putative α -glucosidase	SS, GPI	+368	Unknown	4
An13g03710	<i>agdD</i>	GH13	putative α -glucosidase			Unknown	
An01g13610	<i>amyD</i>	GH13	putative α -amylase		+504; -32	Cell wall α -glucan synthesis	
An09g03110	<i>amyE</i>	GH13	putative α -amylase		-76	Cell wall α -glucan synthesis	
An01g06120	<i>gdbA</i>	GH13	glycogen debranching enzyme		-487; +393	Glycogen metabolism	
An14g04190	<i>gbeA</i>	GH13	glycogen branching enzyme			Glycogen metabolism	
An04g09890	<i>agsA</i>	GH13	putative α -glucan synthase	SS		Cell wall α -glucan synthesis	5
An15g07810	<i>agsB</i>	GH13	putative α -glucan synthase	SS	+287	Cell wall α -glucan synthesis	5
An12g02450	<i>agsC</i>	GH13	putative α -glucan synthase	SS	-973; +622, -185	Cell wall α -glucan synthesis	5
An02g03260	<i>agsD</i>	GH13	putative α -glucan synthase	SS		Cell wall α -glucan synthesis	5
An09g03070	<i>agsE</i>	GH13	putative α -glucan synthase	SS		Cell wall α -glucan synthesis	5
An03g06550	<i>glcA</i>	GH15	glucoamylase	SS, SBD	-792; -669; +423; -301	Starch degradation	6
An12g03070	<i>glcB</i>	GH15	putative glucoamylase		-878	Unknown	
An04g06920	<i>agdA</i>	GH31	α -glucosidase	SS	+574; +191, +904; -334	Starch degradation	7
An01g10930	<i>agdB</i>	GH31	putative α -glucosidase	SS		Starch degradation	
An09g05880	<i>agdE</i>	GH31	putative α -glucosidase II	SS		Protein glycosylation	
An18g05620	<i>agqF</i>	GH31	unknown			Unknown	
An07g00350	<i>agqG</i>	GH31	unknown	SS	+402	Unknown	
An09g03300	<i>axIA</i>	GH31	putative α -xylosidase	SS	+126	Xyloglucan degradation	
An01g04880	<i>axIB</i>	GH31	putative α -xylosidase		+430; +138,	Xyloglucan degradation	

^a SS = predicted N-terminal Signal Sequence; GPI = predicted Glycosylphosphatidylinositol anchor signal; SBD = predicted starch binding domain.

^b The presence of consensus AmyR binding motif (CGGN8(A/C)GG) was analysed in the promoter region up to 1 kb upstream of the start codon

^c References: 1 (Boel et al. 1990); 2 (Korman et al. 1990); 3 (Van der Kaaij & Yuan. unpubl) ; 4 (van der Kaaij et al. 2007); 5 (Damveld et al. 2005); 6 (Boel et al. 1984); 7 (Nakamura et al. 1997)..

Table 1B. Functionally described family GH13 and GH31 members from other organisms, used for the multiple sequence alignments in Fig. 1. For each *A. niger* protein identified, a functionally or biochemically characterized protein with the highest similarity was used in the phylogenetic analysis

Accession no.	Name	Family	Enzyme activity	Features ^a	Biological function	Organism	Ref ^b
BAA78714	AndGbe1	GH13	glycogen branching enzyme		Glycogen metabolism	<i>A. nidulans</i>	1
BAA34996	ScGdb1	GH13	glycogen debranching enzyme		Glycogen metabolism	<i>S. cerevisiae</i>	2
P19571	BsAmyA	GH13	α -amylase	SS	Starch degradation	<i>Bacillus</i> sp.	3
CAA54266	BsAgIA	GH13	α -glucosidase		Starch degradation	<i>Bacillus</i> sp.	4
CAA21237	SpAah1	GH13	unknown	SS, GPI	α -Glucan biosynthesis	<i>S. pombe</i>	5
CAA91249	SpAah2	GH13	unknown	SS, GPI	α -Glucan biosynthesis	<i>S. pombe</i>	5
CAB40006	SpAah3	GH13	unknown	SS, GPI	α -Glucan biosynthesis	<i>S. pombe</i>	5
CAA16864	SpAah4	GH13	unknown	SS, GPI	α -Glucan biosynthesis	<i>S. pombe</i>	5
ABK62854	HcAmy1	GH13	unknown		α -Glucan biosynthesis	<i>H. capsulatum</i>	6
ABF50883	AN7345.2	GH31	α /B-glucosidase	SS	Starch/cellulose degradation	<i>A. nidulans</i>	7
ABF50846	AN7505.2	GH31	α -xylosidase		Xylan degradation	<i>A. nidulans</i>	7
BAB39856	AndAgdB	GH31	α -glucosidase	SS	Starch degradation	<i>A. nidulans</i>	8
AAU87580	TrAgull	GH31	α -glucosidase II	SS	Protein glycosylation	<i>T. reesei</i>	9
A45249	CAMAL2	GH31	maltase		Maltose degradation	<i>C. albicans</i>	10

^a SS = predicted N-terminal Signal Sequence; GPI = predicted Glycosylphosphatidylinositol anchor signal

^b References: 1 (Sasangka et al. 2002); 2 (Teste et al. 2000); 3 (Tsukamoto et al. 1988); 4 (Nakao et al. 1994); 5 (Morita et al. 2006); 6 (Marion et al. 2006); 7 (Bauer et al. 2006); 8 (Kato et al. 2002b); 9 (Geysens et al. 2005); 10 (Geber et al. 1992).

Identification and transcriptional regulation of GH13 family members

The HMMer search for family GH13 enzymes in the *A. niger* genome resulted in the identification of 18 protein sequences of which 10 had not been identified previously (Table 1A). Table 2 displays the four conserved regions typical for family GH13 proteins as identified in these enzymes. A phylogenetic tree was produced combining the *A. niger* family GH13 enzymes with several functionally characterized GH13 family proteins from other organisms (Fig. 1). The combination of this phylogenetic analysis with a functional annotation of the proteins revealed 6 recognizable subgroups.

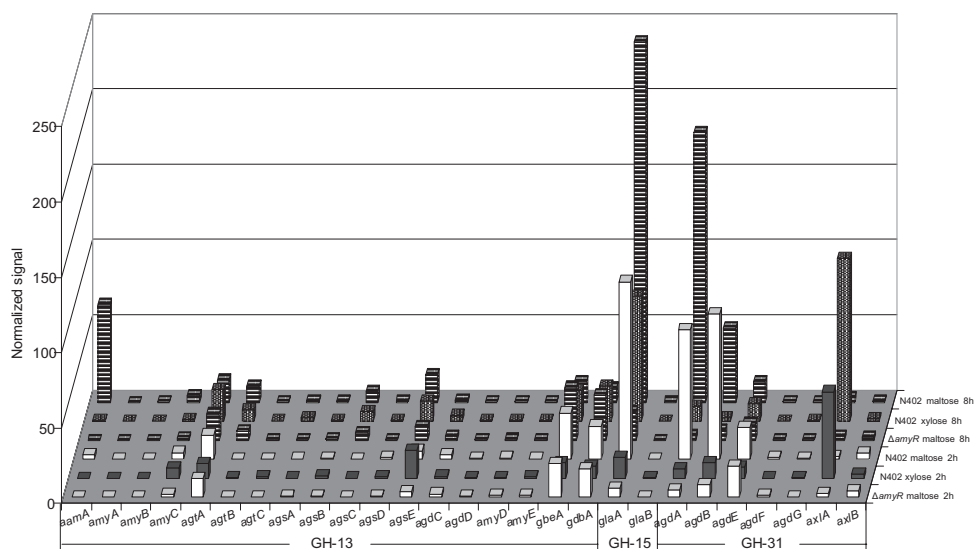


Fig. 2. Expression profiles of *A. niger* family GH13, 15 and 31 enzymes. Accession numbers of the gene names are given in Table 1. Strain and time points after transfer from the preculture are indicated on the right hand side. The numeric values and Present/Absent calls from the expression data are provided as Supplementary Table 1.

Group I consists of 4 extracellular α -amylases. Three of these are the previously characterized extracellular α -amylases acid-amylase (AamA) and α -amylases AmyA and AmyB (Korman et al. 1990). One new extracellular α -amylase was identified and named AmyC. This protein displays high similarity with the known *A. niger* α -amylases (74% identity to AmyA and AmyB, 65% identity to AamA). The *amyC* gene is located in a gene cluster also containing an α -glucosidase gene (*agdA*) and the *amyR* gene encoding the AmyR transcription factor. Expression of *aamA* in *A. niger* N402 was not detectable in xylose media, but was strongly induced in the presence of maltose (Fig. 2). The expression of *aamA* was reduced to a non-detectable level in the Δ *amyR* strain. Expression of *amyA* and *amyB* was not detected in *A. niger* N402 in any of the conditions tested (Fig. 2) (see Discussion). The expression of the newly identified *amyC* gene was relatively low compared to the *aamA* gene. At 2 h after transfer from the preculture, the expression level of *amyC* was

independent of the carbon source, while after 8 h the expression level was reduced 3 fold on xylose compared to maltose. Additionally, the expression on maltose was reduced 2-3 fold in the $\Delta amyR$ strain (Table S1). The presence of three putative AmyR binding elements in the promoter region of *amyC* further suggests that its expression is controlled by AmyR (Table 1A).

The *A. niger* acid amylase (encoded by *aamA*) in the CBS 513.88 strain does not contain a predicted Starch Binding Domain (SBD), while its homologues in *A. nidulans*, *A. fumigatus* and other aspergilli contain a full length SBD (Table S2). In addition, the presence of a functional SBD in the AamA protein of *A. niger* N402 was suggested because of the purification of AamA via its SBD from culture fluid of a N402 $\Delta glaA$ strain, and subsequent demonstration of the SBD with specific antibodies (pers. comm. M.F. Coeffet-Le Gal and D. Archer, University of Nottingham, UK). We therefore PCR amplified the *aamA* gene and its 3' flanking regions using N402 genomic DNA as a template, determined its DNA sequence and compared it to the *aamA* gene and its 3' flanking region from CBS 513.88. The comparison revealed that the *aamA* gene of *A. niger* strain N402 does include a SBD and that the CBS513.88 strain harbours a deletion of 230 nucleotides right after the part encoding the (β/α)₈ barrel, causing a frame shift and the introduction of a stop codon resulting in a truncated protein (Fig. S2). After the deletion, the DNA sequence continues by encoding part of a predicted SBD, but this part is not translated (Fig. S2). The SBD in AamA found in the *A. niger* N402 strain is also predicted to be present in *A. niger* strain ATCC 1015 (Baker, 2006).

Group II contains three putative GPI-anchored enzymes, recently identified as α -glucanotransferases, and named AgtA, AgtB and AgtC, respectively (van der Kaaij et al. 2007a). This subgroup of proteins is characterized by the presence of two hydrophobic signal sequences. The N-terminal signal sequence is predicted to serve for translocation to the endoplasmic reticulum (ER) while the C-terminal sequence is predicted to be replaced by a preassembled glycosylphosphatidylinositol (GPI) anchor in the ER (Orlean 1997). The three enzymes cluster together with the α -amylases in the phylogenetic tree (Fig. 1A), but can be distinguished from the α -amylases by their catalytic domains which are clearly different from the consensus sequence for the α -amylase family. In all three proteins, one or two highly conserved histidines in conserved regions I and IV are replaced by other hydrophilic residues (Table 2). Family GH13 members without these histidine residues, which are part of the active site, are very rare (Uitdehaag et al. 1999). Mutation of these residues generally results in reduced activity or altered reaction specificity of the enzymes (Chang et al. 2003; Leemhuis et al. 2004). Interestingly, both the conserved His-residues in Regions I and IV are also missing in all (putative) α -glucan synthases (Table 2). Other aspergillus genomes harbour two or three Agt homologues (Fig. S1A, Table S2), all sharing the aberrant conserved regions and predicted GPI-anchoring. In the *A. niger* genome, both *agtB* and *agtC* are located next to genes encoding putative α -glucan synthases, and this arrangement of genes is conserved in other aspergilli. The *agtA* gene is constitutively expressed in both the wild type strain N402 and the $\Delta amyR$ strain under all growth

conditions examined (Fig. 2). Expression of *agtB* was only detected 8 h after transfer, regardless of the carbon source and independent of AmyR, while expression of *agtC* was not detected.

Table 2. Alignment of the four conserved regions of all family GH13 enzymes identified in *A. niger*, as well as in four Aah proteins from *S. pombe* and Amy1p from *H. capsulatum*. The seven residues generally conserved in family GH13 are indicated in bold and the three catalytic residues are additionally underlined. The group to which the proteins are assigned, as described in this paper, is indicated.

Enzyme	Region I	Region II	Region III	Region IV
AamA	LMVDVVPNH	DGLRIDSVLE	YCVGEVDN	NFIENHD
AmyA	LMVDVVANH	DGLRIDTVKH	YCIGEVLVD	TFVENHD
AmyB	LMVDVVANH	DGLRIDTVKH	YCIGEVLVD	TFVENHD
AmyC	LMVDVVANH	DGLRVDTVKN	YCIGEVEFD	TFVENHD
AgtA	LLLDVVINN	DGLRIDAAKS	FMTEGVMD	NFIEDQD
AgtB	LMLDIVVGD	DGLRIDSVLN	FTVGEAGAT	TFTANQD
AgtC	LMMDTVINN	DGLRIDAAKH	FMTEGVLDQ	SFSENHD
AmyD	IYWDVLNH	SGMRIDAVKH	FIVGEYWK	TFVANHD
AmyE	VLWDVLNH	SGMRIDAAKH	FVIGEYWS	TFVTNHD
AgdC	LLMDLVVNH	DGFRMDVINP	FSVGEIMP	LYWENHD
AgdD	LMMDLVVNH	CGFRMDVINP	ITVGEETPY	IFLECHD
GbeA	VLLDVVHSH	DGFRFDGVTS	ITVAEDVS	AYAESH
GdbA	SLTDVVVNH	SGFRIDNCHS	TVFAELFT	FMDCTHD
AgsA	VIMDNTLAT	DGFRFDKAVQ	FLPGEITS	YGVSNQD
AgsB	VLFDNTFGT	DGFRVDKALQ	YIPGEIVS	FGVTNQD
AgsC	VIFDNTLAT	DGFRYDKATQ	FIAGEITG	YGVTNQD
AgsD	VIFDNTLAT	DGFRYDKAIQ	FLPGEITG	YGVTNQD
AgsE	VIFDNTIAT	DGFRYDKATQ	FIAGEITG	YGATNQD
HcAmy1	I IWDTVLNH	SGLRIDAAKH	LLVAEYWK	TFVMNHD
SpAah1	IMFDALANS	DGIRIDAVKQ	FAIGEIMFS	NFLENHD
SpAah2	ILLDVAINS	DGIRFDKAKH	FTIGEYFT	TFIGNHD
SpAah3	VMLDSIVNS	DGLRIDAVKM	YSVGEVFS	TFIENHD
SpAah4	LMVDVAINH	DGIRFDAMGD	FCMGDLKS	NFVENHD

Group III consists of two putative, intracellular α -glucosidases, named AgdC (An02g13240) and AgdD (An13g03710). The protein sequences contained all residues commonly conserved in the α -amylase superfamily (Table 2). The predicted intracellular proteins lack clear similarity to any previously characterized fungal protein, although similar enzymes are predicted in *A. oryzae*, *A. nidulans* and *A. fumigatus* (Fig. S1A). Their most related functionally characterized homologue is an α -glucosidase from *Bacillus* sp. SAM1606 (Nakao et al. 1994). Expression of *agdC* was low and not induced on maltose, and expression of *agdD* was not observed (Fig. 2).

Group IV contains two putative intracellular α -amylases, named AmyD (An01g13610) and AmyE (An09g03110), which share 55% identity. AmyD was recently characterized as an α -amylase with low hydrolyzing activity on starch and related substrates (Van der Kaaij et al. 2007b). AmyD and AmyE are similar to a recently identified protein Amy1p from *H. capsulatum*, which was shown to be involved in cell wall α -glucan synthesis. The latter protein has not been characterized biochemically. Functionally characterized enzymes with similarity to this cluster therefore included only bacterial enzymes, of which maltohexaose-forming α -amylase of alkalophilic *Bacillus* sp. #707 (Tsukamoto et al. 1988) had the highest similarity. No significant expression of *amyD* or *amyE* was detected in our experiments (Fig. 2). Predicted enzymes highly similar to *A. niger*

AmyD and AmyE were also present in other *Aspergillus* species (Fig. S1A, Table S2). Like the *A. niger amyE* gene, the orthologues in *A. nidulans* (AN3309.3) and in *A. oryzae* (AO003001497) are clustered in the genome with genes encoding α -glucanotransferases and α -glucan synthases (Table S2).

The two *A. niger* proteins in Group V could be reliably annotated as enzymes involved in glycogen metabolism: glycogen branching enzyme (GbeA) and glycogen debranching enzyme (GdbA). A homologue for each of these enzymes is present in the other *Aspergillus* genome sequences. Transcriptional analysis in *A. niger* showed that both genes were expressed both on xylose and maltose and that their expression was unaffected in the $\Delta amyR$ strain (Fig. 2). Group VI contains the five predicted α -glucan synthase genes (Damveld et al. 2005). The derived proteins are highly similar (66- 83%) to each other and all contain the two catalytic domains (GT5 and GH13) characteristic for these proteins. Both *agsB* and *agsC* are clustered in the genome with genes encoding the above mentioned α -glucanotransferases *agtC* and *agtB*, respectively. In both cases, the direction of transcription of the pair of genes is such that they can be transcribed from their intergenic region. Expression analysis in our microarray data collection, which includes data from various time points on several carbon sources (glucose, maltose, xylose, inulin, sucrose) indicate that *agsC* and *agtB* are co-expressed at later growth stage, independent of the carbon source (data not shown). Expression of both *agsB* and its neighboring gene *agtC* was not detected in any of the growth conditions. The expression levels of *agsA* and *agsD* were very low, or not detectable. *AgxE* was highly expressed in all experiments independent of AmyR (Fig. 2). As expected, several proteins with high similarity to the *A. niger* α -glucan synthases are predicted from the other *Aspergillus* genomes, although the number of homologues present is highest in *A. niger* (Fig. S1A, Table S2).

Identification and transcriptional regulation of GH15 family members

A HMMer search for GH15 family members in the *A. niger* genome returned two predicted proteins: the previously described glucoamylase GlaA (Boel et al. 1984) and an unknown predicted protein named GlaB (An12g03070). The GlaB protein lacks both an N-terminal signal sequence and an SBD, and displays a low similarity of 26 % with GlaA. These two types of family GH15 enzymes are also recognized in other aspergilli studied (Fig. S1B). All predicted proteins similar to GlaB lack both a signal for secretion and an SBD, features typically present in the fungal GH15 enzymes described to date. *A. niger glaA* was expressed on xylose and strongly induced on maltose, as described previously (Fowler et al. 1990). The induction of *glaA* was AmyR dependent. Expression of *glaB* was not detected in any of the conditions tested (Fig. 2).

Identification and transcriptional regulation of GH31 family members

HMMer searches in the genome of *A. niger* revealed the presence of seven GH31 family members, of which only one was previously identified as an α -glucosidase (*aglA*) (Nakamura et al. 1997). We propose to name all (putative) α -glucosidases in *A. niger* Agd

enzymes, similar to the nomenclature in *A. nidulans*, and to rename AglA as AgdA, to prevent confusion with α -galactosidases (den Herder et al. 1992). The phylogenetic tree (Fig. 1B) shows the presence of (at least) 4 subgroups within family GH31. In the group of the α -glucosidases, AgdA clusters with AgdB, a predicted extracellular protein. AgdB has some similarity to two *A. nidulans* enzymes: AN7345.2 (62% identity) with both α - and β -glucosidase activity (Bauer et al. 2006) and α -glucosidase B (AndAgdB/ AN8953.3, 53% identity), with strong transglycosylation activity (Kato et al. 2002b). However, none of the true orthologues of AgdB has been characterized (Fig. S1C). Transcript analysis revealed that *agdB* was regulated similar to *agdA*, as the strong induction of both genes in the presence of maltose was dependent on the presence of *amyR*. However, where most AmyR regulated genes (*aamA*, *glaA* and *agdA*) reached their highest level of induction after 8 h growth on maltose, the expression level of *agdB* was decreased after 8 h compared to 2 h (Fig. 2).

From the remaining five GH31 family members, AgdE (An09g05880) shows similarity to *Trichoderma reesei* glucosidase II (TrAguII) (Geysens et al. 2005). This type of α -glucosidases is located in the ER where it is involved in the trimming of α -(1,3)-linked glucose residues from N-glycan core structure Glc3Man9GlcNAc2, which may be attached to proteins designated to be secreted (Geysens et al. 2005). AgdE contains a predicted signal sequence which might serve to direct the protein to the ER. The *agdE* gene was constitutively expressed in all conditions tested (Fig. 2), consistent with its predicted function as an α -glucosidase II. Two additional GH31 family members, AgdF (An18g05620) and AgdG (An07g0350), lack similarity to any functionally described proteins. Expression of both genes was not detected in any of the tested growth conditions.

The two final family GH31 members, named AxlA and AxlB, are highly similar to AN7505.2 from *A. nidulans*, which was recently characterized as an α -xylosidase (Bauer et al. 2006). Gene *axlA* was highly expressed in the presence of xylose, while no expression was detected in maltose grown cultures (Fig. 2). The high expression of *axlA* on xylose further supports its putative function as a xylanolytic enzyme. The gene encoding the putatively intracellular AxlB was expressed at a very low level independent of the carbon source (Fig. 2).

A BLASTP search in the genomes of *A. fumigatus*, *A. nidulans* and *A. oryzae* for family GH31 enzymes resulted in a similar collection of enzymes as identified in *A. niger* (Fig. S1C). Several clusters of orthologous proteins are distinguishable, but the assignment of an enzymatic activity to these clusters is not yet possible due to a lack of well studied homologues for these enzymes, thus requiring biochemical studies in our future work.

Genome wide analysis of AmyR dependent maltose induced genes using microarrays

The expression analysis of genes encoding family GH13, GH15 and GH31 enzymes in the *A. niger* genome revealed that the expression of only a limited number of genes was induced by maltose in an AmyR dependent way. In fact, only four genes (*aamA*, *glaA*, *agdA*, and *agdB*) showed the predicted expression pattern for genes encoding enzymes involved in the breakdown of starch (Fig. 2). To identify additional genes with a possible role in starch

metabolism, the expression of all 14,509 predicted open reading frames on the Affymetrix microarray chips was analyzed. Comparison of the transcriptome of the wild type strain (N402), grown for 2 h after transfer on either xylose or maltose, identified 634 genes that were significantly induced by maltose (One-way ANOVA analysis $p < 0.08$ and > 2 -fold change in expression level). A set of 12 genes was expressed > 2 -fold higher in N402 compared to the $\Delta amyR$ strain (One-way ANOVA analysis $p < 0.08$) after the transfer of these strains from xylose to maltose. Combining the two gene sets resulted in a collection of 6 genes that were > 2 -fold induced on maltose in an AmyR dependent manner (Fig. 3, Table 3). Five of these were assigned to the category of carbohydrate transport and metabolism according to FunCat (Ruepp et al. 2004), including the three genes encoding known extracellular starch degrading enzymes (*aamA*, *glaA* and *agdA*) and a putative α -glucosidase (*agdB*) (Table 3A). Gene An15g03940, encoding a protein with high similarity (68%) to a *Candida intermedia* glucose/xylose symporter (Leandro et al. 2006), was 2.3-fold induced by maltose; this induction was not observed in the $\Delta amyR$ strain. The sixth induced gene (An11g02550) encodes a protein highly similar (72%) to *Kluyveromyces lactis* phosphoenolpyruvate carboxykinase (Kitamoto et al. 1998) which functions in gluconeogenesis by catalyzing the conversion of oxaloacetate into phosphoenolpyruvate.

A similar comparative analysis was performed for samples taken for cultures that had grown for 8 h after the transfer. A total of 28 genes were significantly induced by maltose in comparison to xylose, and 161 genes were 2 fold higher expressed in N402 compared to the $\Delta amyR$ strain (One-way ANOVA analysis $p < 0.08$). By combining the two sets, we identified 18 genes which were induced by maltose and whose induction was AmyR dependent (Fig. 3, Table 3B). Nine of these genes encode proteins involved in carbohydrate transport and metabolism, from which 5 were also identified as differentially expressed after 2 h (*aamA*, *glaA*, *agdA*, *agdB*, and An15g03940). Additionally identified genes included *amyC*, encoding an extracellular α -amylase, three genes encoding putative sugar transporters, and several other genes belonging to various functional categories (Table 3). The results from the microarray experiments were validated using Northern blot analysis for a selected number of amylolytic genes. These genes include *glaA*, *aamA*, *agdB* and *amyC*. As shown in supplementary Fig. 3, the results of the Northern hybridizations are in good agreement with the microarray data (Suppl. Table 1). The Northern analysis also revealed an additional mRNA of a larger size for the *amyC* gene, 8 h after transfer. The detection of two different sized mRNAs suggests two different mRNA start sites or different polyadenylation sites, but the exact sequence of the different mRNAs and possible consequences for the gene model have not been addressed so far.

Interestingly, the transcription factor gene *amyR* was also induced 2 and 8 h after transfer to maltose compared to the transfer to xylose (2.6 fold, P-value 0.014; 2.8-fold, P-value 0.027, respectively), indicating transcriptional regulation of the *amyR* gene itself. As the AmyR transcription factor is per se missing in the deletion strain, it is not appropriate to include the *amyR* gene in the group of maltose induced, AmyR dependent genes. However, the 1-kb promoter region of the *amyR* gene contains two AmyR binding motifs, allowing for the possibility that AmyR can induce its own expression.

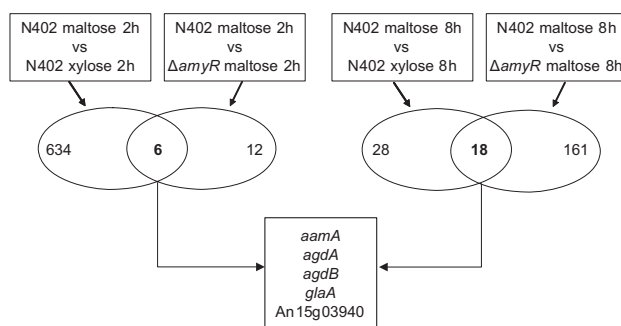


Fig. 3. Results of microarray analysis for maltose induced and AmyR dependent genes. Venn-diagram showing the number of genes induced on maltose compared to xylose in *A. niger* N402, and the number of genes induced in N402 compared to the $\Delta amyR$ deletion strain. The number of genes both induced by maltose and dependent on AmyR is indicated in bold. The maltose induced and AmyR dependent genes which are present in both 2 h and 8 h after transfer from a preculture are shown in the box.

Discussion

In the present study, we have mined the full genome sequence of *A. niger* strain CBS 513.88 (Pel et al. 2007) for the presence of glycoside hydrolases belonging to families GH13, GH15 and GH31. Members of these protein families in aspergilli, including *A. niger*, have been studied extensively, mainly because of their industrial relevance. Nevertheless, our study revealed the presence of several groups of enzymes that had not been identified previously. These novel enzymes are conserved among several *Aspergillus* species as well as other Ascomycetes, indicating that they may play an important role in fungal metabolism.

The GH13 family in *A. niger* contains six separate groups of amylase-type enzymes (Fig. 1), of which three groups had not been described thus far. The best described group comprises the extracellular α -amylases, and is now extended with AmyC in addition to the three known α -amylases. Overexpression of the *amyC* gene in *A. niger* resulted in increased levels of endo- α -amylase activity in the medium, indicating that this gene indeed encodes an extracellular α -amylase (Van der Kaaij and Yuan, unpublished results). The relatively low expression of *amyC* compared to other starch degrading enzymes may explain why this protein has not been identified previously. The localization of the *amyC* gene in the genome is noteworthy, as it is part of a small cluster of amylolytic genes (with *agdA*) and their transcriptional regulator gene *amyR*. A similar organization is observed in the genomes of other aspergilli such as *A. nidulans* and *A. oryzae* RIB40 (Gomi et al. 2000). In *A. fumigatus* the same cluster is extended with a glucoamylase (similar to GlaA). Gene clusters of transcriptionally co-regulated genes in filamentous fungi are often involved in the same process, e.g. secondary metabolite production (Woloshuk et al. 1995) or catabolism of amino acids (Hull et al. 1989). A possible function of AmyC might be to act as a scouting enzyme for the presence of starch, resulting in the subsequent activation of AmyR by starch derived molecules (maltose or isomaltose). Alternatively, AmyC might be regulated by another system additional to AmyR, e.g. pH regulated expression, or its expression might be upregulated locally, e.g. in hyphal tips.

Table 3A: Overview of 6 maltose induced and AmyR dependent genes in *A. niger* at 2 h after transfer from preculture.

ORF no. ^a	N402 maltose ^b	call ^c	N402 xylose	call	Δ amyR maltose	call	Fold induction maltose vs xylose	N402 vs Δ amyR	P-value	Putative function
<i>Carbohydrate transport and metabolism</i>										
An11g03340	3.15 $\pm$ 0.12	P	0.08 $\pm$ 0.04	A	0.13 $\pm$ 0.09	A	38.6	24.2	0.080	acid alpha-amylase AamA
An04g06920	85.96 $\pm$ 22.54	P	6.24 $\pm$ 0.98	P	5.03 $\pm$ 0.03	P	13.8	17.1	0.064	extracellular alpha-glucosidase AgdA
An01g10930	96.50 $\pm$ 31.50	P	10.47 $\pm$ 2.59	P	8.35 $\pm$ 2.13	P	9.2	11.5	0.062	extracellular alpha-glucosidase AgdB
An03g06550	117.80 $\pm$ 43.16	P	13.78 $\pm$ 4.14	P	5.85 $\pm$ 0.47	P	8.5	20.1	0.072	glucoamylase GlaA
An15g03940	152.30 $\pm$ 19.00	P	65.74 $\pm$ 3.04	P	70.59 $\pm$ 2.88	P	2.3	2.2	0.074	putative monosaccharide transporter
<i>Energy</i>										
An11g02550	58.87 $\pm$ 3.86	P	10.13 $\pm$ 1.41	P	25.27 $\pm$ 5.35	P	5.8	2.3	0.033	putative phosphoenolpyruvate carboxykinase

^a Gene name in bold indicates that the gene is >2-fold induced on maltose in an AmyR dependant way after both 2 and 8 hours.

^b The expression level was based on the geometric mean value of the duplicate samples and the deviation values between the duplicate samples are indicated. The P-value was based on 1-way ANOVA analysis

^c P, M or A represent detection calls for present, marginal or absent, respectively. Genes were divided over different functional groups according to FunCat (Ruepp et al. 2004) .

Table 3B. Overview of the 18 maltose induced and AmyR dependent genes in *A. niger* at 8 h after transfer from preculture.

ORF no. ^a	N402 maltose ^b	call ^c	N402 xylose	call	Δ amyR	maltose	call	Fold induction maltose vs xylose	Fold induction N402 vs Δ amyR	P-value	Putative function
<i>Carbohydrate transport and metabolism</i>											
An11g03340	65.26 $\pm$ 7.48	P	0.58 $\pm$ 0.57	A	0.08 $\pm$ 0.01	A	A	113.0	808.4	0.004	acid alpha-amylase Aama
An01g10930	49.46 $\pm$ 3.42	P	2.37 $\pm$ 0.08	P	6.84 $\pm$ 1.99	P	P	20.9	7.2	0.006	extracellular alpha-glucosidase AgdB
An04g06920	179.10 $\pm$ 13.90	P	9.89 $\pm$ 3.63	P	3.36 $\pm$ 0.46	P	P	17.6	51.7	0.012	extracellular alpha-glucosidase AgdA
An02g03540	62.79 $\pm$ 8.36	P	8.81 $\pm$ 0.61	P	0.96 $\pm$ 0.82	P	P,A	6.7	61.9	0.012	putative hexose transport protein MstC
An15g03940	144.40 $\pm$ 6.80	P	35.93 $\pm$ 2.72	P	3.04 $\pm$ 0.03	P	P,M	4.0	47.6	0.003	putative monosaccharide transporter
An03g06550	239.30 $\pm$ 2.05	P	83.03 $\pm$ 4.35	P	19.03 $\pm$ 1.75	P	P	2.9	12.6	0.012	glucoamylase Glaa
An04g06930	4.36 $\pm$ 0.50	P	1.50 $\pm$ 0.16	P	1.40 $\pm$ 0.03	P	P	2.9	3.1	0.062	extracellular alpha-amylase AmyC
An09g04810	18.47 $\pm$ 0.44	P	6.46 $\pm$ 0.35	P	2.52 $\pm$ 0.02	P	P	2.9	7.4	0.003	putative hexose transporter
An12g07450	137.60 $\pm$ 13.00	P	51.57 $\pm$ 11.82	P	52.50 $\pm$ 6.31	P	P	2.7	2.6	0.064	sugar/H ⁺ symporter Msta
<i>Energy</i>											
An03g06270	4.72 $\pm$ 0.29	P	1.59 $\pm$ 0.14	A	0.20 $\pm$ 0.25	A	A	3.3	25.6	0.035	putative isoamyl alcohol oxidase
An16g06010	5.18 $\pm$ 1.55	P	1.80 $\pm$ 0.39	P	0.22 $\pm$ 0.02	A	A	2.9	23.8	0.032	putative phosphoglycerate mutase
<i>Amino acid transport and metabolism</i>											
An03g00280	8.35 $\pm$ 0.90	P	3.57 $\pm$ 0.74	P,A	0.12 $\pm$ 0.06	A	A	2.3	68.0	0.050	similarity to tyrosinase protein
<i>Cell rescue, defence and virulence</i>											
An14g05730	3.99 $\pm$ 0.59	P	0.23 $\pm$ 0.03	A	1.37 $\pm$ 0.73	P,A	P,A	18.8	3.2	0.037	similarity to integral membrane protein
An03g00290	5.96 $\pm$ 0.89	P	1.68 $\pm$ 0.47	P	0.64 $\pm$ 0.21	P	P	3.5	9.3	0.068	similarity to integral membrane protein
An06g00490	8.14 $\pm$ 0.89	P	3.81 $\pm$ 0.30	P,A	2.23 $\pm$ 0.44	A	A	2.1	3.7	0.066	similarity to integral membrane protein
<i>Protein fate</i>											
An18g04260	5.44 $\pm$ 0.40	P	2.57 $\pm$ 0.07	P	1.77 $\pm$ 0.28	M,A	M,A	2.1	3.1	0.056	similarity to UDP-galactose transporter
<i>unclassified</i>											
An16g01290	15.04 $\pm$ 0.28	P	5.22 $\pm$ 1.69	P,A	1.97 $\pm$ 0.26	A	A	2.6	6.9	0.027	unknown
An09g06130	12.62 $\pm$ 2.51	P	5.41 $\pm$ 0.93	P	2.69 $\pm$ 0.30	P,A	P,A	2.3	4.7	0.059	unknown

^{a,b,c} see for footnotes Table 3A

The lack of detectable expression of *amyA* and *amyB* in *A. niger* N402 is puzzling. Both genes were expressed in *A. niger* strain CBS 513.88 in a fed-batch fermentation using glucose as a carbon source (Pel et al. 2007), but batch cultivation of *A. niger* N402 on glucose did not result in detectable expression in microarray studies (E. Martens, H. Kools, P. Schaap, pers. comm.). These results indicate a difference between both strains with regard to either the presence of *amyA/B* in *A. niger* strain N402 or in the transcriptional regulation of *amyA/B*. In the recently released genome sequence of *A. niger* strain ATCC 1015 from the Joint Genome Institute (Baker, 2006) we could not identify orthologues of the AmyA/AmyB proteins suggesting that these genes might be actually absent from the N402 strain. The absence, or lack of expression of *amyA/amyB*, is consistent with the observation that the *A. niger* MGG029 Δ *aamA* strain (N402 background) with gene disruptions for both *aamA* and *glaA* grows very poorly on starch (Weenink et al. 2006), confirming that AmyA and AmyB do not contribute significantly to the amylolytic potential of N402.

In this study we have identified three new groups within the GH13 family that had not been described before in *Aspergilli*: α -amylase-like proteins predicted to be GPI-anchored enzymes (group II) (subfamily GH13_1, as defined by Stam et al. 2006), intracellular α -glucosidases (group III) (unknown subfamily) and intracellular α -amylases (group IV) (subfamily GH13_5). Detailed biochemical studies of two GPI-anchored proteins revealed that they are α -glucanotransferases, acting mainly on maltooligosaccharides (van der Kaaij et al. 2007a). The expression of these genes was not regulated in response to maltose metabolism or AmyR. The Agt enzymes are homologues of a group of four GPI-anchored amylase-like proteins recently identified in *S. pombe* (Aah1-4p), which have similarly aberrant conserved regions (Table 2). A deletion of one of these homologues resulted in aberrant morphology of the cell and an increased sensitivity towards cell wall degrading enzymes (Morita et al. 2006). A similar phenotype was observed for an *A. niger* knockout of *agtA* (van der Kaaij et al. 2007a). Such a phenotype can be explained if α -glucan synthesis is affected. The clustering of *agt* genes with genes encoding α -glucan synthases in all four *Aspergillus* species studied, and the expression pattern of the *A. niger agt* genes both support the hypothesis that Agt enzymes play a role in α -glucan synthesis.

A. niger contains two putative intracellular α -amylases, AmyD and AmyE, that belong to the recently identified subfamily GH13_5 (Stam et al. 2006). A homologous gene in *H. capsulatum*, *AMY1*, was shown to be essential for the formation of α -(1,3)-glucan in the cell wall (Marion et al. 2006). The possibility of a role for subfamily GH13_5 enzymes in the cell wall of *Aspergilli* is strengthened by their location in the genome: *A. niger*, *A. nidulans* and *A. oryzae* all contain a cluster of genes, containing a homologue of *AMY1* as well as *ags* and *agt* genes (Table S2). Both the GPI-anchored α -glucanotransferases and the family GH13_5 α -amylase like proteins are only present in fungi with cell wall α -glucan and not in the true yeasts, which lack this type of cell wall glucans. However, a study with heterologously expressed AmyD did not confirm a role for this enzyme in α -(1,3)-glucan formation, as the enzyme only showed a (low) hydrolyzing activity on substrates with α -(1,4) glycosidic bonds *in vitro* (van der Kaaij et al. 2007b). The expression levels of *amyD*

and *amyE* were below detection level, and thus could not confirm nor refute the possible involvement of these proteins in cell wall α -glucan synthesis in *A. niger*.

At this moment, no function can be assigned to the predicted intracellular α -glucosidases AgdC and AgdD and their homologues identified in other fungi. The enzymes show similarity to a *Bacillus* α -glucosidase as well as to a maltase from *C. albicans*, and might therefore be involved in the intracellular hydrolysis of maltose. Alternatively, these enzymes could play a role in trehalose metabolism, as they show some similarity to bacterial trehalose-6-phosphate hydrolases. Trehalose is known to be a common reserve carbohydrate in fungi including *A. niger* (Wolschek and Kubicek 1997; Arguelles 2000).

A surprising finding from the genome mining of *A. niger* was the high number of GH enzymes with a predicted intracellular location. This group included not only the starch-acting enzymes described here, but also other predicted intracellular glycoside hydrolases such as β -glucosidase, chitinase, β -mannosidase, β -xylosidase, rhamnosidase and invertase, present in various *Aspergillus* genomes (Pel et al. 2007; Goosen et al. 2007). Most of these proteins were predicted to be exo-acting, i.e. releasing terminal mono- or disaccharide residues from the non-reducing end of the substrate. In *A. niger*, the genes encoding these proteins are expressed at a relatively low level compared to the extracellular starch degrading enzymes, sometimes even below the detection limit. In a recent publication, expression of intracellular chitinases in *A. nidulans* was found to be induced during the autolysis phase of the culture (Yamazaki et al. 2007). A role for these and other predicted intracellular proteins during the autolysis phase would explain their low expression levels observed in *A. niger* during vegetative growth, as well as their lack of secretion signals.

The genome wide expression analysis identified only five genes that were >2-fold induced by maltose in an AmyR dependent way, 2 and 8 h after the transfer to maltose. Among them are genes encoding known starch degrading enzymes (*aamA*, *glaA*), two (putative) α -glucosidases (*agdA*, *agdB*), and a putative sugar transporter protein (An15g03940). The specificity of the putative sugar transporter encoded by the An15g03940 gene has not been examined, but its expression profile suggests that it might function as a glucose or maltose transporter. Three additional sugar transporters were higher expressed on maltose than on xylose after 8 h (Table 3B). One of them, An12g07450/MstA has been functionally characterized as a high affinity glucose transporter (Vankuyk et al. 2004) which is in line with its expression profile. The two additional sugar transporters encoding genes (An02g03540/*mstC* and An09g04810) showed an interesting expression pattern, with induced expression on maltose but also relatively high expression levels on xylose (Table 3B). In the $\Delta amyR$ strain both their expression levels were strongly reduced, also relative to the xylose conditions, indicating that the deletion of AmyR might have had an indirect effect. A possible explanation is that disruption of the AmyR transcription factor resulted in low levels of extracellular enzymes converting maltose to glucose, and consequently stress due to low glucose levels. The low availability of glucose might be a signal to downregulate *mstC* and An09g04810 if these genes encode low affinity sugar transporters. Interestingly, the *amyR* gene itself was also induced by the presence of maltose

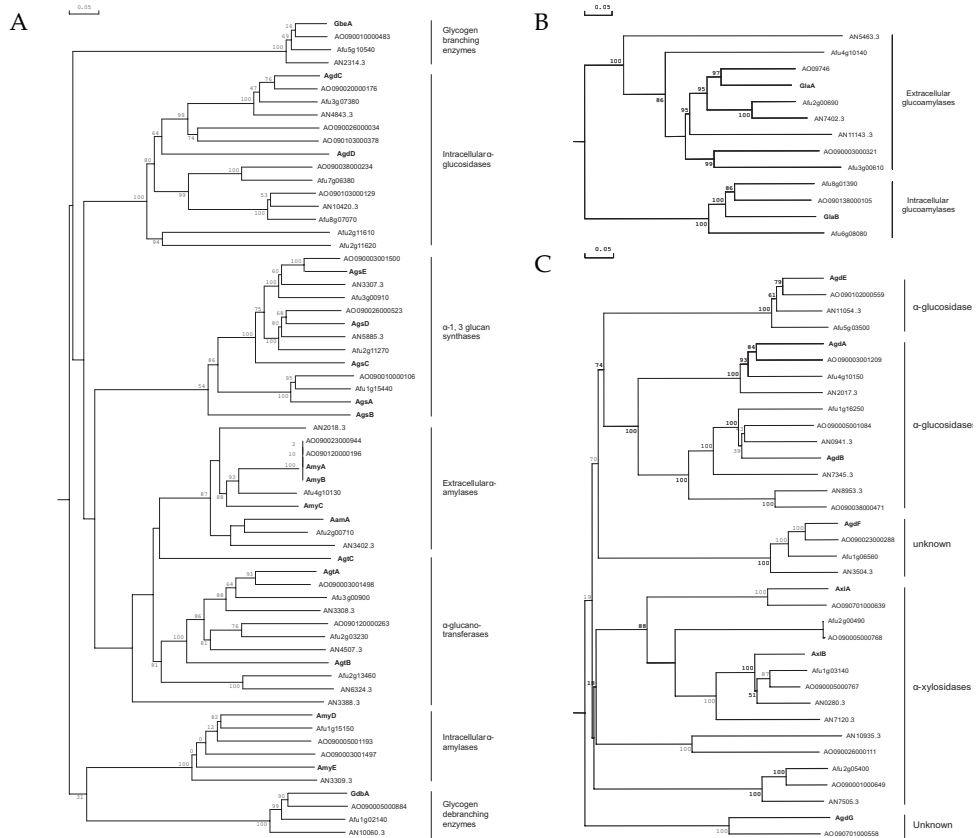
suggesting that its regulation takes place, at least partly, at the transcriptional level. Furthermore, conserved AmyR binding motifs are present in the promoters of all co-regulated genes, including the 1-kb *amyR* promoter region, with the exception of *aamA* (Table 1A). Possibly the *aamA* promoter contains a currently unknown AmyR binding motif.

The results presented in this paper suggest that *A. niger* can metabolize maltose by inducing the expression of a limited amount of enzymes and sugar transporters. Other members of the GH13, GH15 and GH31 protein families might function in starch metabolism in different conditions (solid state growth, or influenced by pH, nitrogen metabolism, etc.), or they may be involved in other processes such as the synthesis of cell wall α -glucan, or the glycosylation of proteins. To determine the exact function of these enzymes will require further biochemical characterization in combination with the detailed analysis of gene deletion mutants.

Acknowledgements

We thank Harrie Kools (Wageningen University, The Netherlands) for performing the HMM searches, Patricia van Kuyk (Wageningen University, The Netherlands) for the *amyR* plasmid, Marie-Francoise Coeffet-Le Gal and David Archer for sharing unpublished results. This work was financially supported by SenterNovem in the framework of the Dutch IOP-Genomics program (project IGE1021).

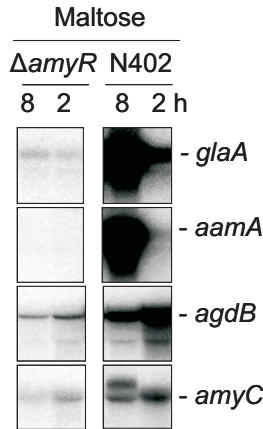
Supplementary Data



Suppl. Fig. 1. Bootstrapped phylogenetic trees of family GH13 (A), GH15 (B), and GH31 (C) enzymes from *A. niger*, *A. fumigatus*, *A. nidulans* and *A. oryzae*. Bootstrap values are indicated on the node of each branch. The tree was created with DNAMAN 4.0 using gap and extension penalties of 10 and 0.5 respectively. The scale bar corresponds to a genetic distance of 0.05 substitution per position. A description of each protein is listed in Supplementary Table 2.

A	<i>A.foetidis</i>	ACGTCGGGCACGAAGCTCATCGAAGCGTACACCTGCACGTCCTGACGGTGGACTCGAAC
	<i>A.vadensis</i>	ACGTCGGGCACGAAGCTCATCGAAGCGTACACCTGCACGTCCTGACGGTGGACTCGAAC
	<i>A.niger</i> N402	ACATCCGGGCACGAAGCTGATCGAAGCGTATACATGCACGTCCTGACGGTGGACTCGAAGC
	<i>A.niger</i> DSM	ACATCCGGGCACGAAGCTGATCGAAGCGTACACATGCACATCCGTGACCGTGGACTCGAGC
	<i>A.niger</i>	ACATCCGGGCACGAAGCTGATCGAAGCGTACACATGCACATCCGTGACCGTGGACTCGAGC
	<i>A.foetidis</i>	GGGGATATCCCTGTGCCGATGGCTTCGGGGATTACCTAGAGTTCTCTCCCTGCTTCGGTG
	<i>A.vadensis</i>	GGGGATATCCCTGTGCCGATGGCTTCGGGGATTACCAAGGGTTCTTCTGCTGCTTCGGTG
	<i>A.niger</i> N402	GGCGATATCCCGGTGCCGATGGCGTCGGGATTACCGAGAGTTCTTCTCCCGCGTCGCTC
	<i>A.niger</i> DSM	GGCGATATTCCTCGGTGCCGATGGCGTCGGGATTACCGAGAGTTCTTCTGCGCGCTCCGTC
	<i>A.niger</i>	GGCGATATTCCTCGGTGCCGATGGCGTCGGGATTACCGAGAGTTCTTCTGCGCGCTCCGTC
	<i>A.foetidis</i>	GTTGATAGTCTCTCGCTTTGTGGGGGAGTGGTA.....ACACAACCAACGACCAAACT
	<i>A.vadensis</i>	GTTGATAGTCTCTCGCTTTGTGGGGGAGTGGTA.....CACAACTACAACC
	<i>A.niger</i> N402	GTCGATAGCTCTTCGCTCTGTGGCGGGAGCGGAAGTAATCTCAACTACAACCAACA
	<i>A.niger</i> DSM	GTCGATAGCTCTTCGCTCTGTGGCGGGAGCGGAAG.....
	<i>A.niger</i>	GTCGATAGCTCTTCGCTCTGTGGCGGGAGCGGAAG.....
	<i>A.foetidis</i>	GCTGCTACCTCCACATCCAAAGCCACACCTCCTCTTCTTCTTCTGCTGCTGCTACT
	<i>A.vadensis</i>	ACAACCTACCTCATCTTCCACGGCTACCTCCACCTCCACATCAAAACCAACACAT.....
	<i>A.niger</i> N402	ACAGCTACCTCATCTTCCACTCGACATCCAAATCGCATCAACCTCGTCTA.....
	<i>A.niger</i> DSM	
	<i>A.niger</i>	
	<i>A.foetidis</i>	ACTTCTTCATCATGCACCGCAACAAGCACACCTTCCCATCACTTCGAAAGACTCGTC
	<i>A.vadensis</i>	.CCTCTTCATCATGCACCGCAACAAGCACACCTTCCCATCACTTCGAAAGACTCGTA
	<i>A.niger</i> N402	.CGTCGACGGCATGCACAGCTACCTTACCTCCCTCGCGGTACGTTTCGAAAGACTCGTC
	<i>A.niger</i> DSM	
	<i>A.niger</i>	
	<i>A.foetidis</i>	ACCACCTACCTACGGGAAGAAGTCTACCTCAGCGGATCTATCTCCAGCTCGGAGAGTGG
	<i>A.vadensis</i>	ACCACACCTACGGGAAGAAGTCTACCTGAGCGGATCTATCTCCAGCTCGGGAGTGG
	<i>A.niger</i> N402	ACGACTACCTACGGGAAGAAATCTACCTGAGCGGATCGATCTCCAGCTTCGGGACTGG
	<i>A.niger</i> DSM	
	<i>A.niger</i>	
	<i>A.foetidis</i>	GATACGAGTGACCGGTGAAGTTGTCCGCGGATGATTATACCTCGAGTAACCCGAGTGG
	<i>A.vadensis</i>	GATACAAGTGACCGGTGAAGTTGTCCGCGAATGATTATACCTCGAGTAACCCGAGTGG
	<i>A.niger</i> N402	GATACGAGTGATCGGTGAAGTGTCCGCGGATGATTATACCTCGAGTAATCCGAGTGG
	<i>A.niger</i> DSM	ATTATACGTCGAGTAATCCGAGTGG
	<i>A.niger</i>	ATTATACGTCGAGTAATCCGAGTGG
	<i>A.foetidis</i>	TCGTGTACTGTGCTGTGCCGGTGGGGACGACCTTCGAGTATAAGTTTATTAAGTCGAT
	<i>A.vadensis</i>	TATGTGTACTGTGCTGTGCCGGTGGGGACGACCTTCGAGTATAAGTTTATTAAGTCGAG
	<i>A.niger</i> N402	TCGGTTACTGTGACGTTGCCGGTGGGGACACCTTCGAGTATAAGTTTATTAAGGTGAG
	<i>A.niger</i> DSM	TCGGTTACTGTGACGTTGCCGGTGGGGACCACTTCGAGTATAAGTTTATTAAGTGGAG
	<i>A.niger</i>	TCGGTTACTGTGACGTTGCCGGTGGGGACCACTTCGAGTATAAGTTTATTAAGTGGAG
	<i>A.foetidis</i>	GAGGCTGGAAGCTGTGACTTGGGAAGTGATCCGAATAGGAGTATACGTGCTCCTGAATGT
	<i>A.vadensis</i>	GAGGATGGGAGTGTACTTGGGAAGTGATCCGAATAGGAGTATACGTGCTCCTGAATGT
	<i>A.niger</i> N402	TCGGATGGGAGTGTACTTGGGAGAGTGATCCGAATCGGAGTATACGGTCCCGAGTGT
	<i>A.niger</i> DSM	<u>TCGGATGGGAGTGTACTTGGGAAAGTGATTGAATCGGAGTATACGGTCCCGAGTGT</u>
	<i>A.niger</i>	<u>TCGGATGGGAGTGTACTTGGGAAAGTGATTGAATCGGAGTATACGGTCCCGAGTGT</u>
	<i>A.foetidis</i>	GGGAGTGGGAGTGGGGAGACGGTGTGTTGATACGTGGAGGTAG
	<i>A.vadensis</i>	GGGAGTGGG.....GAGACGGTGTGTTGATACGTGGAGGTAG
	<i>A.niger</i> N402	GGGAGTGGG.....GAGACGGTGTGTTGATACGTGGAGGTAG
	<i>A.niger</i> DSM	<u>GGGAGTGGG.....GAGACGGTGTGTTGATACGTGGAGGTAG</u>
	<i>A.niger</i>	<u>GGGAGTGGG.....GAGACGGTGTGTTGATACGTGGAGGTAG</u>
B	<i>A.foetidis</i>	TSGTKLIEAYTCTSVTVDSNGDIPVPMASGLPRVLLPASVVDSSSLCGSGS.N..TTTT
	<i>A.vadensis</i>	TSGTKLIEAYTCTSVTVDSNGDIPVPMASGLPRVLLPASVVDSSSLCGSGS...STTTTT
	<i>A.niger</i> N402	TSGTKLIEAYTCTSVTVDSNGDIPVPMASGLPRVLLPASVVDSSSLCGSGSNSSTTTTT
	<i>A.niger</i> DSM	TSGTKLIEAYTCTSVTVDSNGDIPVPMASGLPRVLLPASVVDSSSLCGSGSRLIVE*..
	<i>A.niger</i>	TSGTKLIEAYTCTSVTVDSNGDIPVPMASGLPRVLLPASVVDSSSLCGSGSRLIVE*..
	<i>A.foetidis</i>	TAATSTSKATTSSSSSSAAATSSSCTATSTTLPTFEELVTTTYGEEVYLSGISQLGE
	<i>A.vadensis</i>	T..TSSSTAT..STSTSKTTTSSSCTATSTTLPTFEELVTTTYGEEVYLSGISQLGE
	<i>A.niger</i> N402	T.ATSSSTAT..SKSAS..TSTSTACTATSTSLAVTFEELVTTTYGEEVYLSGISQLGD
	<i>A.foetidis</i>	WDTSDAVKLSADDYTSNPEWSVTVSLPVGTTFEYKFIKVEDGGSVTWESDPNREYTVPE
	<i>A.vadensis</i>	WDTSDAVKLSANDYTSNPEWYVTVSLPVGTTFEYKFIKVEDGGSVTWESDPNREYTVPE
	<i>A.niger</i> N402	WDTSDAVKMSADDYTSNPEWSVTVTLPVGTTFEYKFIKVESDGTVTWESDPNREYTVPE
	<i>A.foetidis</i>	CGSGSGETVFDTW*
	<i>A.vadensis</i>	CGSG..ETVFDTW*
	<i>A.niger</i> N402	CGSG..ETVVDTW*

Suppl. Fig. 2. Alignments showing different AmyA homologues from *A.foetidis* (CBS115.52), *A.vadensis* (CBS113.365), *A.niger* N402 (CBS120.49, ATCC1015), *A.niger* DSM (CBS513.88) and *A.niger* (NRRL3122). (A) Alignment of the 3' end of *aamA* and homologues. The stop codon is indicated in bold, non-coding DNA sequences are underlined. (B) Alignment of the C-terminal part of AamA and homologous proteins. The C-terminal end is indicated with *.



Suppl. Fig. 3. Northern blot analysis of representative amylolytic genes in wild-type (N402) and the $\Delta amyR$ strain from mycelia transferred from xylose (preculture) to maltose. Total RNA was isolated from mycelia growing in maltose 2 h and 8 h after the transfer as indicated.

Suppl. Table 1. The numeric values, Present/Absent (P/A) calls, and calculated fold-changes between different conditions for all genes encoding family GH13, 15 and 31 enzymes identified in *A. niger*.

	N402 (mal2h)	P	N402 (xyl2h)	P	$\Delta amyR$ (mal2h)	P	Fold induction maltose vs		N402 (mal8h)	P	N402 (xyl8h)	P	$\Delta amyR$ (mal8h)	P	Fold induction maltose vs		N402 vs $\Delta amyR$
							xylose	$\Delta amyR$							xylose	$\Delta amyR$	
<i>aamA</i>	3.15 ± 0.12	P	0.08 ± 0.04	A	0.13 ± 0.09	A	38.6	24.2	65.26 ± 7.48	P	0.58 ± 0.57	A	0.08 ± 0.01	A	112.9	807.7	
<i>amyA</i>	0.10 ± 0.03	A	0.10 ± 0.03	A	0.07 ± 0.02	A	1.0	1.4	0.12 ± 0.02	A	0.11 ± 0.01	A	0.09 ± 0.05	A	1.0	1.2	
<i>amyB</i>	0.08 ± 0.00	A	0.08 ± 0.01	A	0.06 ± 0.00	A	1.0	1.3	0.26 ± 0.17	A	0.10 ± 0.01	A	0.08 ± 0.03	A	2.5	3.1	
<i>amyC</i>	4.58 ± 1.61	P	6.96 ± 0.73	P	2.12 ± 0.05	P	0.7	2.2	4.36 ± 0.50	P	1.50 ± 0.16	P	1.40 ± 0.03	P	2.9	3.1	
<i>agsA</i>	16.24 ± 4.45	P	9.73 ± 1.70	P	12.44 ± 0.15	P	1.7	1.3	14.75 ± 1.16	P	21.38 ± 1.06	P	17.71 ± 0.93	P	0.7	0.8	
<i>agsB</i>	0.26 ± 0.13	A	0.21 ± 0.01	A	0.16 ± 0.08	A	1.3	1.7	10.76 ± 2.23	P	7.56 ± 0.38	P	6.15 ± 0.45	P	1.4	1.8	
<i>agsC</i>	0.10 ± 0.01	A	1.04 ± 0.05	A	0.08 ± 0.00	A	0.1	1.3	0.37 ± 0.26	A	0.43 ± 0.27	A	0.43 ± 0.00	A	0.9	0.9	
<i>agsA</i>	0.47 ± 0.11	A	0.20 ± 0.01	A	0.64 ± 0.02	A	2.4	0.7	0.90 ± 0.16	A	2.42 ± 0.63	P,A	0.50 ± 0.16	A	0.4	1.8	
<i>agsB</i>	0.76 ± 0.13	A	1.51 ± 0.12	P,A	0.79 ± 0.13	A	0.5	1.0	1.18 ± 0.10	A	0.86 ± 0.14	A	0.88 ± 0.11	A	1.4	1.3	
<i>agsC</i>	0.13 ± 0.02	A	0.17 ± 0.02	A	0.12 ± 0.03	A	0.8	1.1	7.70 ± 0.06	P	6.76 ± 0.43	P	5.20 ± 0.42	P	1.1	1.5	
<i>agsD</i>	1.21 ± 0.07	A	0.82 ± 0.25	M,A	0.67 ± 0.14	A	1.5	1.8	1.17 ± 0.07	P,A	0.84 ± 0.13	M,A	0.52 ± 0.17	A	1.4	2.2	
<i>agsE</i>	5.70 ± 3.40	P	18.55 ± 0.24	P	3.71 ± 0.40	P	0.3	1.5	18.61 ± 3.68	P	13.81 ± 0.93	P	8.99 ± 2.97	P	1.3	2.1	
<i>amyD</i>	1.01 ± 0.11	A	0.24 ± 0.17	A	1.08 ± 0.09	A	4.1	0.9	0.33 ± 0.01	A	0.75 ± 0.22	A	0.70 ± 0.31	A	0.4	0.5	
<i>amyE</i>	0.42 ± 0.51	P,A	1.14 ± 0.08	A	1.17 ± 0.15	P	0.4	0.4	0.37 ± 0.03	A	0.80 ± 0.03	A	0.62 ± 0.20	A	0.5	0.6	
<i>gbeA</i>	30.82 ± 12.98	P	10.48 ± 1.61	P	22.34 ± 4.05	P	2.9	1.4	13.69 ± 0.35	P	20.42 ± 2.81	P	35.19 ± 2.50	P	0.7	0.4	
<i>gdbA</i>	22.08 ± 6.02	P	7.86 ± 1.10	P	19.06 ± 0.33	P	2.8	1.2	9.49 ± 1.83	P	23.58 ± 8.80	P	30.27 ± 4.89	P	0.4	0.3	
<i>agdA</i>	85.96 ± 22.54	P	6.24 ± 0.98	P	5.03 ± 0.03	P	13.8	17.1	179.10 ± 13.90	P	9.89 ± 3.63	P	3.36 ± 0.46	P	18.1	53.3	
<i>agsB</i>	96.50 ± 31.50	P	10.47 ± 2.59	P	8.35 ± 2.13	P	9.2	11.6	49.46 ± 3.42	P	2.37 ± 0.08	P	6.84 ± 1.99	P	20.9	7.2	
<i>agdC</i>	2.93 ± 0.61	P	1.34 ± 0.02	M,A	2.16 ± 0.18	P	2.2	1.4	1.40 ± 0.17	A	3.96 ± 0.35	P	2.71 ± 0.72	P,M	0.4	0.5	
<i>agdD</i>	0.74 ± 0.01	A	0.48 ± 0.04	A	0.65 ± 0.09	A	1.5	1.1	0.59 ± 0.12	A	0.76 ± 0.50	A	0.64 ± 0.25	A	0.8	0.9	
<i>agsE</i>	21.31 ± 1.96	P	7.61 ± 0.83	P	20.41 ± 2.83	P	2.8	1.0	14.10 ± 1.27	P	11.62 ± 1.83	P	11.28 ± 0.07	P	1.2	1.3	
<i>agsF</i>	1.32 ± 0.22	P	0.13 ± 0.02	A	1.46 ± 0.19	P,M	10.4	0.9	0.28 ± 0.03	A	0.11 ± 0.01	A	0.65 ± 0.40	P,A	2.5	0.4	
<i>agsG</i>	0.73 ± 0.00	A	0.48 ± 0.03	A	0.32 ± 0.05	A	1.5	2.3	0.34 ± 0.36	A	0.57 ± 0.04	A	0.28 ± 0.03	A	0.6	1.2	
<i>glaA</i>	117.80 ± 43.16	P	13.78 ± 4.14	P	5.85 ± 0.47	P	8.5	20.1	239.30 ± 2.05	P	83.03 ± 4.35	P	19.03 ± 1.75	P	2.9	12.6	
<i>glaB</i>	0.13 ± 0.05	A	0.30 ± 0.07	A	0.07 ± 0.00	A	0.4	1.9	0.15 ± 0.06	A	0.29 ± 0.04	A	0.06 ± 0.09	A	0.5	2.4	
<i>axiA</i>	1.95 ± 0.05	A	57.37 ± 10.72	P	2.27 ± 0.15	A	0.0	0.9	3.13 ± 0.88	M,A	108.40 ± 3.85	P	2.56 ± 0.23	A	0.0	1.2	
<i>axiB</i>	4.38 ± 2.11	P	2.69 ± 0.25	P	4.23 ± 0.76	P	1.6	1.0	0.86 ± 0.27	A	1.89 ± 0.13	P,A	2.24 ± 0.15	P	0.5	0.4	

Suppl. Table 2. Inventory of all family GH13, 15 and 31 enzymes in *A. fumigatus*, *A. nidulans* and *A. oryzae*.

Aspergillus		ORF no.	Accession no.	Closest <i>A. niger</i> homologue ^a	Features ^b	
GH13 family						
<i>A. fumigatus</i>	Group I	Afu4g10130	EAL89775	AmyA; 5.6E-214	SS	
		Afu2g00710	EAL87170	AamA; 1E-196	SS, SBD	
	Group II	Afu3g00900 ^c	EAL86373	AgTA; 2.4E-204	SS	
		Afu2g03230	EAL87420	AgTA; 3.5E-182	SS, GPI	
	Group III	Afu2g13460	EAL93641	AgTA; 9.1E-127	SS, GPI	
		Afu3g07380	EAL92833	AgdC; 1.5E-275		
		Afu7g06380	EAL86834	AgdC; 1.4E-146		
		Afu8g07070	EAL85467	AgdC; 1.1E-139		
	Group IV	Afu2g11610	EAL93455	AgdC; 3.8E-128		
		Afu2g11620	EAL93456	AgdC; 1.9E-117		
	Group V	Afu1g15150	EAL90846	AmyD; 2.2E-203		
		Afu1g02140	EAL87926	GdbA; 0.0		
	Group VI	Afu5g10540	EAL91552	GbeA; 0.0		
		Afu2g11270	EAL93422	AgdD; 0.0	SS	
		Afu3g00910	EAL86374	AgSA; 0.0	SS	
		Afu1g15440	EAL90874	AgSt; 0.0		
<i>A. nidulans</i>	Group I	AN2018.3	EAA6580	AmyA; 0.0	SS	
		AN3388.3	EAA63356	AmyA; E-104	SS	
		AN3402.3	EAA63370	AamA; 0.0	SS, SBD	
	Group II	AN3308.3 ^d	EAA63276	AgTA; 0.0	SS, GPI	
		AN4507.3 ^{e,f}	EAA60850	AgTA; E-179	SS, GPI	
		AN6324.3	EAA58708	AgTA; E-138	SS, GPI	
	Group III	AN4843.2	ABF50870	AgdC; 0.0		
		AN10420.3 ^d	EAA59077 ^d	AgdC; 0.0		
	Group IV	AN3309.3	EAA63277	AmyE; 0.0		
	Group V	AN10060.3 ^d	EAA65758 ^d	GdbA; 0.0		
		AN2314.3	EAA64425	GbeA; 0.0		
	Group VI	AN5885.3	EAA58394	AgdD; 0.0	SS	
		AN3307.3	EAA63275	AgSE; 0.0	SS	
	<i>A. oryzae</i>	Group I	AO090120000196	POC183	AmyA; 0.0	SS
			AO090023000944	POC183	AmyA; 0.0	SS
		Group II	AO090003001498 ^c	BAE58540	AgTA; 0.0	SS, GPI ^g
AO090120000263			BAE62865	AgTA; 5.00E-72	SS, GPI ^g	
Group III		AO090020000176	BAE63400	AgdC; 0.0		
		AO090026000034	BAE59534	AgdC; 0.0		
		AO090103000378	BAE65811	AgdC; 0.0		
Group IV		AO090038000234	BAE64051	AgdC; e-156		
		AO090103000129	BAE65599	AgdC; e-149		
		AO090003001497	BAE58539	AmyD; 0.0		
Group V		AO090005001193	BAE56147	AmyD; 0.0		
		AO090005000884	BAE55891	GdbA; 0.0		
Group VI		AO090010000483	Q96VA4	GbeA; 0.0		
		AO090003001500	BAE58541	AgSE; 0.0	SS	
		AO090010000106	BAE66017	AgSA; 0.0	SS	
		AO090026000523	BAE59959	AgdD; 0.0	SS	
GH15 family						
<i>A. fumigatus</i>		Afu2g00690	EAL87168	GlaA; 1.1E-219	SS, SBD	
		Afu3g00610	EAL86343	GlaA; 8.9E-160	SS, SBD	
		Afu4g10140 ^g	EAL89774	GlaA; 1E-63	SS, SBD	
		Afu8g01390	EAL85029	GlaB; 2.5E-245		
		Afu6g08080	EAL88661	GlaB; 2.7E-208		
		AN7402.3 ^d	EAA61773	GlaA; 0.0	SS, SBD	
<i>A. nidulans</i>		AN11143.3 ^d	EAA64118 ^d	GlaA; 0.0	SS, SBD	
		AN5463.2	EAA62623	GlaA; 2E-32	SS, SBD	
		AO090010000746	P36914	GlaA; 0.0	SS, SBD	
<i>A. oryzae</i>		AO090003000321	BAE25205	GlaA; e-164	SS, SBD	
		AO090138000105	BAE64481	GlaB; 0.0		
GH31 family						
<i>A. fumigatus</i>		Afu4g10150	EAL89773	agdA; 0.0	SS	
		Afu1g16250	EAL90953	agdB; 0.0	SS	
		Afu5g03500	EAL85956	AgLIA; 0.0	SS	
		Afu1g03140	EAL88026	AxIB; 0.0		
		Afu2g05400	EAL87632	AxIB; 4.2E-47		
		Afu1g06560	EAL88367	AgLIB; 0.0		
		Afu2g00490	EAL87147	AxIA; 1.9E-48		
	<i>A. nidulans</i>		AN2017.3	EAA64849	agdA; 0.0	SS
			AN7345.3	EAA61716	agdB; 0.0	SS
			AN0941.3	EAA65970	agdB; 0.0	SS
			AN8953.3	EAA63748	agdB; 0.0	SS
			AN11054.3 ^d	EAA58873 ^d	AgLIA; 0.0	SS
			AN7120.3	EAA61325	AxIB; 0.0	
			AN0280.3	EAA66153	AxIB; 0.0	
			AN7505.3	EAA62085	AxIB; 2E-52	
			AN10935.3 ^d	EAA61743 ^d	AxIB; 2E-15	
			AN3504.3	EAA59065	AgLIB; 0.0	
	<i>A. oryzae</i>		AO090003001209	BAE58289	agdA; 0.0	SS
		AO090005001084	BAE56058	agdB; 0.0	SS	
		AO090038000471	BAE64257	agdB; 0.0	SS	
		AO090102000559	BAE61578	AgLIA; 0.0	SS	
		AO090701000558	BAE62256	AgLIC; 0.0	SS	
		AO090701000639	BAE62323	AxIA; 0.0	SS	
		AO090005000767	BAE55790	AxIB; 0.0		
		AO090005000768	BAE55791	AxIB; E-153		
		AO090001000649	BAE57147	AxIB; 3E-50		
		AO090026000111	BAE59606	AxIB; 2E-30		
	AO090023000288	BAE58852	AgLIB; 0.0			

^a Closest homologues from other aspergilli were obtained by using full length sequences of *A. niger* GH13, 15 and GH31 family proteins in a BlastP search

^b SS = predicted N-terminal Signal Sequence; GPI = predicted Glycosylphosphatidylinositol anchor signal; SBD = predicted starch binding domain.

^c These proteins are clustered with α -glucan synthases in the genome.

^d Annotation improved manually.

^e An earlier start codon was identified resulting in a predicted signal sequence for excretion which was lacking in the original annotation.

^f GPI-anchoring was not predicted by the online tool of Eisenhaber et al. (2004) but according to the consensus sequence described by De Groot et al. (2003) , a GPI-anchoring signal may be present.

Chapter 4

Two Novel, Putatively Cell Wall Associated and GPI-Anchored, α -Glucanotransferase Enzymes of *Aspergillus niger*

*R. M. van der Kaaij, X.-L. Yuan, A. Franken, A. F. J. Ram, P. J. Punt,
M. J. E. C. van der Maarel and L. Dijkhuizen*

Published in Eukaryot Cell. 2007 Jul; 6(7):1178-1188

Abstract

In the genome sequence of *Aspergillus niger* CBS 513.88, three genes were identified with high similarity to fungal α -amylases. The protein sequences derived from these genes were different in two ways from all described fungal α -amylases: they were predicted to be glycosylphosphatidylinositol anchored, and some highly conserved amino acids of enzymes in the α -amylase family were absent. We expressed two of these enzymes in a suitable *A. niger* strain and characterized the purified proteins. Both enzymes showed transglycosylation activity on donor substrates with α -(1,4)-glycosidic bonds and at least five anhydroglucose units. The enzymes, designated AgtA and AgtB, produced new α -(1,4)-glycosidic bonds and therefore belong to the group of the 4- α -glucanotransferases (EC 2.4.1.25). Their reaction products reached a degree of polymerization of at least 30. Maltose and larger maltooligosaccharides were the most efficient acceptor substrates, although AgtA also used small nigerooligosaccharides containing α -(1,3)-glycosidic bonds as acceptor substrate. An *agtA* knockout of *A. niger* showed an increased susceptibility towards the cell wall-disrupting compound calcofluor white, indicating a cell wall integrity defect in this strain. Homologues of AgtA and AgtB are present in other fungal species with α -glucans in their cell walls, but not in yeast species lacking cell wall α -glucan. Possible roles for these enzymes in the synthesis and/or maintenance of the fungal cell wall are discussed.

Introduction

Aspergillus niger is a filamentous ascomycete fungus with a worldwide distribution. As a saprophyte, the fungus produces and secretes a large variety of extracellular enzymes, especially proteases and polysaccharide hydrolases to convert plant cell walls and storage compounds into growth substrates (Duarte et al., 1994 and Martens-Uzunova et al., 2006). This quality is exploited for the production of enzymes for the food and feed industry on a large scale. Recently, the full genome sequence of *A. niger* CBS 513.88 was determined and annotated (Pel et al., 2007). A high level of synteny was observed between *A. niger* and other sequenced aspergilli, although more extracellular hydrolytic enzymes were annotated for *A. niger*. A detailed full genome search showed the presence of a considerable number of previously unknown, predicted enzymes belonging to the α -amylase superfamily.

The α -amylase superfamily (Kuriki and Imanaka, 1999) comprises a large variety of enzymes that are active towards polysaccharides with α -glycosidic linkages, such as starch and glycogen (MacGregor et al., 2001). Most members of this family are involved in either production of storage compounds, such as glycogen or starch, or degradation of these compounds as extracellular carbon and energy sources. The tertiary structure of these enzymes is characterized by a $(\beta/\alpha)_8$ barrel containing four highly conserved amino acid regions that form the catalytic site (MacGregor et al., 2001) (see also the CAZy website at <http://afmb.cnrs-mrs.fr/CAZY/>). Based on sequence similarity, members of the α -amylase superfamily are divided over glycoside hydrolase (GH) families 13, 70, and 77 (GH13, GH70,

and GH77, respectively). Here, we focus on family GH13, which mostly contains enzymes that perform a hydrolytic reaction, i.e., they cleave an α -glycosidic linkage using water as an acceptor molecule. The best known hydrolytic enzyme of GH13 is α -amylase (EC 3.2.1.1), which hydrolyzes the internal α -(1,4)-glycosidic bonds in starch, glycogen, and maltooligosaccharides (α -(1,4)-linked glucose oligomers), producing shorter maltooligosaccharides and maltose (O- α -D-Glc-(1,4)- α -D-Glc). Other GH13 members perform a transglycosylation (or glucanotransferase) reaction in which they cleave an α -(1,4) glycosidic bond in a donor substrate and subsequently do not use water but instead use another oligosaccharide as the acceptor substrate to form a new α -glycosidic linkage.

A. niger produces a number of extracellular enzymes classified as members of GH13, which are involved in the degradation of starch. These include acid amylase, which is well-known for its stability at low pH (Boel et al., 1990) and two almost identical α -amylase enzymes (AmyA/B) (Korman et al., 1990). We have identified several additional members of GH13 in the *A. niger* genome sequence, three of which clustered together in the phylogenetic tree of the GH13 members and showed a typical glycosylphosphatidylinositol (GPI) anchoring signal at the protein C terminus. A GPI anchor serves as a targeting signal to the cell membrane and/or the cell wall.

The cell walls of aspergilli have been shown to contain four major classes of polysaccharides: chitin, α -glucan, α -(1,3)-glucan, and galactomannan (Bardalaye et al., 1977; Blumenthal et al., 1957; Johnston, 1965; Stagg and Feather, 1973). In addition, it contains covalently attached cell wall proteins (Brul et al., 1997). The α -glucan fraction is composed of an α -(1,3)-glucan with 3 to 10% α -(1,4)-glycosidic linkages (Johnston, 1965; Horisberger et al., 1972), and nigeran, a glucan with alternating α -(1,3) and α -(1,4) glycosidic bonds (Bobbitt et al., 1977). α -(1,3)-Glucan synthases have been identified and functionally studied in a number of different fungi (Hochstenbach et al., 1998; Reese and Doering, 2003; Beauvais et al., 2005; Damveld et al., 2005b). These enzymes are large proteins (~2,400 amino acids long) consisting of three conserved domains which are predicted to be involved in the synthesis, transport, and cross-linking of the α -(1,3)-glucan (Hochstenbach et al., 1998). Detailed structural studies in *Schizosaccharomyces pombe* have revealed that the α -glucan is a linear glucose polymer of 260 residues in length consisting of two α -(1,3)-glucan chains that are interconnected via α -(1,4)-linked glucose residues (Grün et al., 2005). A mutation in the N-terminal part of the α -(1,3)-glucan synthases (the proposed cross-linking domain) abolished the linkage between the two α -(1,3)-glucan chains, indicating that this part of the protein acts as a glucanotransferase, connecting the glucan chains.

Recently, two types of putative GH13 enzymes have been shown to play a role in fungal cell wall α -(1,3)-glucan formation. Marion et al. (2006) showed the involvement of a putative α -amylase (Amy1p) in the formation of α -(1,3)-glucan in the cell wall of *Histoplasma capsulatum*. In this pathogenic fungus, α -(1,3)-glucan is known to play an important role in virulence (Rappleye et al., 2004). A functional knockout of *AMY1* resulted in a lack of α -(1,3)-glucan formation and decreased virulence. The second α -amylase homologue, Aah3p, was studied in *S. pombe* (Morita et al., 2006). A knockout strain of this GPI-anchored protein was hypersensitive towards cell wall-degrading enzymes and

showed aberrant cell shape. The enzymatic activities of Amy1p and Aah3p have not been studied.

In this paper, we report the first biochemical characterization of two GH13 enzymes putatively involved in α -(1,3)-glucan formation. We expressed and purified two GPI-anchored enzymes from *A. niger*, both homologues of Aah3p from *S. pombe*. The biochemical characterization showed that the two *A. niger* enzymes are GH13 α -glucanotransferases, making them the first of their kind to be described for fungi. A gene knockout of one of the enzymes in *A. niger* resulted in increased sensitivity towards calcofluor white (CFW), a cell wall disrupting compound.

Materials and methods

Bioinformatics tools

The full genome sequence of *A. niger* strain CBS 513.88 was provided by DSM (a biotechnology company based in The Netherlands) (Pel et al., 2007). A Hidden Markov model (HMM) profile was built using the HMMR package (Durbin and Eddy, 1998) based on the amino acid sequences of described α -amylases, which were retrieved from the CAZY website (<http://afmb.cnrs-mrs.fr/CAZY/>) (Coutinho and Henrissat, 1999). The obtained profile was used to screen the *A. niger* CBS 513.88 genomic database using the WISE 2 package (Birney et al., 2004). The presence of a signal peptidase cleavage site and a GPI attachment site were predicted by web-based search tools (<http://www.cbs.dtu.dk/services/SignalP/>) (Bendtsen et al., 2004) and http://mendel.imp.univie.ac.at/sat/gpi/gpi_server.html (Eisenhaber et al., 2004), respectively). The GPI attachment prediction was confirmed by a manual comparison of the protein sequences with the consensus sequence for yeast GPI proteins as described by De Groot et al. (2003). Amino acid sequence alignments and phylogenetic analysis were performed using MEGA3.1 (Kumar et al., 1993) and adjusted manually if necessary. Sequences from other fungal genomes were retrieved via the option genomic BLAST at NCBI (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?organism_fungi).

Aspergillus niger strains, growth conditions and transformation

A. niger strain N402 (*cspA1* derivative of ATCC 9029 (Bos et al., 1988)) mRNA was used for the construction of a cDNA library. Genomic DNA from *A. niger* NRRL3122 and N402 strains was isolated and used as a template in PCRs. *A. niger* strain MGG029- Δ *aamA* (Weenink et al., 2006) was used as a host for protein overexpression. This strain, derived from strain MGG029 (*prtT glaA::fleo^r pyrG*), is deficient in the expression of several extracellular proteases, and it has no glucoamylase gene (*glaA*) and acid amylase gene (*aamA*), resulting in very poor growth on starch (Weenink et al., 2006). Strain MA70.15 (*pyrG⁻ ku70::amdS*) (Meyer et al., 2007) was used for disruption of the *agtA* gene.

Aspergillus strains were grown in *Aspergillus* minimal medium (MM) or complete medium (CM) which is MM with the addition of 0.1% Casamino Acids and 0.5% yeast extract (Oxoid, Basingstoke, United Kingdom) (Bennett and Lasure, 1991). Cultures for protein production were grown in CMS (CM supplemented with 2% (wt/vol) sucrose and 1% (wt/vol) glucose). Spores were obtained by growing *A. niger* on CM with 2% (wt/vol) agar for 4 days and scraping off the spores in 0.9% (wt/vol) NaCl. Liquid cultures were inoculated with 10^6 spores liter⁻¹ medium and subsequently grown at 30°C while shaking at 280 rpm. Transformation of *A. niger* was performed as described previously (Punt and van den Hondel, 1992) using lysing enzymes (Sigma, Zwijndrecht, The Netherlands). Selection of positive clones was performed on the basis of their ability to grow successively on MM containing 15 mM CsCl and 10 mM acetamide or acrylamide as the sole nitrogen source, brought about by the expression of the *amdS* gene (Kelly and Hynes, 1985).

Cloning procedures

All basic molecular techniques were performed according to standard procedures (Sambrook et al., 1989). *E. coli* TOP10 (Invitrogen, Carlsbad, U.S.A.) or DH5 α (Stratagene, La Jolla, U.S.A.) were used for transformation and amplification of recombinant DNA. The primers used were obtained from Eurogentec (Seraing, Belgium) or Biolegio (Nijmegen, The Netherlands). All steps during the construction of the overexpression vectors were checked by restriction analysis, and the final constructs were checked by sequencing (GATC Biotech AG, Konstanz, Germany). Genomic DNA was isolated from *A. niger* N402 and NRRL3122 as described by Kolar *et al.* (1988). All PCR reactions were performed with 2.5 units of *Pwo* DNA polymerase (Roche, Indianapolis, U.S.A.), 1x buffer and 1 mM of each dNTP in a total volume of 25 μ l. A cDNA library was produced from *A. niger* N402 grown on MM with addition of starch as sole carbon source. Primers used are indicated in Table 1.

The overexpression vector for the transformation of *A. niger* was provided by Dr. J. Benen (Wageningen University, Wageningen, The Netherlands), and was produced as follows: Gene *pgaII* (encoding polygalacturonase II from *A. niger*) was cloned into pPROM-S (Benen et al., 1999) using *NsiI* and *KpnI* restriction sites. A *NotI* site was generated immediately downstream of the stop codon of the *pgaII* gene by site directed mutagenesis. The gene encoding acetamidase (*amdS* (Kolar et al., 1988)) was amplified by PCR with specific primers from plasmid p3SR2 (Wernars et al., 1985) and cloned in front of the *pki*-promoter region (Parenicova et al., 1998) using *XbaI* restriction sites, resulting in vector P*pki-pgaII-amdS*. The construct for the overexpression of the chimaeric protein AgtA-SBD (a fusion between AgtA and the Starch Binding Domain (SBD) of *A. niger* glucoamylase GlaA) was made as follows: Genomic DNA of *A. niger* N402 was used as a template in a PCR reaction with specific primers to generate the DNA fragment encoding the SBD, including the linker region.

Table 1. Primers used for the production of plasmids for overexpression and deletion in *A. niger*. The restriction sites used for cloning are underlined.

Primer name	Primer sequence	Restriction enzyme
<i>Primers used for construction of pΔagtA</i>		
AgtAP1for	ataagaatg <u>cgggccgc</u> TGTCCTGTGTTCACGCT	<i>NotI</i>
AgtAP2rev	gctctagaAATGATCAAGGGTTGCGTACA	<i>XbaI</i>
AgtAP3for	gctctagaTATGCTGATAGCTACAGATGG	<i>XbaI</i>
AgtAP4rev	cgggatccGGAGTGGATAGCTGGTAAGGC	<i>BamHI</i>
<i>Primers used for construction of Ppki-agtA-SBD-amdS</i>		
SBD-fw	ggccctatgcatggccct <u>gggccc</u> ACCTGTGCGGCCACATCTGC	<i>NsiI</i> , <i>ApaI</i>
SBD-rev	cccgtg <u>cgggccgc</u> CTACCGCCAGGTGTCAGTCAC	<i>NotI</i>
AgtA-SBD-fw	ggccctatgcATATGGTCTCAATGTCGCCCCGTC	<i>NsiI</i>
AgtA-SBD-rev	ccgggagggcccTCCGCACAGCCCACTGCC	<i>ApaI</i>
<i>Primers used for construction of Ppki-agtA-amdS and Ppki-agtB-amdS</i>		
AgtA-fw	ggccctatgcatGTCTCAATGTCGCCCCCTGC	<i>NsiI</i>
AgtA-rev	cccgtg <u>cgggccgc</u> TTACCACATCCCCACAATCA	<i>NotI</i>
AgtB-for	ggccctatgcatTTTCGAAAATCCGCTTCCCTC	<i>NsiI</i>
AgtB-rev	cccgtg <u>cgggccgc</u> TTATATCCGGAATGCCAAAAAT	<i>NotI</i>

The primers were designed to amplify nucleotide 3643 to 4149 (numbering according to *glaA* coding sequence EMBL ID AY250996). An *ApaI* restriction site was built into the forward primer to allow the subsequent cloning of the *agtA* gene fragment in frame with the SBD. The SBD-encoding fragment was cloned into vector *Ppki-pgaII-amdS* using *NsiI* and *NotI* restriction sites, thereby replacing the *pgaII* gene. The cDNA fragment encoding *agtA* was generated by PCR with specific primers on the cDNA library. The primers were designed to amplify the gene up to nucleotide 1766 (in gene sequence), which does not include the C-terminal GPI-anchoring part. The *AgtA* cDNA was cloned in frame N-terminally of the SBD using *NsiI* and *ApaI*, resulting in the expression vector *Ppki-agtA-SBD-amdS*. Sequencing of the construct revealed one point mutation compared to the original gene sequence in the genomic database. Nucleotide 1420 (in the coding sequence of *agtA*) was changed from A to G, resulting in Ala474 instead of Thr474 in the derived amino acid sequence. This mutation was consistent in several independent clones and was therefore considered to be representing a strain difference between N402 and CBS 513.88. The same mutation was found in the equivalent protein sequence published by the DOE Joint Genome Institute (<http://genome.jgi-psf.org/Aspni1/Aspni1.home.html>).

The constructs for overexpression of *agtA* and *agtB* were produced as follows: the complete gene sequences of genes *agtA* and *agtB* were amplified with specific primers from genomic DNA isolated from *A. niger* NRRL3122 (van Dijck et al., 2003). The primers contained restriction sites for *NsiI* and *NotI*, which were used to clone the gene fragments into vector *Ppki-pgaII-amdS*, thereby replacing the *pgaII* gene. This resulted in the vectors *Ppki-agtA-amdS* and *Ppki-agtB-amdS*.

Protein production, purification and detection procedures

Production of AgtA-SBD

Several stable transformants were checked for their level of extracellular production of the chimeric protein AgtA-SBD by Western blotting with polyclonal antiserum raised against purified SBD (antiserum kindly provided by D. Archer and D. McKensie (University of Nottingham & IFR Norwich, United Kingdom)) (Le Gal-Coeffet et al., 1995). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting were performed according to standard protocols. Immobilon-P (Millipore, Billerica, MA) was used as the blotting membrane. The untransformed strain was included as a negative control, while purified SBD (M. J. van der Maarel et al., unpublished data) served as a positive control for immunodetection. The *A. niger* transformant producing the highest levels of AgtA-SBD was grown in a 5-liter batch fermentor (New Brunswick Scientific, Edison, NJ) inoculated with 100 ml culture pregrown on potato dextrose broth (Difco). The medium used for batch fermentation consisted of the following components: 3% (wt/vol) glucose, 117 mM NaNO₃, 3.25 mM MgSO₄·7H₂O, 14.7 mM KH₂PO₄, 0.69 mM CaCl₂·H₂O, 0.5% Yeast Extract and spore elements. Fermentor conditions were pH 4.5, a temperature of 30°C and aeration at 1.2 liter min⁻¹. The feed contained 20% (w/v) glucose, 74 mM KH₂PO₄, 350 mM NaNO₃, 1% yeast extract and 1% tryptone and was added at a rate of 5 ml h⁻¹.

Three days after the feed was started, the growth medium was collected by filtration over miracloth (Calbiochem, EMD Biosciences, La Jolla, CA). The pH of the medium was subsequently adjusted to pH 6 with 1 M NaOH, and AgtA-SBD was extracted from the medium via binding of the SBD to starch granules based on the procedure described by Paldi et al. (Paldi et al., 2003). One liter of medium was added to 13 g waxy maize starch that was prewashed with elution buffer (10 mM sodium acetate (NaAc), pH 6) and incubated for 2 h at 4°C while shaking gently. The starch with bound proteins was collected by centrifugation (15 min at 5000 × g) and washed once with ice-cold elution buffer followed by another centrifugation step. AgtA-SBD was eluted from the starch granules by the addition of 25 ml of elution buffer per batch of 13 g starch and subsequent incubation at 40°C for 1 h while shaking gently. After the starch granules were removed by centrifugation, the protein was collected from the supernatant. A second round of binding was performed to remove residual proteins and components of the medium. The protein was concentrated, and the buffer was changed to 20 mM Tris-HCl, pH 8, using a Centriprep YM50 column (Millipore, Bedford, MA). The sample was applied to an anion-exchange column (ResourceQ (1 ml); Amersham Biosciences, Piscataway, NJ) equilibrated with 20 mM Tris-HCl, pH 8. Proteins were eluted with a NaCl gradient (0 to 1 M NaCl) at a flow rate of 1 ml min⁻¹. AgtA-SBD was eluted as a single activity peak at 280 mM NaCl. Purification of AgtA-SBD was confirmed by Western blot analysis. Deglycosylation was performed with 850 U of endoglycosidase H (Endo H) (New England Biolabs, Ipswich, MA) on 1.5 µg of purified protein for 20 h in a total volume of 15 µl according to the manufacturer's instructions.

Production of AgtA and AgtB

After transformation of each of the plasmids *Ppki-agtA-amdS* and *Ppki-agtB-amdS* into *A. niger* MGG029- Δ *amA*, nine transformants were selected which showed the best growth on selective medium. The best transformant for overexpression of each protein was selected by growth in liquid CMS and visual inspection of protein production on SDS-polyacrylamide gels. The selected transformants were grown in CMS for 3 days at 30°C and 200 rpm. Mycelium was removed from the culture medium by filtration over miracloth. The medium was concentrated over a Centrprep YM-50 membrane filter, and the concentrated protein was taken up in 20 mM Tris-HCl buffer, pH 8. The proteins were purified via anion-exchange chromatography as described above. Both AgtA and AgtB were eluted as a single activity peak at a concentration of 150 mM NaCl. At each stage of the protein purification, the protein amount was measured using the Bradford method with reagent from Bio-Rad (Hercules, CA), and purity was checked using SDS-PAGE analysis (Laemmli, 1970) and staining with Biosafe Coomassie (Bio-Rad).

Enzymatic assays

All oligosaccharides used were obtained from Sigma, except for nigerotriose (O- α -D-Glc-(1,3)- α -D-Glc), which was purchased from Dextra Laboratories (Reading, United Kingdom), and nigerose (O- α -D-Glc-(1,3)- α -D-Glc), which was a kind gift from Nihon Shokuhin Kako Co. Ltd. (Shizuoka, Japan). *Lactobacillus reuteri* polysaccharide was a gift from S. Kralj (University of Groningen, The Netherlands), and α -(1,3)-glucan isolated from *A. nidulans* was a kind gift from B. J. Zonneveld (Leiden University, Leiden, The Netherlands). As soluble starch, Paselli SA2 with an average degree of polymerization (DP) of 50 (AVEBE, Foxhol, The Netherlands), was used. All reactions were performed at 37°C. To determine the optimum pH for activity of AgtA and AgtB, 0.5 μ g purified enzyme was incubated with 20 mM maltopentaose (a maltooligosaccharide with a DP of 5) in a 20 μ l reaction volume for 30 min at 11 different pH values. The reaction was buffered by either 30 mM NaAc buffer at pH 4.2 to 7.0 or K₂HPO₄/KH₂PO₄ buffer at pH 6.3 to 8.0. Subsequently, 2 μ l of the reaction mixture was spotted onto a thin-layer chromatography (TLC) plate (silica gel 60 F254; Merck, Darmstadt, Germany), and after the plate was dried, it was run for 6 h in 75 ml running buffer (butanol-ethanol-MilliQ, 5/5/3 (vol/vol/vol)) in a container of 22 by 6 by 22 cm. After the plate was run, it was dried and sprayed with 50% sulfuric acid in methanol and left to develop for 10 min at 110°C.

Hydrolyzing activity on potato starch was determined by the incubation of 1 μ g of purified enzyme with 600 μ l of 0.02%, 0.2%, or 2% (wt/vol) dissolved potato starch in 50 mM NaAc buffer containing 1 mM CaCl₂. Reactions were buffered at pH 5.5 (AgtA) or pH 4.8 (AgtB) and performed in duplicate. Samples of 50 μ l were taken after several time intervals up to 4 h and used for the determination of reducing ends and glucose. The formation of reducing ends was measured with the bicinchoninic acid method (Meeuwsen et al., 2000), and glucose formation was measured with the glucose GOD-PAP assay (Roche, Mannheim, Germany). Appropriate calibration curves and negative controls were included for all assays and reactions.

Standard assay conditions for all further enzymatic reactions were as follows: 0.4 μ g of purified enzyme was incubated in 20 μ l of 25 mM NaAc buffer (pH 5.5) containing 1 mM CaCl₂ and 0.01% sodium azide in the presence of 20 mM oligosaccharide substrate and/or 4% (wt/vol) soluble starch or other polysaccharide, except for nigeran and α -(1,3)-glucan. Nigeran and α -(1,3)-glucan were dissolved in 1 M NaOH, after which the pH of a 1% solution was adjusted to pH 5.5 with hydrogen acetate. The final concentration used in reaction mixtures was 0.5% (wt/vol). Reaction products were detected either by TLC (as described above) or high-performance liquid chromatography (HPLC) (Dionex) analysis. For HPLC analysis, 5 μ l of the reaction mixture was diluted in 1.5 ml 90% dimethyl sulfoxide. Separation of oligosaccharides was achieved by the method of Kralj et al. (2004b).

Disruption of the *agtA* gene

The plasmid used to disrupt the *agtA* gene was constructed as follows. The DNA fragments flanking the *agtA* gene were amplified by PCR using N402 genomic DNA as template: 1.5 kb of 5' flanking DNA and 1.0 kb of 3' flanking DNA were amplified by PCR using primers AgtAP1f and AgtAP2r, AgtAP3f and AgtAP4r (Table 1), respectively. Each primer was adapted with a restriction site for further cloning. The amplified PCR fragments were digested with *NotI* and *XbaI* or *XbaI* and *BamHI*, respectively, and cloned in a three way ligation into *NotI* and *BamHI* digested pBlue-ScriptII SK to obtain plasmid pAgtAF53. Subsequently, pAgtAF53 was digested with *XbaI*, and ligated with the 2.7 kb *XbaI* fragment containing the *A. oryzae* *pyrG* gene, obtained from plasmid pAO4-13 (de Ruiter-Jacobs et al., 1989) which resulted in the *agtA* disruption plasmid, p Δ *agtA*. Before transformation to MA70.15, p Δ *agtA* was linearized with *NotI*. Uridine prototrophic transformants were selected by incubating protoplasts on agar plates containing MM without uridine. Transformants were purified and genomic DNA was isolated and analyzed by Southern Blot analysis to identify possible Δ *agtA* strains. Deletion of the *agtA* locus by homologous recombination was expected to result in appearance of a 4.8 kb fragment after digestion of the genomic DNA with *KpnI* and the loss of a 2.0 kb fragment which was expected in the wild-type. For the hybridization, the *NotI*-*XbaI* fragment (containing the 5' flanking region of the *agtA* gene) was used. Independently obtained transformants were purified and strains with the expected hybridization pattern (MA71.1, MA71.3, MA71.4, MA71.7) were used for phenotypic analysis.

Phenotypic characterization of the Δ *agtA* strain and AgtA/AgtB overexpression strains.

Sensitivity towards CFW was assayed as described (Ram and Klis, 2006). Conidiospores from the control strains (MA70.15 transformed with pAO4-13 containing *A. oryzae* *pyrG*, and MGG029- Δ *amA*), the Δ *agtA* strains (MA71.1, MA71.3, MA71.4, MA71.7), and the *agtA* and *agtB* overexpression strains were spotted on CFW-plates. All strains were also checked for their ability to grow on starch as sole carbon source by the inoculation of 3 μ l of spore solution containing 10⁶, 10⁵, 10⁴ or 10³ spores on plates containing MM with either 1% (w/v) potato starch or 1% (w/v) glucose as carbon source. Colony growth was followed daily.

were found in the genome sequences of members of the subphylum *Saccharomycotina*, such as *Candida albicans*, *Kluyveromyces lactis*, and *Saccharomyces cerevisiae*.

The amino acid sequences of AgtA, AgtB, and AgtC show 54 to 56% similarity to that of *A. niger* acid amylase (Boel et al., 1990). A phylogenetic analysis of the Agt proteins compared to the *A. niger* α -amylases, the homologous proteins in *A. oryzae*, and the Aah proteins in *S. pombe* is given in **Fig. 2**. The Agt proteins contain most of the amino acids generally conserved in GH13 (Table 2). However, in all three protein sequences, the highly conserved His143, which is part of conserved region I, is absent (numbering used is according to acid α -amylase of *A. niger*, unless indicated otherwise). In AgtB and AgtC, a second conserved histidine in conserved region IV is replaced by glutamate (Table 2).

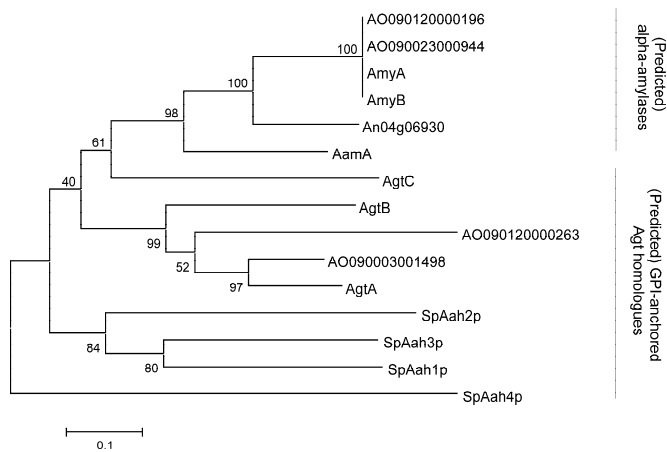


Fig. 2. Bootstrapped phylogenetic tree of the *A. niger* extracellular α -amylases (acid amylase AamA, α -amylase A and B, and putative α -amylase An04g06930) and α -glucanotransferases (AgtA/B/C), the homologous putative proteins identified in the genome of *A. oryzae* and the *S. pombe* Aah proteins (SpAah1p to SpAah4p). The alignment and phylogenetic analysis were performed with MEGA version 3.1 using default settings. A bootstrapped tree was constructed with the neighbor-joining method using 500 replicates.

Table 2. Alignment of the generally conserved regions of the α -amylase family as present in *A. niger* acid amylase and α -amylase AmyA compared to homologous regions in AgtA, AgtB and AgtC from *A. niger* and the homologous protein Aah3p from *S. pombe*. Catalytic residues are underlined, generally conserved residues are indicated in bold.

Enzyme	Region I	Region II	Region III	Region IV
Acid amylase	LMVDVVPNH	DGLRIDSVLE	YCVGEVDN	NFIENHD
AmyA	LMVDVVANH	DGLRIDTVKH	YCIGEVLD	TFVENHD
AgtA	LMMDTVINN	DGLRIDAAKH	FMTGEVLQ	SFSENHD
AgtB	LLLDVVINN	DGLRIDAAKS	FMTGEVMD	NFIEDQD
AgtC	LMLDIVVGD	DGLRIDSVLN	FTVGEGAT	TFTANQD
Aah3p (<i>S. pombe</i>)	VMLDSIVNS	DGLRIDAVKM	YSVGEVFS	TFIENHD
Amino acid numbering according to <i>A. niger</i> acid amylase	135 – 143	222 - 231	247 - 254	312 - 318

Purification and enzyme activity of AgtA fused to a SBD

The *agtA* gene, lacking its C-terminal anchoring domain, was fused to the SBD of the *A. niger* glucoamylase gene serving as a protein tag, and the fusion construct was transformed into *A. niger* MGG029- $\Delta aamA$, resulting in 25 stable transformants.

The transformant with the highest expression level of the AgtA-SBD fusion protein in the culture medium as determined by Western blotting with polyclonal antiserum raised against purified SBD (Le Gal-Coeffet et al., 1995) was selected for larger-scale protein production. AgtA-SBD was isolated from the culture medium by binding it to waxy maize starch granules, and after further purification by anion-exchange chromatography, a single protein band with an apparent molecular mass of about 130 kDa was obtained (Fig. 3). This band was recognized by anti-SBD antibodies, a strong indication that this band contained the AgtA-SBD fusion protein (calculated molecular mass of 73.6 kDa). Treatment with Endo H to remove N glycosylation resulted in a small decrease in the apparent molecular mass, indicating that the protein was N glycosylated (Fig. 3). The high molecular mass of the fusion protein after Endo H treatment could be caused by heavy O glycosylation of the linker region of the SBD (Williamson et al., 1992).

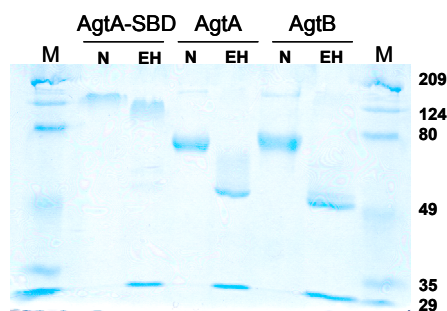


Fig. 3. SDS-PAGE analysis of purified AgtA-SBD, AgtA and AgtB in native form (N) and after treatment with Endo H (EH). The M lanes contain molecular mass markers. The positions of molecular mass markers (in kilodaltons) are indicated to the right of the gel.

Purified AgtA-SBD was incubated with potato starch to determine its hydrolyzing activity. Hydrolysis of starch may result in the formation of glucose, maltose, or longer maltooligosaccharides, the formation of which can be quantified by the measurement of the reducing ends formed during the reaction. For AgtA-SBD, a low hydrolyzing activity of $0.46 \pm 0.02 \mu\text{mol}$ of reducing ends $\text{mg}^{-1} \text{min}^{-1}$ was detected when incubated with 0.2% starch, and no formation of glucose was observed. TLC analysis of the reaction mixture showed that short oligosaccharides were formed in small amounts from soluble starch (Fig. 4, lane 1). Incubation with both maltose and soluble starch resulted in the formation of more oligosaccharides than made from starch alone (Fig. 4, lane 2). No products were formed from maltose alone (result not shown). When AgtA-SBD was incubated with maltopentaose or maltoheptaose (maltooligosaccharide with a DP of 7), a variety of oligosaccharides

ranging from maltose to oligosaccharides with a DP of at least 13 to 18, respectively, were formed (Fig. 4, lanes 3 and 4). This result indicated that AgtA-SBD hydrolyzed starch to some extent but mainly acted as a glucanotransferase, transferring parts of a donor oligosaccharide, which might be starch, to an acceptor substrate, e.g., maltose, thereby producing a variety of oligosaccharides of different lengths.

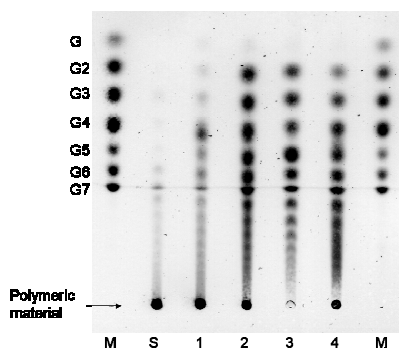


Fig. 4. TLC analysis of reaction products of AgtA-SBD from different substrates. Enzyme (0.4 μ g) was incubated with 20 mM oligosaccharide and/or 4% soluble starch for 1 h at 37°C. The arrow indicates where the following samples were loaded: molecular size markers (lanes M) containing a mix of maltooligosaccharides ranging from glucose (G1) to maltoheptaose (G7), unmodified soluble starch (lane S), and reaction products of AgtA-SBD incubated with soluble starch (lane 1), maltose and soluble starch (lane 2), maltopentaose (lane 3), or maltoheptaose (lane 4).

Production and purification of AgtA and AgtB

To rule out any effect of the incorporated SBD on the enzymatic activity of AgtA, both AgtA and AgtB were overexpressed in *A. niger* in their native form for further biochemical analysis. All selected transformants overproduced a protein with an estimated molecular mass of 85 kDa not observed in the untransformed strain. In a similar procedure of cloning and transformation, we also attempted to produce the AgtC protein. Although insertion of the overexpression construct in selected transformants was confirmed by Southern blot analysis, none of the transformants overproduced the AgtC protein. The characterization of AgtC was therefore not included in this study.

Proteins expressed in the culture medium of *A. niger* MGG029- Δ amA-AgtA and of *A. niger* MGG029- Δ amA-AgtB were concentrated and subsequently submitted to anionexchange chromatography, resulting in the purification of AgtA and AgtB. Both proteins had an apparent molecular mass of approximately 70 kDa, but after removal of N glycosylation, the apparent protein masses decreased to approximately 55 kDa, close to their theoretical masses (58.8 kDa for AgtA and 57.7 kDa for AgtB, after removal of the C-terminal end for GPI anchoring) (Fig. 3).

Glucanotransferase activity of AgtA and AgtB on maltooligosaccharides

Both AgtA and AgtB were incubated with maltooligosaccharides ranging in size from maltose to maltohexaose (a maltooligosaccharide with a DP of 6), and the products were analyzed by TLC. With maltopentaose or maltohexaose as substrates, both enzymes produced a range of oligosaccharides with a DP of 15 or larger, similar to what was observed previously for AgtA-SBD (Fig. 5, lanes 3, 4, 7, and 8). Incubation with maltopentaose at lower concentrations (2 or 10 mM) also resulted in the formation of products with a DP of 6 and larger (results not shown). Activity of both enzymes on the smaller maltooligosaccharides was limited (Fig. 5, lanes 1 and 2 and lanes 5 and 6). Neither of the enzymes produced glucose in detectable amounts. When incubated with dissolved potato starch, both enzymes produced small amounts of reducing ends (hydrolysis on 2% starch $0.55 \pm 0.19 \mu\text{mol mg}^{-1} \text{min}^{-1}$ for AgtA and $0.26 \pm 0.16 \mu\text{mol mg}^{-1} \text{min}^{-1}$ for AgtB; the hydrolysis rate was approximately four times lower when measured on 0.2% starch). TLC analysis of the products produced from maltoheptaose at different pH values indicated that AgtA was active between pH 4.5 and pH 6, and AgtB showed activity between pH 4 and pH 5.5 (data not shown). Glucose polymers with other types of glycosidic linkages, like dextran (a glucan polymer with α -(1,6)-glycosidic bonds), nigeran, *A. nidulans* α -(1,3) glucan, an *L. reuteri* polysaccharide containing α -(1,3) and α -(1,6) glycosidic bonds (Kralj et al., 2004a), and cellulose (α -(1,4) glycosidic bonds) were tested alone or in combination with maltose as an acceptor substrate. None of these polysaccharides acted as a substrate for AgtA or AgtB (data not shown).

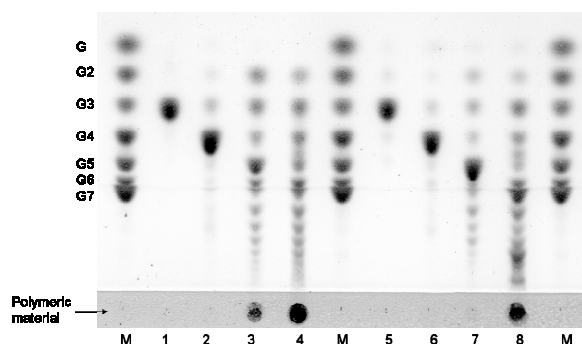


Fig. 5. TLC analysis of the reaction products of AgtA (left side) and AgtB (right side) incubated with different substrates. Purified enzyme ($0.4 \mu\text{g}$) was incubated with 20 mM substrate for 1 h at 37°C . The spots where the samples are applied are indicated by the arrow. The molecular size marker (M) lanes contain a mix of maltooligosaccharides ranging from glucose (G1) to maltoheptaose (G7). The figure shows reaction products of AgtA (lanes 1 to 4) or AgtB (lanes 5 to 8) incubated with maltotriose (lanes 1 and 5), maltotetraose (lanes 2 and 6), maltopentaose (lanes 3 and 7), or maltohexaose (lanes 4 and 8).

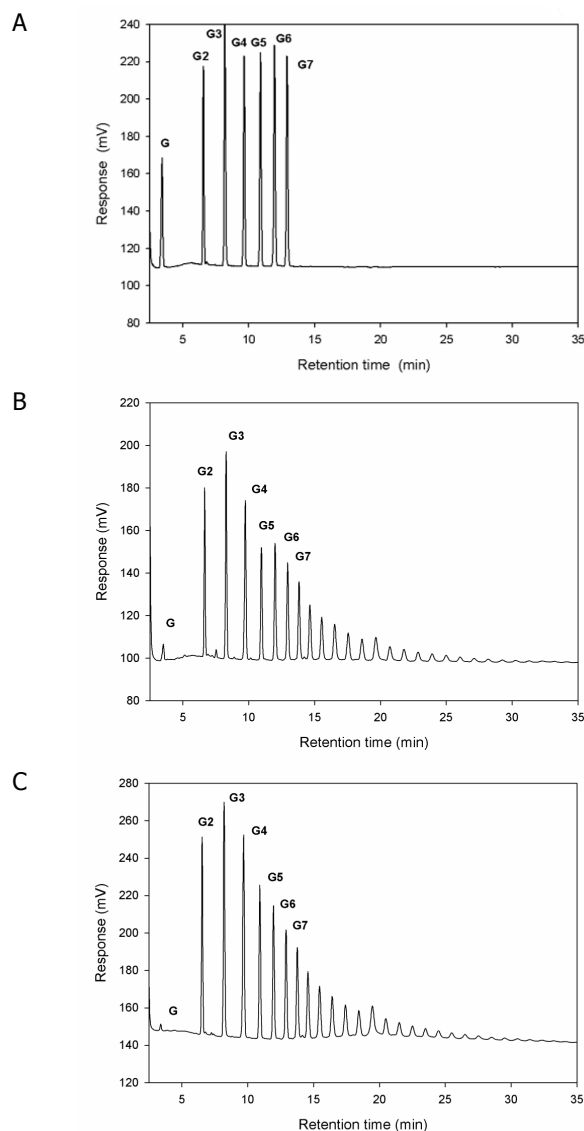


Fig. 6. HPLC analysis of the reaction products formed by incubation of AgtA and AgtB on maltoheptaose. (A) Elution profile of a standard mixture of maltooligosaccharides containing glucose (G1) to maltoheptaose (G7). (B and C) Reaction products of AgtA (B) and AgtB (C) after incubation of 0.4 μ g purified enzyme with 20 mM maltoheptaose at 37°C for 1 h.

AgtA and AgtB can use α -(1,3)-glucooligosaccharides as acceptor

The presence of the putative GPI-anchoring sequence in both AgtA and AgtB indicate that both enzymes are possibly present in the cell wall or cell membrane. The α -glucans present in the cell wall of *A. niger* are α -(1,3)-glucan (Horisberger et al., 1972) and nigeran (Bobbitt et al., 1977). In the previous paragraph we showed that neither nigeran nor α -(1,3)-glucan

were used as (donor) substrate by AgtA and AgtB. To investigate whether small substrates with α -(1,3)-glycosidic bonds could be used as an acceptor substrate, both enzymes were incubated with soluble starch as donor substrate, combined with glucose, maltose, nigerose or nigerotriose as acceptor substrates. Analysis of the reaction products by TLC and HPLC revealed that glucose was not used as acceptor substrate by either of the enzymes; maltose was an efficient acceptor as shown previously for AgtA-SBD (results not shown). AgtA also formed a series of oligosaccharides using either nigerose or nigerotriose as acceptor substrates, although the amount of products formed was smaller than when maltose acted as acceptor substrate (Fig. 7A). AgtB did not use nigerose or nigerotriose efficiently as acceptor substrate (Fig. 7B). No activity of AgtA or AgtB was observed on nigerose or nigerotriose as the sole substrate (result not shown). These results indicated that small α -(1,3)-linked oligosaccharides can be used as acceptor substrate by AgtA, and to a very limited extent by AgtB, but only in combination with an α -(1,4)-linked donor molecule.

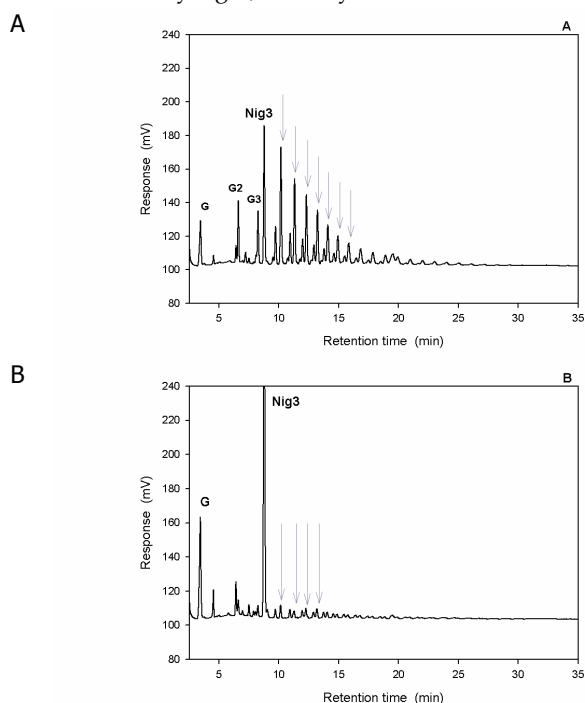


Fig. 7. HPLC analysis of the reaction products formed by AgtA (A) and AgtB (B) upon incubation with soluble starch and nigerotriose (Nig3). G, G2, and G3 indicate peaks representing glucose, maltose, and maltotriose, respectively. Peaks representing products most likely containing α -(1,3) glycosidic bonds, resulting from the use of nigerotriose as an acceptor substrate, are indicated with gray arrows. Enzyme (0.4 μ g) was incubated with 20 mM nigerotriose and 4% soluble starch at 37°C for 18 h.

Δ agtA and agtA/agtB overexpression strains are Calcofluor White hypersensitive

To examine the consequence of the loss of the *agtA* gene in *A. niger* and to analyze the physiological role of this enzyme, a deletion mutant of the *agtA* gene was constructed. Ten

randomly chosen p Δ agtA transformants were subjected to Southern blot analysis which revealed that in 8 out of the 10 transformants, the agtA gene was properly deleted (data not shown). Phenotypic analysis of several Δ agtA strains revealed that their growth rate on solid media was slightly reduced, but no changes in the morphology of the hyphae or conidia were observed. We analyzed the sensitivity of the Δ agtA strain to the cell wall-disturbing compound CFW. Hypersensitivity towards CFW has been shown to be indicative of mutants with impaired cell wall strength (Ram et al., 1994; Damveld et al., 2005a). As shown in Fig. 8A, the Δ agtA strains showed an increased sensitivity towards CFW. The observed sensitivity is not as strong as for the deletion of other *A. niger* cell wall-related proteins, such as α -(1,3)-glucan synthase A or the cell wall protein A (Damveld et al., 2005a, 2005b). The overexpression of agtA and agtB also resulted in an increased sensitivity towards CFW (Fig. 8B). All deletion and overexpression strains were also tested for their ability to grow on starch as the sole carbon source. Deletion of agtA and overexpression of AgtA or AgtB had no significant effect on the ability of *A. niger* to grow on starch compared to the untransformed strains.

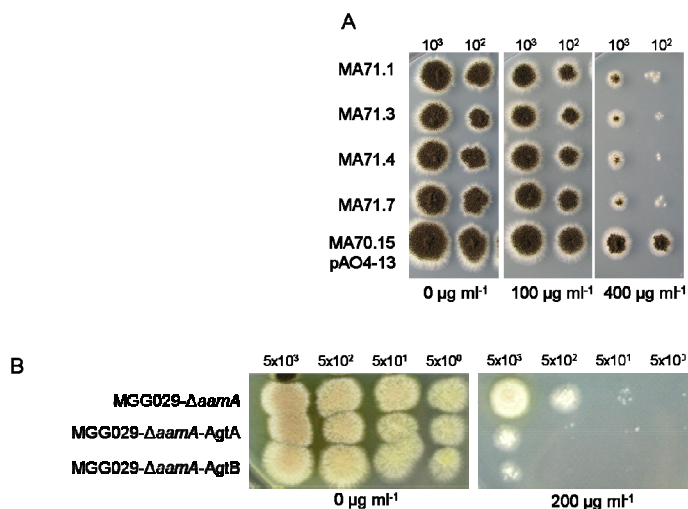


Fig. 8. Effect of agtA deletion (A) or agtA and agtB overexpression (B) on the susceptibility of the resulting strains towards CFW-induced cell wall stress. (A) *A. niger* Δ agtA strains (MA71.1, MA71.3, MA71.4, and MA71.7) and the control strain (MA70.15 transformed with pAO4-13), grown on complete medium containing 0, 100, or 400 µg ml⁻¹ CFW for 96 h. (B) *A. niger* strains overexpressing AgtA or AgtB and the parental strain (MGG029- Δ aamA) grown on complete medium containing 0 or 200 µg ml⁻¹ CFW for 72 h. The number of spores applied per spot is indicated above the panels.

Discussion

All putative GPI-anchored GH13 enzymes identified in the genome sequences of four aspergilli, as well as *N. crassa*, *M. grisea*, and *S. pombe*, were originally annotated as α -amylases, because of their high similarity to known extracellular fungal α -amylases. However, most of the protein sequences missed the commonly conserved His143 in region I

(Jespersen et al., 1991). Mutation of this residue, which is located in the active site (Uitdehaag et al., 1999), may result in a strongly reduced activity in α -amylases (Chang et al., 2003; Nakamura et al., 1993) or in altered reaction specificities in other GH13 enzymes (see Nakamura et al., 1993 and Leemhuis et al., 2004). Despite the missing His143 residue, the *A. niger* Agt enzymes and their homologues are clearly members of GH13, based on their high similarity with known proteins in this family. The second generally conserved His residue, His317 in conserved region IV, is replaced by Gln in AgtB and AgtC. Although His317 is overall highly conserved in the α -amylase family, this residue appears to be less important for the determination of the catalytic activity because some α -amylases are known which also do not possess a His in this position (Hoshiko et al., 1987; Kang et al., 2004). In conclusion, the Agt enzymes in *A. niger* and homologues in other fungi are highly similar to the well described fungal α -amylases, but aberrant conserved regions combined with the presence of a GPI-anchoring signal make them clearly distinguishable. We produced two of the *A. niger* Agt enzymes and determined their biochemical activities. As the enzymatic activity of the Agt proteins was not known at the start of this research, it was decided to start with the production of the enzyme encoded by gene An09g03100 (AgtA) fused with a clearly recognizable tag, the SBD. This allowed isolation and purification of the protein from media via binding to starch granules, as well as recognition of the purified protein using antibodies. Subsequently, both AgtA and AgtB (encoded by gene An12g02460) were produced in their native form for further biochemical characterization. A host *A. niger* strain with very low α -amylase activity was chosen to strongly reduce interference by native activities when searching for the biochemical activity of the investigated enzymes. AgtA-SBD as well as native AgtA and AgtB had a very low hydrolyzing activity on starch but clearly showed α -glucanotransferase activity on maltooligosaccharides alone and on maltooligosaccharides plus starch. We therefore propose to name the *A. niger* GPI-anchored GH13 enzymes AgtA and B (Agt for α -glucanotransferase). Since the predicted protein encoded by gene An15g07800 has the same sequence characteristics, we propose naming it AgtC.

AgtA and AgtB were overproduced from their entire predicted coding sequences, which included the predicted GPIanchoring signal. The presence of both AgtA and AgtB in the medium suggested that these proteins were not fully retained at the plasma membrane or cell wall but that at least part of the proteins was released into the medium. This might be an indication that the proteins are not GPI anchored, although their C-terminal signal sequence as well as the investigation of the homologous Aah3p from *S. pombe* (Morita et al., 2006) suggest otherwise. Another explanation is that the enzymes had been released by endogenous phospholipase C activity, as was previously shown to occur with GPI-anchored proteins in *A. niger* and *S. cerevisiae* (Brul et al., 1997).

AgtA and AgtB produced similar ranges of products consisting of maltooligosaccharides, indicating that both enzymes formed α -(1,4) glycosidic bonds and can therefore be classified as 4- α -glucanotransferases (EC 2.4.1.25). Also, a small amount of panose was produced by AgtA, indicating an ability to synthesize α -(1,6) linkages. Alternatively, panose may have been produced by a minor contamination of α -glucosidase,

which is known to produce α -(1,6) linkages (Kato et al., 2002). The enzymatic activities of AgtA and AgtB are unique among the α -glucanotransferases from bacteria as well as eukarya. The α -glucanotransferases that have been described until now usually release one glucose molecule for every transfer event (Takaha et al., 1993), but AgtA and AgtB did not produce glucose in significant amounts. Additionally, bacterial amylomaltases use very small donor and

acceptor molecules (maltotriose and glucose, respectively) (Kaper et al., 2004), while AgtA and AgtB prefer longer donor molecules with a minimum length of five glucose residues and maltose as the smallest possible acceptor substrate. The use of the α -(1,3)-linked oligosaccharides nigerose and nigerotriose as acceptor substrates by GH13 α -glucanotransferases has not been reported before. We conclude that the *A. niger* α -glucanotransferases represent a new subgroup of GH13 in view of their atypical donor and acceptor profiles and their C-terminal GPI anchoring sequences. Based on their common putative cell wall-associated location and amino acid sequences, it is expected that the closely related GPI-anchored GH13 proteins in other fungi will show similar glucanotransferase activities, although their precise substrate and product profiles remain to be determined.

Most extracellular members of GH13 are involved in the degradation of starch to supply energy and carbon to the cells. There are strong indications that this is not the case for AgtA and AgtB. This study describes that neither a knockout of *agtA* nor overexpression of AgtA or AgtB had an effect on the ability of *A. niger* strains to grow on starch, even if the parental strain was severely hampered in this trait. In another study, we have also shown that expression of the *agtA* and *agtB* genes is not regulated by AmyR, the general regulator for starch-processing enzymes in aspergilli (Yuan et al., submitted) (Petersen et al., 1999), and similar results were found for the homologous genes in *Aspergillus nidulans* (Nakamura et al., 2006). Taken together, these data indicate that the Agt proteins most likely are not involved in starch catabolism. An alternative function could be the production or modification of α -glucans in the fungal cell wall, which was suggested for one of the GPI-anchored Agt homologues in the fission yeast *S. pombe*, Aah3p, in a functional study (Morita et al., 2006). Deletion of *aah3* resulted in a morphological defect and hypersensitivity towards cell wall-degrading enzymes. The knockout could not be rescued by transformation with the *aah3* gene in which the catalytic residues had been mutated, showing the importance of the enzymatic activity rather than the structural properties of the protein. Our finding of a clear enzymatic activity for AgtA and AgtB confirms the importance of the catalytic residues for their physiological role. The proposed role for AgtA and its homologues in cell wall α -glucan production or maintenance is strengthened by the analysis of *agtA* knockout strains, which showed increased sensitivity towards CFW. The overexpression of AgtA and AgtB in *A. niger* caused a similar effect, which might be an indication that the unnaturally large amounts of these enzymes have a negative effect on cell wall strength.

The function of the Agt proteins in the fungal cell wall might be analogous to the function of GPI-anchored α -glucanosyltransferases. These enzymes, identified in several

yeasts and fungi, including aspergilli, play a role in the cross-linking of cell wall α -glucan (Mouyna et al., 2000; Popolo et al., 1993). In *Aspergillus fumigatus*, the α -glucan component of the cell wall is used as a target for antifungal drugs (Beauvais et al., 2001). A knockout of one of its GPI-anchored α -glucanosyltransferases, Gel2p, resulted in an altered cell wall composition, increased sensitivity for CFW, and reduced virulence (Mouyna et al., 2005). We tested the abilities of AgtA and AgtB to process the two α -glucan cell wall components, nigeran and α -(1,3) glucan, but no activity was detected. Because of their poor solubility at low pH, these substrates were offered at a relatively low concentration and partly in crystalline form, which might prevent the enzymes from acting on these cell wall components in vitro. It was shown, however, that AgtA performed a transglycosylation reaction involving an α -(1,4)-linked donor substrate and an α -(1,3)-linked acceptor substrate. A similar reaction was thought to occur in *S. pombe* cell walls, where two linear polysaccharide chains of α -(1,3)-glucan with several α -(1,4) linkages at the reducing end were interconnected by a transglycosylation reaction (Grun et al., 2005). Although this process was suggested to be performed by the transferase domain of α -(1,3)-glucan synthase Ags1p, a similar cross-linking reaction could be performed by AgtA. This would also explain the clustering of *agt* and *ags* (α -glucan synthase) genes conserved in many ascomycetes. Clustering of genes involved in the same metabolic pathway is well described in fungi (Keller et al., 1997).

To conclude, we have studied two novel putatively GPI anchored GH13 enzymes of *A. niger*, with homologues in many other fungi. The AgtA and AgtB enzymes both showed a unique type of α -(1,4) glucanotransferase activity, and our functional characterization indicated that their involvement in the facilitation of growth on starch is unlikely. The characterization of a knockout of *agtA* suggested that this enzyme could be involved in cell wall α -glucan synthesis, which is in line with the results on a knockout of a homologous protein from *S. pombe* (Morita et al., 2006). More study is needed to confirm the proposed physiological role for these α -glucanotransferases and to identify their exact in vivo reaction.

Acknowledgements

We are very grateful to Dr. Jacques Benen and Harrie Kools of Wageningen University for providing the *pki-pgaII* plasmid and the help with the HMMR searches. We also acknowledge the contribution of Peter Sanders (TNO Quality of Life, The Netherlands) for the HPLC analysis and Mark Arentshorst (Leiden University) for technical assistance with the knockout. We thank DSM for providing the *A. niger* genome sequence.

This work was supported by SenterNovem in the framework of the IOP-Genomics program (project IGE1021).

Chapter 5

Database mining and transcriptional analysis of genes encoding inulin modifying enzymes of *Aspergillus niger*

*Xiao-Lian Yuan**, *Coenie Goosen**, *Harrie Kools*, *Marc J.E.C. van der Maarel*,
Cees A.M.J. van den Hondel, *Lubbert Dijkhuizen* and *Arthur F.J. Ram*

** These authors contributed equally to this work*

Abstract

As a soil fungus, *Aspergillus niger* is able to metabolize a wide variety of carbon sources, employing sets of enzymes able to degrade plant derived polysaccharides. In this study, we have surveyed the genome sequence of *A. niger* strain CBS513.88, to analyze the gene/enzyme network involved in utilization of the plant storage polymer inulin, and of sucrose, the substrate for inulin synthesis in plants. In addition to three known activities, encoded by the *suc1* gene (invertase activity; designated *sucA*), the *inuE* gene (exo-inulinase activity) and the *inuA/inuB* genes (endo-inulinase activity), two new putative invertase-like proteins were identified. These two putative proteins lack N-terminal signal sequences and therefore are expected to be intracellular enzymes. One of these two genes, designated *sucB*, is expressed at a low level, and its expression is up-regulated when *A. niger* is grown on sucrose or inulin-containing media. Transcriptional analysis of the genes encoding the sucrose (*sucA*) and inulin hydrolyzing enzymes (*inuA* and *inuE*) indicated that they are similarly regulated and all strongly induced on sucrose and inulin. Analysis of a $\Delta creA$ mutant strain of *A. niger* revealed that expression of the extracellular inulinolytic enzymes is under control of the catabolite repressor CreA. Expression of the inulinolytic enzymes was not induced by fructose, not even in the $\Delta creA$ background, indicating that fructose did not act as an inducer. We provide evidence that sucrose, or a sucrose-derived intermediate, but not fructose, acts as an inducer for the expression of inulinolytic genes in *A. niger*.

Introduction

Inulins are linear polymers of fructose residues (fructans), which are primarily linked by β -2,1-glycosidic bonds, and usually followed by a terminal glucose moiety. Inulin is present as storage polysaccharide in roots and tubers of plants such as Jerusalem artichoke, chicory and dahlia (Cairns, 2003). Its presence has also been implicated in protection against water deficit in dry and cold conditions (Hendry & Wallace, 1993; Pilon-Smits *et al.*, 1995). Inulin in plants is synthesized by the concerted action of two fructosyltransferases, with sucrose as the primary fructosyl donor (see for review Ritsema & Smeekens, 2003). Inulin has attracted considerable research attention because it is a relatively inexpensive and abundant substrate for the production of fructose-rich syrups, as well as a source for the production of fructo-oligosaccharides (FOS). Both fructose syrups and FOS are regarded as 'functional foods' since they positively influence the composition of the intestinal microflora (Yun, 1996; Roberfroid & Delzenne, 1998; Kaplan & Hutkins, 2003).

Yeasts and filamentous fungi have been shown to employ various enzymes to degrade inulin and sucrose (Pandey *et al.*, 1999). Apart from displaying substrate hydrolysis, some of these enzymes can also perform transfructosylation reactions, producing the trisaccharide 1-kestose from sucrose (Rehm *et al.*, 1998; Sangeetha *et al.*, 2004; Yanai *et al.*, 2001) and even longer fructo-oligosaccharides (Heyer & Wendenburg, 2001). Currently, all known fungal inulin modifying enzymes are grouped together in family 32 of glycoside hydrolases (GH32) (<http://afmb.cnrs-mrs.fr/CAZY/index.html>, Coutinho &

Henrissat, 1999). Members of this family GH32 share conserved amino acid motifs and possess a similar three-dimensional protein structure (Pons *et al.*, 1998; Alberto *et al.*, 2004; Nagem *et al.*, 2004).

A. niger degrades inulin using both endo-inulinases (EC 3.2.1.7), encoded by the *inuA* and *inuB* genes (Ohta *et al.*, 1998; Akimoto *et al.*, 1999), and an exo-inulinase (EC 3.2.1.80), encoded by the *inuE* gene (Moriyama *et al.*, 2003). Endo-inulinase hydrolyzes inulin internally to produce mainly inulotriose and -tetraose (Akimoto *et al.*, 1999), whereas exo-inulinase hydrolyzes the terminal β -2,1-fructosidic bonds in both sucrose and inulin (Arand *et al.*, 2002; Kulminskaya *et al.*, 2003; Moriyama *et al.*, 2003). Invertase (β -fructofuranosidase, EC 3.2.1.26), encoded by the *suc1* gene (Boddy *et al.*, 1993), hydrolyzes the β -2,1-glycosidic bond in sucrose to produce fructose and glucose (L'Hocine *et al.*, 2000). A specific β -fructosyltransferase activity (EC 2.4.1.9) without significant invertase activity has been purified from *A. niger* strain AS0023. This enzyme transfers fructose residues from the non-reducing terminal β -2,1-glycosidic bond in sucrose to another sucrose or inulin molecule to form kestose or higher fructo-oligosaccharides (L'Hocine *et al.*, 2000). Unfortunately, the gene encoding this enzyme activity has not been identified and characterized yet.

Recent advances made in the area of genome sequencing of *A. niger* opened possibilities to further exploit this fungus to identify additional inulin modifying enzymes. The full genomic sequence of *A. niger* was made available to us by DSM Food Specialties (<http://www.dsm.com>). Based on deduced amino acid similarities, we have identified six putative proteins that belong to family GH32. Apart from the three known fungal enzymes (InuA/B, InuE and Suc1), three new putative inulin modifying enzymes were identified. One of them appears to be a pseudogene (*inuQ*), while the other two encode intracellular invertase-like proteins and were named SucB and SucC. The transcriptional regulation of these five putative inulin/sucrose modifying proteins in relation to various carbon sources has been studied in further detail.

Materials and methods

Strains and culture conditions

A. niger strain N402 used in this study was derived from the wild-type strain *A. niger* van Tieghem (CBS 120.49, ATCC 9029) (Bos *et al.*, 1988). The *A. niger* strain used for the sequencing of the genome by DSM is CBS513.88 (a natural derivative of strain NRRL3122). Strain AB4.1 is a *pyrG* derivative of N402 (van Hartingsveldt *et al.*, 1987) and was used to construct the *creA* deletion strain. *A. niger* strains were grown in Minimal Medium (MM) (Bennett & Lasure, 1991) containing 7 mM KCl, 11 mM KH₂PO₄, 70 mM NaNO₃; 2mM MgSO₄, 76 nM ZnSO₄, 178 nM H₃BO₃, 25 nM MnCl₂, 18 nM FeSO₄, 7.1 nM, CoCl₂, 6.4 nM CuSO₄, 6.2 nM Na₂MoO₄ and 174 nM EDTA. Erlenmeyer flasks of 300-ml were inoculated with 2×10^6 spores/ml and incubated at 30 °C in a rotary shaker at 300 rpm for 21 h. Each flask contained 100 ml of MM (pH 6.5) supplemented with 0.1% (w/v) casamino acids and

2% (w/v) carbon source. Glucose, sucrose (BDH Chemicals Ltd), xylose, fructose and maltose (Sigma-Aldrich), inulin (Sensus, Frutafit® TEX, Cosun, The Netherlands) and starch (Windmill starch, Avebe, The Netherlands) were used as carbon sources. For transfer experiments, N402 was pregrown in MM supplemented with 2% (w/v) xylose or 2% (v/v) glycerol and 0.1% (w/v) casamino acids for 18 h at 30 °C on a rotary shaker at 300 rpm. Mycelium was harvested by suction over a nylon membrane and washed with MM without carbon source. Aliquots of 1.5 g wet weight of mycelium were transferred to 70 ml of MM containing 1% (w/v) various carbon sources and incubated at 30 °C with agitation. Mycelial samples were taken at different time points by harvesting over a Myracloth filter followed by freezing in liquid nitrogen. The samples were stored at -80 °C prior to the isolation of total RNA. Conidiospores were obtained by harvesting spores from a complete medium plate (minimal medium with 0.5% (w/v) yeast extract and 0.1% (w/v) casamino acids) containing 1% (w/v) glucose, after 4-6 days of growth at 30 °C, using a 0.9% (w/v) NaCl solution.

Transformation of *A. niger* AB4.1 was as described by Punt & van den Hondel (1992) using lysing enzymes (L1412, Sigma-Aldrich) for protoplast formation. The bacterial strain used for transformation and amplification of recombinant DNA was *Escherichia coli* XL1-Blue (Stratagene, U.S.A). Transformation of XL1-Blue was performed according to the heat shock protocol as described by Inoue *et al.*, 1990.

Table 1. Fungal family GH32 proteins used for multiple sequence alignment in Figs. 1 and 2

Name	Main activity	Acc. No.	Organism	Reference
AngSucAp	Invertase	DQ233218	<i>Aspergillus niger</i> CBS 513.88	This study
AngSuc1p	Invertase	S33920	<i>Aspergillus niger</i> B60	Boddy <i>et al.</i> , 1993
AsySftAp	Fructosyltransferase	CAB89083	<i>Aspergillus sydowii</i> IAM 2544	Heyer & Wendenburg, 2001
AjaFopAp	Fructosyltransferase	BAB67771	<i>Aspergillus japonicus</i> ATCC 20611	Yanai <i>et al.</i> , 2001
AngSucBp	Putative invertase	DQ233219	<i>Aspergillus niger</i> CBS 513.88	This study
AngSucCp	Putative invertase	DQ233220	<i>Aspergillus niger</i> CBS 513.88	This study
AngInuEp	Exo-inulinase	DQ233222	<i>Aspergillus niger</i> CBS 513.88	This study
Afo1-Sstp	Fructosyltransferase	CAA04131	<i>Aspergillus foetidus</i> NRRL 337	Rehm <i>et al.</i> , 1998
Ang12InuEp	Exo-inulinase	BAD01476	<i>Aspergillus niger</i> 12	Moriyama <i>et al.</i> , 2003
Aawinu1p	Exo-inulinase	CAC44220	<i>Aspergillus awamori</i> var. 2250	Arand <i>et al.</i> , 2002
PspInuDp	Exo-inulinase	BAC16218	<i>Penicillium</i> sp. TN-88	Moriyama <i>et al.</i> , 2002
AngInuAp	Endo-inulinase	DQ233221	<i>Aspergillus niger</i> CBS 513.88	This study
Afilnu2p	Endo-inulinase	CAA07345	<i>Aspergillus ficuum</i> ATCC 16882	Uhm <i>et al.</i> , 1998
Ang12InuAp	Endo-inulinase	BAA33797	<i>Aspergillus niger</i> 12	Ohta <i>et al.</i> , 1998
Ang12InuBp	Endo-inulinase	BAA33798	<i>Aspergillus niger</i> 12	Ohta <i>et al.</i> , 1998
PpulnuAp	Endo-inulinase	BAA12321	<i>Penicillium purpurogenum</i>	Onodera <i>et al.</i> , 1996
PspInuCp	Endo-inulinase	BAB19132	<i>Penicillium</i> sp. TN-88	Akimoto <i>et al.</i> , 2000
Kmalnu1p	Inulinase	CAA40488	<i>Kluyveromyces marxianus</i> ATCC 12424	Laloux <i>et al.</i> , 1991
SceSuc2p	Invertase	P00724	<i>Saccharomyces cerevisiae</i>	Taussig & Carlson, 1983
Pjalnv1p	Invertase	CAA73208	<i>Pichia jadinii</i> NRRL-Y1084	Chávez <i>et al.</i> , 1998
PanInv1p	Invertase	CAA56684	<i>Pichia anomala</i> CBS5759	Pérez <i>et al.</i> , 1996
Spolnv1p	Invertase	BAA25684	<i>Schizosaccharomyces pombe</i> TP4	Tanaka <i>et al.</i> , 1998

Database mining of *A. niger* genome

The *A. niger* CBS513.88 genome has been determined by random sequencing of selected BACs to a 7.5-fold coverage. The resulting genome sequence (35.9 Mb) consists of approximately 400 contigs, which are assembled into 19 supercontigs (Dr. N. van Peij, DSM,

personal communication). The sequence of the *A. niger* genome is currently available to academic groups and non-profit organizations on request (hans.roubos@dsm.com) after signing a Material Transfer Agreement and will be generally available after publication of the *A. niger* genome sequence (H. Pel *et al.*, in preparation). Accession numbers of currently described families GH32 and GH68 members were selected from the Carbohydrate-Active Enzymes server at URL: <http://afmb.cnrs-mrs.fr/CAZY/> (Coutinho & Henrissat, 1999), and the corresponding protein sequences were extracted from the GenBank/GenPept database and Swiss-Prot database released at URL: <http://www.ncbi.nlm.nih.gov/entrez/> and URL: <http://www.expasy.org/sprot/>. Sequences were aligned with the ClustalW program (Thompson *et al.*, 1994; Chenna *et al.*, 2003) and transformed in a Hidden Markov Model (HMM) profile (Eddy, 1998) with the HMMbuild program from the HMMer package obtained at URL: <http://hmmer.wustl.edu/>. Subsequently the *A. niger* genome was searched using the HMM profiles and the wise2 package from URL: <http://www.ebi.ac.uk/Wise2/>. Multiple sequence alignment of known fungal fructan modifying enzymes, based on full length predicted protein sequences (Table 1), was performed using the ClustalW interface in Mega 3.1 (www.megasoftware.com) with gap-opening and extension penalties of 10 and 0.2, respectively. Bootstrap test of phylogeny was performed by the Neighbor Joining method using 1000 replicates.

Northern analysis

Total RNA was isolated by grinding frozen (-80 °C) mycelium in liquid nitrogen with a pestle and mortar. Powdered mycelium (200 mg) was extracted with 1 ml TRIzol Reagent (Invitrogen, U.S.A) in accordance with the supplier's instructions. For Northern analysis, 5 µg of total RNA was incubated with 3.3 µl 6 M glyoxal, 10µl DMSO and 2 µl 0.1M phosphate buffer (pH 7) in a total volume of 20 µl for 1 h at 50 °C to denature RNA. RNA electrophoresis was performed in a SEA-2000 electrophoresis apparatus (Elchrom Scientific, Switzerland) at 10 °C. The RNA samples were separated on 1.5% (w/v) agarose gel using 0.01M phosphate buffer (pH 5) and transferred to Hybond-N filters (Amersham, UK) by capillary blotting. Filters were prehybridized at 65 °C for 2h in a solution of 0.9M NaCl, 90 mM Na₃-citrate, 1.0% (w/v) ficoll, 1.0% (w/v) polyvinylpyrrolidone, 1.0% (w/v) bovine serum albumin, 10mM EDTA, 0.5% (w/v) SDS and 100 µg/ml single stranded herring sperm DNA. Hybridizations were performed at 42 °C for 18 h in a solution of 50% (v/v) formamide, 10% (w/v) dextran sulphate, 0.9M NaCl, 90 mM Na₃-citrate, 0.2% (w/v) ficoll, 0.4% (w/v) polyvinylpyrrolidone, 0.4% (w/v) bovine serum albumin, 0.4% (w/v) SDS and 100 µg/ml single stranded herring sperm DNA. Blots were washed twice in high stringency washing buffer (30 mM NaCl, 3 mM Na₃-citrate and 0.5% (w/v) SDS) for 20 min at 65 °C. Probes for the detection of the six (putative) sucrose and fructan modifying enzymes of *A. niger*, were generated using six pairs of oligonucleotide primers by PCR using *A. niger* N402 genomic DNA as template (Supplementary Table 1). The PCR-amplified fragments were run on an agarose gel and purified from the gel. The purified DNA fragments were cloned into plasmid pGEMT-easy and sequenced to confirm their identity. Probes were generated by digestion of the pGEMT-easy vector containing the inulinolytic enzyme-encoding gene with

*Eco*RI. Fragments were purified from gel and [α -³²P]dCTP labeled probes were synthesized using Rediprime II DNA labeling System (Amersham Pharmacia Biotech, UK) according to the instructions of the manufacturer.

Disruption of the carbon catabolite repressor CreA in *A. niger*

The plasmid used to disrupt the *creA* gene was constructed as follows. The DNA fragments flanking the *creA* ORF were amplified by PCR using N402 genomic DNA as template: 1.4 kb of 5' flanking DNA and 0.9 kb of 3' flanking DNA was amplified by PCR using primers CreAP1f and CreAP2r, CreAP3f and CreAP4r (Supplementary Table 1), respectively. Each primer was adapted with a restriction site for further cloning. The amplified PCR fragments were digested with *Not*I and *Bam*HI or *Bam*HI and *Kpn*I respectively, and cloned into pBlue-ScriptII SK to obtain plasmids pF5 and pF3. Subsequently, pF3 was digested with *Bam*HI and *Kpn*I, and the obtained fragment was ligated into *Bam*HI and *Kpn*I digested pF5 to give pF53. pF53 was digested with *Sal*I and *Bam*HI and inserted with the *Sal*I-*Bam*HI fragment containing *A. oryzae* *pyrG* gene, obtained from plasmid pAO4-13 (de Ruiter-Jacobs *et al.*, 1989) and resulted in the *creA* disruption plasmid, pXY1.1. The plasmid pXY1.1 was linearized with *Not*I and transformed to AB4.1. Uridine prototrophic transformants were selected by incubating protoplasts on agar plates containing MM without uridine. Transformants were purified and genomic DNA was isolated and analyzed by PCR to identify possible $\Delta creA$ strains. Primer pairs used to identify homologous recombination of the *creA* deletion construct on the *creA* locus were creAP5f and PAO10 or PAO9 and creA6f. Primer pairs used in the PCR to analyze the presence of the wild type *creA* gene were creAP5f and creAP7r, creAP8f and creAP6r (Supplementary Table 1). Three independent *creA* deletion strains with identical phenotypes were obtained and designated XY1.1, XY1.2 and XY1.3. Strain XY1.1 was further used for analysis of the expression of inulin modifying enzymes and we will refer to this strain as the $\Delta creA$ strain in the remaining of the paper. For complementation of the $\Delta creA$ strain, the *creA* gene, including 1.3 kb promoter and 0.8 kb of terminator sequences, was amplified by PCR using primers creAP1f and creAP6r. The PCR product of 3.5 kb was cloned into pGEMT-easy (Promega) and co-transformed with pAN7.1 (Punt *et al.*, 1987) to $\Delta creA$ strain XY1.1 to generate XY1.1-CreA.

Nucleotide accession numbers

The *A. niger* CBS513.88 DNA sequences encoding family GH32 members, including 1000 bp up- and downstream of open reading frame and their predicted protein sequences were obtained from DSM (Dr. G. Groot). The sequence data have been submitted to the GenBank database under accession numbers DQ233218 (*sucA*), DQ233219 (*sucB*), DQ233220 (*sucC*), DQ233221 (*inuA*), DQ23322 (*inuE*) and DQ233223 (*inuQ*).

Results

Identification of Glycoside Hydrolase family 32 members in the *A. niger* genome

Glycoside hydrolase families GH32 and GH68 include invertase, levanase, inulinase and levansucrase enzymes from bacterial, fungal and plant origin (Coutinho & Henrissat 1999; Pons *et al.*, 1998). Both families are structurally similar (clan GH-J), sharing a similar 5-fold β -propeller fold (Meng & Futterer, 2003; Nagem *et al.*, 2004). Protein sequences from members of families GH32 and GH68 were extracted from GenBank/GenPept and SwissProt databases and were used to construct HMM profiles to identify additional members in the genome of *A. niger* CBS513.88. Family GH68 profiles did not return any significant matches. The family GH32 profile did return five significant sequences. In addition to three family GH32 members already described for *A. niger* (invertase (Suc1, Boddy *et al.*, 1993), exo-inulinase (InuE, Moriyama *et al.*, 2003) and endo-inulinases (InuA/InuB, Ohta *et al.*, 1998), two new members were identified which were named SucB and SucC.

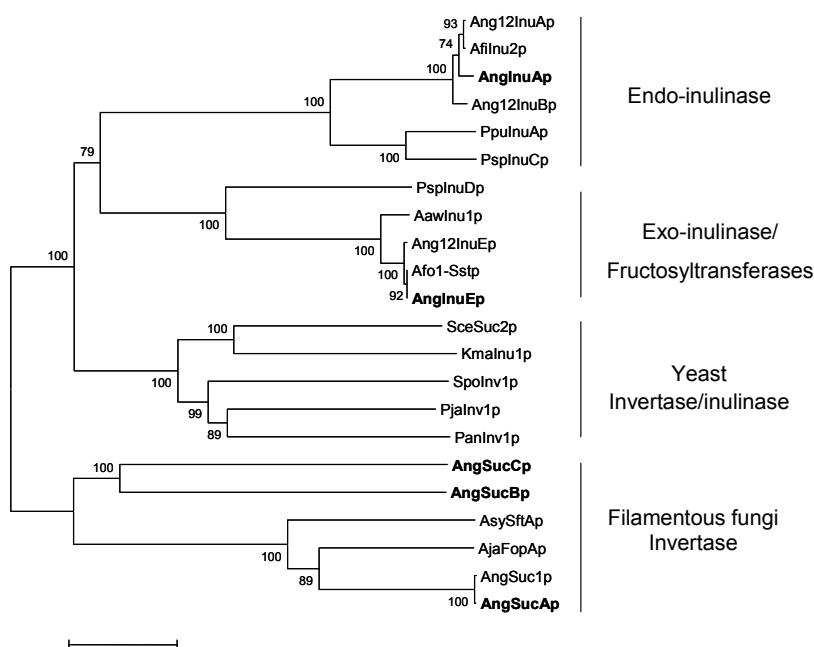


Fig. 1. Neighbour-joining tree of functionally characterized GH32 family members from filamentous fungi and yeast species. GH32 proteins identified in the genome of *A. niger* CBS513.88 are shown in bold. The proteins, their main activities, accession numbers, and the source (organisms) of the protein sequences used in this alignment, are listed in Table 1. Bootstrap values are indicated on the node of each branch. The tree was created with Mega 3.1 using default settings for gap and extension penalties. Bar indicates 10% amino acid sequence difference.

A neighbor-joining tree based on phylogenetic analysis of all currently available functionally described family GH32 fungal protein sequences was constructed. As shown in Fig. 1, four sub-groups of fungal inulinolytic enzymes can be distinguished: i) exo-inulinases/ fructosyltransferases, ii) endo-inulinases, iii) yeast invertases/inulinases and iv) invertases from filamentous fungi. The complete inventory and the notion that these 4 groups are evident within the fungal GH32 enzymes had not been noticed in earlier studies in which phylogenetic trees of GH32 family proteins had been constructed (Pons *et al.*, 1998; Moriyama *et al.*, 2003). In the secondary structure of microbial GH32 proteins eight well conserved domains (A, B, B1, C, D, E, F and G, respectively) can be distinguished (Pons *et al.*, 1998; Ohta *et al.*, 1998). Domains A, D, and E (designated block I, II and III in Pons *et al.*, 2004), contain highly conserved acidic residues that are located in the active site of members of family GH32. These highly conserved acidic domains, as well as the other conserved domains, are also present in most of the fungal fructan modifying enzymes (Fig. 2).

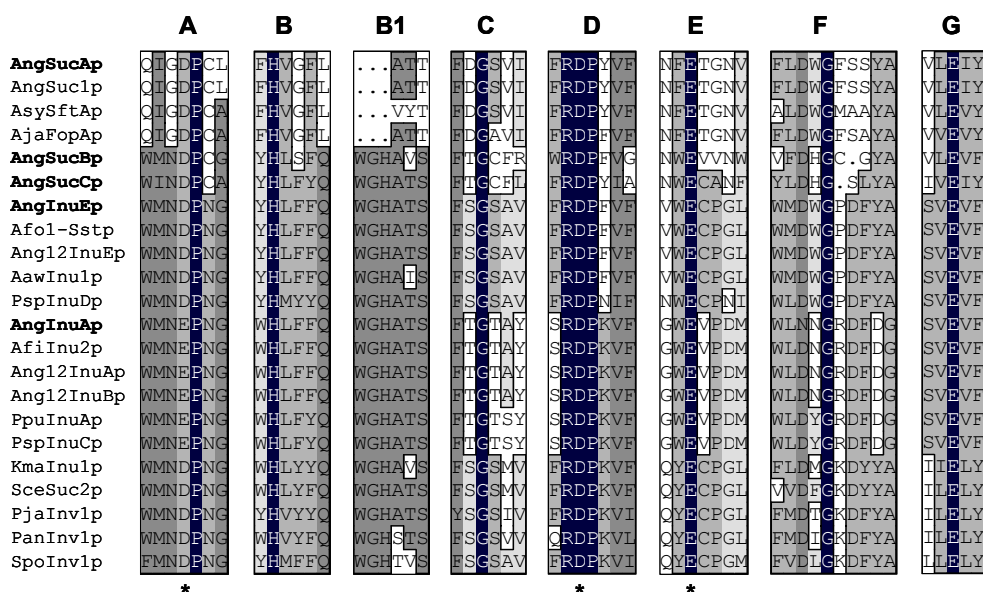


Fig. 2. Conserved domain alignment of family GH32 enzymes from filamentous fungi and yeast species. Conserved residues are shaded by different intensities based on homology level in the alignment. Black, 100% homology; dark grey, $\geq 75\%$ homology; light grey, $\geq 50\%$ homology; very light grey, $\geq 33\%$ homology. Conserved acidic residues are indicated with asterisk (*). The conserved 8 domains (A, B, B1, C, D, E, F and G) are indicated at the top of the alignment. Proteins used are also shown in Fig. 1 and listed in Table 1.

It has been previously reported that *A. niger* contains two genes that encode endo-inulinases that differ by only 8 of the 516 amino acids (Ohta *et al.*, 1998). In the genome sequence of *A. niger* strain CBS513.88, a single gene encoding an endo-inulinase could be identified. The presence of single-copy endo-inulinase genes in *A. niger* strain CBS513.88 and in

A. niger strain N402 was confirmed by Southern blot analysis (data not shown). Sequence comparison of the endo-inulinase of CBS513.88 and the endo-inulinase of *A. niger* strain 12 revealed that the endo-inulinase from strain CBS513.88 displayed higher similarity to InuA (9 amino acid differences) than to InuB (15 amino acid differences) of *A. niger* strain 12. The single endo-inulinase in strain CBS513.88 has been designated InuA.

Besides the three known inulinolytic activities, two previously unknown GH32 family members were identified in the *A. niger* CBS513.88 genome sequence. Unlike the SucA, InuA and InuE proteins, SucB and SucC lack an N-terminal signal sequence, or any other recognizable targeting signal, which suggests that the proteins are localized intracellularly, in the cytosol. SucB and SucC protein sequences were analyzed using SecretomeP 1.0b prediction (<http://www.cbs.dtu.dk/services/SecretomeP-1.0>) (Bendtsen *et al.*, 2004) to assess if these enzymes might be secreted via a non-classical secretion pathway. Both enzymes have a NN-score close to the threshold value of 0.6 (0.651 for SucB and 0.586 for SucC) which does not exclude the possibility that the two proteins are secreted via a non-classical secretion pathway. This data should be interpreted with care, as the program has been trained using sequences of human non-classical exported proteins.

Phylogenetic analysis indicated that SucB and SucC group together with fungal invertases (Fig.1). Comparison of the deduced amino acid sequence of SucB with all functionally described GH32 family proteins revealed highest identity to the *A. niger* SucA protein (24% identity, 41% positives, e-value of 4×10^{-32}). SucC also displays the highest identity to the *A. niger* SucA protein (28% identity, 42% positives, e-value of 2×10^{-43}). Pairwise comparison indicated that SucB and SucC are more identical to each other (35% identity, 52% positives, e-value of 2×10^{-97}) than to any other functionally annotated GH32 family member (Fig. 1). The *sucB* and *sucC* genes encode proteins of 617 and 601 amino acids respectively, and contain all conserved domains (A-G), including the conserved acidic residues in domains A, D and E (Fig. 2).

Identification of Glycoside Hydrolase family 32 members in other fungal genomes.

The genomes of the ascomycete fungi *A. nidulans*, *A. fumigatus*, *Neurospora crassa*, *G. zeae*, *Magnaporthe grisea* and the basidiomycete fungus *Ustilago maydis* were analyzed for the presence of GH32 family members. The results are summarized in Table 2 and a phylogenetic tree is given as a supplementary figure (Suppl. Fig. 1). *A. nidulans* contains two genes that encode proteins belonging to the GH32 family. AN5012.2 displays the highest identity to the *A. niger* InuE and probably represents the extracellular *A. nidulans* sucrose-hydrolyzing activity described by Vainstein & Peberdy (1990; 1991). An invertase, similar to the *A. niger* SucA protein appears to be absent in *A. nidulans*. The second member of the GH32 family (AN3837.2) is highly identical to the *A. niger* SucB protein (Table 2). Like in *A. niger*, the protein is predicted to be intracellular in *A. nidulans*, as an N-terminal signal sequence is not present. Orthologs of an endo-inulinase (InuA-like) or a second intracellular invertase-like protein (SucC-like) as found in *A. niger*, were not found in the *A. nidulans* genome. *A. fumigatus* contains four genes that encode proteins belonging to the GH32 family. Based on the sequence alignments, Afu5g00530 and Afu5g00480 are likely to encode

the endo- and exo-inulinases, respectively. Afu6g05000 is highly similar to the *A. niger* SucB protein and also predicted to be intracellularly localized. The fourth GH32 family member, Afu2g01240 shows the highest sequence identity to yeast-like invertases, also reflected in the sequence alignment (Suppl. Fig. 1). *N. crassa* contains only a single GH32 family member. This protein displays the highest sequence identity with beta-fructofuranosidases from bacterial origin (*Bacillus megaterium* FruA; e-value $3e^{-72}$) (Chiou *et al.*, 2002) and groups in the fungal tree together with the yeast-like invertases (Suppl. Fig. 1). *G. zeae* (anamorph *Fusarium graminearum*) contains five GH32 family members. FG08415.1 is the putative ortholog of the *A. niger* SucA protein. In the original annotation, the protein lacks an N-terminal signal sequence. However, deleting the first 81 amino acids of the predicted protein sequence renders a protein of 619 amino acids, containing a predicted signal sequence. FG02339.1 is most homologous to the *S. cerevisiae* Suc2 protein (e-value $3e^{-82}$) and clusters together with yeast-like invertases. FG03288.1 encodes a protein that is highly identical to the SucC protein of *A. niger*. The protein is also predicted to be intracellularly localized. FG06451.1 shows the highest identity (e-value $1e^{-72}$) to the *Bacillus subtilis* SacC protein which has levanase activity (Martin *et al.*, 1987). FG02067.1 is an interesting member of the GH32 gene family as it shows very limited sequence identity to the known enzymes. It is predicted to be an intracellular enzyme and shows the highest identity (e-value $2e^{-09}$) to a cycloinulooligosaccharide fructanotransferase from *Bacillus macerans* (Kim & Choi, 2001). *M. grisea* contains five GH32 enzymes, four of them cluster in a separate branch different from the other fungal proteins (Suppl. Fig. 1). One of the predicted enzymes (MG10748.4) is N-terminally truncated because the ORF is at the end of a contig. The annotation of MG10767.4 has been improved, resulting in a protein that contains all the conserved GH32 domains. MG02507.4 clusters in the group of yeast-like invertases (Table 2, Suppl. Fig. 1), whereas the remaining four form a distinct group with high sequence identities to each other (e-value $1e^{-112}$ or lower). Three of the enzymes are predicted to be secreted, for the fourth protein this is not known (N-terminally truncated because this ORF is at the end of a contig). It will be of interest to characterize the biochemical properties of those enzymes. Finally, *Ustilago maydis* contains two GH32 family members. UM01945.1 encodes a protein that is most similar to yeast-like invertases, while UM03605.1 encodes a predicted intracellular protein mostly related to the *A. niger* SucC protein.

The comparison of the GH32 enzymes from various filamentous fungi thus reveals that different fungi contain different repertoires for the utilization of inulin and or sucrose. Intracellularly located GH32 enzymes appear not unique for *A. niger* as the presence of these proteins is also predicted in the other fungi (Table 2). With one exception (FG02067), these enzymes cluster as a distinct group in the phylogenetic tree (Suppl. Fig. 1).

Table 2. Inventory of GH32 proteins in genomes of filamentous fungi.

Fungus	ORF number	Accession number	Closest <i>A. niger</i> homolog ^{*2}	Predicted localization ^{*3}
<i>A. nidulans</i>	AN3837.2	EAA59102	SucB; 1e-144	Intracellular
	AN5012.2 ^{*1}	EAA61090	InuE; 4e-166	Extracellular
<i>A. fumigatus</i>	Afu2g01240	EAL87222	InuE; 2e-70	Extracellular
	Afu5g00530	EAL86248	InuA; 4e-183	Extracellular
	Afu5g00480	EAL86253	InuE; 3e-169	Extracellular
	Afu6g05000	EAL85540	SucB; 1e-182	Intracellular
<i>N. crassa</i>	NCU04265.2	CAC28747	InuE; 3e-41	Extracellular
<i>G. zeae</i>	FG02067.1	EAA69589	InuA; 5e-05	Intracellular
	FG02339.1	EAA69879	InuE; 7e-59	Extracellular
	FG03288.1	EAA72735	SucC; e-136	Intracellular
	FG06451.1	EAA78236	InuE; 4e-61	Extracellular
	FG08415.1 ^{*1}	EAA71892	SucA; 0.0	Extracellular
<i>M. griseae</i>	MG02507.4	XP_365805	InuE; 1e-55	Extracellular
	MG05785.4	XP_369679	SucA; 2e-49	Extracellular
	MG07837.4	XP_367933	SucB; 8e-36	Extracellular
	MG10748.4	XP_360436	SucA; 1e-08	Truncated protein; Domains A, B1, B, C, D, F are missing
	MG10767.4 ^{*1}	XP_360455	SucA; 2e-14	Extracellular
<i>U. maydis</i>	UM01945.1	EAK82399	InuE; 1e-66	Extracellular
	UM03605.1	EAK84508	SucB; 2e-89	Intracellular

*1 Annotation improved manually.

*2 BlastP searches were performed with the full length protein sequences

*3 Localization was predicted using SignalP 3.0 Prediction program at <http://www.cbs.dtu.dk/services/SignalP/>. Proteins with a predicted N-terminal signal sequence were considered as extracellular proteins. Proteins lacking an N-terminal signal sequence were considered as intracellular enzymes.

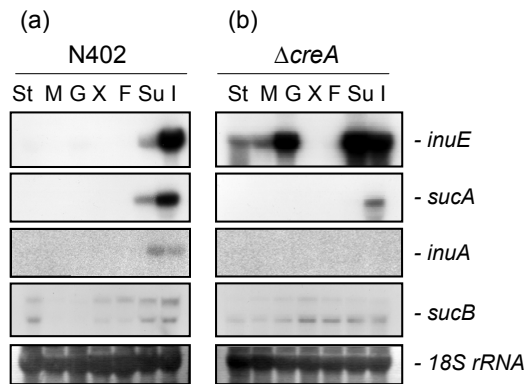


Fig. 3. Expression analysis of *A. niger* inulinolytic genes in wild type N402 (A) and the $\Delta creA$ strain (B). Total RNA was extracted from mycelia grown for 21 h at 30 °C in MM containing 2% (w/v) of different carbon sources. RNA (10 μ g) was separated by agarose gel electrophoresis, blotted, and hybridized with ³²P-labeled probes specific for inulinolytic genes. St, starch; M, maltose; G, glucose; X, xylose; F, fructose; Su, sucrose; I, inulin. Ethidium Bromide staining of 18S rRNA was used as loading control.

Transcriptional regulation of inulin modifying enzymes in *A. niger*

The expression of the five putative inulin/sucrose modifying enzymes identified in the genome of *A. niger* in relation to the presence of different carbon sources was studied by Northern blot analysis. RNA was extracted from *A. niger* N402 mycelia grown in minimal medium containing xylose, glucose, maltose, starch, fructose, sucrose or inulin (all 2% (w/v)) as sole carbon source. Expression of the *inuE*, *sucA* and *inuA* genes could be detected only on sucrose and inulin (Fig. 3A). *sucB* was not only expressed in the presence of sucrose and inulin; weak expression was also detected on other carbon sources (Fig. 3A). *SucB* expression on glucose and maltose was detected after longer exposure time of Northern blot (not shown). The detection of two different sized mRNAs for the *sucB* gene suggests two different mRNA start sites, or two different poly-adenylation sites. The presence or absence of one of the two mRNAs was not correlated with a particular carbon source. Expression of *sucC* was not detected on any of the carbon sources tested (results not shown). The differences in expression level of the various genes (relatively low on sucrose and high on inulin) might be caused by differences in the growth stage of the cultures. Growth of *A. niger* on sucrose is much faster than on inulin. After 21 h, the sucrose-grown culture was in the mid-exponential phase of growth, while the inulin-grown culture was still in the early-exponential phase (data not shown).

The expression of genes encoding inulinolytic enzymes was also studied using transfer experiments. *A. niger* strain N402 was pre-grown in 2% (w/v) xylose medium for 18 h and mycelium was transferred to either 1% (w/v) inulin, 1% (w/v) sucrose or 1% (w/v) maltose medium and grown further for 2, 4, 8 and 24 h. Since the inulinolytic genes were not expressed on xylose (Fig. 3A), this carbon source was chosen for pre-culturing. Mycelia were isolated before the transfer ($t = 0$ h) and at specific time points after the transfer (2, 4, 8, and 24 h) and total RNA was isolated and subjected to Northern analysis. As expected, no expression of inulinolytic enzymes was observed during growth on xylose (Fig. 4, $t = 0$ h). The expression of *inuE*, *sucA*, *inuA*, and *sucB* were all induced after transfer from xylose to sucrose and inulin (Fig. 4). No induction of any of the inulinolytic genes was observed on maltose although a low level of expression of the *inuE* could be observed. The induction of genes on sucrose was much faster and more dramatic compared to the response to inulin, reflecting the much faster growth on sucrose than on inulin (see above). The expression of the *inuE*, *sucA* and *inuA* genes decreased after 4 h growth on sucrose, probably due to the rapid utilization or conversion of the available sucrose. As inulin was utilized more slowly, expression levels remained higher for a longer period. The results from Fig. 4 also indicate that, although the different genes are co-regulated and all induced on sucrose or inulin, slight differences in expression pattern were observed. Whereas the *inuE* transcript remains present most abundantly at 24 h after transfer to inulin, the levels of *sucA* and *inuA* mRNA had already decreased. We have no conclusive explanation for these observations; these differences maybe due to differences in mRNA stability, or that expression is also under control of multiple regulatory mechanisms such as pH regulation, mediated by the PacC transcriptional regulator or carbon catabolite repression mediated by the CreA repressor

protein. The presence of PacC (5'-GCCARG-3') (see for review Penalva & Arst, 2002) and CreA binding-sites (5'-SYGGRG-3') (see for review Ruijter & Visser, 1997) in the promoters of most of these genes might support the latter explanation.

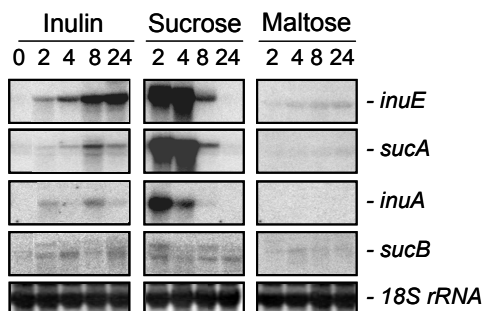


Fig. 4. Northern blot analysis of inulinolytic genes in *A. niger* wild type strain after transfer from xylose to different carbon sources. *A. niger* was pregrown in MM containing 2% (w/v) xylose and 0.1% (w/v) casamino acids at 30 °C for 18 h. Total RNA was prepared from mycelia grown for 2, 4, 8 and 24 h after transfer from pregrown culture to fresh MM containing 1% (w/v) inulin, 1% (w/v) sucrose or 1% (w/v) maltose. 5 µg of total RNA was loaded on the gel. Ethidium Bromide staining of 18S rRNA was used as loading control.

Expression analysis of the inulinolytic genes in a $\Delta creA$ background.

In the yeast *Kluyveromyces fragilis*, fructose has been described as an inducer of inulinase expression (Grootwassink & Hewitt, 1983). The results shown in Fig.3A indicate that the genes encoding the inulinolytic enzymes are not expressed on fructose in *A. niger*. This might be caused by the high concentration of fructose used, resulting in carbon catabolite repression mediated by CreA. Therefore, the expression of genes encoding inulinolytic enzymes was also examined in the $\Delta creA$ mutant of *A. niger* strain N402. A $\Delta creA$ null mutant was generated and verified as described in Materials and Methods. The phenotype of the $\Delta creA$ strain was similar to that described for severe *creA* mutants in other *A. niger* strains and in *A. nidulans* (Ruijter & Visser, 1997; Shroff *et al.*, 1997), including reduced radial growth and reduced conidiation (data not shown). Transformation of the $\Delta creA$ strain with the wild type *creA* gene fully complemented the reduced growth and reduced conidiation phenotype of the $\Delta creA$ (data not shown).

The expression analysis of the *inuE* gene in the $\Delta creA$ strain showed that the expression of *inuE* gene was higher than in the wild type strain, indicating that this gene is under control of carbon metabolite repression. Expression of *inuE* was found in the $\Delta creA$ strain after growth on glucose, and to some extent on maltose and starch. This is in contrast to the expression of *inuE* on xylose and fructose which is non-detectable in both the wild type and the $\Delta creA$ strain. Apparently, the expression of the *inuE* gene requires an activator or inducer molecule and is not expressed under derepressing ($\Delta creA$) conditions. Inspection of the 1-kb promoter sequence of the *inuE* gene revealed the presence of 13 putative CreA binding sites that might be involved in mediating repression, but their functionality has not been studied. Analysis of the expression of the other genes encoding the inulinolytic

enzymes showed a different expression pattern in the $\Delta creA$ strain. *SucA* and *inuA*, both expressed specifically on inulin and sucrose in the wild type strain, were lower expressed in the $\Delta creA$ strain (*sucA*), or not detectable in the $\Delta creA$ strain (*inuA*) (Fig. 3B). Expression of *sucA* and *inuA* was not detected when grown on xylose and fructose, similar as observed for the *inuE* gene. In contrast to the expression profile of the *inuE* gene, no expression was detected of the *sucA* and *inuA* gene after growth on starch, maltose and glucose. Thus, although CreA binding sites are present in the 1 kb promoter regions of the *sucA* and the *inuA* genes (6 and 5 sites, respectively), the expression of both genes does not seem to be directly controlled by CreA. Disruption of catabolite repression in the $\Delta creA$ strain could lead to the inactivation of complex pathways, which might bring about decreased or total loss of expression. The expression pattern of the *sucB* gene also suggests that this gene is, like *inuE*, under carbon catabolite repression control. *SucB* is not exclusively expressed on inulin and sucrose, but also on starch, xylose and fructose, and to a lesser extent on maltose and glucose (Figs. 3 and 4). Expression of the *sucB* gene in the $\Delta creA$ background showed two remarkable features i) expression of *sucB* in the $\Delta creA$ strain was also detected on maltose and glucose, in addition to the other carbon sources, indicating that CreA mediated catabolite repression of *sucB* on maltose and glucose in the wild type strain, ii) expression of *sucB* in the $\Delta creA$ background favors the transcription of the smaller sized mRNA molecule. The 1-kb promoter region of *sucB* contains 7 putative CreA binding sites. Expression of the *sucC* gene was also not detected in the $\Delta creA$ mutant (data not shown). Clearly, the different responses of the *inuE* gene, the *sucB* gene and the *sucA/inuA* genes in a $\Delta creA$ mutant background suggest the involvement of additional factors other than the presence of an inducer and repression via CreA. The additional factors might include environmental factors such as pH, nitrogen availability and temperature.

Sucrose acts as an inducer of the inulinolytic system in *A. niger*

Fructose has been shown to induce expression of inulinase in the yeast *K. fragilis* (Grootwassink & Hewitt, 1983). The expression analysis of the inulinolytic genes in the wild-type *A. niger* strain and the $\Delta creA$ strain after growth on fructose did not result in detectable expression of any of the genes, indicating that fructose is not the inducing molecule for expression (Fig. 3).

Further evidence that fructose did not act as an inducer for the expression of the inulinolytic genes was obtained in a transfer experiment. Wild-type *A. niger* strain N402 was pre-grown in 2% (w/v) glycerol minimal medium for 18 h and mycelium was transferred to minimal medium containing decreasing concentrations of fructose: 50 mM, 5 mM, 500 μ M, 50 μ M, 5 μ M, 500 nM, 50 nM, 5 nM, and no carbon sources. As shown in Fig. 5A, this gradual decrease in fructose did not result in expression of the different inulinolytic genes, not even after 4 h of growth. A similar transfer experiment was performed to medium containing sucrose at an identical concentration series. As expected, sucrose induced expression of the genes encoding the inulinolytic enzymes (Fig. 5B). The induction of *inuE*, *sucA*, *inuA* and *sucB* reached the highest level at a concentration of 50 μ M sucrose, indicating that some form of repression at high sucrose concentrations may exist, i.e.

catabolite repression by released glucose from sucrose hydrolysis. The addition of 50 μ M or 5 μ M of glucose, 1-kestose or 1-nystose did not trigger induction of the inulinolytic system (data not shown).

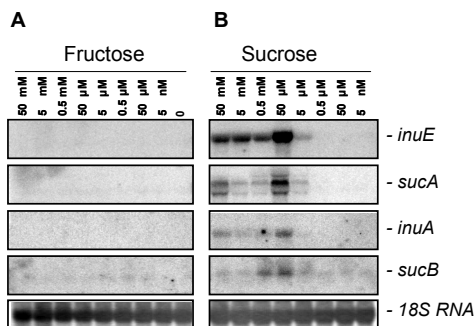


Fig. 5. Northern blot analysis of genes encoding inulinolytic enzymes of *A. niger* N402 using carbon-shift experiments. *A. niger* was pregrown in MM containing 2% (w/v) glycerol at 30 °C for 18 h. Total RNA was prepared from mycelia grown for 4 h after transfer to fresh MM containing concentrations of fructose (A) or sucrose (B) as indicated. 5 μ g of total RNA was loaded on the gel. Ethidium bromide staining of 18S rRNA was used as loading control.

Discussion

The availability of the *A. niger* genome sequence (http://www.dsm.com/en_US/html/dfs/genomics_aniger.htm) and subsequent search for inulinolytic enzymes using HMM profiles generated for GH32 family members, has allowed identification of five putative enzymes acting on inulin and/or sucrose in the genome of *A. niger*. A sixth gene was identified, but considered as a pseudogene because of the various stop codons that were present in the predicted open reading frame. In addition to three extracellular inulinolytic activities of *A. niger*, the invertase encoded by the *suc1* gene, the exo-inulinase encoded by the *inuE* gene, and the endo-inulinase activity encoded by the *inuA* gene, two additional genes, *sucB* and *sucC* were identified. We will briefly discuss the different enzymes in the genome of *A. niger* strain CBS513.88 in relation to previously identified enzymatic activities in *Aspergilli* and in relation to their activity. The *A. niger* invertase activity has been shown to be encoded by the *suc1* gene (Boddy *et al.*, 1993). The Suc1 protein from *A. niger* strain B60 has been described as a 566 amino acid protein; its protein sequence is identical to that of the SucA protein encoded by the genome of *A. niger* CBS513.88. The Suc1 protein groups in the phylogenetic tree with the previously described fructosyltransferase FopA (Yanai *et al.*, 2001). FopA originally was published as an enzyme from *A. niger* strain ATCC20611, but this strain has been reclassified as an *A. japonicus* strain (ATTC culture collection database; <http://www.lgcpromochem-atcc.com/common/catalog/numSearch/numResults.cfm?atccNum=20611>). Compared to the Suc1 protein sequence (Boddy *et al.*, 1993), FopA has an extra C-terminal extension consisting of 38 amino acids (Yanai *et al.*, 2001). We have re-examined the *suc1* coding region and from that analysis we predict the existence of a 55-bp long intron sequence 1 bp

upstream of the termination codon. Taking the intron sequence into account, the Suc1 protein is predicted to become 39 amino acids longer than the sequence published by Boddy *et al.*, 1993. To verify our prediction, the *suc1* cDNA was amplified using RT-PCR and sequenced, which confirmed the presence of the intron (data not shown). We have renamed the *suc1* gene as *sucA*, to meet the general nomenclature rules for *A. niger*. Although the *A. niger* SucA protein is 67% identical to the FopA protein, and both proteins cluster in the same branch of the phylogenetic tree, the enzymatic activities of the two proteins appear to be different. FopA displays a much higher fructosyltransferase activity than Suc1 from *A. niger* strain B60 (Yanai *et al.*, 2001). SucA, which is identical to Suc1, therefore is likely to encode a fructofuranosidase lacking any detectable fructosyltransferase activity under the given assay conditions. A second fungal fructosyltransferase encoding gene was identified in *A. sydowi* (*sftA*) (Heyer & Wendenburg, 2001), which clusters with FopA and SucA (Fig. 1). The *sftA* gene was only expressed in the conidia of *A. sydowi*, and its product SftA is capable of producing fructo-oligosaccharides up to 40 fructose units long (*in vitro*). High molecular weight polymers were detected when intact conidia were incubated with sucrose as substrate. Unexpectedly, these two transferases (SftA and FopA) did not show a higher sequence identity to each other than to SucA. Our sequence alignment did not reveal any obvious differences between the proteins that could explain their different reaction specificities. SucA orthologs were not identified in the genome of *A. nidulans* and *A. fumigatus*.

A surprising finding was the presence of only one copy of an endo-inulinase gene in the *A. niger* genomes of strain CBS513.88 and N402. Previous studies with *A. niger* strain 12 revealed the presence of two very similar genes, *inuA* and *inuB*, both encoding endo-inulinase (Ohta *et al.*, 1998). Expression analysis revealed that only the *inuB* gene is transcribed actively (Akimoto *et al.*, 1999). Most likely, the presence of the two genes is a result of a recent duplication event that is specific for *A. niger* strain 12 and has not occurred in *A. niger* strains CBS513.88 and N402.

The *inuE* gene encodes a third known inulinolytic enzyme, an exo-inulinase, characterized from *A. niger* strain 12 (Moriyama *et al.*, 2003). The putative exo-inulinase from CBS513.88 is 99% identical to InuE from *A. niger* 12 and 100% identical to *A. foetidus* fructosyltransferase 1-Sst (Rehm *et al.*, 1998); the latter was shown to produce 1-kestose in the presence of high concentrations of sucrose. However, Moriyama *et al.* (2003) did not detect transfructosylation activity in the culture filtrate of a *Pichia pastoris* strain expressing *inuE*, although the enzyme was incubated at substrate concentrations where transfructosylation should occur (150 mM), as reported by Rehm *et al.* (1998). It appears unlikely that the three amino acid differences between 1-Sst and InuE affect reaction specificity since the differences involve similar amino acid residues: His199Gln, Gly476Ser and Thr499Ser. Moriyama *et al.* (2003) suggested that 1-Sst might be an exo-inulinase that possessed an additional fructosyltransferase activity in the presence of high concentrations of sucrose. Both proteins contain signal sequences at their N-terminal ends and are secreted as extracellular proteins. *InuE* was specifically induced on sucrose and inulin and repressed on fructose and glucose (Moriyama *et al.*, 2003). In view of the 100% sequence identity of the

A. niger CBS513.88 InuE and 1-Sst proteins, it is most likely that the *A. niger* InuE protein also has fructosyltransferase activity. However, further biochemical data about InuE are required. The InuE protein is also 91% identical to the *A. awamori* exo-inulinase Inu1 protein (Arand *et al.*, 2002). *A. awamori* Inu1 is the only enzyme in this branch of the phylogenetic tree with levanase (hydrolysis of β -2,6 glycosidic linkages) activity. No transfructosylation activity has been reported for the *A. awamori* Inu1 protein. The different enzymatic properties of the enzymes in this group are currently not well understood and need further biochemical investigation.

Two new invertase-like proteins (SucB and SucC) were identified in the genome of *A. niger*. As the two proteins contain all the conserved domains and the catalytic residues of GH32 family members it is very likely that these proteins contain sucrolytic or inulinolytic activities. Preliminary biochemical analysis of the recombinant SucB protein expressed in *E. coli* indeed indicates that SucB has hydrolyzing activity on sucrose, 1-kestose and nystose as well as transfructosylation activity resulting in the formation of 1-kestose and nystose from sucrose and 1-ketose, respectively (C. Goosen PhD thesis 2006, unpublished results).

Both SucB and SucC are predicted to be intracellularly localized. The current gene models do not indicate the presence of typical hydrophobic signal sequences for targeting the protein into the ER in order to secrete the proteins via the secretory pathway. The algorithm (SecretomeP 1.0b prediction) used to predict the probability that SucB and SucC are exported via a non-classical secretion pathway, was not conclusive as the prediction scores of the proteins were close to the threshold values. Moreover, these values should be interpreted with care, as the program has been trained using sequences of human non-classical exported proteins as no non-classical protein export has been shown to be present in filamentous fungi. Also in the genomes of other filamentous fungi GH32 family members without a predicted N-terminal signal sequence were identified (Table 2), indicating that the presence of these intracellular enzymes is wide-spread among filamentous fungi and not specific for *A. niger*.

The presence of carbohydrate degrading enzymes that do not contain a signal sequence are not limited to inulinolytic enzymes alone. In the genome of *A. nidulans*, 41 polysaccharide degrading enzymes were predicted that lacked a signal sequence (de Vries *et al.*, 2005b). The presence of these intracellular enzymes strongly suggests that fungi are capable of transporting oligosaccharides into the cell which subsequently are hydrolyzed by the intracellular enzymes. Alternatively, these intracellular enzymes may possess also transglycosylation activity, e.g. for the synthesis of inducer molecules that activate transcription factors.

The genome of *A. niger* also contained an open reading frame which showed homology to the group of exo-inulinases. However, to assemble this ORF of 137 amino acids, containing only a single conserved domain (domain C) characteristic for family GH32 members, three putative frame shifts had to be corrected. Missing domains were not found in sequences adjacent to the predicted ORF. The genomic DNA region containing this putative inulinase (*inuQ*) was amplified by PCR from CBS513.88 and N402 and re-sequenced which confirmed that the original DNA sequence was the correct one (data

not shown). A similar organization of the *inuQ* locus is also present in the *A. niger* genome sequence of strain ATCC 1015. From this we concluded that *inuQ* is probably a pseudogene and not producing a functional protein. Northern analysis of mRNA isolated from cells grown on a variety of different carbon sources did not result in detection of *inuQ* mRNA, indicating that the gene is not transcribed under the conditions tested.

Expression analysis of the five genes revealed that the genes encoding the extracellular enzymes (SucA, InuE and InuA) are co-regulated and specifically expressed on sucrose and inulin. It is rather surprising that the *inuA* gene is induced by sucrose. Physiologically, there is no reason for the fungus to secrete this enzyme during growth with sucrose, since the enzyme does not hydrolyze the fructose-glucose disaccharide. However, a beneficial mechanism might have evolved as sucrose and inulin might often be present together in plant material. Inulin is broken down primarily into fructose residues by the action of the exo-inulinase gene product and into inulo-oligosaccharides by the action of the endo-inulinase. The specific expression of *sucA* and *inuE* on inulin and sucrose is in agreement with previous observations (Wallis *et al.*, 1997; Moriyama *et al.*, 2003; Rehm *et al.*, 1998). The expression of the single endo-inulinase (*inuA*) gene in *A. niger* N402 on sucrose and inulin is different from that described for *inuA* (no expression observed) and *inuB* (constitutively expressed on inulin, fructose and glucose) in *A. niger* strain 12 (Akimoto *et al.*, 1999). Analysis of the promoter sequences of the *inuA* genes of the *A. niger* CBS513.88 and N402 strains revealed that they contain 5 putative CreA binding motifs (SYGGRG) in their first 1000 base pairs. *CreA* encodes a wide-domain regulatory protein that binds to the promoter of target genes to prevent or decrease expression if a favorable carbon source (such as glucose or fructose) is present (Dowzer & Kelly, 1991; Ruijter & Visser, 1997). No such binding sites were observed when analyzing the upstream sequence region of the *A. niger* 12 *inuB* gene. This difference in the promoter sequences of *inuA* (*A. niger* CBS513.88 and N402) and of *inuB* (*A. niger* 12) may be responsible for their different expression patterns in relation to different carbon sources. Thus, *A. niger* strains CBS513.88, N402 and *A. niger* 12 differ both in the number of genes encoding endo-inulinases and in the way their expression is regulated.

Two lines of evidence led us to propose that sucrose, and not fructose, acts as an inducer for the expression of the genes encoding the inulinolytic enzymes. First, we showed that a low concentration of sucrose (50 μ M initial concentration) induced the expression of the genes encoding the inulinolytic enzymes. The lower expression of the inulinolytic genes at higher sucrose concentrations suggested that sucrose (or the hydrolysis products of sucrose, glucose and fructose) caused carbon catabolite repression, possibly via the repressor protein CreA. The addition of 5 μ M sucrose resulted in relatively low levels of expression. Lowering the sucrose concentration even further did not result in detectable mRNA levels of any of the genes after 4 h of growth. At this point we can not rule out the possibility that also these lower concentrations of sucrose might have induced expression of the inulinolytic enzymes at earlier time points. Assuming that the mRNAs might not be very stable, we might have missed the induction as we have analyzed the expression only after 4 h. In an identical experimental set-up, also fructose was tested as an inducer,

previously reported to act as an inducer for inulinase expression in the yeast *K. lactis* (Grootwassink & Hewitt, 1983). We obtained no evidence that fructose acted as an inducer for the expression of the inulinolytic enzymes in *A. niger*. Formally, it is possible that the inulinolytic genes are induced by low influx levels of fructose which escaped detection after 4 h of growth. The addition of low concentrations fructose may have resulted in starvation and the inability of the fungus to induce expression by a lack of energy. One could postulate that fructose can act as an inducer at low concentrations but repress expression at higher concentrations. If this is the case, growth of the $\Delta creA$ strain on fructose should lead to a high level of expression of the genes encoding the inulinolytic enzymes. However, in Fig. 3B, we show that the inulinolytic genes are not expressed in the $\Delta creA$ strain, giving additional support that fructose does not act as an inducer for the expression of the inulinolytic enzymes. The results obtained from our transcriptional study fit well with the early observations by Vainstein & Peberdy (1991) that the invertase production in *A. nidulans* was the highest in sucrose medium and low in the culture fluid of fructose-grown mycelia. These findings indicate that the expression of inulinolytic enzymes is similarly regulated in *A. nidulans* and *A. niger*. As also glucose, 1-kestose or 1-nystose did not induce expression of the inulinolytic genes, we suggest a mechanism by which sucrose is transported across the plasma membrane. Once intracellular, the sucrose molecule, or a derivative of sucrose, acts as an inducer to activate a transcription factor to drive the expression of the inulinolytic genes. Current research is aimed at the identification of transcriptional activator(s) involved in the activation of expression of inulinolytic and or sucrolytic enzymes, and to determine the possible role of the intracellular *sucB* gene during growth on inulin and sucrose.

Acknowledgements

The authors thank Dr. J. Visser for helpful discussions and Angelique Franken for technical assistance. We thank DSM for providing access to the *A. niger* genome sequence, Sensus and Cosun for providing inulin, and Avebe for providing starch. This research was supported by Senter IOP (Innovatief Onderzoeks Programma), The Netherlands (IGE01021).

Chapter 6

Molecular and biochemical characterization of a novel intracellular invertase from *Aspergillus niger* with transfructosylating activity

Coenie Goosen*, Xiao-Lian Yuan*, Jolanda M van Munster, Arthur F J Ram,
Marc J E C van der Maarel, Lubbert Dijkhuizen

* These authors contributed equally to this work

Abstract

A novel sub-family of putative intracellular invertase enzymes (glycoside hydrolase family 32) has previously been identified in fungal genomes. Here we report phylogenetic, molecular and biochemical characteristics of SucB, one of two novel intracellular invertases identified in *Aspergillus niger*. The *sucB* gene was expressed in *Escherichia coli* and an invertase negative strain of *Saccharomyces cerevisiae*. Enzyme purified from *E. coli* lysate displayed a molecular weight of 75 kDa, judging from SDS-PAGE analysis. Its optimum pH and temperature for sucrose hydrolysis were determined to be 5.0 and 37-40°C, respectively. In addition to sucrose, the enzyme hydrolyzed 1-kestose, nystose and raffinose, but not inulin and levan. SucB produced 1-kestose and nystose from sucrose and 1-kestose, respectively. With nystose as substrate, products up to a degree of polymerization (DP) of 4 were observed. SucB displayed typical Michaelis-Menten kinetics with substrate inhibition on sucrose (apparent K_m , K_i , and V_{max} of 2.0 (± 0.2) mM, 268.1 (± 18.1) mM and 6.6 (± 0.2) $\mu\text{mol min}^{-1} \text{mg}^{-1}$ of protein (total activity) respectively). At sucrose concentrations up to 400 mM, FTF activity contributed approximately 20-30% to total activity. At higher sucrose concentrations, FTF activity increased up to 50% of total activity. Disruption of *sucB* in *A. niger* resulted in earlier onset of sporulation on solid media containing various carbon sources, whereas no alteration of growth in liquid culture media was observed. SucB thus does not play an essential role in inulin or sucrose catabolism in *A. niger*, but may be needed for intracellular conversion of sucrose to fructose, glucose, and small oligosaccharides.

Introduction

Fructans and fructooligosaccharides (FOS) consist of a chain of fructose molecules linked to a terminal glucose residue. These fructose monomers are linked by either β -2,1- (inulin) or β -2,6- (levan) glycosidic bonds. Inulin and levan have several favorable properties which make them commercially interesting for applications in pharmaceutical- and food industries (Vijn & Smeekens, 1999). In the human digestive track, FOS are almost exclusively fermented by bifidobacteria and lactobacilli, which have beneficial health effects (May *et al.*, 1994; Tannock, 1997; Sghir *et al.*, 1998). Commercially, FOS are produced by the enzymatic hydrolysis of inulin isolated from plants, primarily chicory and Jerusalem artichoke (Vijn & Smeekens, 1999). Alternatively, sucrose can be converted into FOS, using a range of different transfructosylating enzymes, originating from plants, bacteria and fungi (Vijn & Smeekens, 1999). FTF and hydrolytic enzymes belong to glycoside hydrolase families (GH) 32 and 68 (Coutinho & Henrissat, 1999) constituting enzyme clan GH-J, based on shared conserved domains (Nagem *et al.*, 2004). These enzymes have been reported to be present in a variety of plants, bacteria and fungi (Vijn & Smeekens, 1999). FOS synthesis has been reported for the commercially important fungus *Aspergillus niger*, reflecting a side reaction of an invertase (EC. 3.2.1.26; (Somieri *et al.*, 1997; this chapter)) or as the result of the activity of a specific fructosyltransferase (EC. 2.4.1.9; (L'Hocine *et al.*, 2000)). Nguyen

et al. (1999) reported the presence of an intracellular invertase in *A. niger* IMI303386, grown on sucrose or inulin as sole carbon source. The purified enzyme produced free glucose and fructose from sucrose hydrolysis, as well as 1-kestose and nystose from sucrose FTF. However, the gene encoding this enzyme has not been identified and characterized. Yanai *et al.* (2001) reported characteristics of an extracellular β -fructofuranosidase from *A. niger* 20611. This enzyme displayed increased FTF activity compared to other known *Aspergillus* invertases. The strain was, however, later reclassified as *Aspergillus japonicus* ATCC 20611. The true identity, diversity, and characteristics of invertases and FTF enzymes present in *A. niger* thus remained to be determined.

Recently, the complete genome sequence of *A. niger* has become available (Pel *et al.*, 2007) and was analyzed for putative sucrose and fructan-modifying enzymes (Yuan *et al.*, 2006). In addition to *sucA*, encoding the previously characterized extracellular invertase (Boddy *et al.*, 1993), two novel putative invertase genes were identified (*sucB* and *sucC*) (Fig. 1). The *sucB* (but not the *sucC*) gene was (constitutively) expressed at a low level on starch and xylose, and up-regulated in the presence of sucrose and inulin (Yuan *et al.*, 2006). Orthologues of these genes have also been identified in other fungal genomes (see below). We have cloned and heterologously expressed the *A. niger sucB* gene, allowing a biochemical characterization of the purified enzyme. An *A. niger sucB* gene disruption strain was constructed to determine whether this novel intracellular invertase enzyme plays a significant role in growth on sucrose and inulin. This paper reports on the phylogenetic, molecular and biochemical characterization of SucB. The data show that in addition to invertase activity, SucB displays transfructosylating activity.

Materials and methods

Phylogenetic analysis

The complete amino acid sequence of SucB (DQ233219) (39) was blasted against the protein database at Swissprot (<http://www.ncbi.nlm.nih.gov/BLAST/>). Identified sequences containing family GH32 domains (<http://afmb.cnrs-mrs.fr/CAZY/>) were aligned with SucB (CLUSTALW interface in MEGA 3.1, <http://www.megasoftware.com>) followed by Bootstrap test of phylogeny (gap-opening, 10; extension penalties, 0.2; 1000 replicates). Sequence logos were created by SequenceLogo (<http://bio.cam.ac.uk/seqlogo/>).

Strains, plasmids, media and growth conditions

A. niger strains N402 (Bos *et al.*, 1988), NRRL3122 (Pel *et al.*, 2007) and AB4.1 (van Hartingsveldt *et al.*, 1987) and *Escherichia coli* strains XL1-Blue (Stratagene, La Jolla, Ca), TOP 10 and BL21 (DE3) STAR (Invitrogen, Carlsbad, Ca) were used in this study. The *A. niger* genome sequence derived from strain CBS513.88 (a natural derivative of strain NRRL3122). *A. niger* strains were grown in *Aspergillus* Minimal Medium (MM, (Bennet *et al.*, 1991)) or Complete medium (CM; MM supplemented with 0.5% (w/v) yeast extract and

0.1% (w/v) casamino acids). Conidiospores were obtained by harvesting spores from a CM plate containing 1% (w/v) glucose, after 4-6 days of growth at 30°C, using a 0.9% (w/v) NaCl solution. Transformation of *A. niger* AB4.1 was as described previously (Punt & van den Hondel, 1992). Cloning of *sucB* was performed using Gateway cloning technology (Invitrogen) and the integrity of constructs were verified by DNA sequencing (Baseclear, Leiden, the Netherlands). The Gateway expression vectors pDEST17 and pYES-DEST52 were used for expression in *E. coli* and the invertase negative strain of *S. cerevisiae* (BY4743Δ*suc2*).

S. cerevisiae strains were grown aerobically at 30°C in SC plus glucose media (1.7 g.l⁻¹ yeast nitrogen base, 5 g.l⁻¹ ammonium sulphate, 2.5 g.l⁻¹ sodium succinate, 5 g.l⁻¹ Casamino acids, 0.1 g.l⁻¹ tryptophan, 20 g.l⁻¹ glucose), followed by induction of expression in SC media plus 20 g.l⁻¹ galactose.

Cloning and purification of SucB

The coding sequence of *sucB* was amplified in a two step procedure. The first two exons were amplified using primers sets SucBGATEF, SucBDNAP1 and SucBDNAP2, SucBGATER, respectively, followed by joining of the two exons in a single PCR reaction together with outside primers SucBGATEF and SucBGATER (Table 1). Amplifications were performed in a GeneAmp PCR system 2700 thermal cycler (Applied Biosystems, Foster City, Ca) using the Expand High Fidelity PCR System (Roche Diagnostics Corporation, Indianapolis, IN.) under the following conditions: initial denaturation of 2 min at 94°C, 30 cycles of 15 s denaturation at 94°C, annealing at 55°C for 30 s and elongation at 72°C for 90 s, followed by a final elongation step of 7 min at 72°C. Gateway cloning of the fragments was performed according to the manufacturer's instructions (Invitrogen.) to create constructs pDEST17-*sucB* and pYES-DEST52-*sucB*, respectively.

Table 1. Primers used during this study

Name	Sequence
SucBGATEF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATAATGGAGCGGCAAACTAGCCCTCAG
SucBGATER	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCACTCGGCATCGACTTCTTCC
SucBcDNAP1	TGTTGGTGGATGGTGCCATGGAGGACTCTAT
SucBcDNAP2	CATGGCACCATCCACCAACACAACACAGCCCT
SucBP1f	TAGCGGGCCGCATCGCGACTGTCTCATACA
SucBP2r	CAGCGGCTTGGGTCTAGAATCATCCGAATATTCTCGATT
SucBP3f	ATCCTCTAGATCGACCTGCAGTCGATCGCGAGTGAGAT
SucBP4r	TGGAATTCTCATTTACGTCATCGTCGGCGA

Cultures of *E. coli* BL21 (DE3) STAR containing pDEST17-*sucB* were inoculated into fresh LB media containing 100 µg ml⁻¹ ampicillin. Soluble expression of SucB was achieved at 18°C (optical density at 600nm of approximately 0.4), followed by induction at 18°C for 6 h (1 mM isopropyl-β-D-thiogalactopyranoside [IPTG]). Cells were harvested by centrifugation (10 min, 4°C, 4000 x g) and cell pellets were resuspended in 5 ml 50 mM sodium phosphate buffer (pH 8) containing 250 mM NaCl, 10 mM imidazole and 5mM beta-mercaptoethanol. Following sonification (seven times for 15 sec at 8 µm with 30 s intervals), cell-free lysate was obtained by centrifugation (20 min at 4°C, 10,000 x g). SucB

was purified from the cell-free lysates using His-tag affinity chromatography according to the manufacturer's instructions (Sigma-Aldrich, St. Louis, Mo.). Protein concentration, size, and purity were determined using the Bradford reagent (Bio-Rad, Hercules, Ca.), sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and Biosafe Coomassie staining (Bio-Rad).

Cultures of *S. cerevisiae* BY4743 Δ suc2 cells containing the expression vector pYES-DEST52-sucB were used to inoculate 50 ml of fresh medium containing 2% galactose to induce protein expression. After 5 h of growth, cells were pelleted by centrifugation and washed with sterile demineralized water. Y-PER (Pierce Biotechnology, Rockford, IL) was used for cell lysis (according to the manufacturer's recommendations), followed by the preparation of cell-free lysate as described above.

Activity assays

Enzymatic activity was quantified spectrophotometrically by separate measurements of the released glucose and fructose from sucrose (D-glucose/D-fructose kit, Roche Diagnostics Corporation). The pH and temperature optima were determined by measuring initial rates in 50 mM phosphate-citrate buffer containing 100 mM sucrose using a pH range of 4.0 to 7.5 and a temperature range of 25 to 60°C. The effect of sucrose concentration on enzyme activity was determined using two independent experiments performed in triplicate by measuring initial rates over a substrate range of 12 sucrose concentrations (2.5 mM to 1 M) in 50 mM acetate buffer at 37°C, pH 5.0. Corrections for background glucose and fructose values at high substrate concentrations were made accordingly. Non-linear regression curve fitting was done using Sigma Plot (Systat Software, Richmond, Ca) applying the Michaelis-Menten formula for substrate inhibition ($y = V_{max} * [S] / (K_m + [S] + ([S]^2 / K_i))$, where y equals the specific activity ($\mu\text{mol} \cdot \text{mg}^{-1} \cdot \text{min}^{-1}$).

Substrate specificity and product range

In order to identify substrate specificity of SucB, and the products made after incubation, 14 μg of purified SucB was incubated overnight in 50 mM acetate buffer (pH 5.0) at 37°C with a range of substrates. The effect of reducing agents dithiothreitol (DTT) and β -mercaptoethanol (BME) on SucB activity was also tested by incubating the enzyme at optimal conditions overnight in the presence of these agents at concentrations up to 100 mM. Substrate conversion and product formation were analyzed by thin layer- (TLC) (aluminum sheets silica gel 60 F 254, Merck and Co., Whitehouse station, NJ.) and high performance anion exchange chromatography (HPAEC) Dionex Corporation, Sunnyvale, Ca.). Sample separation on TLC was performed using a mixture of butanol, ethanol and water (3.8:3.8:2.4 [vol/vol/vol]) or ethyl acetate, 2-propanol and water (6:3:1 [vol/vol/vol]). Subsequently, plates were dried and sprayed with developer solution [95% methanol, 5% sulfuric acid and 3 g l⁻¹ 2-(1-naphthylamino) ethylamine dihydrochloride]. Product formation by SucB was confirmed by HPAEC as described previously (Ozimek *et al.*, 2006).

Construction of the *sucB::pyrG* gene deletion strain

A *sucB* deletion cassette was constructed by PCR amplification of 1.0 kb of 5' and 3' DNA flanking regions of the *sucB* gene using primers SucBP1-SucBP4 (Table 1). Both fragments were cloned into pBlue-ScriptII (Stratagene) using appropriate restriction enzymes (Table 1). The *A. oryzae pyrG* gene from pAO4-13 (de Ruiter-Jacobs *et al.*, 1989) was isolated as a 2.7 kb *Xba*I fragment and cloned between the 5' and 3' *sucB* flanking regions to give $p\Delta sucB$.

Prior to transformation into AB4.1, $p\Delta sucB$ was linearized with *Eco*RI. Uridine prototrophic transformants were purified and screened for *sucB* deletion by Southern blot analysis (Sambrook *et al.*, 1989). Genomic DNA was isolated and digested with *Xho*I and the 3' region flanking the *sucB* gene was used as a probe. As predicted, a 2.2-kb hybridizing DNA fragment was observed in the wild type strain, whereas a 4.0-kb DNA fragment was detected in *sucB* deletion strains (data not shown). Several *sucB* deletions strains were independently obtained, and strain NC1.1 ($\Delta sucB$) was used throughout this study.

Microtiter plate growth assay

Growth of *A. niger* strains N402 and NC1.1 was determined using a HTS7000 BioAssay Reader (Perkin Elmer Life and Analytical Sciences, Inc., Wellesley, MA.). Spores (1×10^4) were inoculated in each well of a 96-wells microtiter plate (Nalge Nunc International, Rochester, NY.) and incubated at 32°C for 56 h. Each well contained 200 μ l of MM with 1% (wt/vol) of one of the various carbon sources, supplemented with 0.1% (wt/vol) casamino acids to stimulate spore germination. Six replicates of each condition were made. Growth was monitored by measuring the optical density at 595 nm (OD_{595}) every 2 h.

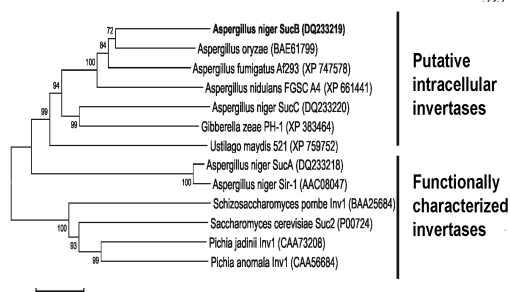


Fig. 1. Neighbor-joining tree of functionally characterized and putative fungal invertases. Bootstrap values are indicated on the node of each branch. The tree was created with MEGA 3.1 using default settings for gap and extension penalties. The bar indicates 10% amino acid sequence difference.

Results

Sequence analysis

Using the predicted amino acid sequence of SucB in phylogenetic analysis, we have identified SucB orthologues in various other fungal species. *In silico* analysis indicated that

these putative invertases also lack any recognizable signal peptide sequences, as has been reported for SucB (Yuan *et al.*, 2006). Multiple sequence alignment between these new group of putative intracellular invertases and known fungal invertases indicate that they cluster together, in a separate subfamily, clearly distinct from known extracellular fungal and yeast invertase proteins (Fig. 1). Table 2 depicts sequence logos (Crooks *et al.*, 2004) constructed from alignments of the SucB subfamily members, revealing the presence of all eight conserved domains characteristic for family GH32.

Table 2. Sequence logo depiction of conserved motifs identified in invertases with known function and SucB orthologues *

Domain	Functionally characterized invertases	Putative intracellular invertases
A	WMINDPNG	W _Y NDPCG _A
B	HVGFQ	YHL _Y Q
B1	WGHAITS	WGHAITS
C	ESGS _Y	FTGCE _E
D	FRDP _K V _F	WRDP _Y _F _S
E	QYECPCG _N	NWE _Y _N
F	FDG _{KD} YA	FDHG _C YA
G	LE _Y	VLE _Y E

* See Yuan *et al.*, 2006. Conserved motifs of glycoside hydrolase family 32 are indicated on the left. Sequences used to construct the logos are the same as depicted in Fig. 1.

Cloning and purification of SucB

Initial attempts to clone the *sucB* from a cDNA library constructed from an inulin growing *A niger* strain N402 were unsuccessful, probably because of the relatively low level of expression of the *sucB* gene (Yuan *et al.*, 2006). Following amplification, the full coding region of *sucB* (1,854 bp) was obtained. The same procedure was followed to clone the

second putative intracellular invertase identified in *A. niger* (*sucC*); however, we failed to obtain functional expression in either *E. coli* or *S. cerevisiae*.

Purification of SucB from *E. coli* was confirmed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis, where the protein size was estimated at 75 kDa (70 kDa, calculated). C-terminally His-tagged SucB expressed in *S. cerevisiae* strain BY4743 Δ *suc2* could not be sufficiently purified using Ni-nitrilotriacetic acid affinity chromatography. Therefore, the cell-free lysate was used for comparative studies. SucB expressed in both *E. coli* and *S. cerevisiae* displayed similar characteristics, whereas no activity could be detected in *S. cerevisiae* BY4743 Δ *suc2* containing the empty expression vector. SucB in the cell-free extract of *E. coli* or *S. cerevisiae*, as well as the affinity purified SucB from *E. coli*, only displayed activity for a maximum storage time of three days (4°C or -20°C in 20% glycerol). Thus, for all subsequent analysis, the enzyme was used directly after purification.

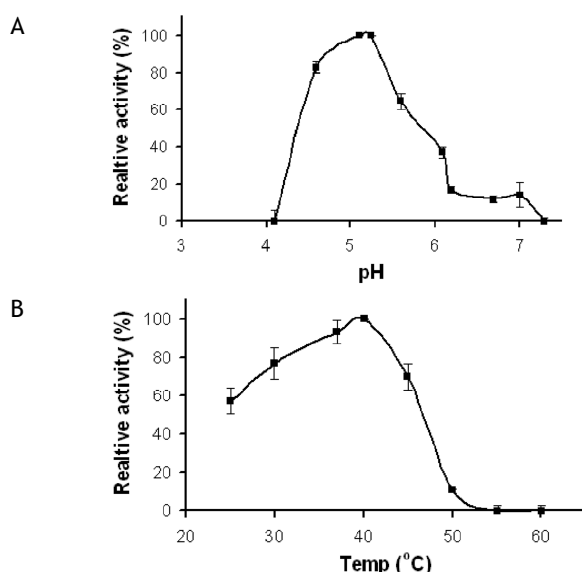


Fig. 2. Effect of pH (A) and temperature (B) on SucB activity. The enzymatic activity was determined by measuring the amount of glucose released from the initial reaction of SucB (14 μ g) incubated with 200 μ l of 100 mM sucrose in citrate-acetate buffer at 37°C. Values depicted are the mean of duplicates ($\pm$ SEM), based on at least two independent experiments.

Influence of pH and temperature on SucB enzyme activity

The optimal pH and temperature conditions for SucB activity with sucrose as substrate were determined by measuring the amount of released glucose enzymatically (total activity). SucB activity could be detected from pH 4.5 to 7, albeit at very low activity levels at pH values above 6.3. SucB displayed maximal activity at pH 5.0 (Fig. 2A). The optimal temperature for SucB total activity is in the range 37 to 40°C (Fig. 2B). At temperatures of

50°C or higher, no activity could be detected, whereas at lower temperatures, the SucB specific activity remained relatively high, with 50% activity remaining at 25°C.

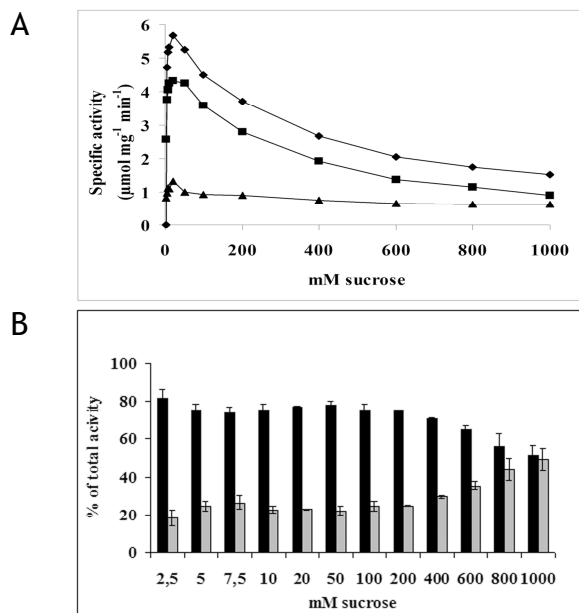


Fig. 3. Effect of sucrose concentration on SucB activity. (A) Total activity was determined by measuring the amount of glucose released from the initial reaction of SucB incubated with 12 sucrose concentrations ranging from 2.5 mM to 1 M in 50 mM acetate buffer pH 5.0 at 37°C. Values depicted were calculated from triplicate measurements. (♦) Total activity; (■) Invertase activity; (▲) FTF activity. (B) Percentage of either hydrolytic (black bars) or FTF (grey bars) activity compared to the total activity of SucB, displayed for a range of sucrose concentrations. Measurements are the mean of duplicates ($\pm$ SEM), based on two independent initial measurements. Approximately 14 μ g of purified protein was used in each case.

Kinetic analysis of SucB activity

Incubation of SucB with increasing sucrose concentrations lead to a typical Michaelis Menten-type kinetics with substrate inhibition (apparent K_m , K_i and V_{max} of 2.0 ± 0.2 mM, 268.1 ± 18.1 mM and 6.6 ± 0.2 μ mol $\text{mg}^{-1} \text{min}^{-1}$ respectively). Hydrolysis and FTF reactions displayed similar patterns, with apparent K_m values of 2.5 ± 0.2 and 0.9 ± 0.5 mM, apparent V_{max} values of $5.5 (\pm 0.1)$ and $1.2 (\pm 0.1)$ μ mol $\text{mg}^{-1} \text{min}^{-1}$, and apparent K_i values of $206.3 (\pm 12.8)$ and $797.6 (\pm 196.5)$ mM, respectively. Increasing the sucrose concentration from 2.5 mM to 1 M resulted in a decreased hydrolysis activity and in an increased (20% to 50%) FTF activity in relation to total SucB activity (Fig. 3). These observations were confirmed by the TLC product analysis, showing increased 1-kestose and decreased free fructose synthesis at higher sucrose concentrations (Fig. 4). HPAEC analysis showed that apart from the formation of 1-kestose as the major FTF product, minor amounts of nystose were also produced (result not shown).

SucB substrate specificity and product formation

Substrate specificity analysis showed that SucB is able to hydrolyze sucrose, raffinose and the inulin-type oligosaccharides 1-kestose and nystose (releasing fructose in each case), as well as to perform oligomerization reactions (Fig. 5 A). Incubation of SucB with 100 mM of 1-kestose in the presence or absence of 100 mM of sucrose produced free fructose, sucrose and nystose, indicating that 1-kestose could be used as both donor and acceptor substrates.

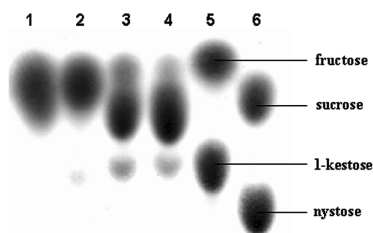


Fig. 4. TLC analysis of the reaction products after overnight incubation of SucB at 37°C with 50 mM (lane 1), 100 mM (lane 2), 500 mM (lane 3) and 1 M sucrose (lane 4). Lane 5 shows the standards fructose and kestose and lane 6 sucrose and nystose. Approximately 14 µg of purified protein was used in all incubations. All samples were equilibrated to the same concentration before spotting on the plate.

Alternatively, the sucrose formed subsequently was used as donor substrate for transfer of free fructose to 1-kestose. Incubations of SucB with nystose alone or nystose plus sucrose, yielded only free fructose, sucrose and 1-kestose (Fig. 5 A). HPAEC analysis confirmed that nystose was the largest product produced from incubation of SucB with kestose, and that a minor amount of pentakestose (degree of polymerization of 4) was produced after overnight incubation. SucB incubation in the presence of sucrose plus 1-kestose or nystose did not facilitate product diversification, but resulted in an increase in concentration of the observed products only. Furthermore, SucB incubation with galactose and sucrose did not yield any other products than observed for sucrose alone (Fig. 5 A).

SucB hydrolysed the sucrose moieties of the sugars raffinose [α -D-galactose-(1,6)- α -D-glucose-(1,2)- β -D-fructose] (Fig. 5 A) and stachyose [α -D-galactose-(1,6)- α -D-galactose-(1,6)- α -D-glucose-(1,2)- β -D-fructose] (Fig 5 B). No hydrolysis of any of the α -glycosidic linkage sugars {trehalose [α -D-glucose-(1,1)- α -D-glucose], turanose [α -D-glucose-(1,3)- β -D-fructose], palatinose [α -D-glucose-(1,6)- β -D-fructose] or melizitose [α -D-glucose-(1,2)- β -D-fructose-(1,3)- α -D-glucose]} was observed. Using trehalose alone or in combination with sucrose (ratios 5:1 to 1:5) as substrate, only 1-kestose formation from sucrose could be observed as FTF product (results not shown). None of these α -glycosidic bond substrates were used as donor/acceptor substrates in FTF reactions with sucrose (ratios 5:1 to 1:5). Hydrolysis of larger polysaccharides such as inulin or levan could not be detected, not even after overnight incubation (Fig. 5 A). Similar results were obtained when cell free extracts from the recombinant *S. cerevisiae* strain carrying the construct pYES-DEST52-*sucB* were used.

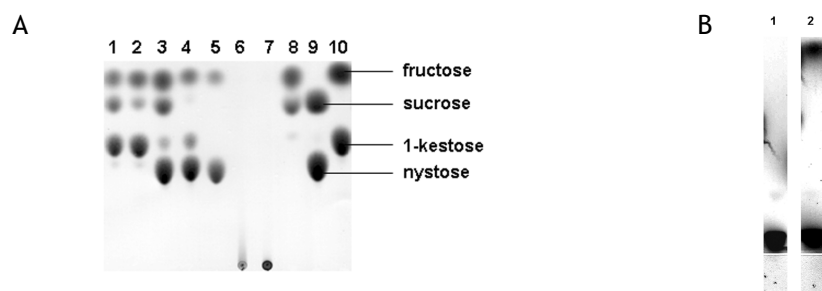


Fig. 5. TLC analysis of substrate specificity of SucB and reaction products after overnight incubation at pH 5.0 and 37°C. (A) 100 mM kestose with sucrose (lane 1), kestose (lane 2), nystose with sucrose (lane 3), nystose (lane 4), raffinose (lane 5), 1% inulin (chicory, see text, lane 6), levan (*Bacillus subtilis* produced, gift from Cosun Food Technology, the Netherlands, lane 7) and 100 mM galactose with sucrose (lane 8). Lanes 9 and 10 contain the standards fructose, sucrose, 1-kestose and nystose. (B) 100 mM stachyose incubated without (lane 1) and with (lane 2) SucB. Released fructose is indicated. Approximately 14 µg of purified protein was used in each case.

In silico translation of the SucB reading frame revealed the presence of 13 cystein residues, which can potentially form disulfide bridges. Addition of up to 100 mM of the reducing agents DTT or BME did not influence the sucrose hydrolysis or FTF activities of SucB indicating the absence of any structurally important disulfide bridges.

Disruption of the *sucB* gene

Disruption of *sucB* did not result in a significant change in growth rates and yields of *A. niger* in liquid media containing sucrose, inulin, glucose, fructose, xylose, maltose or starch (data not shown). Interestingly, growth of the *sucB* deletion mutant strains (e.g. strain NC1.1) on solid media containing these substrates resulted in an earlier onset (approx. 1 day) of sporulation compared to that of the wild type *A. niger*. The inclusion of additional uridine in the culture media to exclude suboptimal complementation by PyrG did not rescue the NC1.1 strain from this sporulation effect. No difference in colony diameter was observed between the NC1.1 and the wild type strains on the various carbon sources, indicating that the growth of the Δ *sucB* strain was not affected.

Discussion

The recent availability of the complete genome sequence of *A. niger* (Pel *et al.*, 2007) enabled identification of two novel putative intracellular invertases (*sucB* and *sucC*) (Yuan *et al.*, 2006). Although these genes share conserved amino acid residues with other family GH32 members, phylogenetically they cluster together with other putative intracellular invertases from fungal origin in a new distinct group (Fig. 1). These new putative invertases also contain all the conserved catalytic residues, as depicted in the sequence logos (Table 2). The only exception is Sir-1, which is missing the catalytic aspartate in domain A. This invertase was isolated from *A. niger* strain IBT10sb and displayed increased FTF properties (Somari *et al.*, 1997).

We have previously shown that SucB expression was upregulated by sucrose and inulin, whereas the enzyme was constitutively expressed at a low level with all other substrates used (Yuan *et al.*, 2006). The *sucB* gene also appears to be under catabolite repression control, evident from expression profiling in a *creA* deletion strain (Yuan *et al.*, 2006). With regards to *sucC*, no expression could be detected in mycelial mRNA under the same conditions. Using chromosomal DNA to construct the predicted open reading frame of SucC also failed to produce functional protein in both *E. coli* and *S. cerevisiae*.

In silico analysis of the SucB sequence revealed the absence of any recognizable signal peptide sequences for protein secretion, indicating that it may play a role in *A. niger* intracellular metabolism. In view of the low levels of expression observed for SucB in *A. niger*, and to avoid simultaneous separation with other invertases/fructosyltransferases present in *A. niger* (L'Hocine *et al.*, 2000), we overproduced the SucB enzyme in *E. coli* as well as in an invertase-negative strain of *S. cerevisiae* and subsequently determined its biochemical properties.

SucB clearly acts as an invertase, able to hydrolyze the glucose-fructose glycosidic linkage in the smaller fructose-containing oligosaccharides sucrose, kestose and nystose. Oligosaccharides larger than nystose, including polymeric inulin and levan could not be hydrolyzed. The enzyme was unable to hydrolyze α -glycosidic bonds in substrates or to use these compounds as donor/acceptor substrates in FTF reactions. Weak hydrolysis of the sucrose moiety of stachyose, but not melizitose, indicates that the fructose of the sucrose moiety should thus be positioned terminally to enable correct orientation and binding in the active site. A similar observation was made for the β -fructosidase from *Thermatoga maritima*, where it was shown that hydrolysis occur in a typical exo-fashion (Liebl *et al.*, 1998). Taking into account the diversity of substrates hydrolyzed, one could assume that the -1 subsite in the active site cleft (for numbering see Davies *et al.*, 1997) can accommodate fructose (Alberto *et al.*, 2004; Nagem *et al.*, 2004; Alberto *et al.*, 2006), and that the enzyme most probably does not contain multiple binding sites for fructose (as in the case for the exo-inulinase of *A. awamori*) (Kulminskaya *et al.*, 2003). In the three-dimensional structure of the *Cichorium intybus* fructan 1-exohydrolase, Verhaest *et al.* (2005) observed the presence of multiple glycerol molecules bound in the cavity between the N- and C-terminal domains of the protein. This forms an open cleft which is emerging from the active site, and is believed to be responsible for the binding of inulin or higher molecular weight fructans. Obstruction of this cleft could possibly influence the binding of high molecular weight inulins, limiting the enzyme to hydrolysis of small oligosaccharides only (Verhaest *et al.*, 2005; Alberto *et al.*, 2006). In view of the low amino acid similarity between SucB and other characterized invertases and the absence of structural data, we can only speculate that the same obstructing feature is present in SucB. This feature could explain the inability of the enzyme to bind and hydrolyze larger oligo- and polymeric fructans.

Apart from the hydrolytic activity observed for SucB, the enzyme was also able to perform FTF reactions. This activity was already detected at sucrose concentrations as low as 2.5 mM, with 1-kestose as the major oligomerization product (20-30% of total activity). At 1M, SucB displayed approximately 50% FTF activity, largely due to a decrease of hydrolytic

activity. Nystose was also produced as a minor product in the FTF reaction when 1-kestose was used a substrate. Also, the presence of a minor amount of pentakestose was observed after overnight incubation with nystose. Larger SucB products have never been observed. The data thus indicate that SucB is responsible for the intracellular production of small inulin-type oligosaccharides. In 1995, Muramatsu and Nakakuki previously described the purification and characterization of an intracellular beta-fructofuranosidase from *Aspergillus sydowi* that could transfer fructose from sucrose to trehalose, thus creating novel oligofructosyl trehaloses. However, when SucB was incubated with trehalose, no novel oligosaccharides were observed.

The SucB characteristics differ from the previously published data on the extracellular *A. niger* invertase Suc1/SucA/INV enzyme (Boddy *et al.*, 1993; Wallis *et al.*, 1997; L'Hocine *et al.*, 2000), and other invertases, in a number of aspects. SucB displayed an apparent K_m of 2.0 ± 0.2 mM, for sucrose, which is substantially lower than reported previously (30 and 160 mM for Suc1 and 35.67 mM for INV), but comparable to that of the extracellular acid invertase of *Fusarium solani* (3.57 mM) (Bhatti *et al.*, 2006). Extracellular invertases from both fungal as well as bacterial origin generally display lower affinity for sucrose than was observed for SucB (Gascon *et al.*, 1968; Reddy & Maley, 1996; Wallis *et al.*, 1997; Liebl *et al.*, 1998). However, Rubio and Maldonado (1995) described the purification and characterization of an invertase from an *A. niger* strain isolated from lemons. The invertase was purified from the mycelial lysate, and displayed a substrate affinity of 0.0625 mM for sucrose and a temperature optimum of 60°C. These characteristics clearly deviate from what has been observed for SucB, indicating that this protein might either be another intracellular invertase, or an isoform of SucB produced intracellularly in *A. niger* and not during recombinant expression in *E. coli* or yeast.

Using sucrose as substrate, SucB displays an apparent V_{max} (total activity) of 6.6 ± 0.2 $\mu\text{mol mg}^{-1} \text{min}^{-1}$ (this study). This figure is more than a thousand fold lower than the V_{max} for the extracellular invertase in *A. niger* AS0023 (7,758.3 $\mu\text{mol mg}^{-1} \text{min}^{-1}$) (L'Hocine *et al.*, 2000). The purified extracellular *A. niger* invertase also did not display any detectable FTF activity, not even at a sucrose concentration as high as 2.2 M (L'Hocine *et al.*, 2000), suggesting that the FTF described previously in literature could have derived from contaminating fructosyltransferases (Hirayama *et al.*, 2006).

Compared to the extracellular invertase of *A. niger* (Suc1), SucB displayed a lower optimum temperature value (37 to 40°C versus 50°C, respectively), whereas a comparable pH optimum was determined (Boddy *et al.*, 1993; Wallis *et al.*, 1997). However, SucB was only active in a narrow pH range (above pH 4, below pH 7 (Fig. 2A) compared to that observed for Suc1 (above pH 3, below pH10) (Boddy *et al.*, 1993; Wallis *et al.*, 1997). In the extracellular environment, Suc1 should be able to function in fluctuating pH conditions, where the extracellular pH could vary between 1.5 and 7.0 (Hesse *et al.*, 2002). In the intracellular environment, however, the cytoplasmic- and vacuolar pHs of *A. niger* is kept constant at 7.6 and 6.2, respectively. This balance is maintained in order to control pH sensitive processes such as DNA transcription and protein synthesis (Hesse *et al.*, 2002). Taking these facts into consideration, and in the absence of any detectable sequence for

protein export, we conclude that SucB functions sub-optimally in the intracellular environment.

Although intracellular invertases have been identified in fungi before (Gascon *et al.*, 1968; Muramatsu & Nakakuki, 1995; Nguyen *et al.*, 1999), little is known about the role they play in the intracellular environment. To determine whether SucB plays a crucial role in the metabolism of *A. niger*, a SucB disruption mutant strain was constructed. No difference in growth rate, yield, and morphology was observed between the *sucB* disruptant and the wild type *A. niger* N402 using liquid minimal media with sucrose or inulin as carbon sources. When the NC1.1 *sucB* disruptant strain was grown on solid minimal media containing various substrates, an earlier onset of sporulation was observed compared to the wild type. However, as observed in liquid media, no difference in growth was observed since the colony diameter was equal to that of wild type *A. niger*. Supplementing the culture media with uridine to minimize suboptimal complementation by the inserted *pyrG* gene did not alleviate the observed effect, suggesting that SucB (in)directly plays a role in the sporulation of *A. niger*.

Intracellular proteins usually do not contain disulfide bridges, which play a crucial role in structure and function of extracellular proteins (Raina & Missiakas, 1997). These bridges could be disrupted by the addition of reducing agents, e.g. DTT and BME, which in turn may cause loss of activity, or decreased enzyme stability. The inability of high concentrations of either DTT or BME to disrupt SucB activity gives a further indication that no disulfide bridges crucial to activity or structural integrity are present. This further supports the view that SucB is functioning in the intracellular environment in *A. niger*.

Considering the high affinity for sucrose, the narrow functional pH range, and the absence of an export signal sequence and functionally important disulfide bridges, we speculate that SucB plays an intracellular role in salvaging low concentrations of sucrose, kestose or nystose into fructose and glucose as energy sources. SucB may also function in transfer of fructose units from sucrose to fructan or unknown acceptor molecules. These molecules may be responsible for energy storage, the regulation of osmolarity or play a role in the induction of other proteins involved in the modification of fructans.

Further analysis and complementation studies should be conducted to determine the effect of *sucB* gene disruption on the expression of other fructan modifying enzymes, and the possible role it could play in initiating the earlier onset of sporulation.

Acknowledgements

We thank N. Carvalho for her assistance in generating the *A. niger sucB* gene deletion strain and Dr. P. van Heusden for providing us with the *S. cerevisiae suc1* deletion strain. COSUN Food Technology Centre (CFTC, Roosendaal, the Netherlands) is thankfully acknowledged for providing substrates, HPAEC analysis, and stimulating discussions. The national IOP program (The Netherlands) is acknowledged for funding this project (project code IGE 1021). This project is part of the CarbNet program (Carbohydrate modifying enzyme network of *Aspergillus niger*).

Chapter 7

Identification of InuR, a new Zn(II)₂Cys₆ transcriptional activator involved in the regulation of inulinolytic genes in *Aspergillus niger*

Xiao-Lian Yuan, Johannes A. Roubos, Cees A. M. J. J. van den Hondel, Arthur F. J. Ram

Abstract

The expression of inulinolytic genes in *Aspergillus niger* is co-regulated and induced by inulin and sucrose. In this paper, we have identified a positive acting transcription factor, InuR, which is required for the induced expression of inulinolytic genes. InuR is a member of the fungal specific class of transcription factors of the Zn(II)2Cys6 type. Involvement of InuR in inulin and sucrose metabolism was suspected because of the clustering of *inuR* gene with *sucB*, which encodes an intracellular invertase with transfructosylation activity and a putative sugar transporter encoding gene (An15g00310). Deletion of the *inuR* gene resulted in a strain displaying a severe reduction in growth on inulin and sucrose medium. Northern analysis revealed that expression of inulinolytic and sucrolytic genes e.g. *inuE*, *inuA*, *sucA*, as well as the putative sugar transporter gene (An15g00310) is dependent on InuR. Genome wide expression analysis revealed, three additional putative sugar transporters encoding genes (An15g04060, An15g03940 and An17g01710) which were strongly induced by sucrose in an InuR dependent way. *In silico* analysis of the strongly InuR regulated genes revealed a putative binding site for InuR consisting of two CGG triplets separated by 8 nucleotides.

Introduction

Some plants store part of their carbon as sucrose and as polymers of fructose (fructans). Plant fructans, often referred to as inulins, have predominantly β -2,1 linkages between the fructose moieties and terminate with a sucrosyl residue. Inulin is present in a range of plant species. In some species such as chicory, garlic, Jerusalem artichoke and Dahlia tubers about 15-20% of the plant fresh weight consists of inulin (Gupta and Kaur, 1997; Kaur and Gupta, 2002). Apart from the use of inulin for the production of fructose rich syrups, inulin and fructo-oligosaccharides (FOS) have attracted considerable research attention because of the large number of health benefits obtained from inulin and FOS in the diet humans (see for a review Kaur and Gupta, 2002).

Aspergillus niger is a saprophytic fungus, mainly present in the soil, feeding preferably on organic matter such as plant cell wall polysaccharides (cellulose, hemicellulose and pectin), and on plant storage polysaccharides (starch and inulin). *A. niger* is able to produce various inulin degrading enzymes including an exo-inulinase, an endo-inulinase and an invertase (Moriyama et al. 2003; Ohta et al. 1998, Akimoto et al. 1999; Bergès et al 1993, Boddy et al. 1993; L'Hocine et al. 2000). The inulinolytic enzymes are simultaneously induced in the presence of inulin or sucrose, indicating that the expression of the genes encoding these enzymes is co-regulated and controlled by a common transcription factor (Yuan et al. 2006).

Transcriptional activation of catabolic enzyme networks is often accomplished through transcription factors of the Zn(II)2Cys6 type. This type of transcription factor is unique to fungi. Gal4 acts as a transcriptional activator for the expression of genes involved in galactose utilization, and is the best characterized zinc cluster protein. In general, Zn(II)2Cys6 transcription factors possess a well conserved N-terminal localized DNA binding motif (CX₂CX₆CX₅₋₁₆CX₂CX₆₋₈), followed by a less well defined conserved domain

known as a fungal specific transcription factor domain (Pfam04082). The C-terminal part of the Zn(II)2Cys6 transcription factor normally contains the activation domain. In the genome sequence of *Saccharomyces cerevisiae* 54 putative Zn(II)2Cys6 transcription factors have been identified (Akache et al. 2001). They are involved in the transcriptional control of a wide variety of cellular processes (Todd and Andrianopoulos, 1997; Akache et al. 2001) including genes involved in amino acid metabolism (ArgRp, Crabeel et al. 1995; Leu3p, Kohlhaw, 2003; Lys14p, El Alami et al. 2000), sugar metabolism (Mal63p, Ni and Needleman, 1990), pyrimidine biosynthesis (Ppr1p, Flynn and Reece, 1999). In the genome of filamentous fungi even more Zn(II)2Cys6 transcription factors are present. In the recently released genome sequence of *A. niger* 296 ORFs were identified that contain the Zn(II)2Cys6 motif and similar numbers are predicted from the genomes of *A. fumigatus* and *A. nidulans* (Pel et al. 2007). The role of only a few of the Zn(II)2Cys6 transcription factors in filamentous fungi have been studied in detail. Filamentous fungal Zn(II)2Cys6 proteins are involved in the regulation of various processes which include secondary metabolite production (e.g. AflR, Woloshuk et al. 1994), pigmentation (e.g. Cmr1, Tsuji et al. 2000), nitrogen metabolism (NirA, Burger et al. 1991), proline utilization (PrnA, Cazelle et al. 1998; Leucine biosynthesis (Leu3p, Kirkpatrick and Schimmel, 1995), alcohol utilization (AlcR, Kulmberg et al. 1992) and sexual development (RosA, Vienken et al. 2005). Well studied examples of fungal Zn(II)2Cys6 transcription factors required for polysaccharide catabolism include AmyR, a transcriptional activator of starch degrading enzymes (Petersen et al. 1999; Gomi et al. 2000; Tani et al. 2001a), and XlnR, a transcriptional activator for xylanolytic enzymes (van Peij et al. 1998a; van Peij et al. 1998b; Gielkens et al. 1999; Hasper et al 2004).

It is well established that the genes encoding many proteins required for secondary metabolite production are clustered in the genome (Woloshuk et al. 1994; Walton, 2000). In addition to several biosynthetic enzymes that synthesize the secondary metabolites, such clusters also contain a Zn(II)2Cys6 transcription factor which is required for the transcriptional activation of the genes within the cluster. Gene clustering is not limited to secondary metabolite pathways, but also found for the proline utilization genes (PrnA, Gomez et al. 2002) and for some of the amylolytic genes under the control of the regulator AmyR. Adjacent to the gene for the AmyR transcription factor are genes encoding alpha-glucosidase (AgdA) and an alpha-amylase (AmyA) (Gomi et al. 2000). Also in *A. parasiticus*, a sugar utilization gene cluster has been identified (Yu et al. 2000). However, it is clear that clustering of genes with related functions is not a general rule, since the genes adjacent to the XlnR transcription factor do not seem to be involved in xylan degradation.

In a previous study, we showed that the expression of the genes encoding inulin degrading enzymes is co-regulated and specifically induced in the presence of inulin and sucrose (Yuan et al. 2006). The availability of the genome sequence of *A. niger* allowed the rapid identification of Zn(II)2Cys6 transcription factors adjacent to or in close proximity to the genes encoding (putative) enzymes involved in inulin catabolism. We have identified a Zn(II)-cluster transcription factor (InuR) which is adjacent to a gene encoding a sugar transporter and a recently identified intracellular invertase (SucB) with transfructosylation activity (Goosen et al. 2007).

In this paper we show that deletion of the *inuR* gene results in strongly impaired growth on inulin and sucrose, and also abolishes the induction of the genes encoding the extracellular enzymes involved in inulin and sucrose degradation. The gene encoding the putative sugar transporter located next to *inuR* is co-regulated with the genes encoding the inulinolytic enzymes. These results show that InuR acts as a positive acting transcriptional activator for the induced expression of genes involved in the breakdown of inulin and sucrose and the uptake of inulin breakdown products.

Materials and methods

Strains, culture conditions and transformation

A. niger strain N402 used in this study was derived from the wild-type strain *A. niger* van Tieghem (CBS 120.49, ATCC 9029) (Bos et al. 1988). The *A. niger* strain used by DSM to sequence the genome is NRRL 3122. Strain AB4.1 is a *pyrG* negative derivative of N402 (van Hartingsveldt et al. 1987) and was used to construct disruption strains. *A. niger* strains were grown in *Aspergillus* minimal medium (MM) (Bennett and Lasure, 1991), or *Aspergillus* complete medium (CM) consisting of MM medium with the addition of 0.5% (w/v) yeast extract and 0.1% (w/v) casamino acids. Growth medium was supplemented with 10 mM uridine (Serva) when required. For shake flask cultures, *A. niger* strains were grown in MM supplemented with 1% (w/v) carbon source and 0.1% (w/v) casamino acids. The conidiospores were inoculated at a concentration of 2×10^6 spores per ml. Glucose and sucrose (BDH chemicals), maltodextrin (Avebe), starch (Windmill Starch, Avebe), xylose, fructose and maltose (Sigma-Aldrich), raffinose (Sigma chemicals) and inulin (Sensus Frutafit, Cosun) were used as carbon sources.

For transfer experiments, *A. niger* strains were pre-grown in MM supplemented with 2% (w/v) xylose and 0.1% (w/v) casamino acids for 18 h at 30 °C on a rotary shaker at 300 rpm. Then mycelium was harvested by suction over a nylon membrane and washed with MM without carbon source. Aliquots of 1.6 g wet weight of mycelium were transferred to 300-ml Erlenmeyer flasks containing MM supplemented with 1% (w/v) carbon source and incubated at 30 °C for the time indicated in the text. The mycelium was harvested over Myracloth (Calbiochem) and frozen in liquid nitrogen, followed by storage at -80 °C prior to the isolation of total RNA. Conidiospores were obtained by harvesting spores from CM plates after 4-6 days of growth at 30 °C, using a 0.9% NaCl solution. Transformation of *A. niger* AB4.1 was done as described by Punt and van den Hondel (1992) using lysing enzymes (L1412, Sigma) for protoplastation. The bacterial strain used for transformation and amplification of recombinant DNA was *Escherichia coli* XL1-Blue (Stratagene). Transformation of XL1-Blue was performed according to the heat shock protocol as described by Inoue et al. (1990).

Construction of the *inuR::pyrG* deletion strain

The *A. niger inuR* gene was deleted by the replacement of the complete open reading frame of the *inuR* gene with the *Aspergillus oryzae pyrG* gene. The plasmid used to disrupt *inuR*

was constructed as follows. The DNA fragments flanking the *inuR* ORF were amplified by PCR using N402 genomic DNA as template. Fragments with a length of 0.8 kb of 5' flanking DNA and 1.0 kb of 3' flanking DNA were amplified using primers *inuRP1* and *inuRP2*, *inuRP3* and *inuRP4* (Suppl. Table 1), respectively. Each primer was adapted with restriction sites as indicated for further cloning. The amplified PCR fragments were digested with *NotI* and *XhoI* or *XhoI* and *KpnI*, respectively and cloned into pBlue-Script II to obtain plasmid *pInuRF5* and *pInuRF3*. Subsequently, the 0.8 kb *NotI-XhoI* fragment containing 5' flanking of *inuR* from *pInuRF5*, was ligated into *NotI/XhoI* digested *pInuRF3* to give *pInuRF53*. The *A. oryzae pyrG* gene was isolated as a 2.8 kb *Sall-BamHI* fragment, obtained from plasmid *pAO4-13* (de Ruiter-Jacobs et al. 1989) and ligated into *XhoI/BamHI* digested *pInuRF53* to obtain *pΔinuR* (*pXY3.1*). Before transformation, the plasmid *pΔinuR* was linearized with *NotI* and transformed into *A. niger pyrG⁻* strain AB4.1. Uridine prototrophic transformants were selected by incubating protoplasts on agar plates containing MM without uridine. Transformants were screened by PCR using primers *PAO9* and *inuRP5* or *PAO10* and *inuRP6* (Suppl. Table 1). Only transformants with a targeted deletion of the *inuR* gene should result in the amplification of a 1.2 or 1.4 kb PCR fragment, respectively. PCR positive transformants were verified by Southern blot analysis as described by Sambrook et al. (1989). Chromosomal DNA from positive PCR transformants was isolated as described by Kolar et al. (1988). 10 µl of genomic DNA was digested for 3h with 10U of *XhoI* or *EcoRI*, respectively. The 1.0 kb of the 3' flanking DNA fragment of *inuR* was used as probe for the detection of *inuR* disruptants. A probe was generated by digestion of the *pInuRF3* plasmid containing the *inuR* 3' flanking fragment with *XhoI* and *BamHI*. Fragments were purified from gel and [α -³²P]dCTP labeled probes were synthesized using Rediprime II DNA labeling System (Amersham Pharmacia Biotech) according to the instructions of the manufacturer.

Complementation of *inuR* gene

The *A. niger inuR* gene was amplified by PCR using primers *inuRP7* and *inuRP8* (Table 1) and N402 genomic DNA as a template. The primers corresponded to 1 kb upstream and 0.5 kb downstream of the *A. niger inuR* gene respectively. PCR was performed using Phusion™ High-Fidelity DNA Polymerase (Finnzymes) according to the manufacturer's instruction. The amplified 4.1 kb PCR fragment was cloned into pGEMT-easy vector (Promega) to obtain the *inuR* complementation plasmid *pXY5.1*. *pXY5.1* was then co-transformed with *pAN7.1* harboring hygromycin gene as selection marker (Punt et al. 1987) into *A. niger ΔinuR* strain XY3.1 (see in the text). Hygromycin resistant transformants were selected by incubating protoplasts on agar plates containing MM with 200 mg/ml of hygromycin. Southern blot analysis was performed on selected transformants as described by Sambrook et al (1989).

Construction of the $\Delta inuR \Delta amyR$ double deletion strain

To make a *pyrG*⁻ strain of the $\Delta inuR$ strain (XY3.1) this strain was incubated on MM plates containing 1 mg/ml of 5'FOA and 10 mM uracil. 5'FOA resistant, uracil requiring mutants were transformed with pAB4.1 harboring the *A. niger pyrG* or with pPyrE harboring the *A. niger pyrE* gene on a 4.3 kb SstII subclone, respectively. One of the selected 5'FOA resistant mutants, XY4.1, was complemented with the *pyrG* gene, indicating that the uracil auxotrophy was caused by a mutation in the *pyrG* gene. The strain XY4.1 was then used to construct a $\Delta inuR \Delta amyR$ double disruption strain by transforming plasmid p $\Delta amyR$ which also contains the *pyrG* selection marker (Yuan et al. submitted). Uridine prototrophic transformants were selected by incubating protoplasts on agar plates containing MM without uridine. Transformants that showed normal growth on glucose but not on starch were selected. One possible $\Delta inuR \Delta amyR$ double mutant (XY5.1) was verified by Southern analysis and used for further analysis.

Northern blot analysis

Total RNA isolation, Northern analysis and synthesis of DNA probes was performed as described in Yuan et al. (2006). The primers used to generate probe An15g00310 are shown in Suppl. Table 1.

Microtiter plate growth assay

Growth of *A. niger* strains were determined using a HTS7000 BioAssay Reader (Perkin Elmer Life Sciences). Spores (1×10^4) were inoculated in each well of a 96-well microtiter plate (Nalge Nunc International, USA) and incubated at 32°C for 56 h. Each well contained 200 μ l of MM containing 1% (w/v) of one of the various carbon sources each supplemented with 0.01% (w/v) casamino acids to stimulate spore germination. Six replicates of each condition were made. Growth was monitored by measuring turbidity (OD_{595}) at 2 h intervals.

Microarray experiments and data analysis

RNA extracted from the *A. niger* $\Delta inuR$ strain and wild-type strain (N402) grown on different carbon sources were used for microarray experiments using custom made 'dsmM_ANIGERa_coll' Affymetrix GeneChip® Microarrays kindly provided by DSM Food Specialties (Delft, The Netherlands). For microarray experiments, mycelia were isolated from grown cultures by transfer experiment (see above) and each growth condition was performed in duplicate as independent biological experiments.

Total RNA was isolated from mycelia using TRIzol reagent (Invitrogen) and RNA quality was verified by analyzing aliquots with glyoxal/DMSO gel electrophoresis and Agilent Bioanalyzer "Lab on chip" system (Agilent Technologies, U.S.A.). Processing, labeling and hybridization of cRNA to *A. niger* Affymetrix GeneChips were performed according to the corresponding Affymetrix protocols for "Eukaryotic Target Preparation"

and “Eukaryotic Target hybridization”. For probe array washing and staining, the protocol “Antibody Amplification for Eukaryotic Targets” was followed. Hybridized probe array slides were scanned with Agilent technologies G2500A Gene Array Scanner at a 3 µm resolution and a wavelength of 570 nm. Affymetrix Microarray Suite software MAS5.0 was used to calculate the signal and p-values and to set the algorithm’s absolute call flag, which indicates the reliability of the data points according to P (present), M (marginal) and A (absent). The data on each chip were globally scaled to an arbitrary target gene intensity of 500.

The pre-scaled chip data from each hybridization experiment were normalized using Genespring 7.0 software (Silicon Genetics). Normalization with default parameters in Genespring software (Per Chip: Normalize to 50th percentile, Per Gene: Normalize to median) was used. For this study, we focused on sucrose induced genes, therefore pre-filtering of data was performed to select for genes whose detection calls are present in both sucrose duplicate samples of wild-type strain N402. The selected data set was further performed for One-way ANOVA analysis. Fold changes in expression levels between two different conditions were then computed for genes with $p < 0.05$ based on One-way ANOVA analysis and changes by more than 2-fold were considered significant and are reported here. Microarray data were deposited into ArrayExpress with an accession code A-MEXP-848 at <http://www.ebi.ac.uk/miamexpress>.

***In silico* analysis of regulatory elements of co-regulated genes**

The 1.0 kb promoter sequences of the seven sucrose induced, InuR dependant genes: An12g08280 (InuE), An11g03200 (InuA), An08g11070 (SucA), An15g04060, An15g00310, An15g03940, and An17g01710 were obtained from the CBS513.88 strain. The 1.0 kb promoter regions were analyzed on both strands using DNA Analyzer module of PhyloSopher 6.5.1 (Genedata A.G., Basel, Switzerland) using the Gibbs Sampling Strategy for Multiple Alignment (Lawrence et al. 1993).

Results

Identification of potential transcriptional regulators involved in inulin catabolism

A. niger is able to produce various inulinolytic enzymes involved in the modification or degradation of inulin and sucrose. The transcriptional regulation of the genes encoding the inulinolytic enzymes of *A. niger*, including an exo-inulinase (InuE), an endo-inulinase (InuA), and an invertase (SucA), has been studied previously (Moriyama et al. 2003; Yuan et al. 2006). In addition to the extracellular enzymes, two additional potential intracellular inulinolytic enzymes (SucB and SucC) were predicted from the *A. niger* genome sequence (Yuan et al. 2006). These proteins contain all the conserved domains of the GH32 family to which inulinolytic enzymes belong. Finally, a potential ORF was found in the genome which showed homology to the group of exo-inulinases. Because the predicted ORF lacked several of the conserved domains and contained several frame shift mutations, this gene (*inuQ*) was considered to be a pseudogene (Yuan et al. 2006).

The expression of the genes encoding the extracellular inulinolytic enzymes is co-regulated and induced on inulin and sucrose, suggesting that these genes are under control of a single transcription factor (Yuan et al. 2006). Some transcriptional activators involved in sugar catabolism are clustered in the genome with their target genes (Gomi et al. 2000; Yu et al. 2000). To identify possible candidate transcription factors involved in the regulation of inulinolytic genes, we searched the *A. niger* genome sequence for transcription factors located close to the genes encoding the inulinolytic enzymes. We found that *inuA* (An11g03200) and *inuQ* (An11g03210) are located next to each other and that also a Zn(II)2Cys6 transcription factor encoding gene (An11g03220) was located adjacent to *inuQ*. The orientation of the three ORFs is in the same direction indicating that they do not share a common promoter region (Fig. 1).

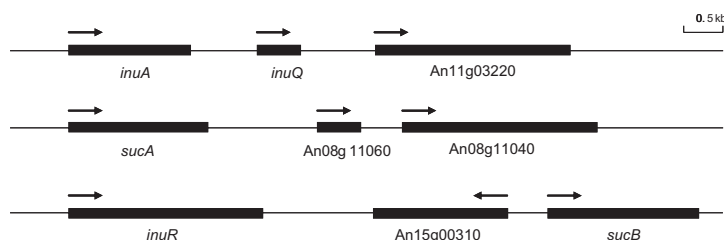


Fig. 1. Schematic representation of the clustering of putative transcription factors with inulinolytic genes in the genome of *A. niger*. An11g03220 and An08g11040 and *inuR* encode putative transcription factors. The genes related to inulin degradation include: *inuA* (endo-inulinase), *inuQ* (pseudogene), *sucA* (invertase), *sucB* (homologous to *sucA*) and An15g00310 (putative sugar transporter). An08g00160 encodes a hypothetical protein without known function. The arrows indicate the transcriptional orientations of the genes.

In the proximity of *sucA* (An08g11070), a second transcription factor encoding gene was found. This transcription factor contains a Cys2His2 zinc finger DNA binding motif. Between this gene, An08g11040 and *sucA* an additional ORF (An08g11060) is present. This 273 amino acid protein encodes a hypothetical protein without any conserved motifs. A BlastP search revealed that this protein has no close orthologs in other fungal genomes. The transcription factor, encoded by An08g11040, shows strong sequence similarity to the *S. cerevisiae* Zpr1p (e-value 1e-100), an essential transcription factor that contributes to normal cell proliferation (Gangwani et al. 1998).

Clustered with *sucB* (An15g00320) a third putative transcription factor encoding gene (An15g00300) was found which also belonged to the Zn(II)2Cys6 transcription factor family. Between *sucB* and An15g00300, an additional ORF (An15g00310) is located (Fig. 1). The protein encoded by this gene has all the characteristics of a sugar transporter protein. In the direct proximity of the two remaining genes encoding inulinolytic enzymes (*InuE* and *SucC*) no potential transcription factors were found. The three transcription factors identified via clustering were considered to be good candidates to be involved in the transcriptional regulation of the inulinolytic system of *A. niger*. In order to study their possible role in inulin utilization, deletion strains of these transcription factors were

constructed. Deletion of An11g03220 did not have any measurable effect on growth on inulin or sucrose in comparison with the wild-type strain, which indicates that this gene is not required for inulin utilization. Growth of the deletion strain on xylose, glucose, fructose, starch and maltose was also identical to the growth of the wild-type strain (data not shown). A deletion strain of An08g11040 was never obtained although over 500 transformants were screened. A possible explanation might be that An08g11040 is an essential gene. As indicated above, An08g11040 is most homologous to the essential Zpr1p transcription factor of *S. cerevisiae*. Since the homology with Zpr1p does not indicate involvement of this transcription factor in inulin utilization, no further attempts were made to obtain the deletion strain. As shown in detail below, deletion of An15g00300 resulted in a strain which showed a severe growth defect on inulin and on sucrose (Fig. 4 and 5). For the remainder of the paper, we will refer to An15g00300 as *inuR*.

InuR homologs were also identified in the genomes of other aspergilli. The alignment of the InuR proteins is given in Fig. 2. The strong conservation of the transcription factor among aspergilli, suggest that the regulation of inulinolytic gene expression is mediated in the various aspergilli by the InuR proteins present in the different fungi.

InuR encodes a Zn(II)2Cys6 transcription factor that is required for growth on inulin

Based on the analysis of the CBS513.88 genomic sequence, the predicted open reading frame of the *inuR* encoding gene is 2466 bp long and interrupted by four introns with sizes of 153 bp, 80 bp, 58 bp and 48 bp, respectively. The predicted InuR protein sequence is comprised of 709 amino acid residues which would result in a protein with a calculated molecular mass of 78.3 kDa. Analysis of the predicted InuR protein indicated that InuR contains two conserved domains. One is the Zn(II)2Cys6 (CX₂CX₆CX₅CX₂CX₆C) binuclear cluster which represents the DNA-binding domain at the NH₂-terminal end (residues 35-61). The domain is very well conserved (Pfam00172) and binds two Zn atoms which coordinate folding of the domain. A BlastP search revealed orthologous transcription factors in the genome of *A. nidulans*, *A. oryzae* and *A. fumigatus* (see above, Fig. 2). The most similar transcription factors to InuR which have been functionally characterized, apart from the proposed orthologs in *A. oryzae*, *A. fumigatus* and *A. nidulans*, are the AmyR transcription factors from the different *Aspergillus* species. Subsequent phylogenetic analysis indicated that the subgroup of InuR transcription factors is most closely related to the group of AmyR transcription factors (data not shown). Besides the Zn(II)2Cys6 motif, a second conserved domain, known as the middle homology region (MHR) or Pfam04082 domain, is present in Zn(II)2Cys6 transcription factors. The region is thought to assist the Zn(II)2Cys6 cluster in DNA target discrimination (Schjerling and Holmberg, 1996). In the *A. niger* InuR protein, the MHR/Pfam04082 domain comprises aa 353-425, and the domain is also present in the InuR proteins of the other aspergilli.

AgInuRp	MNFNSGAATTSAMTCAIRTSRRKSSGA.....IARRACDCKFRKIKCSLSQPCAGNS..NGFECHFFQCKRGT	73
AoInuRp	MNPFSDAITTPMTCAIRTESRRKTSGA.....IARRACDCKFRKIKCSLSQPCAGNS..NGFECHFFQCKRGT	72
AfInuRp	MNPFSDAITTPMTCAIRTESRRKTSGA.....IARRACDCKFRKIKCSLSQPCAGNS..NGFECHFFQCKRGT	56
AdInuRp	..MFFS.....SRQSRSGGSQAGPKCATRACDCKFRKIKCSLSQPCAGNS..NGFECHFFQCKRGT	62
AgInuRp	GRVSVQIRQ..QOTN...L..PALNSSCSQVAFNPLGAFAYVOTESP..SRGDCNSNVH...P..IQDAAHFM..H.N.EA	139
AoInuRp	GRVSVQIRQ..QOTN...L..PALNSSCSQVAFNPLGAFAYVOTESP..SRGDCNSNVH...P..IQDAAHFM..H.N.EA	133
AfInuRp	GRVSVQIRQ..QOTN...L..PALNSSCSQVAFNPLGAFAYVOTESP..SRGDCNSNVH...P..IQDAAHFM..H.N.EA	118
AdInuRp	..ARVQIRQ..QOTN...L..PALNSSCSQVAFNPLGAFAYVOTESP..SRGDCNSNVH...P..IQDAAHFM..H.N.EA	103
AgInuRp	SVFVEL...SHQPFPAPFAPNWSG...ASQLETHSNQVSWNPRNDYEWLPDLDLQ..VVFYDFGCSNLYLKASLFS	212
AoInuRp	SVFVEL...SHQPFPAPFAPNWSG...ASQLETHSNQVSWNPRNDYEWLPDLDLQ..VVFYDFGCSNLYLKASLFS	205
AfInuRp	SVFVEL...SHQPFPAPFAPNWSG...ASQLETHSNQVSWNPRNDYEWLPDLDLQ..VVFYDFGCSNLYLKASLFS	191
AdInuRp	SVFVEL...SHQPFPAPFAPNWSG...ASQLETHSNQVSWNPRNDYEWLPDLDLQ..VVFYDFGCSNLYLKASLFS	163
AgInuRp	ITQIAAFAGLQSHPAEASVPMSPGV..L.HPP..PNEENTLMESSIVEANMIPWLVYDFDLRHLPTLVNRSLSIFIRML	289
AoInuRp	ITQIAAFAGLQSHPAEASVPMSPGV..L.HPP..PNEENTLMESSIVEANMIPWLVYDFDLRHLPTLVNRSLSIFIRML	284
AfInuRp	ITQIAAFAGLQSHPAEASVPMSPGV..L.HPP..PNEENTLMESSIVEANMIPWLVYDFDLRHLPTLVNRSLSIFIRML	268
AdInuRp	ITQIAAFAGLQSHPAEASVPMSPGV..L.HPP..PNEENTLMESSIVEANMIPWLVYDFDLRHLPTLVNRSLSIFIRML	238
AgInuRp	SCEHRNPFQGMGLISICAFSLTQFIEDEPPTSSSEAQAQAFSMNTEATMRSCSDFGCEPFEAVLTSFFLFGCLFGSN	369
AoInuRp	SCEHRNPFQGMGLISICAFSLTQFIEDEPPTSSSEAQAQAFSMNTEATMRSCSDFGCEPFEAVLTSFFLFGCLFGSN	364
AfInuRp	SCEHRNPFQGMGLISICAFSLTQFIEDEPPTSSSEAQAQAFSMNTEATMRSCSDFGCEPFEAVLTSFFLFGCLFGSN	348
AdInuRp	SCEHRNPFQGMGLISICAFSLTQFIEDEPPTSSSEAQAQAFSMNTEATMRSCSDFGCEPFEAVLTSFFLFGCLFGSN	318
AgInuRp	QHNAAFLRLREAUCLAPITGLNNHDSYDLFSDERGCGLRTYLVLSITERAYALQRHFITFEGSGEGRVSVDFIHNA	449
AoInuRp	QHNAAFLRLREAUCLAPITGLNNHDSYDLFSDERGCGLRTYLVLSITERAYALQRHFITFEGSGEGRVSVDFIHNA	444
AfInuRp	QHNAAFLRLREAUCLAPITGLNNHDSYDLFSDERGCGLRTYLVLSITERAYALQRHFITFEGSGEGRVSVDFIHNA	428
AdInuRp	QHNAAFLRLREAUCLAPITGLNNHDSYDLFSDERGCGLRTYLVLSITERAYALQRHFITFEGSGEGRVSVDFIHNA	397
AgInuRp	TRSVVSGIVVREKDAPEMGILARIMELFDALPEEDVILCWNRONINAG..YCKRLTEAKPLAHONISRVSESERYK..	525
AoInuRp	TRSVVSGIVVREKDAPEMGILARIMELFDALPEEDVILCWNRONINAG..YCKRLTEAKPLAHONISRVSESERYK..	520
AfInuRp	TRSVVSGIVVREKDAPEMGILARIMELFDALPEEDVILCWNRONINAG..YCKRLTEAKPLAHONISRVSESERYK..	504
AdInuRp	TRSVVSGIVVREKDAPEMGILARIMELFDALPEEDVILCWNRONINAG..YCKRLTEAKPLAHONISRVSESERYK..	477
AgInuRp	..GQDPERSK.....STLGDNG.....N.FEAGLETCQCADVFTRKWLQNFVWLCSTHGLLRPS..EHHEL	587
AoInuRp	..GQDPERSK.....STLGDNG.....N.FEAGLETCQCADVFTRKWLQNFVWLCSTHGLLRPS..EHHEL	582
AfInuRp	..GQDPERSK.....STLGDNG.....N.FEAGLETCQCADVFTRKWLQNFVWLCSTHGLLRPS..EHHEL	566
AdInuRp	GHGSDPFDSDPQEPDSASSGQNGRELIESRALRDELSETQCADVILVQKWLQNFVWLCSTHGLLRPS..EHHEL	557
AgInuRp	GLHNVSVAKETIDICRSLRSSVEAHGICGAEKLYDIATVATSVLNV..QMPLAKGTILASP..ARAM..P.N....PQI	658
AoInuRp	GLHNVSVAKETIDICRSLRSSVEAHGICGAEKLYDIATVATSVLNV..QMPLAKGTILASP..ARAM..P.N....PQI	659
AfInuRp	GLHNVSVAKETIDICRSLRSSVEAHGICGAEKLYDIATVATSVLNV..QMPLAKGTILASP..ARAM..P.N....PQI	638
AdInuRp	GLHNVSVAKETIDICRSLRSSVEAHGICGAEKLYDIATVATSVLNV..QMPLAKGTILASP..ARAM..P.N....PQI	612
AgInuRp	P..M.AGS.Q..I..V.S.KT..HEDFILINSLRGGNHFEYRYRTLS..S.L..DGIGHA.WGSO	709
AoInuRp	P..M.AGS.Q..I..V.S.KT..HEDFILINSLRGGNHFEYRYRTLS..S.L..DGIGHA.WGSO	726
AfInuRp	P..M.AGS.Q..I..V.S.KT..HEDFILINSLRGGNHFEYRYRTLS..S.L..DGIGHA.WGSO	701
AdInuRp	TIQ.....V....ANNYKILNLTIRGGNHFEYRYRTLS..S.L..DGIGHA.WGSO	652

Fig. 2. Alignment of InuRp and its orthologs in aspergilli. The amino acid sequences of the Zn₂Cys₆ binuclear DNA-binding motif (Pfam00172) and the Fungal specific transcription factor domain (Pfam04082) are underlined. The conserved six cystein residues are indicated by stars. AoInuRp, AgInuRp, AfInuRp and AdInuRp are InuR protein orthologs from *A. oryzae*, *A. niger*, *A. fumigatus* and *A. nidulans*, respectively.

Construction and growth properties of the *inuR* disruptant in *A. niger* N402

An *inuR* disruptant mutant was constructed as described in the Materials and Methods. Two putative *inuR* deletions strains, XY3.1 and XY3.2, which were identified using PCR screening, were further examined by Southern blot analysis to verify proper deletion of the *inuR* gene. Using a 3' flanking fragment of *inuR* as a probe, *EcoRI* digested genomic DNA showed a shift from a 4.3 kb hybridizing DNA fragment in the wild-type strain N402 to a 5.7 kb fragment in both deletion mutants as expected. *XhoI* digested genomic DNA showed a shift from an anticipated 1.6 kb hybridizing fragment in WT to a fragment of more than 6.0 kb in the deletion mutants (Fig. 3). Strain XY3.1 was used to further characterize the InuR function and we will refer to this strain as the Δ *inuR* strain.

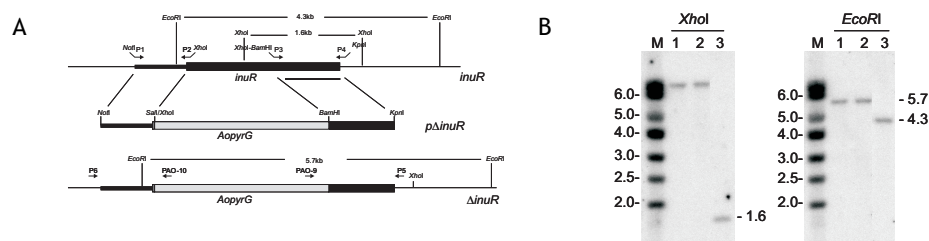


Fig. 3. Strategy of construction of the *inuR* gene disruptants in *A. niger*. (A) Schematic representation of the *inuR* wild-type locus, the *inuR* disruption plasmid (pΔ*inuR*), and the predicted *inuR* locus after deletion (Δ*inuR*). Primers used to make the disruption cassette and primers used for PCR screening are indicated. The *Bam*HI-*Kpn*I fragment underlined was used as a probe. (B) Southern blot analysis of the *A. niger* wild-type and *inuR* disruption strains. Genomic DNA was digested with *Xho*I or *Eco*RI. Lane 1 and lane 2 are *inuR* disruptant strains; Lane 3 is a wild-type strain. The two blots were hybridized with the *Bam*HI-*Kpn*I fragment as indicated.

The effect of the disruption of the *inuR* gene on the utilization of various carbon sources was analyzed on agar plates. Growth and conidiation of the Δ*inuR* strain was strongly reduced on plates containing inulin in comparison to the wild-type strain (Fig. 4). As shown, growth of the Δ*inuR* strain on inulin was not completely abolished. After seven days of growth, a thin colony with a reduced degree of conidiation was formed. Some effect on growth and conidiation was also observed for the Δ*inuR* strain growing on plates containing sucrose and to a lesser extent on raffinose. The differences in growth on sucrose and raffinose were most clear after three days of growth (Fig. 4). Both the colony size and the degree of conidiation of the Δ*inuR* strain and the wild-type strain after seven days of growth were very similar (data not shown). The Δ*inuR* strain showed a similar growth phenotype to the wild-type on MM plates containing fructose, glucose, maltose, starch, xylose, sorbitol or glycerol as sole carbon source (Fig. 4 and data not shown).

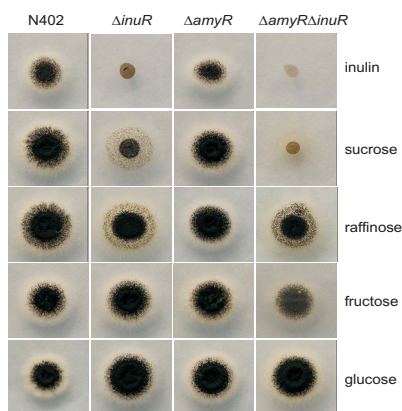


Fig. 4. Effects of *A. niger* strains on the utilization of different carbon sources. The wild-type (N402) and *inuR* disruptant (Δ*inuR*), *amyR* disruptant (Δ*amyR*) and double mutant (Δ*amyR*Δ*inuR*) strains were grown on MM plates containing 1 % different carbon sources at 30°C for 3 days.

Growth of the $\Delta inuR$ strain on different carbon sources was also determined by analyzing growth in a microtiter plate well containing MM with various carbon sources. Equal numbers of spores of the wild-type strain (N402) and $\Delta inuR$ strain were inoculated and plates were incubated at 32°C. For each condition at least four replicates were made and the average value of the replicates and its standard deviation value were determined as shown in Fig. 5. Whereas the wild-type strain is able to grow on inulin, the $\Delta inuR$ showed a clear growth defect. After an initial increase in optical density in the first 16 h, the OD₅₉₅ of the $\Delta inuR$ strain remained constant after further incubation whereas the wild-type strain continued to grow. Similar growth defect of $\Delta inuR$ strain was also observed on sucrose as sole carbon source although the $\Delta inuR$ strain showed a slightly higher initial growth than on inulin (Fig. 5). A small, but reproducible growth difference in growth on raffinose on agar plates was also observed in the microtiter plate analysis. Initially, the growth curves of the wild-type and the $\Delta inuR$ strain on raffinose were identical, but after 24 to 32 hours growth of the $\Delta inuR$ strain seem to be retarded (Fig. 5). Growth of the $\Delta inuR$ strain was also measured in microtiter plate wells containing MM with maltodextrin, maltose, glucose, xylose or fructose as sole carbon source. The $\Delta inuR$ strain grew equally well as the wild type strain on these carbon sources (Fig. 5). In conclusion, the growth analysis of the $\Delta inuR$ strain indicates that the Zn(II)2Cys6 transcription factor is required for the efficient utilization of inulin. The results also indicate that the InuR affects the utilization of sucrose and to a lesser extent raffinose, and it may therefore act as a positive transcriptional activator of the genes encoding the inulinolytic enzymes. To further substantiate that the defective growth of the $\Delta inuR$ strain on inulin was caused by the disruption of the *inuR* gene, a plasmid containing the *inuR* gene with 1 kb upstream and 0.5 kb downstream region was constructed and co-transformed into the $\Delta inuR$ strain with pAN7.1 (Punt et al. 1987).

Hygromycin resistant transformants were analyzed for growth on inulin and subjected to Southern blot analysis. Transformants that showed restored growth on inulin all harbor the *inuR* gene, while transformants showing defective growth on inulin did not contain the *inuR* gene (data not shown). The restoration of the wild-type growth on inulin upon retransformation of the $\Delta inuR$ strain with a functional copy of the *inuR* gene indicated that the inability to grow on inulin was indeed a consequence of *inuR* disruption.

InuR is required for the induction of inulinolytic genes

To further establish the role of InuR as a positive acting transcriptional regulator for inulinolytic gene expression, Northern blot analysis was performed to determine the mRNA level of the various inulinolytic enzyme encoding genes both in the $\Delta inuR$ strain and the WT strain. We previously showed that the expression of the genes encoding the exo-inulinase (*inuE*), endo-inulinase (*inuA*), invertase (*sucA*), was detected only in the presence of inulin and sucrose and not on other carbon sources tested such as xylose, glucose, maltose or starch. Expression of the *sucB* gene was low on all the carbon sources tested and slightly induced in the presence of inulin and sucrose (Yuan et al. 2006).

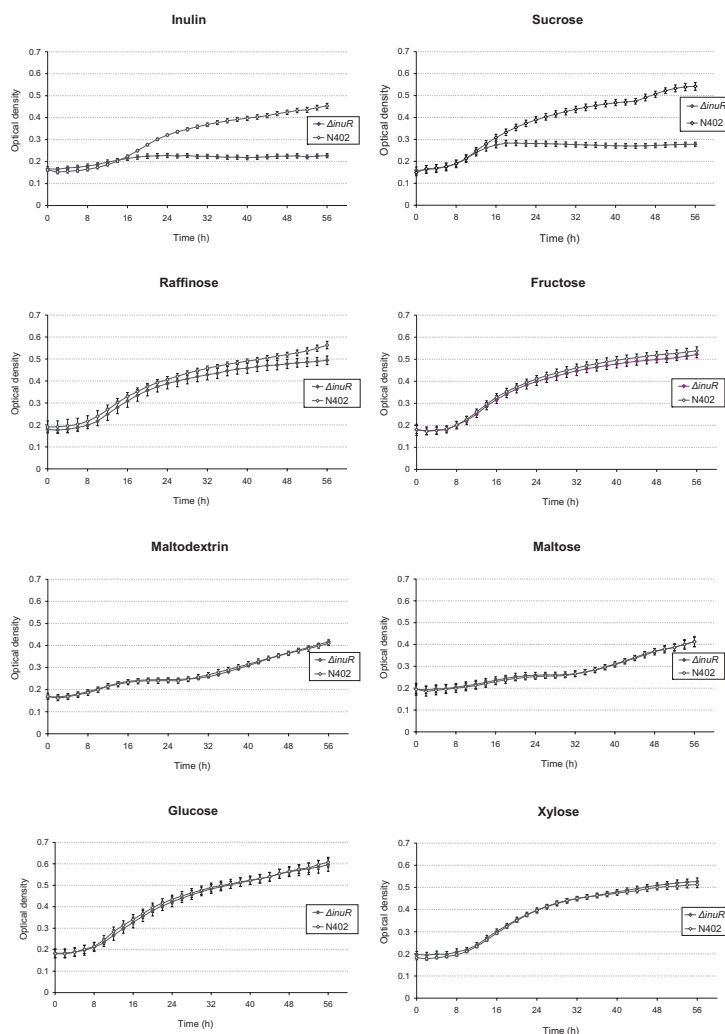


Fig. 5. Microtiter plate assay of growth of *inuR* disruptant on different carbon sources. *A. niger* wild type (N402; open circles) strain and *inuR* disruptant ($\Delta inuR$; solid circles) were inoculated in microtiter plate wells and each well was inoculated with 1×10^4 spores in 200 μ l of MM containing 1 % carbon source and 0.1% casamino acids. The growth was monitored by measuring culture turbidity (OD₅₉₅) and read every 2 hours. Each condition was performed in six replicates. The standard deviation value is indicated.

In order to clarify whether the expression of these inulinolytic genes was controlled by *InuR*, both the wild-type strain and the $\Delta inuR$ strain were pre-grown for 18 h in MM containing xylose as a carbon source. After the pre-growth, equal amounts of mycelium was transferred to fresh MM containing 1% (w/v) inulin or 1% (w/v) sucrose as a carbon source. At different time points after the transfer RNA was extracted from mycelia and subjected to Northern blot analysis (Fig. 6). The transfer to inulin and sucrose resulted in an induction of the inulinolytic enzymes in the wild-type strain. In the $\Delta inuR$ strain, no induction of

expression was found for the *inuE* and *inuA* genes after transfer to inulin and a very low level of *sucA* mRNA was detected after the transfer to inulin. The transfer to sucrose did not result in a detectable induction of any of the three inulinolytic genes in the $\Delta inuR$ strain. A slight induction of expression of *sucB* in the presence of inulin and sucrose seems also dependent on InuR (Fig. 6). The expression of the gene encoding a putative sugar transporter, An15g00310 in relation to inulin and sucrose was also examined. As discussed above, An15g00310 is clustered and located between the *inuR* and the *sucB* genes (Fig. 1). Expression of An15g00310 is strongly induced in response to the presence of inulin and sucrose and dependant on the InuR transcription factor (Fig. 6). No or very low expression levels of An15g00310 were detected when grown on maltose, glucose, fructose or xylose or after a transfer to maltose indicating that An15g00310 expression is specific for growth on inulin and sucrose (data not shown). The *inuR* gene was constantly expressed at a relatively low level on the various carbon sources tested (inulin, sucrose, maltose, fructose or glucose, data not shown), suggesting that the activation of the InuR transcription factor is regulated post-transcriptionally.

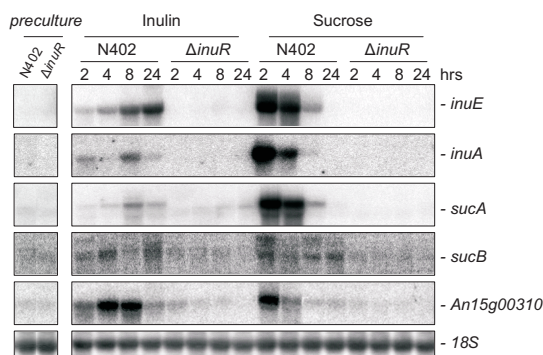


Fig. 6. Northern blot analysis of inulinolytic genes in wild-type (N402) and the $\Delta inuR$ strain from mycelia transferred from xylose (preculture) to inulin or sucrose. Total RNA was isolated from mycelia growing in inulin or sucrose at different time points (2, 4, 8, and 24 hours) as indicated after the transfer

Genome wide transcriptional profiling of the InuR regulon in *A. niger*

The experiments described above suggest that InuR is a transcriptional activator of inulinolytic genes and required for the efficient utilization of inulin and sucrose. In addition to regulating the inulinolytic genes, it is possible that InuR also regulates other genes in the genome, through direct or indirect processes in the pathway of inulin or sucrose utilization. To identify additional genes with a possible role in inulin or sucrose metabolism, we performed genome wide expression analysis using *A. niger* Affymetrix Gene Chip Arrays. The *A. niger* wild-type strain N402 and the $\Delta inuR$ strain were pre-grown for 18 h in 2 % (w/v) xylose and transferred to fresh medium containing 1 % (w/v) sucrose as carbon source and grown for two hours on sucrose. As a control, the mycelium from the N402 strain was also transferred to fresh 1 % (w/v) xylose medium. For each culture condition a duplicate

experiment was carried out. RNA was isolated from each of the samples and subjected to microarray analysis (see Materials and Methods).

To identify genes that were induced by sucrose, we first compared the expression profile of *A. niger* wild-type strain N402 grown in sucrose to that of *A. niger* wild-type strain N402 grown in xylose 2 h after transfer. The expression of 97 genes was more than 2-fold increased by sucrose (One-way ANOVA analysis $p < 0.05$) (Suppl. Table 2). To identify genes that were induced on sucrose in an InuR dependant way, the expression profile of the *A. niger* wild-type strain N402 grown in sucrose was compared to that of the $\Delta inuR$ strain after the 2h transfer to sucrose. A total of 40 genes was identified which expression was significantly increased on sucrose in the N402 strain, but not in the $\Delta inuR$ strain (One-way ANOVA analysis $p < 0.05$ and > 2 fold change) (Suppl. Table 3.) Finally, by combining these two data sets, 26 genes were identified for which expression was induced by sucrose in an InuR dependent manner (Fig. 7, Table 1).

The predicted protein sequences of these 26 genes were assigned to various functional categories (Ruepp et al. 2004, Mewes et al. 1997) and shown in Table 2. Interestingly, genes with the highest fold changes (> 10 -fold change) almost all fall into the category of carbohydrate transport and metabolism. Among them are the genes encoding the extracellular inulinolytic enzymes, *sucA*, *inuE* and *inuA*, as well as the above mentioned putative sugar transporter gene (An15g00310), supporting the previous Northern analysis. An additional gene that showed a very high fold-change and belongs to the category of carbohydrate transport and metabolism is An15g04060. An15g04060 encodes a putative sugar transporter which is over 200-fold induced by sucrose in a wild-type strain N402 when compared to xylose and the induction was greatly reduced by 127-fold when *inuR* gene was deleted. Blastp search revealed that the transporter displayed the highest identity a fructose transporter protein (Frz1) from *Zygosaccharomyces bailii* (Pina et al. 2004). An additional putative sugar transporter gene (An15g03940) was induced on sucrose in an InuR dependant way. In contrast to the other two transporters (An15g00310 and An15g04060) this transporter is relatively highly expressed on both xylose and in the $\Delta inuR$ mutant, resulting in relative small fold-changes. An additional gene (An17g01710) was strongly induced on sucrose (about 19-fold), in an InuR dependant way. The protein encoded by An17g01710 is annotated as a multidrug resistance (MDR) protein because of its highest identity to other MDR proteins of the Major Facilitator Superfamily of transporters. Since the substrates of MDR proteins are hard to predict on the basis of the amino acid sequence, the annotation should be interpreted with care and does not exclude the possibility that this protein has a role in sugar transport. Apart from the An15g04060 transporter which might function as a fructose transporter based on the strong homology towards Frz1p (Pina et al. 2004), a useful prediction of substrates of the other putative transporters can not be made.

In the list of induced genes (Table 1), a considerable number are related to transcription, translation and amino acid metabolism. The increase in expression of those genes might indicate an overall increase of metabolic activity on sucrose compared to xylose and a subsequent higher expression of the protein transcription/translation machinery.

The genome wide analysis indicated that three extracellular inulinolytic genes and four putative sugar transporter genes were strongly induced on sucrose in an InuR dependant way. To identify possible elements in the promoter regions of these genes an *in silico* analysis of 1-kb upstream promoter sequences of these seven genes was performed. As shown in Fig. 7, optimal alignment of the promoter sequences revealed a consensus sequence (CGGN8CGG) which is present at least once in the promoter region of all genes. At least one of the triplets consistently is preceded by a T, resulting in a conserved TCGG motif in every proposed element. Within the second triplet, the conservation of the GG is consistent, but conservation of the C is not. This putative binding site of the InuR protein (CGGN8CGG) is identical to the binding site of the AmyR transcription factor which binds and induces the expression of amylolytic genes (Petersen et al. 1999). To examine whether AmyR is also involved in the regulation of inulinolytic genes, the *_amyR* mutant was constructed and analysed for growth on various carbon sources. As shown in Fig. 4, deletion of the *amyR* transcription factor did not result in a growth defect on inulin, sucrose or other carbon sources. The $\Delta amyR$ strain did show a strong growth reduction on maltose (data not shown). To examine a possible overlapping role of the InuR and AmyR transcription factors for inulin and/or sucrose utilization, the $\Delta amyR/\Delta inuR$ double mutant was constructed. Growth analysis of the double mutant on inulin showed a minor further reduction in growth and sporulation on inulin and sucrose (Fig. 4) indicating AmyR and InuR have independent functions.

gene	orientation	start	5' sequence 3'									
<i>sucA</i>	-	-385	C	C	G	G	AGGTTTG	T	C	G	G	T -400
An15g04060	-	-637	A	G	G	G	ATGATCC	T	C	G	G	A -652
An15g00310	+	-600	C	C	G	G	AGGATCT	T	C	G	G	C -585
<i>inuE</i>	-	-624	G	C	G	G	AGAAACA	T	C	G	G	G -639
<i>inuE</i>	-	-429	A	C	G	G	ACCTCAT	T	C	G	G	G -444
An17g01710	-	-5	G	T	G	G	ATATGAC	T	C	G	G	G -20
<i>inuA</i>	+	-197	C	G	G	G	AGAAATT	T	C	G	G	G -182
<i>sucA</i>	+	-243	C	G	G	G	GGAAATC	T	C	G	G	G -228
An15g04060	-	-805	T	C	G	G	GTTGACC	T	C	G	G	G -820
An15g03940	+	-568	T	C	G	G	GATGATC	T	C	G	G	C -553
An17g01710	+	-393	T	C	G	G	CTGCCCG	A	C	G	G	C -378
An15g03940	+	-396	T	C	G	G	GCACTGA	G	T	G	G	T -411
<i>sucA</i>	-	-161	T	C	G	G	GGAAATT	G	A	G	G	A -176
An15g04060	-	-421	T	C	G	G	TTGGTTT	G	G	G	G	T -436
<i>inuE</i>	-	-569	T	C	G	G	TATACCC	C	T	G	G	T -554

Fig. 7. In silico prediction of putative InuR binding sites in the promoter of sucrose induced genes. Alignment of conserved promoter regions containing a perfect CGG-CGG direct repeat are shown in bold. Numbers beside the sequences indicate the position relative to the ATG of the downstream located gene. Note that the elements in the *inuE*, *sucA* (first element), An15g04060, and An17g01710 (second element) are presented in the reversed orientation.

Table 1. Description of sucrose induced and InuR dependent genes in *A. niger* at 2 h after transfer from preculture

ORF no.	N402 sucrose	N402 xylose	<i>inuR</i> sucrose	Fold induction sucrose vs xylose	Fold induction N402 vs Δ <i>inuR</i>	P value	Function
Carbohydrate transport and metabolism							
An08g11070	103.40 $\pm$ 21.54 ^a	P 0.11 $\pm$ 0.01	A 0.26 $\pm$ 0.45	A 930.3	391.1	0.008	invertase
An15g04060	36.82 $\pm$ 1.36	P 0.18 $\pm$ 0.03	A 0.29 $\pm$ 0.05	A 204.2	126.6	0.009	putative fructose facilitator
An12g08280	145.30 $\pm$ 27.90	P 1.81 $\pm$ 0.28	P 1.97 $\pm$ 0.21	P 80.3	73.8	0.008	exo-inulinase
An11g03200	29.27 $\pm$ 0.80	P 0.63 $\pm$ 0.23	A 0.74 $\pm$ 0.06	A 46.2	39.6	0.007	endo-inulinase
An15g00310	20.64 $\pm$ 2.95	P 1.34 $\pm$ 0.17	A 1.76 $\pm$ 0.97	P 15.4	11.7	0.018	putative monosaccharide transporter
An15g03940	167.30 $\pm$ 1.20	P 65.74 $\pm$ 3.04	P 61.68 $\pm$ 47.66	P 2.5	2.7	0.027	putative monosaccharide transporter
Amino acid transport and metabolism							
An04g01340	44.31 $\pm$ 3.39	P 16.25 $\pm$ 1.27	P 10.16 $\pm$ 14.17	P 2.7	4.4	0.041	putative asparagine synthase
An14g02720	7.75 $\pm$ 0.04	P 2.37 $\pm$ 0.17	P 2.88 $\pm$ 1.17	P 3.3	2.7	0.037	putative neutral amino acid permease
Lipid transport and metabolism							
An08g03670	10.76 $\pm$ 0.09	P 4.45 $\pm$ 0.14	P 3.91 $\pm$ 2.36	P 2.4	2.8	0.015	putative ethanolamine kinase
An11g06190	2.42 $\pm$ 0.14	P 1.16 $\pm$ 0.07	P 0.97 $\pm$ 0.19	A 2.1	2.5	0.039	putative glycerophosphoinositol transporter
Phosphate transport and metabolism							
An02g00180	8.04 $\pm$ 0.06	P 2.85 $\pm$ 0.03	P 0.50 $\pm$ 0.22	A 2.8	16.2	0.001	putative phosphate-repressible phosphate permease
Nucleotide transport and metabolism							
An02g09960	41.96 $\pm$ 4.97	P 8.61 $\pm$ 0.73	P 16.54 $\pm$ 7.61	P 4.9	2.5	0.033	putative phosphoribosyl(formylglycinamide) synthase
Other transport mechanism							
An17g01710	54.97 $\pm$ 15.19	P 2.83 $\pm$ 0.65	P 5.71 $\pm$ 10.10	P 19.4	9.6	0.049	putative multidrug transporter
An08g02330	26.99 $\pm$ 4.40	P 5.73 $\pm$ 0.31	P 12.12 $\pm$ 0.63	P 4.7	2.2	0.024	putative ABC multidrug transporter
An17g00240	40.24 $\pm$ 3.78	P 12.14 $\pm$ 0.61	P 19.54 $\pm$ 8.12	P 3.3	2.1	0.038	putative ABC transporter
Transcription							
An02g05460	9.25 $\pm$ 1.05	P 2.60 $\pm$ 0.23	A 4.26 $\pm$ 0.83	M 3.6	2.2	0.040	putative DNA-directed RNA polymerase
An02g12610	3.81 $\pm$ 0.25	P 1.55 $\pm$ 0.02	A 1.21 $\pm$ 0.89	P 2.5	3.2	0.043	putative Ran-GTPase-activating protein
Protein synthesis							
An02g12640	48.21 $\pm$ 6.42	P 9.16 $\pm$ 1.17	P 18.71 $\pm$ 1.32	P 5.3	2.6	0.041	putative RNA-binding protein
An17g01815	55.05 $\pm$ 1.42	P 15.21 $\pm$ 0.37	P 24.59 $\pm$ 6.90	P 3.6	2.2	0.004	putative translation initiation factor
An01g06040	10.17 $\pm$ 2.79	P 3.21 $\pm$ 0.04	P 3.88 $\pm$ 0.04	P 3.2	2.6	0.025	putative 60S ribosome biogenesis protein
An02g14080	23.20 $\pm$ 0.69	P 9.94 $\pm$ 0.25	P 10.87 $\pm$ 7.87	P 2.3	2.1	0.010	putative asparaginyl-tRNA synthetase
Protein fate							
An11g04980	36.47 $\pm$ 1.45	P 9.59 $\pm$ 1.06	P 12.82 $\pm$ 0.53	P 3.8	2.8	0.010	putative importin unit beta-4
Unclassified							
An01g00340	9.56 $\pm$ 0.97	P 2.26 $\pm$ 0.42	P 3.13 $\pm$ 0.32	P 4.2	3.1	0.036	unknown protein
An15g03150	13.42 $\pm$ 0.68	P 4.01 $\pm$ 0.16	P 4.02 $\pm$ 0.26	P 3.4	3.3	0.008	unknown protein
An07g02690	5.27 $\pm$ 0.03	P 1.69 $\pm$ 0.02	A 2.22 $\pm$ 0.27	P 3.1	2.4	0.001	unknown protein
An01g05500	3.77 $\pm$ 0.06	P 1.25 $\pm$ 0.03	P 1.46 $\pm$ 0.35	P 3.0	2.6	0.005	unknown protein

^a The expression level was based on the geometric mean value of the duplicate samples and the deviation values between the duplicate samples are indicated. P, M or A, representing detection calls for present, marginal or absent respectively. The P-value was based on 1-way ANOVA analysis. Genes were divided over different functional groups according to FunCat (Ruepp et al. 2004).

Discussion

We previously showed that the genes encoding extracellular inulinolytic enzymes in *A. niger* are co-regulated at the transcription level and identified sucrose, but not fructose or glucose as an efficient inducer of inulinolytic genes. It was proposed that there might be a common transcription factor to activate the expression of inulinolytic genes in response to the presence of inducing carbon sources (Yuan et al. 2006), similar to that which has been found in other carbohydrate degrading enzymes networks such as the amylolytic genes (AmyR) or the xylanolytic genes (XlnR). In this study we reported the successful identification of an inulinolytic transcriptional activator, InuR.. Previous studies to isolate the AmyR and XlnR transcription factors were based on genetic screens to identify mutants with impaired expression of amylolytic and xylanolytic enzymes (Petersen et al. 1999; Gomi et al. 2000; van Peij et al. 1998a). Here, we used a different approach which made use of the full genome sequence of *A. niger*. The approach is based on the observation that in some cases, e.g. in the case of the AmyR transcription factor, the target genes of the transcription factor are adjacent to the transcription factor itself. In the case of the inulinolytic genes, three possible inulinolytic gene clusters were identified, each of which contains a gene encoding a putative transcription factor (Fig. 1). Disruption mutants were successfully obtained for two of the three transcription factors and subsequent phenotypic analysis indicated that only InuR (An15g00300) was involved in the inulin utilization pathway. The Δ *inuR* mutant showed severe reduced growth on agar plates and in submerged cultures containing sucrose or inulin as sole carbon source. Furthermore, we showed that the induction of the inulinolytic genes by inulin and sucrose is dependant on the presence of InuR. In addition to the transcription factor itself the inulinolytic gene cluster identified in *A. niger* consists of a sugar transporter and *sucB* which encodes an intracellular invertase with transfructosylating activity (Goosen et al. 2007).

To examine whether the position and clustering of the InuR orthologs was conserved among aspergilli, we compared the genes adjacent to the InuR homologous transcription factors in *A. nidulans*, *A. oryzae* and *A. fumigatus*. The clustering of the three genes, *inuR* (An15g00300), An15g00310 (putative sugar transporter) and *sucB* (An15g00320) in *A. niger* was conserved in *A. nidulans* (AN3835; e-value 5e-144, AN3836; e-value 7e-158 and AN3837; e-value 5e-142), and *A. oryzae* (AO090701000036; e-value 0.0, AO090701000037; e-value 0.0 and AO090701000038; e-value 0.0, respectively). In *A. fumigatus*, the InuR orthologue (Afu6g05010; e-value 0.0) is located next to the SucB orthologue (Afu6g05000; e-value 0.0), but no homolog to the putative sugar transporter (An15g00310) was present in the gene cluster in *A. fumigatus*. The initial gene model of the *inuR* gene in the *A. niger* annotation was lacking the ZnII(2)Cys(6) binuclear cluster domain. Recent analysis of the intron/exon positions in N-terminal region resulted in the prediction of a 709 amino acid long protein containing a complete Zn(2)Cys(6) binuclear cluster domain. BlastP searches with the *A. niger* InuR protein revealed orthologous genes in the three *Aspergillus* species examined (*A. nidulans*, *A. oryzae* and *A. fumigatus*). The closest homologs in *A. oryzae* and *A. nidulans* did not contain a Zn(II)₂Cys₆ domain and the *A. fumigatus* homolog only a

partial ZnII(2)Cys6 cluster. The gene models of the other aspergilli were manually inspected and by improving the gene model a Zn(II)2Cys6 domain was found to be present in the proteins. Interestingly, it turned out that the gene model of *A. niger* was completely conserved in the other aspergilli. Thus, the gene model of the *A. niger inuR* gene could be used as a template for the prediction of the *inuR* gene models in the other aspergilli. Both the numbers of introns as well as the intron/exon positions were conserved among the four Aspergilli. This result implies that gene model prediction in aspergilli can be improved by comparison of the gene models of orthologous genes.

The *inuR* gene is situated in the genome next to a putative sugar transporter gene (An15g00310) and *sucB*. We recently showed that the *sucB* gene encodes an intracellular invertase with transglycosylating activity (Goosen et al. 2007). As depicted in Fig. 1, the putative sugar transporter (An15g00310) in the cluster shares the same promoter region with the *sucB* gene. The Northern analysis and microarray analysis showed that the two genes are not coregulated indicating that the common promoter region does not act bidirectionally as has been observed e.g. in the case of the *niiA* and *niaD* genes (Punt et al. 1995). Expression of the gene An15g00310 was highly induced by sucrose or inulin and the induction was dependent on InuR (Fig. 3), while the *sucB* gene was lowly expressed on various carbon sources and hardly induced on sucrose and inulin. The function of An15g00310 remains to be elucidated but it is tempting to suggest that it involved in the transport of sucrose, or its hydrolysis products fructose or glucose.

The function of SucB in relation to inulin or sucrose metabolism is also currently unknown. We showed that deletion of the *sucB* did not result in a phenotype related to growth on inulin or sucrose (Goosen et al. 2007). The constitutive low level of expression of the *sucB* gene might suggest that SucB acts as a scouting enzyme to 'sense' the presence of a useful carbon source by generating an inducer molecule which can further activate transcription of additional inulinolytic genes. A possible explanation for the lack of phenotype of the *sucB* deletion mutant is the presence of an additional intracellular invertase, encoded by the *sucC* gene. Both the enzymatic activity of the SucC protein as well as the effect of deletion of *sucC* and simultaneously deletion of *sucB* and *sucC* will be the subject of future research.

The genome wide transcription analysis using Affymetrix microarrays identified a limited number of genes that were strongly induced by sucrose in an InuR dependent way. These genes encode the extracellular inulinolytic enzymes and four genes encoding putative transporters. The specificity of the sugar transporters should be determined biochemically as has been done previously for *A. niger* sugar transporters (vanKuyk et al. 2004) since phylogenetic comparison to suggest possible substrates is not reliable (Patricia vanKuyk, pers. communication). The limited number of genes that are induced by sucrose in an InuR dependant way suggests a relatively simple regulatory mechanism regarding sucrose catabolism in *A. niger*. Unpublished additional microarray data showed that the same three extracellular enzymes were induced on inulin in an *inuR* dependant way. The expression of the putative sugar transporters on inulin differed from the pattern observed on sucrose.

Two of the four transporters (An17g01710 and An15g04060) were not induced on inulin after 2 and 8 hours of growth on inulin. The transporter encoded by the An15g03940 gene was 4 to 5 fold induced on inulin compared to xylose after 8 hours of the transfer. Expression of this sugar transporter is also high on xylose (Table 1) and maltose (Yuan et al. unpublished results), indicating that the expression of this sugar transporter is not very tightly regulated and suggest a broad specificity of the sugar transporter. The fourth transporter (An15g00310) was more highly expressed on inulin 8 hours after the transfer than on maltose and xylose (4 to 5 fold induction), and the induction was InuR dependant. The difference between the observed expression patterns of the extracellular enzymes and the putative sugar transporters suggest that under the different growth conditions the fungus expresses the same set of extracellular enzymes, but that the fungus fine-tunes the expression of the sugar transporters in order to allow efficient uptake. Such a fine-tuning mechanism could explain why such a large number of sugar transporters is present in the genome of *Aspergillus niger* as 461 proteins have been identified that belong to the Major Facilitator Superfamily of transporters (Pel et al. 2007). Such a mechanism allows the fungus to express the sugar transporters with the optimal properties in relation to substrate diversity, substrate concentration, environmental pH, and other environmental factors and nutrient conditions.

The InuR transcription factor is a member of the Zn(II)2-Cys6 DNA transcription factor family. It is a fungal specific type of transcription factor. In aspergilli, this family of transcription factors has expanded significantly when genome data is compared to that of the yeast *S. cerevisiae* or *Neurospora crassa*. In *S. cerevisiae* and *N. crassa* 54 and 89 genes respectively (Akache et al. 2001; Borkovich et al. 2004), have been identified that contain the Zn(II)2-Cys6 DNA binding domain, the genomes of aspergilli contain 216 (*A. nidulans*), 188 (*A. fumigatus*) or 296 (*A. niger*) Zn(II)2-Cys6 DNA transcription factors. The actual number of proteins containing a Zn(II)2-Cys6 DNA binding domain might still increase in aspergilli and other fungi. We showed that the predicted gene model for the *inuR* gene in several aspergilli was incorrect, resulting in a protein lacking the Zn(II)2-Cys6 DNA binding domain. Closer examination of the *A. nidulans* and *A. niger* genome annotations revealed that the difficulty of the automatic annotation in predicting the correct gene model for this class of transcription factors is not limited to *inuR* gene, but is a more general problem. Manual improvement of the gene models identified 326 Zn(II)2-Cys6 DNA binding domain containing proteins in *A. nidulans* and 480 Zn(II)2-Cys6 DNA binding domain containing proteins in *A. niger* (Ram and Punt, unpublished data). Thus, it is likely that other fungal genomes contain more as yet unidentified Zn(II)2-Cys6 DNA binding domain containing proteins, which could be revealed with improved gene models or manual annotation.

Transcription factors containing the Zn(II)2-Cys6 DNA binding domain usually bind the promoter region of their target genes via a conserved DNA binding site. Often, but not always, the binding site consists of conserved terminal trinucleotides, usually in symmetrical configuration (direct repeat, everted repeat or an inverted repeat) and spaced by an internal variable sequence of a defined length. Examples are the Gal4p and Lac9p transcription factors that bind to CCGN11CCG, Ppr1p and UaY that bind to CCGN6CGG,

and Put3 that binds to CGGN10CCG (Todd and Andrianopoulos, 1997). The binding sites for the AmyR transcription factor have been studied in detail. Several promoter regions of amylolytic genes contain a CGGN8(A/C)GG site in their promoter region to which AmyR can bind and activate transcription (Tani et al. 2001b; Ito et al. 2004).

Comparison of the 1-kb upstream promoters of the strongly InuR dependent sucrose induced genes revealed that a common sequence (C/G)GGN8CGG is present in the promoters of all genes. The proposed InuR binding site is very similar to the AmyR binding site which is present in the promoter of many amylolytic genes (Yuan et al. unpublished results) and raises immediately the question of how the specificity of the InuR and AmyR transcription factors is accomplished. After establishing that the conserved (C/G)GGN8CGG motif in the promoter of inulinolytic genes is involved in the binding and activation of expression of the target genes, the AmyR and the InuR transcription factors are interesting subjects for further research aimed to understand how Zn(II)2-Cys6 transcription factors recognize their target genes and what determines the specificity of the binding. Analysis of the $\Delta amyR \Delta inuR$ double mutant indicated that both transcription factors have independent functions. The slightly decreased growth and sporulation of the $\Delta amyR \Delta inuR$ double mutant on sucrose compared to growth of the $\Delta inuR$ mutant on sucrose might be explained by the activity of α -glucosidases that are under control of the AmyR transcription factor. It is well established that sucrose can be hydrolyzed either by a β -fructosidase, that splits sucrose by attacking the terminal, non-reducing $\alpha 1, \beta 2$ linked fructose moiety, but also by an α -glucosidase catalyzing the hydrolysis of sucrose at the terminal non-reducing $\alpha 1-4$ linked glucose residue (Geber et al. 1992; Kelly et al. 1992; Alberto et al. 2004). *A. niger* contains several α -glucosidase encoding gene of which two of them are under control of the AmyR transcription factor. Disruption of the *amyR* transcription factor further decreases the level of *agdA* and *agdB* transcripts (Yuan et al. unpublished results), which might result in lower α -glucosidase activity and a subsequent decrease in the efficiency by which sucrose can be hydrolyzed.

In conclusion, we have identified InuR as the transcriptional activator required for utilization of inulin and sucrose. The results in this paper suggest that *A. niger* metabolizes inulin and sucrose by inducing the expression of a limited number of extracellular enzymes and sugar transporters that are under transcriptional control of a single transcription factor, InuR. The mechanism by which the InuR transcription factor is activated in response to the presence of inulin and sucrose and the mechanism that determines the specificity of the binding to the InuR target genes are the subjects of future research.

Acknowledgements.

We thank Dr. H. Roubos (DSM, the Netherlands) for performing promoter searches, Dr. J. Visser and Dr. P. vanKuyk for helpful discussions and comments on the manuscript. This work was financially supported by SenterNovem in the framework of the IOP (Innovative Onderzoeks Programma)-Genomics program (project IGE1021).

Supplementary Data

Suppl. Table 1. Primers used in this study

Primer name	5'Sequence 3'
<i>inuRP1</i>	ATAAGAAATGCGGCGCATCTGCTAGTGC GG GGCCTA
<i>inuRP2</i>	CCGCTCGAGTTG CAGCGGCACCAGAAT
<i>inuRP3</i>	CCGCTCGAGCGGTTGGATCCAAGGACAACGCTTACGGACT
<i>inuRP4</i>	GGGTTACCCATGCATGCCGGCAGTATG
<i>inuRP5</i>	TTCGATGCCATAGCGTGGT
<i>inuRP6</i>	GCTCCACGGCTGAACACT
<i>inuRP7</i>	CAATGGTACCAGCTGGTTCCTAGAGCACGC
<i>inuRP8</i>	GAACCTAGACCAAGCAACCGCTGTGAT
pAO-9	AATGTCAATTCCAGCAGCG
pAO-10	TTCGCGAGACTGAATGCGG
An15g00310-P1	TGCGTTCTGTTTCTGGTCG
An15g00310-P2	TGGAGGCTTTGGAGACTGG

^aRestriction sites that were used for cloning are underlined.

Suppl. Table 2. Description of sucrose induced genes in *A. niger* at 2 h after transfer from a xylose preculture.

Order no.	ORF no.	N402 surose	N402 xylose	Fold change sucrose vs xylose	Description		
1	An08g11070	103.40	P	0.11	A	930.3	beta-fructofuranosidase precursor suc1 - <i>Aspergillus niger</i>
2	An15g04060	36.82	P	0.18	A	204.2	strong similarity to transport protein Flr1 - <i>Saccharomyces cerevisiae</i>
3	An12g08280	145.30	P	1.81	P,M	80.3	exo-inulinase inu1 - <i>Aspergillus niger</i>
4	An01g08010	8.21	P	0.13	A	63.7	strong similarity to iron and manganese transporter CCC1 - <i>Saccharomyces cerevisiae</i>
5	An11g03200	29.27	P	0.63	A	46.2	inulinase inuA - <i>Aspergillus niger</i>
6	An17g01710	54.97	P	2.83	P	19.4	similarity to multidrug resistance protein QDR1 - <i>Saccharomyces cerevisiae</i>
7	An01g09040	6.83	P	0.39	A	17.5	strong similarity to DEAD box protein MrDb - <i>Homo sapiens</i>
8	An15g00310	20.64	P	1.34	A	15.4	strong similarity to monosaccharide transporter Mst-1 - <i>Amanita muscaria</i>
9	An01g07300	4.84	P	0.88	A	5.5	strong similarity to gluconokinase gntV - <i>Escherichia coli</i>
10	An02g12640	48.21	P	9.16	P	5.3	strong similarity to nuclear protein gar2 - <i>Schizosaccharomyces pombe</i>
11	An02g14950	19.68	P	3.79	A	5.2	strong similarity to ATP-dependent RNA helicase CA3 - <i>Saccharomyces cerevisiae</i>
12	An02g13890	7.16	P	1.39	A	5.2	weak similarity to calmodulin- and actin-binding protein h-caldesmon - <i>Gallus gallus</i>
13	An02g09960	41.96	P	8.61	P	4.9	strong similarity to phosphoribosylformylglycinamide synthase purL - <i>Escherichia coli</i>
14	An08g02330	26.99	P	5.73	P	4.7	strong similarity to multidrug resistance protein MLP-2 - <i>Rattus norvegicus</i>
15	An08g02300	55.96	P	12.73	P	4.4	weak similarity to enniatin synthetase - <i>Fusarium scripi</i> [truncated ORF]
16	An01g00340	9.56	P	2.26	P	4.2	strong similarity to hypothetical protein SPCC320.08 - <i>Schizosaccharomyces pombe</i>
17	An17g01470	8.89	P	2.26	A	3.9	similarity to 49 kD subunit of DNA-directed RNA polymerase I RPA49 - <i>Saccharomyces cerevisiae</i>
18	An11g04980	36.47	P	9.59	P	3.8	similarity to RAN-binding protein KAP123 - <i>Saccharomyces cerevisiae</i> [truncated ORF]
19	An05g01310	0.51	P	0.14	A	3.8	hypothetical protein
20	An08g07420	34.17	P	9.11	P	3.7	strong similarity to multifunctional pyrimidine biosynthesis enzyme pyrABCN - <i>Aspergillus nidulans</i>
21	An15g06390	11.00	P	2.97	P	3.7	strong similarity to ribosomal RNA processing protein RRP5 - <i>Saccharomyces cerevisiae</i>
22	An02g09500	18.54	P	5.02	P	3.7	similarity to integral membrane protein Pth11 - <i>Magnaporthe grisea</i> strain G-11
23	An17g01815	55.05	P	15.21	P	3.6	strong similarity to translation initiation factor e1f6 - <i>Saccharomyces cerevisiae</i>
24	An02g05460	9.25	P	2.60	A	3.6	strong similarity to 40 kD subunit of DNA-directed RNA polymerase I, III RPA40 - <i>Saccharomyces cerevisiae</i>
25	An08g00720	6.52	P	1.85	A	3.5	strong similarity to conserved eukaryotic protein involved in ribosome biogenesis ERB1 - <i>Saccharomyces cerevisiae</i>
26	An15g03150	13.42	P	4.01	P	3.4	strong similarity to hypothetical protein SPCC1183.11 - <i>Schizosaccharomyces pombe</i>
27	An04g02000	11.07	P	3.32	P	3.3	strong similarity to hypothetical microtubule-interacting protein YTM1 - <i>Saccharomyces cerevisiae</i>
28	An17g00240	40.24	P	12.14	P	3.3	strong similarity to hypothetical protein YER036c - <i>Saccharomyces cerevisiae</i>
29	An14g02720	7.75	P	2.37	P	3.3	strong similarity to neutral amino acid permease nit - <i>Neurospora crassa</i>
30	An14g01560	35.09	P	10.78	P	3.3	strong similarity to zootin ZUO1 - <i>Saccharomyces cerevisiae</i>
31	An16g08030	7.64	P	2.35	P	3.2	strong similarity to ovarian tissue derived protein 18 of patent DE19816395-A1 - <i>Homo sapiens</i>
32	An02g05650	4.00	P	1.24	A	3.2	strong similarity to hypothetical protein YPL217c - <i>Saccharomyces cerevisiae</i>
33	An01g06040	10.17	P	3.21	P	3.2	strong similarity to probable membrane protein YQL077c - <i>Saccharomyces cerevisiae</i>
34	An01g00150	17.88	P	5.64	P	3.2	strong similarity to nuclear protein ENP1 - <i>Saccharomyces cerevisiae</i>
35	An11g01300	11.51	P	3.66	P	3.1	strong similarity to protein involved in ribosome assembly RSA2 - <i>Saccharomyces cerevisiae</i>
36	An07g02690	5.27	P	1.69	A	3.1	similarity to hypothetical regulatory protein PBK1 - <i>Homo sapiens</i>
37	An01g06330	4.57	P	1.48	A	3.1	similarity to transcriptional activator UPC2 - <i>Saccharomyces cerevisiae</i>
38	An08g03460	4.74	P	1.54	P,M	3.1	strong similarity to protein involved in silencing Sas10p - <i>Saccharomyces cerevisiae</i>
39	An02g00890	21.40	P	6.99	P	3.1	strong similarity to phosphoserine transaminase SER1 - <i>Saccharomyces cerevisiae</i>
40	An01g05500	3.77	P	1.25	P	3.0	strong similarity to hypothetical trp asp repeat protein - <i>Schizosaccharomyces pombe</i>
41	An18g04230	9.85	P	3.36	P	2.9	strong similarity to a protein involved in ribosome biogenesis Kri1 - <i>Saccharomyces cerevisiae</i>
42	An08g07050	26.14	P	9.13	P	2.9	strong similarity to 189 kD subunit of DNA-directed RNA polymerase I RPA190 - <i>Schizosaccharomyces pombe</i>
43	An02g00180	8.04	P	2.85	P	2.8	strong similarity to phosphate-repressible phosphate permease pho-4 - <i>Neurospora crassa</i>
44	An08g00480	29.56	P	10.63	P	2.8	strong similarity to hypothetical protein BAB23171.1 - <i>Mus musculus</i>
45	An04g06120	2.68	P	0.97	A	2.8	strong similarity to hypothetical protein YDR398w - <i>Saccharomyces cerevisiae</i>
46	An16g08640	4.77	P	1.73	P	2.8	strong similarity to hypothetical RNA-binding protein SPAC16E8.06c - <i>Schizosaccharomyces pombe</i>
47	An04g01340	44.31	P	16.25	P	2.7	strong similarity to asparagine synthase An1 - <i>Saccharomyces cerevisiae</i>
48	An15g00040	10.31	P	3.79	P	2.7	similarity to hypothetical protein YBL028c - <i>Saccharomyces cerevisiae</i>
49	An09g03230	2.95	P	1.09	P	2.7	similarity to integral membrane protein Pth11 - <i>Magnaporthe grisea</i> strain G-207
50	An01g03760	8.04	P	2.99	P,A	2.7	strong similarity to hypothetical protein pI029 - <i>Schizosaccharomyces pombe</i>
51	An09g05310	9.30	P	3.52	A	2.6	strong similarity to GTP cyclohydrolase I FOL2 - <i>Saccharomyces cerevisiae</i>
52	An01g10270	4.22	P	1.65	P,A	2.6	similarity to hypothetical protein YOR243c - <i>Saccharomyces cerevisiae</i>
53	An05g00960	5.64	P	2.20	P	2.6	strong similarity to U3 small nucleolar ribonucleoprotein imp3 - <i>Saccharomyces cerevisiae</i>
54	An15g03940	167.30	P	65.74	P	2.5	strong similarity to monosaccharide transporter Mst-1 - <i>Amanita muscaria</i>
55	An09g06310	5.01	P	1.99	A	2.5	similarity to hypothetical protein YPR143w - <i>Saccharomyces cerevisiae</i>
56	An18g06360	8.84	P	3.52	P	2.5	similarity to the mycelial surface antigen Csa1 - <i>Candida albicans</i>
57	An16g03090	17.82	P	7.13	P	2.5	strong similarity to small nucleolar RNP component NOP56 - <i>Saccharomyces cerevisiae</i>
58	An02g03110	11.69	P	4.69	P	2.5	similarity to mitochondrial ribosomal protein of the large subunit Yml4 - <i>Saccharomyces cerevisiae</i>
59	An02g03520	15.52	P	6.27	P	2.5	strong similarity to G protein-binding protein CRFG - <i>Homo sapiens</i>
60	An09g06050	28.79	P	11.64	P	2.5	strong similarity to S-adenosylmethionine decarboxylase spe-2 - <i>Neurospora crassa</i>

Order no.	ORF no.	N402 sucrose	N402 xylose	Fold change sucrose vs xylose	Description
61	An07g02650	150.80	P	60.99	P 2.5 strong similarity to translation elongation factor 3 YEF3 - Saccharomyces cerevisiae
62	An02g12610	3.81	P	1.55	A 2.5 strong similarity to Ran-GTPase-activating protein 1 RNA1 - Schizosaccharomyces pombe
63	An07g03670	5.41	P	2.20	P 2.5 strong similarity to plasma membrane receptor SYG1 - Saccharomyces cerevisiae
64	An08g03670	10.76	P	4.45	P 2.4 strong similarity to ethanolamine kinase EK1 - Homo sapiens
65	An04g00680	13.45	P	5.59	P 2.4 similarity to nucleolar protein gar2 - Schizosaccharomyces pombe
66	An02g10390	23.75	P	9.99	P 2.4 strong similarity to eukaryotic translation initiation factor eIF-4A II - Mus musculus
67	An02g02530	37.14	P	15.64	P 2.4 strong similarity to ATP-dependent RNA helicase p68 - Homo sapiens
68	An08g08860	5.57	P	2.38	P 2.3 weak similarity to period protein per - Drosophila suxiae
69	An02g14080	23.20	P	9.94	P 2.3 strong similarity to asparagine--tRNA ligase ASNS - Thermus aquaticus sspc.thermophilus
70	An14g03280	34.65	P	14.90	P 2.3 strong similarity to dihydroxy-acid dehydratase ILV-3 - Saccharomyces cerevisiae
71	An15g00030	6.10	P	2.63	P 2.3 strong similarity to member of the major facilitator superfamily LAS21 - Saccharomyces cerevisiae
72	An07g06520	20.29	P	8.78	P 2.3 strong similarity to heat shock protein 70 homolog pdrl3 - Saccharomyces cerevisiae
73	An12g08670	5.53	P	2.49	P 2.2 similarity to the 3 - 5 exonuclease involved in RNA processing Rrp42 - Saccharomyces cerevisiae
74	An14g03660	22.69	P	10.34	P 2.2 strong similarity to the protein fragment SEQ ID NO: 22357 of patent EP1033405-A2 - Arabidopsis thaliana
75	An08g11120	7.56	P	3.53	P 2.1 weak similarity to protein PR04400 of patent WO20073348-A2 - Homo sapiens
76	An05g02250	11.83	P	5.56	P 2.1 similarity to hypothetical serine-threonine rich protein - Schizosaccharomyces pombe
77	An09g06180	6.43	P	3.04	P,A 2.1 similarity to phosphorylation regulatory protein HP-10 - Homo sapiens
78	An01g08850	136.50	P	64.55	P 2.1 strong similarity to Gbeta like protein CpcB - Aspergillus nidulans
79	An07g02960	186.10	P	88.28	P 2.1 strong similarity to cytoplasmic ribosomal protein of the small subunit S13.e - Saccharomyces cerevisiae
80	An07g04240	8.35	P	3.97	P 2.1 strong similarity to low-molecular-weight GTP/GDP-binding protein krev-1 - Neurospora crassa
81	An07g07980	55.63	P	26.43	P 2.1 strong similarity to SV2 related protein SVOP - Rattus norvegicus
82	An16g00850	3.16	P	1.51	P,A 2.1 strong similarity to hypothetical protein SPAC9E9.15 - Schizosaccharomyces pombe
83	An08g05500	33.80	P	16.13	P 2.1 strong similarity to mitochondrial ribosomal protein of the large subunit Yml3 - Saccharomyces cerevisiae
84	An16g07040	135.70	P	64.90	P 2.1 similarity to beta-1,3-glucanotransferase BGT1 - Aspergillus fumigatus [truncated ORF]
85	An02g09260	12.09	P	5.79	P 2.1 strong similarity to nucleolar protein Nop5 - Saccharomyces cerevisiae
86	An16g06190	2.42	P	1.16	P,A 2.1 strong similarity to glycerophosphoinositol transporter GIT1 - Saccharomyces cerevisiae
87	An08g07030	4.08	P	1.97	A 2.1 similarity to hypothetical protein YNL124w - Saccharomyces cerevisiae
88	An02g00840	27.46	P	13.34	P 2.1 similarity to nonribosomal peptide synthetase MacC5 - Stigmatella aurantiaca
89	An16g08920	5.56	P	2.70	A 2.1 similarity to the neurofilament triplet H protein NEFH - Homo sapiens
90	An04g01570	11.10	P	5.41	P 2.1 strong similarity to helicase SKI2W - Homo sapiens
91	An07g09120	9.68	P	4.72	P 2.1 strong similarity to the hypothetical protein encoded by B23L21.370 - Neurospora crassa
92	An14g02470	37.30	P	18.39	P 2.0 strong similarity to the protein PR0304 of patent WO200104311-A1 - Homo sapiens
93	An18g05810	173.20	P	85.68	P 2.0 strong similarity to cytoplasmic ribosomal protein of the small subunit S26 - Homo sapiens
94	An18g03300	22.27	P	11.04	P 2.0 strong similarity to cytosolic phenylalanine--tRNA ligase alpha subunit Frs1 - Saccharomyces cerevisiae
95	An04g03800	8.63	P	4.28	P 2.0 weak similarity to hypothetical protein CAB51071.1 - Homo sapiens
96	An01g09180	9.89	P	4.93	P 2.0 strong similarity to ISW2 - Saccharomyces cerevisiae
97	An07g06510	4.77	P	2.38	P 2.0 strong similarity to the protein involved in ribosomal RNA processing Rrp40 - Saccharomyces cerevisiae

Suppl. Table 3. Description of InuR regulated genes in *A. niger* at 2 h after transfer from a xylose preculture.

Order no.	ORF no.	N402 sucrose	Δ inuR sucrose	Fold change N402 vs Δ inuR	Description
1	An08g11070	103.40	P	0.26	A 391.1 beta-fructofuranosidase precursor suc1 - Aspergillus niger
2	An15g04060	36.82	P	0.29	A 126.6 strong similarity to transport protein Flr1 - Saccharomyces cerevisiae
3	An12g08280	145.30	P	1.97	P,M 73.8 exo-inulinase inu1 - Aspergillus niger
4	An11g03200	29.27	P	0.74	A 39.6 inulinase inuA - Aspergillus niger
5	An02g00180	8.04	P	0.50	A 16.2 strong similarity to phosphate-repressible phosphate permease pho-4 - Neurospora crassa
6	An15g00310	20.64	P	1.76	P 11.7 strong similarity to monosaccharide transporter Mst-1 - Amanita muscaria
7	An17g01710	54.97	P	5.71	P 9.6 similarity to multidrug resistance protein QDR1 - Saccharomyces cerevisiae
8	An02g06910	10.09	P	1.85	P,A 9.6 strong similarity to probable membrane protein YLI108w - Saccharomyces cerevisiae
9	An01g14740	11.11	P	2.05	A 5.4 glucose oxidase precursor goxC - Aspergillus niger [putative sequencing error]
10	An15g00300	2.09	P	0.39	A 5.4 similarity to transcriptional activator amyr - Aspergillus oryzae
11	An04g01340	44.31	P	10.16	P 4.4 strong similarity to asparagine synthase Asn1 - Saccharomyces cerevisiae
12	An08g07150	10.01	P	2.82	M 3.5 similarity to hypothetical protein mlr2056 - Mesorhizobium loti
13	An16g01610	6.53	P	1.94	A 3.4 weak similarity to gamma-butyrobetaine hydroxylase BBH - Rattus norvegicus
14	An15g03150	13.42	P	4.02	P 3.3 strong similarity to hypothetical protein SPCC1183.11 - Schizosaccharomyces pombe
15	An02g12610	3.81	P	1.21	P,A 3.2 strong similarity to Ran-GTPase-activating protein 1 RNA1 - Schizosaccharomyces pombe
16	An13g02880	10.38	P	3.31	P 3.1 strong similarity to probable integral membrane protein SPAC16A10.01 - Schizosaccharomyces pombe
17	An01g00340	9.56	P	3.13	P,A 3.1 strong similarity to hypothetical protein SPCC320.08 - Schizosaccharomyces pombe
18	An01g12890	1.64	P	0.55	A 3.0 hypothetical protein
19	An11g04980	36.47	P	12.82	P 2.8 similarity to RAN-binding protein KAP123 - Saccharomyces cerevisiae [truncated ORF]
20	An08g03670	10.76	P	3.91	P,A 2.8 strong similarity to ethanolamine kinase EK1 - Homo sapiens
21	An15g03940	167.30	P	61.68	P 2.7 strong similarity to monosaccharide transporter Mst-1 - Amanita muscaria
22	An14g02720	7.75	P	2.88	P 2.7 strong similarity to neutral amino acid permease mtr - Neurospora crassa
23	An01g06040	10.17	P	3.88	P 2.6 strong similarity to probable membrane protein YOL077c - Saccharomyces cerevisiae
24	An01g05500	3.77	P	1.46	P 2.6 strong similarity to hypothetical trp asp repeat protein - Schizosaccharomyces pombe
25	An02g06810	9.86	P	3.82	P 2.6 strong similarity to hypothetical protein YGL232w - Saccharomyces cerevisiae
26	An02g12640	48.21	P	18.71	P 2.6 strong similarity to nucleolar protein gar2 - Schizosaccharomyces pombe
27	An02g09960	41.96	P	16.54	P 2.5 strong similarity to phosphoribosylformylglycinamidine synthase purL - Escherichia coli
28	An02g03270	17.62	P	6.94	P,A 2.5 strong similarity to methylenetetrahydrofolate reductase MET13 - Saccharomyces cerevisiae
29	An01g08880	8.21	P	3.26	P 2.5 similarity to mucin-like protein Muc1 - Saccharomyces cerevisiae
30	An16g06190	2.42	P	0.97	A 2.5 strong similarity to glycerophosphoinositol transporter GIT1 - Saccharomyces cerevisiae
31	An07g02690	5.27	P	2.22	P 2.4 similarity to hypothetical regulatory protein PBK1 - Homo sapiens
32	An04g01440	18.32	P	8.13	P 2.3 strong similarity to the pepsin A 3 precursor - Homo sapiens
33	An17g01815	55.05	P	24.59	P 2.2 strong similarity to translation initiation factor 6 eIF6 - Saccharomyces cerevisiae
34	An01g1650	16.81	P	7.52	P 2.2 strong similarity to formyltetrahydrofolate deformylase purU - Corynebacterium sp.
35	An08g02330	26.99	P	12.12	P 2.2 strong similarity to multidrug resistance protein MLP-2 - Rattus norvegicus
36	An02g05460	9.25	P	4.26	M,A 2.2 strong similarity to 40 kD subunit of DNA-directed RNA polymerase I,II,III RPC40 - Saccharomyces cerevisiae
37	An02g14080	23.20	P	10.87	P 2.1 strong similarity to asparagine--tRNA ligase ASNS - Thermus aquaticus sspc.thermophilus
38	An09g00220	11.41	P	5.52	P 2.1 strong similarity to O-methyltransferase B (omtB) - Aspergillus flavus
39	An05g01860	6.21	P	3.01	P 2.1 strong similarity to hypothetical amidase Z37509 - Saccharomyces cerevisiae
40	An17g00240	40.24	P	19.54	P 2.1 strong similarity to hypothetical protein YER036c - Saccharomyces cerevisiae

Chapter 8

A novel screening method for isolation of mutants involved in inulin signalling in *Aspergillus niger*

Xiao-Lian Yuan, Mark Arentshorst, Cees A. M. J. van den Hondel, Arthur F. J. Ram

Abstract

Among the inulinolytic genes present in *A. niger*, we identified the *inuE* gene as the most strongly induced gene in the presence of the pathway inducing poly- and disaccharides inulin and sucrose, respectively. Using the *inuE* promoter region, we present a positive screening method for the isolation of mutants that do not respond to the presence of inulin or sucrose. Reporter strains were constructed in which the *inuE* promoter was fused to the *racA*-G12V mutant gene or to the eGFP reporter. Expression of the *racA*-G12V mutant gene from the *inuE* promoter resulted in reduced growth and reduced conidiation under *inuE* inducing conditions. Expression from the *inuE* promoter was also visible using the *PinuE*-eGFP reporter and resulted in strong GFP expression under inducing conditions. The growth phenotype of the *PinuE*-*racA*-G12V strain allows a genetic screen to identify mutants defective in the activation of inulin signalling pathway. The rationale of this screening is that a mutation in a gene which is involved in an inulin signalling pathway, e.g. the *inuR* transcription factor mutant, abolishes the induction of the *inuE* gene, resulting in normal growth and normal conidiation of the mutant strain in the presence of the inducer. The functionalities of the two reporters were examined separately and can be combined into a single reporter strain in the future to allow the selection for trans-acting mutations in the inulin signalling pathway. Whereas in the wild type strain, induction of the reporters leads to impaired growth and conidiation (*PinuE*-*racA*-G12V) or GFP expression (*PinuE*-eGFP) on inducing carbon sources, both reporters were not induced in the Δ *inuR* background. These results indicate that the two reporter systems are functioning well and can be combined to identify genes involved in an inulin signalling pathway.

Introduction

The filamentous fungus *Aspergillus niger* is well known for secreting a large spectrum of enzymes involved in the degradation of plant polysaccharides such as xylan, pectin, starch, and inulin (De Vries et al., 2003; Yuan et al., 2006; Martens-Uzunova et al., 2006). The expression of these enzymes is controlled by Zn(II)₂Cys₆ type transcription factors. For example, XlnR and AmyR activate the transcription of genes involved in xylan and starch degradation respectively (van Peij 1998; Petersen et al., 1999 and Gomi et al., 2000). A newly isolated InuR transcription factor controls the expression of inulinolytic genes in *A. niger* (Yuan et al., 2008).

Activation of Zn(II)₂Cys₆ transcription factors regulating genes encoding sugar polymer degrading enzymes is triggered by pathway inducing molecules. For example isomaltose triggers AmyR mediated amylase induction in *Aspergillus* species (Tsukagoshi et al., 2001; Kato et al., 2002). Galactose triggers activation of Gal4 mediated transcription of Gal pathway genes in *Saccharomyces cerevisiae* (Sil et al., 1999). The mechanism of action of the inducing molecules is thought to be via inducing conformational changes in the target transcription factors, which leads to a change in their activity or inter- and intramolecular interactions (Sharrocks, 2000). Obviously, activation of transcription factors is a complex

process. It requires proteins to sense the availability of the nutrients, proteins to generate an inducing molecule, proteins to transport the inducer and to promote the assembly of higher-order activator complexes. The Gal4 system is so far the best understood regulatory mechanism related to sugar metabolism. In *S. cerevisiae*, the interactions and associations of three proteins, Gal4, Gal80 and Gal3, control the regulation of Gal pathway genes. The capacity of Gal4 to activate transcription is inhibited by association of the Gal80 protein to the activation domain of Gal4 in the absence of galactose. In the presence of galactose, Gal3 associates with Gal80, relieving Gal80 inhibition of Gal4, thereby switching on the Gal4 activation system (Pilauri et al., 2005; Diep et al., 2006).

In the inulin degradation pathway in *A. niger*, sucrose was identified as a pathway inducing molecule, probably via the activation of the InuR transcription factor (Yuan et al., 2006; 2008). In this study we set up a screening method to identify genes encoding proteins that are involved in the inulin signalling pathway. The screening was based on our observation that exo-inulinase gene *inuE* is pathway specifically regulated: induced on inulin, sucrose and repressed on glucose, not expressed on fructose and xylose (Yuan et al., 2006). The *racA*-G12V is a mutant of a small GTPase involved in actin polarization. Its expression leads to cessation of cell growth (Arentshorst and Ram, unpublished results). A second reporter, eGFP, was also constructed to examine the activity of the *inuE* promoter microscopically. We reasoned that mutants defective in the induction of the *inuE* gene would grow relative normally in the presence of the inducer, while the wild-type strain expressing the *racA*-G12V mutant protein from the *inuE* promoter would have a strong growth phenotype.

The screening method was evaluated using the *inuR* disruption strain. The results demonstrate that this screening method works and can be successful for the identification of genes involved in inulin signalling. Such a screening method might also be useful for the identification of regulators from which the transcription factor has not been successful so far.

Materials and methods

Strains, culture conditions and transformations

A. niger strain N402, a *cspA1* derivative of the wild-type strain *A. niger* van Tieghem (CBS 120.49, ATCC 9029) (Bos et al., 1988) and strain AB4.1, a *pyrG* negative derivative of N402 (van Hartingsveldt et al., 1987) were used in this study. *A. niger* strains were grown in an *Aspergillus* minimal medium (MM) (Bennett and Lasure, 1991), or an *Aspergillus* complete medium (CM) consisting of a MM medium with the addition of 0.5% (w/v) yeast extract and 0.1% (w/v) casamino acids. The growth medium was supplemented with 10mM uridine (Serva) when required. For the submerged culture, *A. niger* strains were grown in MM supplemented with 1% (w/v) carbon source and 0.1% (w/v) casamino acids. The conidiospores were inoculated at a concentration of 2×10^6 spores per ml. Glucose, sucrose and raffinose (BDH chemicals), xylose, fructose and maltose (Sigma-Aldrich) and inulin

(Sensus Frutafit, Cosun) were used as carbon sources. Conidiospores were obtained by harvesting spores from CM plates after 4-6 days of growth at 30 °C, using a 0.9% NaCl solution. Transformation of *A. niger* AB4.1 was as described by Punt and van den Hondel (1992) using lysing enzymes (L1412, Sigma) for protoplast formation. The bacterial strain used for transformation and amplification of recombinant DNA was *Escherichia coli* XL1-Blue (Stratagene, La Jolla, CA). Transformation of XL1-Blue was performed according to the heat shock protocol as described by Inoue *et al.*, (1990).

Microarray experiments and data analysis

The microarray experiments and data analysis were performed as described in Chapter 7 (Yuan *et al.*, 2008).

Molecular biological techniques

Ligation of DNA fragments was performed with the Rapid DNA ligation kit (Roche). Plasmid DNA was isolated from the *E. coli* transformants using Plasmid miniprep or midiprep Kit (Qiagen). Restriction enzyme digestion (Invitrogen) was performed according to the supplier's manual. Fungal chromosomal DNA was isolated as described by Kolar *et al.* (1988). For Southern analysis, 10 µl of genomic DNA was digested with 20U of restriction enzymes at 37 °C for 3h, respectively. [α -³²P]dCTP labelled probes were synthesized using the Rediprime II DNA labelling System (Amersham Pharmacia Biotech, UK) according to the instructions of the manufacturer.

Construction of reporter plasmids and strains

The recombinant plasmid *PinuE-racA-G12V* was constructed by a three way ligation. The first fragment containing the promoter of *inuE* was amplified by PCR using primers PinuEP1f (ataagaatg**cgggccgctt**ccaccacacggacagt with the *NotI* site underlined) and PinuEP2r (catg**ccatggc**taataagaccccttgaga with the *NcoI* site underlined). The resulting PCR fragment was cloned into a *NotI-NcoI* opened pUC21 vector and verified by sequence analysis. A second fragment containing *racA-G12V-TtrpC* was isolated as an *NcoI-XbaI* fragment from plasmid pPglaA-*racA-G12V-TtrpC* (Arentshorst and Ram, unpublished vector). The third fragment containing the vector backbone was isolated as *NotI-XbaI* from pAN52-1 *NotI* (Punt, unpublished vector). The three fragments were ligated together resulting in plasmid p*PinuE-racA-G12V-TtrpC*. The resulting vector was then opened with *XbaI* to introduce a *pyrG** gene (Gouka *et al.*, 1995), isolated as a 3.8-kb *XbaI* fragment from plasmid p*pyrG*-XbaI* (Arentshorst and Ram, unpublished results) to give a final construct of p*PinuE-racA-G12V-pyrG** (named pMA61.1).

To construct plasmid *PinuE-eGFP-TtrpC*, the first fragment containing eGFP-*TtrpC* and the pUC18 backbone sequence was isolated as an *NcoI-NotI* fragment from Pgpda-H2B-GFP-*TtrpC* (Arentshorst and Ram, unpublished vector) to remove Pgpda-H2B. This opened vector was then ligated with a *NotI-NcoI* cut *PinuE* fragment to obtain the plasmid *PinuE-eGFP-TtrpC*. Finally a 3.8 kb *XbaI* fragment containing the *pyrG** gene was inserted

into *Xba*I opened plasmid *PinuE*-GFP-*TtrpC* and resulted in the final plasmid *PinuE*-GFP-*TtrpC*-*pyrG** (named pMA62.1).

The obtained final plasmids p*PinuE*-*racA*-G12V-*pyrG** and p*PinuE*-GFP-*TtrpC*-*pyrG** were independently transformed into an *A. niger pyrG*- strain AB4.1 (van Hartingsveldt et al., 1987) or an *pyrG*- derivative strain of the Δ *inuR* strain (Yuan et al., 2008). Uridine prototrophic transformants were selected by incubating protoplasts on agar plates containing MM without uridine. Transformants were screened by PCR using primers *PinuEP1f* and *SagDP4* (cgccgggaaagcggttg) for *racA*-G12V reporter strains or *pinuEp1f* and *GFPP5* (ggatgttgccgtctctctt) for GFP reporter strains. Only transformants harboring a targeted plasmid should result in the amplification of a 2.1 or 1.7 kb PCR fragment for *pinuE*-*racA*-G12V and *pinuE*-eGFP, respectively. PCR positive transformants were verified by Southern blot analysis as described by Sambrook et al. (1989).

Phenotype analysis of *A. niger* reporter strains

For the growth assays on plates, equal amounts of fresh conidiospores were incubated on agar plates containing MM supplemented with 1% (w/v) various carbon sources and grown at 30°C for 3 days. For cover slip assays, 1x10⁴ fresh conidiospores were inoculated on two glass cover slips in 5-cm Petri dishes. Each Petri-dish contained 8 ml of MM supplemented with 1% (w/v) carbon source and 0.1% (w/v) casamino acids and incubated at 37°C for 6 hours and then switched to 25°C for 16 hours before microscope observation. Light and GFP images of the thin mycelia fragments attached to the cover slips were analyzed by DIC or fluorescent microscopy using a Zeiss Axioplan 2 Optical microscope at a fixed exposure times and standard filter settings.

Results and discussion

Expression analysis of inulinolytic genes using microarrays

A. niger contains five inulinolytic genes including *inuE* (exo-inulinase), *inuA* (endo-inulinase), *sucA* (invertase), *sucB* (intracellular invertase, Goosen et al., 2007) and *sucC* (putative intracellular invertase) in *A. niger* genome (Yuan et al., 2006). In this study the expression of these inulinolytic genes was studied in relation to various carbon sources and the InuR transcriptional activator (Yuan et al., 2008) using Affymetrix microarrays.

A. niger strains were pregrown on MM containing 2% xylose for 18 h and the mycelia were transferred to fresh MM containing 1% various carbon sources. After 2 h or 8 h, mycelia were harvested and the corresponding RNA samples were isolated and used in microarray experiments to monitor the transcription levels of inulinolytic genes. Based on our previous Northern analysis (Yuan et al., 2006), the expression of inulinolytic genes on sucrose was highly induced 2 h after the transfer, and the induction was dramatically reduced after 8 hours. Therefore, for the sucrose culture, RNA from the 2 h growth time points after the transfer was used for microarray analysis. The results are summarized in

Fig. 1. After the transfer to xylose or maltose, the expression of inulinolytic genes were shown to be either very low (*inuE*, *sucB*) or not detectable (*inuA*, *sucA*, *sucC*), similar as in the expression levels in the xylose preculture. In contrast, the transfer to sucrose resulted in a strong induced expression for the *inuE*, *sucA* and *inuA* genes. After the transfer to inulin, *inuE* was modestly expressed at an early stage (2 h) and displayed strong induction (63-fold) at later stage (8 h). The expression of *sucA* was only modestly induced on inulin at later stage (8 h) (Fig. 1). The expression of *inuA* was not induced 2 hours after the transfer to inulin, and only slightly higher after 8 hours. These results support previous data that the induction of the inulinolytic genes on sucrose was very rapid and more rapid than the induction on inulin (Yuan et al., 2006). As expected the expression of the inulinolytic genes is low on maltose and xylose. As reported earlier (Yuan et al., 2006), the expression of *sucB* and *sucC* is low under the various carbon sources and not induced in the presence of sucrose or inulin. These data also indicate that, among the five inulinolytic genes, *inuE* was the most highly expressed and most strongly inducible gene in response to the presence of both sucrose and inulin.

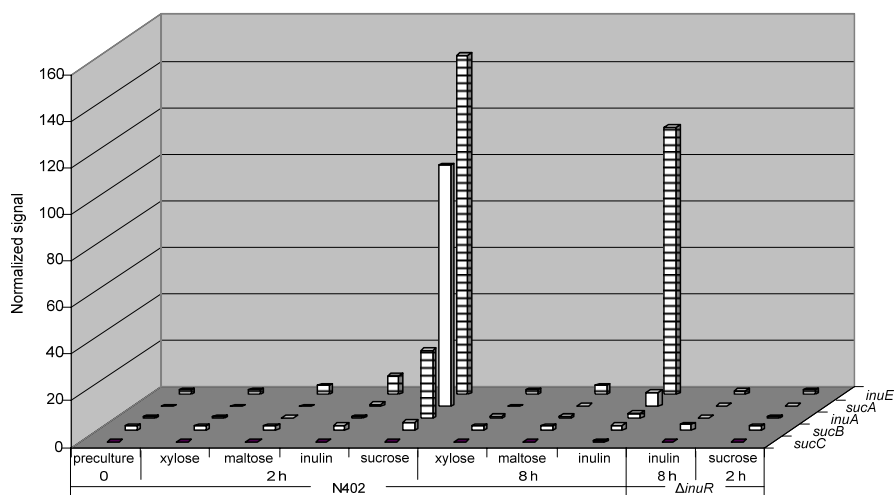


Fig. 1. Transcriptional expression of *A. niger* inulinolytic genes. Accession numbers of the gene names are given in Table 1. Strains and time points after transfer from the preculture are indicated below, inulinolytic genes are indicated on the right. The values and Present/Absent calls from the expression data are provided as Suppl. Table 1.

A similar transfer experiment was also performed for the Δ *inuR* strain. As shown in Fig. 1, the induction of inulinolytic genes on sucrose (for *inuE*, *sucA*, *inuA* at 2 h after transfer) and inulin (for *inuE* and *sucA* at 8 h after transfer) was abolished in the Δ *inuR* mutant and confirms our previous Northern observations that the induced expression of inulinolytic genes requires InuR (Yuan et al., 2008).

Set up of a genetic screen to identify inulin signalling mutants

To screen for mutants affected in the activation of the InuR-mediated induction of inulinolytic genes, a reporter system based on *inuE* promoter was designed. The *inuE* gene was selected because it displayed the strongest induction in expression on inulin and sucrose (Fig. 1). As a reporter for the genetic screen, the *inuE* promoter was cloned into in front of the *A. niger racA*-G12V gene. *racA*-G12V is a dominant active mutant form of the *racA* gene. *racA* encodes a small GTPase involved in actin organization and the induced expression of *racA*-G12V results in the loss of highly polarized tip growth and subsequent swelling of the hyphal tip (Arentshorst and Ram, unpublished data).

Therefore, the reporter strain containing the *PinuE*-*racA*-G12V fusion construct is expected to cause restricted growth on inducing carbon sources for *inuE* expression such as sucrose and inulin, but to display normal growth on non-inducing or repressing carbon sources. Such a screen allows a positive selection to isolate mutants that display normal growth on inducing carbon sources, indicating in this mutant transcription activation of the *inuE* gene is impaired. To test the screening method, we also transformed the reporter construct in a Δ *inuR* strain. Here, the hypothesis is that in the Δ *inuR* background, *inuE* is not induced and therefore this strain is expected to grow normally in the presence of inducing carbon sources. A second reporter gene was constructed in which the *inuE* promoter was cloned upstream to eGFP. In this study, the usefulness of this reporter was evaluated and transformed separately from the *PinuE*-*racA*-G12V construct to both the wild type and the Δ *inuR*. In future work, a strain should be constructed containing both reporter strains which allow the screening for trans-acting mutants, instead of cis-acting mutants. In the following section, the procedure of the construction of the strains is described in detail.

Construction of *A. niger* reporter strains

To construct the *racA*-G12V reporter strain the *PinuE*-*racA*-G12V plasmid was made as described in Materials and methods (Fig. 2A). Briefly, the 1.0 kb promoter of *inuE* was cloned upstream of the *racA*-G12V gene and downstream of the *A. nidulans trpC* gene as a terminator sequence. The *pyrG*^{*} was introduced as a selection marker, allowing the reporter plasmid to be inserted at the *pyrG* locus (Gouka et al., 1995). Since xylose acts as a non-inducing carbon source for *inuE* expression (Yuan et al., 2006), primary transformants were grown on xylose and screened by PCR for the presence of *racA*-G12V (see Materials and methods). Two putative reporter strains MA61.24 and MA61.25 were further examined by Southern blot analysis to verify the proper insertion of the construct. Using an *Xho*I fragment containing an *A. niger pyrG* gene as probe, *Bam*HI digested genomic DNA showed three fragments, 2.3, 2.7 and 9.7 kb in the reporter strain MA61.24 and two fragments 2.1 and 2.3 kb in WT strain N402 as anticipated. Strain M61.25 displayed extra fragments 2.1 and 7.6 kb instead of 9.7 kb, suggesting the *pyrG*^{*} point mutation on the *Bam*HI site was repaired during transformation (Fig. 2C). Strain MA61.24 was used for further characterization.

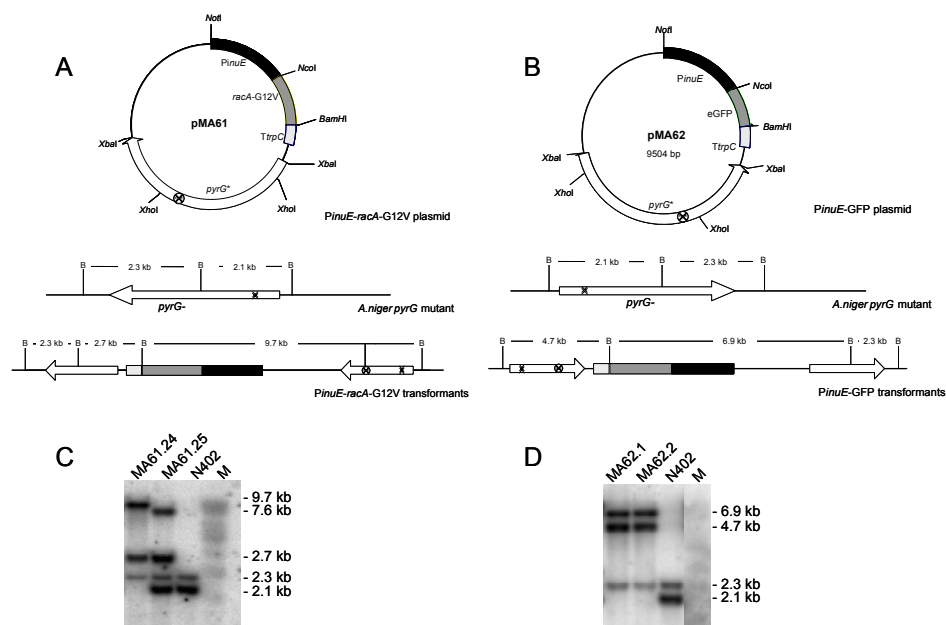


Fig. 2. Schematic representation of the construction of the reporter strains containing construct *PinuE-racA-G12V* (A) or *PinuE-eGFP* (B). Southern analysis of the reporter strains indicated in A and B are shown in C and D, respectively.

The second reporter construct (p*PinuE-eGFP*) was made as described in the materials and method section. Transformants with a targeted integration of the *PinuE-eGFP* construct at the *pyrG* locus were identified by PCR analysis. PCR positive transformants MA61.1 and MA61.2 were verified by Southern analysis as shown in Fig. 2B. Using an *XhoI* digested *A. niger pyrG* gene fragment as a probe, *BamHI* digested genomic DNA showed three fragments 2.3, 4.7 and 6.9 kb in the reporter strain containing *PinuE-eGFP* and two fragments 2.1 and 2.3 kb in WT, as anticipated (Fig. 2D). Strain MA62.1 was used for further characterization.

Phenotype characterization of reporter strain containing *pinuE-racA-G12V*

Colony morphology assay in agar plates

To examine the growth effect of the *PinuE-racA-G12V* reporter strain (MA61.24) on different carbon sources, this strain was grown on MM agar plates containing 1% different carbon sources. As shown in Fig. 3A, both growth and conidiation of the reporter strain MA61.24 were affected on inulin, sucrose, raffinose, and maltose indicating that *racA-G12V* was overexpressed by the activated *inuE* promoter at these carbon sources. The effect was anticipated for inulin and sucrose. The expression of the *inuE* gene has not previously been examined on raffinose, but the results from the reporter strain indicate that the *inuE* gene is

also expressed on raffinose. This is not complete surprising as the action of alpha-galactosidases secreted by *A. niger* might result in the formation of sucrose which can act as an inducer for the inulinolytic enzymes (Yuan et al., 2006). Interestingly, on the maltose plate, the growth of MA61.24 was also affected, suggesting that growth on maltose on plates induces *inuE* expression. Northern and array analysis (Fig. 1) showed a minor expression of the *inuE* on maltose under submerged condition and possibly the expression of the *inuE* gene on maltose is higher on plate. On xylose and glucose MA61.24 growth and conidiation is similar as in the wild type strain (N402), indicating xylose and glucose are repressing carbon sources for *inuE*, which is consistent with previous transcription analysis (Yuan et al., 2006; 2008). Fructose seems to modestly induce *inuE* expression as the growth of reporter strain was slightly affected when compared to N402. On sorbitol and glycerol, which were considered as non-inducing carbon sources, MA61.24 also shows a growth phenotype (reduced conidiation), indicating that these two carbon sources are inducing *inuE* expression as well. Taken together, these data suggest that besides the known inducers (inulin, sucrose) raffinose, maltose, fructose, sorbitol and glycerol also lead to *inuE* expression on solid agar plates. To demonstrate that this reporter system can be used for identifying pathway signalling genes, the reporter construct was also transformed into the Δ *inuR* strain. A *PinuE-racA-G12V* construct (strain XY7.1) was isolated by PCR and confirmed by Southern analysis (data not shown). As shown in Fig. 3B, on all tested conditions, including inducing, non-inducing and repressing carbon sources, no differences of either growth or conidiation was observed between strain XY7.1 containing an *inuE-racA-G12V* construct and the Δ *inuR* strain. This indicates that deleting of *inuR* abolished the expression of *inuE*, thereby resulting in no expression of *racA-G12V*.

Hyphal morphology in submerged culture

We also examined the hyphal morphology of the *PinuE-racA-G12V* reporter strain during spore germination. *A. niger* strains MA61.24 (expressing *PinuE-racA-G12V*) and N402 (wild type strain) were grown on cover slips in a submerged culture (see Materials and methods) containing different carbon sources, followed by microscope observation. As shown in Fig. 3C, under inducing conditions, such as inulin, sucrose and raffinose, the reporter strain MA61.24 containing *PinuE-racA-G12V* showed extreme swelling and short hyphae. This indicates that expression from the *inuE* promoter of *racA-G12V* in the submerged culture occurs already during spore germination and causes defective development of hyphae. This is remarkable different from expression of the *racA-G12V* protein from the glucoamylase promoter which only gives an effect on the morphology of the germlings after 16 hours of germination (Arentshorst and Ram, unpublished results). The morphology of the hyphae of the *racA-G12V* reporter strain on the other carbon sources was normal (Fig. 3C), indicating that *racA-G12V* is not expressed. This result is different from the previous agar plate assay, in which growth on maltose, fructose, sorbitol and glycerol induced *inuE* expression. Obviously there are other factors involved in the physiological growth conditions (e.g. carbon source concentration, pH, age of the mycelium) that affect the inulinolytic system. The results indicate that the expression from the *PinuE* is tightly regulated in

response to carbon sources in a submerged culture during germination and that *racA*-G12V is a good and sensitive indicator to detect expression. As shown in Fig. 3D, reporter strain XY7.1 (Δ *inuR* strain containing *inuE*-*racA*-G12V) displays normal hyphae similar as N402 and the Δ *inuR* mutant, suggesting no expression of *racA*-G12V from the *inuE* promoter in the Δ *inuR* mutant (Fig. 3D).

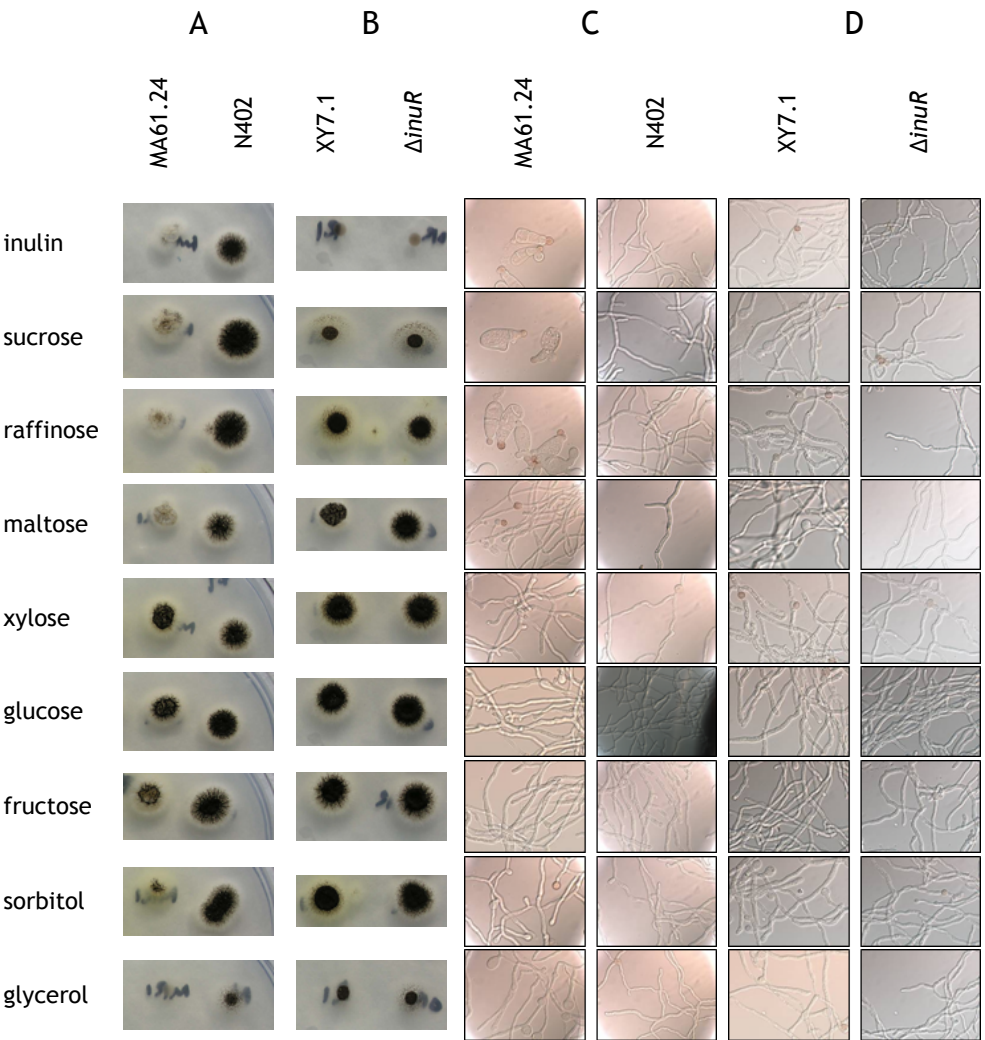


Fig. 3. Effects of overexpression of *racA*-G12V from the *inuE* promoter on *A. niger* growth morphology during conidiation (A and B) and during germination (C and D). MA61.24 is a N402 (*pyrG*⁻) containing a *PinuE*-*racA*-G12V construct; XY7.1 is a Δ *inuR* (*pyrG*⁻) strain containing *PinuE*-*racA*-G12V construct.

From both the agar plate assay for colony morphology and cover slip assay for hyphal morphology, we conclude that inulin, sucrose and raffinose are strong inducers for *inuE* expression and lead to strong phenotypes from the reporter *racA*-G12V. Complicating factor

for a mutant screen is our observation that the $\Delta inuR$ strain can not grow or only shows limited growth on inulin and sucrose plates (Fig. 3B). Interestingly, we found that in the $\Delta inuR$ mutant harboring the *racA*-G12V reporter, the colony size and the degree of conidiation on raffinose were much larger than those on inulin and sucrose. Although growth of N402 on raffinose induces the expression of *inuE*, growth of the $\Delta inuR$ mutant on raffinose still occurs, indicating that *A. niger* has an InuR independent mechanism to metabolize raffinose. As mentioned above, raffinose can be metabolized by alpha-galactosidase to galactose and sucrose. The galactose will assist growth of the potential mutant reporter strains while sucrose induces the InuR mediated expression of *inuE*. From this result, we conclude that raffinose would be a suitable carbon source for mutant screening. By growing the potential mutant on raffinose, we expect that the mutant strain contains the reporter construct to grow normally, whereas the wild type strain containing *inuE*-*racA*-G12V will not grow due to the high expression of *racA*-G12V.

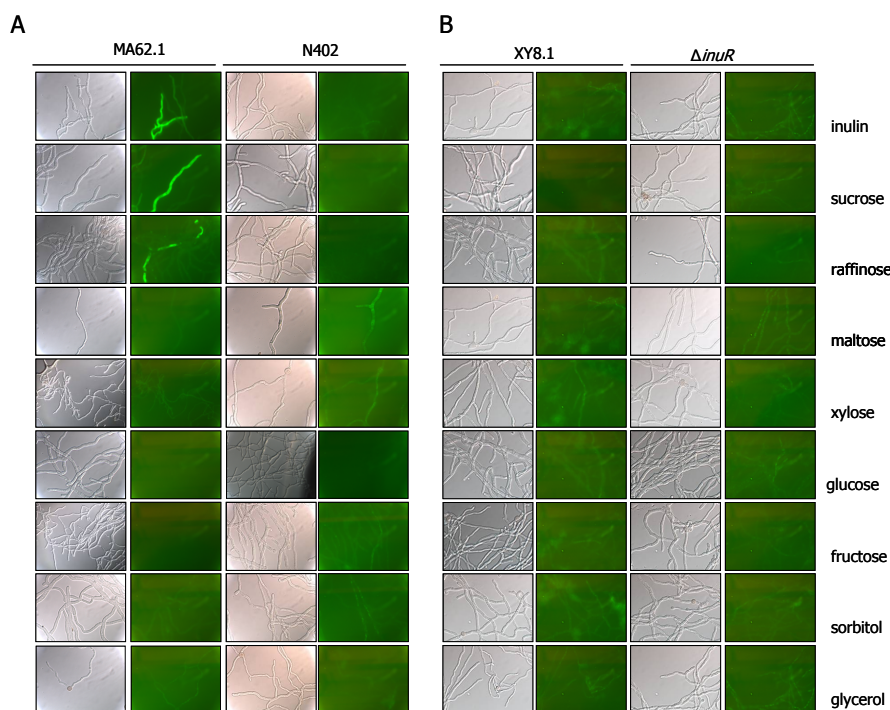


Fig. 4. Effects of overexpression of eGFP from the *inuE* promoter on fluorescence of *A. niger* grown in submerged culture. MA62.1 is a N402 (*pyrG*⁻) containing a *PinuE*-eGFP construct; XY8.1 is a $\Delta inuR$ (*pyrG*⁻) strain containing *PinuE*-eGFP construct.

Microscopic analysis of reporter strains containing *PinuE*-eGFP

To examine the reporter strain containing the *PinuE*-eGFP construct, the strain MA62.1 was grown on cover slips in a submerged culture and observed by fluorescent microscopy.

Similar as observed for the *PinuE-racA-G12V* reporter, inulin, sucrose and raffinose triggered expression from the *inuE* promoter resulting in a strong green fluorescence in the hyphae of GFP reporter strain MA62.1 (Fig. 4A). Due to several layers of hyphae in one image, the fluorescence shown here was only highlighted at one layer with a certain focus. The observed fluorescence was specific for the *PinuE*-GFP construct as very low background fluorescence was observed in the N402 strain. The fluorescence of the reporter strain in the other carbon sources was comparable to the fluorescence in the N402 strain, indicating that these carbon sources do not induce *inuE* under these conditions. These results for the *PinuE*-eGFP are consistent with the results for the expression of the *PinuE-racA-G12V* in germination morphology assay. This result further proves that *PinuE* is tightly regulated and that the *PinuE*-eGFP construct is also a good reporter to detect expression from the *inuE* promoter. The *PinuE*-eGFP reporter construct was also transformed into a Δ *inuR* strain and confirmed by Southern analysis (data not shown). As expected only basal background fluorescence was observed in both the reporter strain XY8.1 and the control strain Δ *inuR* (Fig. 4B).

Taken together our results indicate that both the reporter constructs, *PinuE-racA-G12V* and *PinuE*-eGFP, are suitable for identifying mutants affected in *inuE* expression. By combining these two reporter constructs in a single strain, we can easily discriminate between cis-acting mutants and trans-acting mutants that are affected in the *inuE* expression. From the cis-acting mutants, we might be able to identify InuR binding sites. Trans-acting mutants probably represent mutants in which are disturbed in the activation of the InuR transcription factor or the InuR transcription factor itself. Such a genetic screen and the identification of the mutants would allow the identification of important proteins involved in activation of the inulinolytic system in *A. niger*. These proteins might include proteins that generate inducing molecules, transport or sense the inducing molecules and proteins required for the activation of the InuR transcription factor.

Working model of the inulin signalling pathway in *A. niger*

The recent identification of inulinolytic genes in the *A. niger* genome (Yuan et al., 2006, Chapter 5) and studies on their transcriptional control (Yuan et al., 2006 and 2008, Chapter 5 and 7) as well as the identification of putative sugar transporter encoding genes that are induced in sucrose metabolism (Yuan et al., 2008, Chapter 5), allowed us to build a working model of the inulin signalling pathway which is presented in Fig. 5.

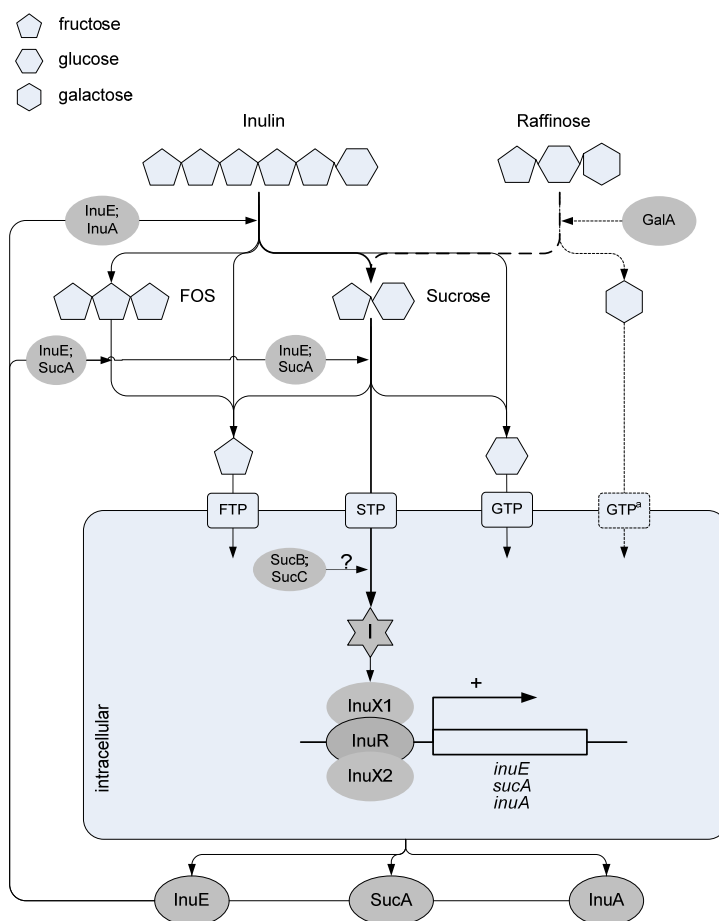


Fig. 5. A working model of the inulin signalling pathway. Solid line indicates inulin degradation pathway, broken line indicates raffinose pathway. Abbreviations stand for: STP, sucrose transporter; FTP, fructose transporter; GTP, glucose transporter; GTP^a, galactose transporter; I, inducer.

We presume that inulin is broken down to oligo-fructoses, sucrose, glucose and fructose by the combined action of InuA (endo-inulinase), InuE (exo-inulinase) and SucA (invertase). These enzymes are likely to be produced in the presence of inulin as the genes encoding these proteins are highly expressed on inulin (Fig. 1 and Yuan et al., 2006). Similarly, raffinose which has been shown to induce *inuE* (this chapter) can be hydrolyzed to generate galactose and sucrose by the action of alpha-galactosidase (GalA). The sucrose can be further hydrolyzed to glucose and fructose by InuE and SucA. In chapter 5 it is shown that sucrose act as an inducer for the inulinolytic genes and we therefore propose that the sucrose which is formed from inulin or raffinose induces expression of the inulinolytic genes. During the growth on sucrose, the sucrose can directly act as an inducer. Although we can not exclude the possibility that extracellular sucrose can activate the InuR

transcription factor, we currently favor the idea that the sucrose is transported into the cell by a proposed sucrose transporter (STP). Candidates for such a sucrose transporter, but also for other predicted transporters (Fructose Transporter (FTP), Glucose Transporters (GTP)), have been identified in the genome wide array analysis but future functional analysis should be performed to show the specificity of the different transporters. The function of *sucB* and *sucC*, both encoding intracellular invertases is not yet clear. In Chapter 6 we show that *sucB* encodes an invertase with transglycosylation activity, and is able to form 1-kestose. Disruption of the *sucB* gene did not affect growth of the disruption strain on inulin or sucrose, indicating that SucB itself is not required to activate the inulinolytic system (Chapter 6). Possibly, SucC and SucB have redundant functions and a definitive conclusion on SucB/SucC functioning awaits the disruption of both genes simultaneously. An attractive possibility is that SucB and/or SucC acts on the intracellular sucrose possibly as transglycosylases forming transglycosylation products (e.g. 1-kestose) which might act as an inducer molecules. Alternatively, SucB/SucC might act as invertases to hydrolyze intracellular sucrose, without having a function in the formation of inducer. Sucrose, or a derivative of sucrose, is thought to activate the InuR transcription factor and this activation might require addition proteins (InuX1-InuX2) which are currently unknown. The genetic screen described in this chapter allows the identification of such proteins as well as proteins that are involved in the transport or formation of the inducer. An important limitation of such a genetic screen is that redundant functions, e.g. the presence of multiple sucrose transporters, or multiple proteins that can form the inducer make the identification of mutants very difficult, if not impossible. The presence of inulin, sucrose or raffinose in the medium and the subsequent induction of the inulinolytic genes are probably accomplished by the constitutive expression of genes encoding proteins that generate the inducer. The methods used in this study to determine expression levels (Northern blot and microarray analysis) are not the preferred methods to quantify expression levels of lowly expressed genes. Quantitative RT-PCR experiments or enzyme activity assays should be performed to determine the genes/enzymes that are constitutively expressed.

Finally, the *racA*-G12V reporter based screening method can be generally used for the identification of transcription factor mutants. Prerequisite for such a screen is the identification of a strongly induced gene, with a relative low basal expression level. Since microarray data with *A. niger* will become more and more available, the identification of such genes is rather straightforward and the strategy discussed in this paper can be an interesting follow-up experiment after performing genome wide array analysis to identify regulatory mutants.

Supplementary Data

Suppl. Table 1: Expression profiles of inulinolytic genes of *A. niger*.

gene	N402 (xyl 18h)			N402 (xyl 2h)			N402 (mal 2h)			N402 (inu 2h)			N402 (suc 2h)			N402 (xyl 8h)			N402 (mal 8h)			N402 (inu 8h)			Δ inuR (inu 8h)			Δ inuR (suc 2h)		
<i>inuE</i>	1.82	P		1.81 $\pm$ 0.28 ^a	P,M		3.93 $\pm$ 2.10	P		7.96 $\pm$ 2.56	P		145.30 $\pm$ 27.90	P		1.89 $\pm$ 0.01	P,M		3.88 $\pm$ 0.22	P		114.60 $\pm$ 2.40	P		1.59 $\pm$ 0.36	P		1.97 $\pm$ 0.21	P,M	
<i>inuA</i>	0.58	A		0.63 $\pm$ 0.23	A		0.26 $\pm$ 0.07	A		0.77 $\pm$ 0.11	A		29.27 $\pm$ 0.80	P		0.89 $\pm$ 0.07	A		0.99 $\pm$ 0.29	A		2.14 $\pm$ 0.18	P		0.50 $\pm$ 0.16	A		0.74 $\pm$ 0.06	A	
<i>sucA</i>	0.14	A		0.11 $\pm$ 0.01	A		0.07 $\pm$ 0.04	A		0.95 $\pm$ 0.65	A		103.40 $\pm$ 21.54	P		0.11 $\pm$ 0.08	A		0.29 $\pm$ 0.60	A		6.07 $\pm$ 0.51	P		0.17 $\pm$ 0.16	A		0.26 $\pm$ 0.45	A	
<i>sucB</i>	2.05	P		1.94 $\pm$ 0.21	P		2.01 $\pm$ 0.17	P		2.23 $\pm$ 0.16	P		3.66 $\pm$ 0.49	P		1.89 $\pm$ 0.06	P		2.06 $\pm$ 0.08	P,M		2.25 $\pm$ 0.43	P		2.55 $\pm$ 0.11	P		2.08 $\pm$ 0.53	P	
<i>sucC</i>	0.45	A		0.34 $\pm$ 0.09	A		0.31 $\pm$ 0.19	A		0.33 $\pm$ 0.16	A		0.34 $\pm$ 0.30	A		0.47 $\pm$ 0.16	A		0.26 $\pm$ 0.02	A		0.66 $\pm$ 0.02	A		0.25 $\pm$ 0.06	A		0.26 $\pm$ 0.31	A	

^aThe expression level was based on the geometric mean value of the duplicate samples and the deviation values between the duplicate samples are indicated. P, M or A, representing detection calls for present, marginal or absent respectively.

Summary

Aspergillus niger is an ubiquitous filamentous fungus and commonly present on decaying plant debris. As a saprophytic fungus, *A. niger* produces a variety of hydrolytic enzymes that are able to break down the plant polysaccharides into smaller molecules to serve as its nutrient and energy sources. Many enzymes secreted by *A. niger* have potential applications in the baking, starch, textile, food and feed industries. This has attracted a lot research attention to explore *A. niger* for better understanding the carbohydrate modifying network in *A. niger*, in order to improve the enzyme production and substrate utilization. In addition, the discovery of the novel enzyme activities produced by *A. niger* is a subject of great interest. Both the discovery of new enzymes in the *A. niger*, as well as a genome-wide analysis of transcriptional networks related to degradation of carbohydrates has become possible through recent sequencing of *A. niger* genome by DSM and DOE Joint Genome Institute (Pel et al., 2007, Baker et al., 2007) and the availability of Affymetrix microarrays (Pel et al., 2007).

In this thesis, we are interested in the molecular mechanisms by which the *A. niger* can sense the presence of different (often complex) carbon sources in its vicinity and how the presence of a certain carbohydrate activates the expression of a network of genes encoding enzymes to degrade the carbon source. Special focus is on the identification of so-called pathway inducing molecules and transcription factors that are required for expression of enzyme networks involved in the breakdown and uptake of starch and inulin, and their derivatives. In addition, attention is also paid to how the expressed enzymes encoded by structural genes react to the carbon sources and to determine their physiological roles in *A. niger*.

Chapter 1 outlines the importance of *A. niger* as an enzyme producer and the current knowledge on enzymes involved in the starch and inulin degradation or modification in *Aspergilli*. The expression and production of enzymes involved in carbohydrate degradation is highly regulated and controlled by various wide-domain transcription factors (CreA, AreA, PacC), and pathway specific transcriptional activators (AmyR and InuR). This Chapter gives the reader an overview and more background information about the various transcription factors that are involved in the regulation of gene expression. The regulation mechanisms of genes encoding the enzymes involved in polysaccharides modification or degradation and the important features of the different transcriptional factors involved in the regulation are described.

To identify novel enzymes acting on starch and inulin, two approaches have been followed. The first approach is described in **Chapter 2** and is based on the construction of starch and inulin specific cDNA expression libraries using Gateway cloning technology. Expression libraries for screening both in bacteria (*E. coli*) and yeast (*S. cerevisiae*) were constructed and characterized in detail. The libraries have been transformed to *E. coli* and *S. cerevisiae*, and transformants have been screened. Due to problems in both the expression level of the proteins encoded by the cDNAs and the viability of the transformants only a limited number of transformants have been screened. We present several suggestions to improve the methods for construction and screening of the cDNA expression libraries.

The second approach to identify new enzymatic activities is by database mining using functionally characterized enzymes related to starch or inulin metabolism. As described in **Chapter 3**, starch can be degraded by the concerted action of α -amylase, glucoamylase and α -glucosidase enzymes, members of the glycoside hydrolase (GH) families 13, 15 and 31, respectively. Thus, a full set of GH13, 15 and 31 family members were recognized in *A. niger* genome and 17 new members were identified. Combination analysis of the genome sequence of *A. niger* CBS 513.88 with transcriptome expression profiles reveals a large number of these genes with unexpected transcriptional regulation: they were neither induced by starch degradation pathway inducer, maltose, nor dependent on the presence of starch utilization activator, AmyR. Only two of the newly identified enzymes, a putative α -glucosidase (AgtB) and an α -amylase (AmyC), were predicted to play a role in starch degradation. The possible physiological functions of other predicted family GH13, GH15 and GH31 enzymes, including predicted intracellular enzymes and cell wall associated proteins, in alternative α -glucan modifying processes are discussed.

Among the newly identified GH13 family enzymes, three members (AgtA, AgtB and AgtC), sharing high similarity to fungal α -amylases were predicted to be glycosylphosphatidylinositol-anchored and lacked some highly conserved amino acids of the α -amylase family. **Chapter 4** describes the functional characterization of the purified proteins (AgtA and AgtB) from *A. niger* overexpression strains, and the phenotypic characterization of gene deletion (*agtA*) and gene overexpression (*agtA* and *agtB*) strains. Both AgtA and AgtB displayed a unique type of α -(1, 4) glucanotransferase (EC 2.4.1.25) activity. Deletion of *agtA* in *A. niger* resulted in an increased susceptibility towards the cell wall disrupting compound Calcofluor White, indicating a cell wall integrity defect in this strain. Interestingly, homologues of AgtA and AgtB are also present in other fungal species with α -glucans in their cell walls, but not in yeast species lacking cell wall α -glucan. Therefore the possible role for these enzymes in the synthesis and/or maintenance of the fungal cell wall is suggested and discussed in this chapter.

In **Chapter 5**, we survey *A. niger* genome for the presence of GH32 family enzymes, which are thought to be involved in utilization of inulin, and its synthetic substrate, sucrose. Two new intracellular proteins SucB and SucC were identified, in addition to three known extracellular inulinolytic enzymes, SucA (invertase), InuE (exo-inulinase) and InuA (endo-inulinase). Transcriptional analysis revealed that the extracellular enzymes are coordinately

regulated and induced by sucrose and inulin. Their transcription was also under control of the carbon catabolite repressor CreA. Further analysis indicated that sucrose, or a sucrose-derived intermediate, but not fructose, acts as an inducer for the expression of inulinolytic genes in *A. niger*.

In **Chapter 6**, we examine if SucB, one of the two novel intracellular invertases, plays an essential role in generating pathway inducing molecule involved in inulin or sucrose catabolism. Therefore the phylogenetic, molecular and biochemical characteristics of SucB were studied. The overexpressed and purified SucB protein was shown to act as an invertase with transfructosylating activity. Disruption of *sucB* in *A. niger* did not result in alteration of growth in inulin or sucrose indicating SucB does not play an essential role in inulin or sucrose catabolism in *A. niger*. However, SucB may be needed for intracellular conversion of sucrose to fructose, glucose, and small oligosaccharides.

As shown in Chapter 5, the extracellular inulinolytic genes were co-regulated and induced by sucrose and inulin, suggesting a common transcriptional activator might exist to activate the expression of the structural genes in the presence of inducing molecules. In **Chapter 7** we describe the identification and characterization of the Zn(II)₂Cys₆ type, inulinolytic transcription activator InuR. The identification of *inuR* was based on inspection of inulinolytic gene clusters in *A. niger* genome. *inuR* gene is clustered with *sucB*, which encodes an intracellular invertase with transfructosylation activity (Chapter 6) and a putative sugar transporter encoding gene (An15g00310). Deletion of the *inuR* gene displayed a severe defective growth on inulin and sucrose as sole carbon source. Induction of the extracellular inulinolytic genes and the sugar transporter gene An15g00310 in the presence of sucrose and inulin was dependent on InuR. Genome wide expression analysis revealed three additional putative sugar transporter encoding genes (An15g04060, An15g03940 and An17g01710), which were strongly induced by sucrose in an InuR dependent way. The specificity of these putative sugar transporters, especially the An15g00310 encoded sugar transporter is an interesting topic and is now under the investigation by biochemical analysis. *In silico* analysis of the strongly InuR regulated genes revealed a putative binding site for InuR consisting of two CGG triplets separated by 8 nucleotides, which is very similar to the AmyR binding site CGGN8(C/A)GG. Analysis of the double mutant ($\Delta amyR \Delta inuR$) revealed their independent regulating function. How the InuR and AmyR recognize their target genes and what determines the specificity of the binding is an interesting topic for further research.

Among the inulinolytic genes present in *A. niger*, we identified the *inuE* gene as the most strongly induced gene in the presence of the pathway inducing poly- and disaccharides inulin and sucrose, respectively. In **Chapter 8**, using *inuE* promoter region combined with *racA*-G12V or GFP reporter genes, we present a novel screening method for the isolation of mutants involved in inulin or sucrose signalling. Such a genetic screen followed by identification of the mutants would allow the identification of proteins that generate inducing molecules, transport or sense the inducing molecules as well as the proteins that activate the pathway transcription factor InuR involved in the inulin catabolism. This method can be generally used for identification of transcription factor

mutants. In addition, during the setup of screening, we found that raffinose is able to induce the expression of *inuE* gene.

Finally, based on last four chapters, Chapter 5, 6, 7 and 8, a speculative working model of inulin signalling pathway was presented at the end of Chapter 8.

Samenvatting

Aspergillus niger is een (veel voorkomende) schimmel die van nature veel voorkomt in de bodem en gespecialiseerd is om te groeien op organisch materiaal van dode planten. Omdat het een saprofytische schimmel is, produceert *A. niger* vele hydrolytische enzymen die plant polysacchariden (lange suikerketens) kunnen afbreken in kleinere moleculen. Deze kunnen (die) door de schimmel (kunnen) worden opgenomen en dienen als energiebron. Veel enzymen die worden uitgescheiden door *A. niger* hebben potentiële toepassingen, en zijn met name interessant voor (in) de voedingsindustrie. Deze interesse is aanleiding geweest om een onderzoeksprogramma te starten waarbij met nieuwe technologieën het enzymnetwerk dat betrokken is bij de afbraak van plant polysacchariden door *A. niger* te bestuderen. Eén van de uiteindelijke doelen van het onderzoek is om het suikerafbrekende enzymnetwerk in *A. niger* beter te begrijpen zodat de enzymproductie en de efficiëntie van de omzetting van bepaalde substraten kunnen worden gecontroleerd en verbeterd. Daarnaast is de ontdekking van de nieuwe enzymactiviteiten van *A. niger* aanleiding geweest voor veel wetenschappelijke en industriële interesse. Zowel voor het ontdekken (de ontdekking) van nieuwe enzymen in *A. niger* als voor het identificeren van genen die mogelijk betrokken zijn bij de afbraak van een bepaald substraat is gebruik gemaakt van nieuwe technologieën. DSM heeft als eerste de complete DNA volgorde van genoom van *A. niger* bepaald. Met deze kennis heeft DSM ook zogenoemde microarrays ontwikkeld, die gebruikt kunnen worden om te meten hoe hoog een bepaald gen tot expressie komt (Pel et al., 2007).

In dit proefschrift zijn we geïnteresseerd in de moleculaire mechanismes waarmee *A. niger* de aanwezigheid van verschillende (vaak complexe) koolstofbronnen in zijn omgeving registreert en hoe de aanwezigheid van sommige suikers de expressie van een netwerk van genen, betrokken bij extracellulaire afbraak, opname en verder metabolisering, kan activeren om een bepaald polysaccharide af te breken. Speciale aandacht is er voor de identificatie van zogenaamde netwerk inducerende suikers (inducers) en transcriptiefactoren die nodig zijn voor de expressie van genen die coderen voor eiwitten die betrokken zijn in het afbreken polysacchariden. In dit proefschrift zijn enzymnetwerken van *A. niger* die betrokken zijn bij de afbraak van zetmeel en inuline bestudeerd. Zowel zetmeel als inuline zijn polysacchariden die door de plant gemaakt worden als reservevoedsel, zgn opslagsuikers. Daarnaast is er ook aandacht besteed aan het bepalen van de biochemische activiteit van nieuw ontdekte enzymen en is geprobeerd (is) om voor deze activiteit een functionele rol in het metabolisme van *A. niger* te bepalen (te plaatsen).

Hoofdstuk 1 laat zien hoe belangrijk *A. niger* is als producent van enzymen en geeft een overzicht van de huidige kennis over de enzymen die betrokken zijn in de afbraak of modificatie van zetmeel en inuline in *Aspergilli*. De expressie en productie van enzymen die zijn betrokken bij polysaccharide afbraak wordt sterk gereguleerd en gecontroleerd door verschillende algemene transcriptiefactoren (CreA, AreA, PacC), en route-specifieke transcriptiefactoren (AmyR and InuR). Dit hoofdstuk geeft de lezer een overzicht en enige achtergrondinformatie met betrekking tot de verschillende transcriptiefactoren in relatie tot genexpressie. Tot slot worden de regulatiemechanismen beschreven van genen die coderen voor de enzymen die betrokken zijn bij de modificatie of afbraak van polysacchariden en de belangrijke aspecten van de verschillende transcriptiefactoren die betrokken zijn bij deze regulatie.

Om nieuwe enzymen te kunnen identificeren die werken op zetmeel en inuline, zijn twee (aanpakken) onderzoekslijnen, methode?? gevolgd. De eerste aanpak is beschreven in **Hoofdstuk 2** en is gebaseerd op de constructie van zetmeel en inuline specifieke cDNA expressie banken door gebruik te maken van Gateway cloning technologie. Expressiebanken voor het screenen van zowel bacteriën (*E. coli*) als gist (*S. cerevisiae*) zijn geconstrueerd en in detail beschreven. De banken zijn getransformeerd naar *E. coli* en *S. cerevisiae* en transformanten zijn geselecteerd. Vanwege problemen in zowel het bereiken van een hoog expressienivo van de cDNAs in de transformanten en in de levensvatbaarheid van de transformanten is slechts een klein aantal van hen geselecteerd. We geven enkele suggesties om de methoden voor constructie en screening van cDNA expressie banken te verbeteren.

Een tweede aanpak om nieuwe enzymatische activiteiten te identificeren is door gebruik te maken van functioneel gekarakteriseerde enzymen die zijn gerelateerd aan het metabolisme van zetmeel en inulin. De aminozuurvolgorde van deze eiwitten worden dan met bepaalde computerprogramma's vergeleken met de aminozuurvolgorde van alle eiwitten zoals die voorspeld kunnen worden op basis van de genomsequentie. Zoals in **Hoofdstuk 3** is beschreven, kan zetmeel worden afgebroken door verschillende enzymen, α -amylase, glucoamylase en α -glucosidase leden van de glycoside hydrolase (GH) families 13, 15 en 31. Deze eiwitfamilies worden gekarakteriseerd door de aanwezigheid van gebieden met sterk geconserveerde aminozuren tegen geconserveerde domeinen. Door het zoeken naar eiwitten met vergelijkbare domeinen, is de volledige verzameling van GH13, 15 en 31 familieleden herkend in het *A. niger* genoom en werden 17 nieuwe leden geïdentificeerd. De expressie-analyse van de genen die coderen voor deze eiwitten laat een groot aantal van deze genen zien met onverwachte expressiepatroon: ze werden noch geïnduceerd door maltose (het inducerende suiker voor de enzymen betrokken bij zetmeelafbraak) noch was de expressie afhankelijk van de aanwezigheid van de AmyR transcriptiefactor. Van slechts twee van de nieuw geïdentificeerde enzymen, een mogelijk α -glucosidase (AgdB) en een mogelijk α -amylase (AmyC), kan op grond van het expressieprofiel geconcludeerd worden dat ze een rol zouden kunnen spelen tijdens de groei van *A. niger* op zetmeel. De mogelijke fysiologische functies van de andere

voorspelde familie GH13, GH15 and GH31 enzymen, inclusief voorspelde intracellulaire enzymen en eiwitten die zijn geassocieerd met het celmembraan of de celwand in alternatieve α -glucan gerelateerde processen, worden besproken.

Van de nieuw geïdentificeerde enzymen uit de GH13 familie werden drie leden (AgtA, AgtB and AgtC) in detail verder bestudeerd. Deze eiwitten die hoge gelijkheid vertonen met schimmel α -amylases, zijn speciaal omdat ze waarschijnlijk via een zgn. GlycosylPhosphatidylInositol-anchor (GPI-anker) aan het plasmamembraan of aan de celwand gebonden zijn. Opvallend was ook dat ze enkele sterk geconserveerde aminozuren uit de α -amylase familie misten. **Hoofdstuk 4** beschrijft de functionele karakterisering van twee van deze eiwitten (AgtA and AgtB). Beide eiwitten werden gezuiverd uit *A. niger* stammen die deze eiwitten tot overexpressie brengen en werden biochemisch gekarakteriseerd. Door het analyseren van het groeidefekt van een *agtA* deletie stam en van stammen die *agtA* of *agtB* tot overexpressie brengen is de functie van de eiwitten verder bestudeerd. Zowel AgtA als AgtB vertonen een uniek type α -(1, 4) glucanotransferase (EC 2.4.1.25) activiteit. Deletie van *agtA* in *A. niger* resulteerde in een stam met een verhoogde gevoeligheid voor een stof die de opbouw van de celwand verstoort, Calcofluor White, wat duidt op een rol van dit eiwit tijdens de vorming van de celwand. Interessant is dat homologen van AgtA and AgtB ook aanwezig zijn in andere schimmelsoorten met α -glucan in hun celwand, maar niet in gist soorten die geen α -glucan in hun celwand hebben. Een mogelijke functie van deze enzymen is dus dat ze een rol spelen bij de synthese van het α -glucan in de celwand.

In **Hoofdstuk 5** onderzoeken we het *A. niger* genoom op de aanwezigheid van enzymen uit de GH32-familie, waarvan wordt vermoed dat ze betrokken zijn bij de afbraak van inuline. Twee nieuwe intracellulaire eiwitten, SucB en SucC, werden geïdentificeerd, naast drie al bekende extracellulaire inulinolitische enzymen, SucA (invertase), InuE (exo-inulinase) en InuA (endo-inulinase). Transcriptie-analyse liet zien dat de extracellulaire enzymen op een gecoördineerde manier tot expressie komen en dat de inductie plaatsvindt op inuline en sucrose. De transcriptie van deze genen was ook onder controle van de cataboliet repressor CreA. Verdere analyse gaf ook aan dat sucrose of een sucrose-afgeleid molecuul, maar niet fructose, als inducer optreedt voor de expressie van inulinolytische genen in *A. niger*.

In **Hoofdstuk 6** onderzoeken we of SucB, één van de twee nieuw ontdekte intracellulaire invertases, een essentiële rol speelt bij het maken van de inducer betrokken bij het inuline of sucrose katabolisme. Hiervoor zijn de phylogenetische, moleculaire and biochemische karakteristieken van SucB bestudeerd. Van het gezuiverde SucB eiwit is aangetoond dat het zich gedraagt als een invertase met transfructosylerende eigenschappen. Deletie van het *sucB* gen in *A. niger* resulteerde niet in verandering van de groei op inulin of sucrose, wat aantoont dat SucB geen essentiële rol lijkt te spelen tijdens inuline of sucrose katabolisme in *A. niger*. SucB zou daar en tegen nodig kunnen zijn voor intracellulaire omzetting van sucrose naar fructose en glucose of voor de hydrolyse van kleine fructo-oligosacchariden.

Zoals in Hoofdstuk 5 is beschreven werd de expressie van de genen die coderen voor de extracellulaire inulinolytische enzymen op eenzelfde wijze gereguleerd en geïnduceerd door sucrose en inuline. Dit suggereert dat er een gemeenschappelijke transcriptiefactor zou kunnen bestaan om de expressie van de genen te activeren in aanwezigheid van een inducer. In **Hoofdstuk 7** beschrijven we de identificatie en karakterisering van een Zn(II)2Cys6 type transcriptiefactor, die we InuR hebben genoemd (**Inu**line **R**egulator). De identificatie van *inuR* was gebaseerd op de inspectie van genclusters in het *A. niger* genoom. Het *inuR* gen is geclusterd met *sucB*, een mogelijk intracellulair invertase (Hoofdstuk 6) en een gen dat waarschijnlijk codeert voor een suikertransporter. Deletie van het *inuR* gen liet een sterke gereduceerde groei zien op inuline en sucrose. Inductie van de genen die coderen voor de extracellulaire inulinolytische enzymen en de suiker transporter naast de *inuR* transcriptiefactor in de aanwezigheid van sucrose en inulin was afhankelijk van InuR. Expressie-analyse van alle genen uit het *A. niger* genoom onthulde de inductie van drie additionele genen die werden geïnduceerd op sucrose op een InuR afhankelijke manier. Deze drie genen coderen voor mogelijke suikertransporters die mogelijk betrokken zijn bij het transport van sucrose of sucrose afgeleide suikers (glucose en fructose) over de plasmamembraan. De specificiteit van deze suikertransporters is een interessant onderwerp en wordt nu verder onderzocht dmv biochemische analyses. *In silico* analyse van de promotergebieden van de door InuR gereguleerde genen onthulde een waarschijnlijke bindingsplaats voor InuR. Deze bestaat uit twee CGG triplets, gescheiden door acht nucleotiden, en komt voorin alle InuR afhankelijke, sucrose geïnduceerde genen. Deze mogelijke bindingsplaats lijkt erg op de al bekende bindingsplaats van de AmyR transcriptiefactor (CGGN8(C/A)GG). Analyse van de dubbelmutant ($\Delta amyR \Delta inuR$) onthulde dat de twee transcriptiefactoren hoogstwaarschijnlijk onafhankelijk van elkaar functioneren. Hoe de InuR en AmyR transcriptiefactoren hun doelgenen bepalen en wat de specificiteit van de binding veroorzaakt is een interessant onderwerp voor verder onderzoek.

Onder de inulinolytische genen in *A. niger* hebben we het *inuE* gen geïdentificeerd als het sterkst geïnduceerde gen in de aanwezigheid van inuline en sucrose. In **Hoofdstuk 8** presenteren we, gebruikmakend van de *inuE* promoter en twee reporter genen (*racA-G12V* en Groen Fluorescerend Proteïne (GFP)), een nieuwe screeningsmethode voor de isolatie van mutanten die betrokken zijn bij inuline of sucrose signalering. Deze genetische screen, gevolgd door de identificatie van de gemuteerde genen, zou de identificatie mogelijk moeten maken van eiwitten die inducer moleculen genereren, transporteren of registreren. Ook zouden genen die coderen voor eiwitten die betrokken zijn bij de activatie van de InuR transcriptiefactor geïdentificeerd kunnen worden. Deze methode kan in het algemeen worden gebruikt voor de identificatie van transcriptionfactor mutanten in *A. niger* en wellicht ook in andere schimmels. Tot slot wordt aan het eind van hoofdstuk 8, gebaseerd op de laatste vier hoofdstukken (5 t/m 8), een speculatief model gepresenteerd waarin de mogelijke functie van verschillende eiwitten die betrokken zijn bij de afbraak en opname van inuline en inuline afbraakproducten door *A. niger* wordt beschreven.

黑曲霉是一种广泛存活于植物残体中的丝菌体式真菌。作为一种腐生菌，它能产生大量不同的水解酶，用来降解环境中植物的高分子多糖成分为小分子化合物，从而为其本身提供能源和碳源。这些酶大都很多（潜在的）用途，比如：在面包业，淀粉业，织物业以及食品和饲料工业等。这些用途近来吸引了大量的研究兴趣。通过遗传、生化、生物信息等途径来开发黑曲霉旨在更加清楚地了解其对糖类物质的降解、修饰网络，从而为提高酶的产量和底物的利用率提供科学依据。另外，从黑曲霉中发现新活性酶也是一个主要因素。近年来由荷兰的生物技术公司 DSM 和美国的 DOE 基因组研究所对黑曲霉的基因组测序的完成，以及黑曲霉的基因芯片技术的具备，使得从黑曲霉中发现新活性酶以及从基因组水平上来分析其机体内的碳水化合物（糖类）降解的转录调控网络成为可能。

这篇论文，我们的主要研究目标是探索黑曲霉降解、修饰碳水化合物的分子机制。比如，黑曲霉是如何感受到周围环境中存在的不同碳源的；某种特定诱导碳源的存在或产生又是如何激活这些降解酶基因的转录表达的。我们的研究重点在于寻找多糖类物质如淀粉、菊糖、以及它们的衍生物在黑曲霉体内的降解途径中出现的诱导分子和转录调节因子。另外，由这些结构基因编码的酶是如何催化不同的碳源的，以及这些酶的生理功能是什么，也是我们研究的重要方面。

本论文的**第一章**概括了黑曲霉作为酶产生菌的重要性以及当前有关淀粉和菊糖降解修饰酶的知识。有关碳水化合物降解的酶的转录表达和其蛋白产量在有机体内是高度调节的。这种调节主要由各种不同的广泛转录因子（如 CreA、AreA、PacC 等），和特定糖类降解途径的转录因子（如 AmyR 和 InuR）来控制的。这一章主要给读者了解有关调节基因表达的转录因子的一些背景信息，同时也详细阐述了与碳水化合物降解有关的酶编码基因的调节机制以及不同转录因子的一些重要特征。

为了发现降解淀粉、菊糖有关的新活性酶，我们采用了两种方法。第一种方法在**第二章**里阐述。该方法是利用克隆技术来构建淀粉和菊糖特定的黑曲霉基因表达文库。这些基因表达文库分别构建在以细菌和酵母为转化体的载体上。分子分析证明这些基因表达文库质量尚好并被转化到细菌和酵母体内。由于 cDNA 翻译的蛋白表达量以及蛋白表达转化体的数量的不足，目前只筛选出少量的转化体。该章末尾提出了几种改善 cDNA 表达文库的构建及筛选方法的建议和措施。

第二种方法是利用已知活性酶的序列来调查黑曲霉的基因组，从而获得相似序列编码的蛋白。在本文的**第三章**，我们介绍了植物淀粉可被 α -淀粉酶，葡萄糖化酶及 α -葡萄糖苷酶降解成小分子的葡萄糖。这些酶分别属于糖苷水解酶系 13、15 和 31 家族，因此在第三章里我们主要调查了黑曲霉基因组中可能编码糖苷水解酶系 13、15 和 31 家族的酶。结果我们

发现了 17 个新蛋白酶。通过基因组序列以及它们的转录表达水平调节的综合分析, 结果显示大部分基因表现出不同寻常的转录调节模式。也就是说, 它们既不受淀粉降解途径的诱导物-麦芽糖的诱导, 也不受淀粉降解调节因子-AmyR 的调控。只有两种酶, α -糖苷酶 AgdB 和 α -淀粉酶 AmyC, 可能具备降解淀粉的活性。其它的蛋白酶包括胞内酶、胞壁相关的酶, 很有可能涉及到细胞壁的组成成分 α -葡聚糖的修饰过程, 而不涉及淀粉的降解过程。这些假设在第三章中得到详细讨论。

在前一章提到的 GH13 谱系的蛋白中, 有三个蛋白酶 (AgtA、AgtB、AgtC) 尽管与真菌淀粉酶类有高度的序列相似性, 但其具有另外的 C 末端信号肽序列, 即糖基化磷脂酰肌醇 (GPI) 的附着信号序列, 并且缺乏某些 α -淀粉酶的高度保守的氨基序列。该论文的第四章描述了这些蛋白酶在黑曲霉中的过量表达、提取和纯化的功能性质, 以及其编码基因的删除和过量表达所表现出的生理表现特性。我们发现 AgtA 和 AgtB 均表现出糖苷转移酶的特征。其中 AgtA 的基因删除表现出对细胞壁干扰物 Calcofluor White (CFW) 强烈的敏感性, 表明该基因的缺失致使细胞壁内部结构的完整性受到破坏。蛋白序列分析显示 AgtA 和 AgtB 的同源蛋白也出现在其它有 α -葡聚糖成份的细胞壁的真菌中, 而不在缺乏 α -葡聚糖的细胞壁的酵母中。这一现象进一步表明 AgtA 蛋白酶活性与细胞壁结构有关。

糖苷水解酶系 32 家族蛋白有可能涉及到菊糖和它的合成底物-蔗糖的降解过程。在第五章里, 我们调查了黑曲霉基因组来寻找属于 GH32 谱系的基因。除了已知的三种胞外酶-SucA, InuE 和 InuA 外, 我们还发现了另外两种胞内酶, SucB 和 SucC。转录分析表明这些胞外酶是协同调节的并受菊糖和蔗糖诱导的。它们的转录水平也受到碳催化代谢抑制因子 CreA 的控制。进一步的实验分析表明在黑曲霉体内, 是蔗糖或它的衍生物, 而不是果糖, 作为诱导物来诱导菊糖分解酶基因的转录表达的。

在第六章里, 我们调查了其中一个 GH32 谱系的胞内降解酶-SucB, 检查其是否在菊糖或蔗糖的降解中起着产生诱导物的作用。我们对 SucB 蛋白的系统进化、分子和生化特性进行了研究。SucB 的过量表达并纯化的蛋白具有胞内降解蔗糖酶的特性并兼有转化果糖的活性。该基因的删除并没有改变黑曲霉在菊糖或蔗糖中生长特征, 表明 SucB 蛋白并不在菊糖或蔗糖的降解中起着产生诱导物的作用。但是 SucB 有可能在细胞内的蔗糖转化成果糖、葡萄糖或其它小分子的寡糖中行使一定的功能。

正如在第五章中提到的, 胞外菊糖分解酶类在转录水平上是协同调节的并受蔗糖和菊糖的诱导, 这个结果提示在菊糖的降解途径中存在着一个共同的转录调节因子, 在诱导物出现的情况下, 该调节因子激活其结构基因的转录表达。在本文的第七章中, 我们详细阐述了这个新的菊糖转录激活因子-InuR 的发现及其特征。InuR 的发现是基于黑曲霉基因组中菊糖酶类基因群的调查。inuR 基因靠近 SucB 基因, 二者之间隔着一个糖类运输蛋白基因 (An15g00310)。inuR 基因的删除导致黑曲霉在菊糖和蔗糖中的严重生长缺陷。菊糖和蔗糖对胞外菊糖酶基因以及这糖运输蛋白基因的诱导是依赖于 InuR 调节因子的。基因组的转录表达分析发现另外三个糖运输蛋白基因 (An15g04060, An15g03940 and An17g01710), 它们也受到蔗糖的强烈诱导并且依赖于 InuR。这些糖运输蛋白基因的具体功能, 特别是 An15g00310 编码的糖运输蛋白是个很有意义的课题并且正在用生化分析的手段研究中。由电脑模拟分析所有 InuR 调控的基因的启动子我们发现一个潜在的 InuR 结合位点-CCGN8CGG, 该序列很类似于淀粉降解调节因子-AmyR 结合位点-CCGN8C/AGG。双重基因的删除分析揭示了 InuR

和 AmyR 的独立调控功能。究竟 InuR 和 AmyR 是如何辨认它们各自的目标基因以及什么因素决定它们的结合特异性将是个有意义的课题并有待于进一步研究。

在这组菊糖酶基因谱中，我们发现 *inuE* 是受到菊糖和蔗糖诱导最强的基因。在本文的**第八章**，我们利用 *inuE* 基因的启动子，分别结合 *racA-G12V* 和 GFP 报告基因，找到了一个新的筛选方法来分离出菊糖或蔗糖信号途径中的突变体。通过这遗传筛选方法结合突变体的辨别我们能够找到那些产生诱导分子，运输或感应诱导分子的蛋白，以及那些激活菊糖降解途径的酶基因转录因子。该方法可以被广泛运用到寻找转录因子的突变体。在建立这个可能性的筛选方法的过程中，我们发现棉籽糖也有诱导菊糖酶基因的特性。

基于本论文最后四个章节的结果，我们提出了一个黑曲霉体内菊糖降解信号途径的假说模型。该**假说模型**具体陈述在第八章的末尾。

References

- Akache B, Wu K, Turcotte B (2001) Phenotypic analysis of genes encoding yeast zinc cluster proteins. *Nucleic Acids Res* 29:2181-2190
- Akimoto H, Kushima T, Nakamura T, Ohta K (1999) Transcriptional analysis of two endoinulinase genes *inuA* and *inuB* in *Aspergillus niger* and nucleotide sequences of their promoter regions. *J Biosci Bioeng* 88:599-604
- Akimoto H, Kiyota N, Kushima T, Nakamura T, Ohta K (2000) Molecular cloning and sequence analysis of an endoinulinase gene from *Penicillium* sp. strain TN-88. *Biosci Biotechnol Biochem* 64:2328-2335
- Alberto F, Bignon C, Sulzenbacher G, Henrissat B, Czjzek M (2004) The three-dimensional crystal structure of invertase (β -fructosidase) from *Thermotoga maritima* reveals a bimodular arrangement and an evolutionary relationship between retaining and inverting glycosidases. *J Biol Chem* 279:18903-18910
- Alberto F, Jordi E, Henrissat B, Czjzek M (2006) Crystal structure of inactivated *Thermotoga maritima* invertase in complex with the trisaccharide substrate raffinose. *Biochem. J.* 395:457-462
- Arand M, Golubev AM, Neto JR, Polikarpov I, Wattiez R, Korneeva OS, Eneyskaya EV, Kulminskaya AA, Shabalin KA, Shishliannikov SM, Chepurnaya OV, Neustroev KN (2002) Purification, characterization, gene cloning and preliminary X-ray data of the exo-inulase from *Aspergillus awamori*. *Biochemistry Journal* 362:131-135
- Archer DB, Peberdy JF (1997) The molecular biology of secreted enzyme production by fungi. *Crit Rev Biotechnol.* 17(4):273-306
- Arguelles JC (2000) Physiological roles of trehalose in bacteria and yeasts: a comparative analysis. *Arch Microbiol* 174:217-224
- Ascone I, Lenouvel F, Sequeval D, Dexpert H, Felenbok B (1997) First experimental evidence of a zinc binuclear cluster in AlcR protein, mutational and X-ray absorption studies. *Biochim Biophys Acta.* 1343(2):211-220
- Baneyx F (1999) Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol.* 10:411-421
- Bardalaye PC, Nordin JH (1977) Chemical structure of the galactomannan from the cell wall of *Aspergillus niger*. *J. Biol. Chem.* 252:2584-2591
- Barker SA, Carrington TR (1953) Studies of *Aspergillus niger*. Part II. Transglycosylation by *Aspergillus niger*. *J Chem Soc* 1953:3588-3593
- Baker SE (2006) *Aspergillus niger* genomics: past, present and into the future. *Med Mycol.* 44: Suppl 1:S17-21. Review
- Bauer S, Vasu P, Persson S, Mort AJ, Somerville CR (2006) Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. *Proc Natl Acad Sci U S A* 103:11417-11422
- Beauvais A, Latgé JP (2001) Membrane and cell wall targets in *Aspergillus fumigatus*. *Drug Resist Updat* 4:38-49
- Beauvais A, Maubon D, Park S, Morelle W, Tanguy M, Huerre M, Perlin DS, Latgé JP (2005) Two α (1-3) glucan synthases with different functions in *Aspergillus fumigatus*. *Appl. Environ. Microbiol.* 71:1531-1538
- Bendtsen JD, Nielsen H, Von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340:783-795
- Bendtsen JD, Jensen LJ, Blom N, Von Heijne G, Brunak S (2004) Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng Des Sel* 17:349-356
- Benen JA, Kester HC, Visser J (1999) Kinetic characterization of *Aspergillus niger* N400 endopolygalacturonases I, II and C. *Eur. J. Biochem.* 259:577-585
- Bennet JW, Lasure LL (1991) Growth media. In: Bennet JW and Lasure LL (eds) *More gene manipulations in fungi*. Academic Press, San Diego, pp. 441-447
- Bennett JW, Lasure LL (1991) *More gene manipulations in fungi*. Academic Press, San Diego
- Berger H, Pachlinger R, Morozov I, Goller S, Narendja F, Caddick M, Strauss J (2006) The GATA factor AreA regulates localization and *in vivo* binding site occupancy of the nitrate activator NirA. *Mol Microbiol.* 59(2):433-446

- Berges T, Barreau C, Peberdy JF, Boddy LM (1993) Cloning of an *Aspergillus niger* invertase gene by expression in *Trichoderma reesei*. *Curr Genet* 24:53-59
- Berka RM, Dunn-Coleman NS, Ward M (1992) Industrial enzymes from Aspergilli. In: Bennett JW, Klich MA (eds) *Aspergillus*, biology and industrial applications. Butterworth-Heinemann, London, pp 155-202
- Bhatti HN, Asgher M, Abbas A, Nawaz R, Sheikh MA (2006) Studies on kinetics and thermostability of a novel acid invertase from *Fusarium solani*. *J. Agric. Food Chem.* 54:4617-4623
- Bigelis R and Lasure LL (1987) Fungal enzymes and primary metabolites used in food processing. In: Food and beverage mycology. Beuchat, L.R (ed.) Van Nostrand Reinhold, New York, USA, pp 473-516
- Birney E, Clamp M, Durbin R (2004) GeneWise and Genomewise. *Genome Res* 14:988-995
- Blumenthal HJ, Roseman S (1957) Quantitative estimation of chitin in fungi. *J. Bacteriol.* 74:222-224
- Bobbitt TF, Nordin JH, Roux M, Revol JF, Marchessault RH (1977) Distribution and conformation of crystalline nigeran in hyphal walls of *Aspergillus niger* and *Aspergillus awamori*. *J. Bacteriol.* 132:691-703
- Boddy LM, Berges T, Barreau C, Vainstein C, Dobson MJ, Balance DJ, Peberdy JF (1993) Purification and characterization of an *Aspergillus niger* invertase and its DNA sequence. *Curr Genet* 24:60-66
- Boel E, Brady L, Brzozowski AM, Derewenda Z, Dodson GG, Jensen VJ, Petersen SB, Swift H, Thim L, Woldike HF (1990) Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-Å resolution of two enzymes from *Aspergillus niger*. *Biochemistry* 29 (26):6244-6249
- Boel E, Hjort I, Svensson B, Norris F, Norris KE, Fiil NP (1984) Glucoamylases G1 and G2 from *Aspergillus niger* are synthesized from two different but closely related mRNAs. *EMBO J* 3:1097-1102
- Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, Seiler S, Bell-Pedersen D, Paietta J, Plesofsky N, Plamann M, Goodrich-Tanrikulu M, Schulte U, Mannhaupt G, Nargang FE, Radford A, Selitrennikoff C, Galagan JE, Dunlap JC, Loros JJ, Catchside D, Inoue H, Aramayo R, Polymenis M, Selker EU, Sachs MS, Marzluf GA, Paulsen I, Davis R, Ebbola DJ, Zelter A, Kalkman ER, O'Rourke R, Bowring F, Yeadon J, Ishii C, Suzuki K, Sakai W, Pratt R (2004) Lessons from the genome sequence of *Neurospora crassa*: tracing the path from genomic blueprint to multicellular organism. *Microbiol Mol Biol Rev* 68:1-108. Review
- Bos CJ, Debets AJ, Swart K, Huybers A, Kobus G, Slakhorst SM (1988) Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in *Aspergillus niger*. *Curr Genet* 14:437-443
- Brady RL, Brzozowski AM, Derewenda ZS, Dodson EJ, Dodson GG (1991) Solution of the structure of *Aspergillus niger* acid alpha-amylase by combined molecular replacement and multiple isomorphous replacement methods. *Acta Crystallogr B.* 47:527-535
- Brul S, King A, van der Vaart JM, Chapman J, Klis F, Verrips CT (1997) The incorporation of mannoproteins in the cell wall of *S. cerevisiae* and filamentous *Ascomycetes*. *Antonie Van Leeuwenhoek* 72:229-237.
- Buckholz RG, Gleeson MA (1991) Yeast systems for the commercial production of heterologous proteins. *Biotechnology (N Y)*. 9(11):1067-1072
- Burger G, Strauss J, Scazzocchio C, Lang BF (1991) *nirA*, the pathway-specific regulatory gene of nitrate assimilation in *Aspergillus nidulans*, encodes a putative GAL4-type zinc finger protein and contains four introns in highly conserved regions. *Mol Cell Biol* 11:5746-5755
- Cahuzac B, Cerdan R, Felenbok B, Guittet E (2001) The solution structure of an AlcR-DNA complex sheds light onto the unique tight and monomeric DNA binding of a Zn(2)Cys(6) protein. *Structure* 9(9):827-836
- Cairns, A. J. (2003). Fructan biosynthesis in transgenic plants. *J Exp Bot* 54, 549-567
- Carey M, Kakidani H, Leatherwood J, Mostashari F, Ptashne M (1989) An amino-terminal fragment of GAL4 binds DNA as a dimer. *J Mol Biol.* 209(3):423-432
- Carpita NC, Kanabus J, Housley TL (1989) Linkage structure of fructans and fructan oligomers from *Triticum aestivum* and *Festuca arundinaceae* leaves. *J Plant Physiol* 134:162-168
- Cazelle B, Pokorska A, Hull E, Green PM, Stanway G, Scazzocchio C (1998) Sequence, exon-intron organization, transcription and mutational analysis of *prnA*, the gene encoding the transcriptional activator of the *prn* gene cluster in *Aspergillus nidulans*. *Mol Microbiol* 28:355-370
- Chang CT, Lo HF, Chi MC, Yao CY, Hsu WH, Lin LL (2003) Identification of essential histidine residues in a recombinant alpha-amylase of thermophilic and alkaliphilic *Bacillus sp.* strain TS-23. *Extremophiles* 7:505-509
- Chavez FP, Pons T, Delgado JM, Rodriguez L (1998) Cloning and sequence analysis of the gene encoding invertase (INV1) from the yeast *Candida utilis*. *Yeast* 14, 1223-1232
- Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* 31:3497-3500

- Chiou CY, Wang HH, Shaw GC (2002) Identification and characterization of the non-PTS *fru* locus of *Bacillus megaterium* ATCC 14581. *Mol Genet Genomics* 268:240-248
- Christgau S, Sandal T, Kofod LV, Dalbøge H (1995) Expression cloning, purification and characterization of a beta-1,4-galactanase from *Aspergillus aculeatus*. *Curr Genet* 27(2):135-141
- Coussement PAA (1999) Inulin and oligofructose: safe intakes and legal status. *J. Nutr.* 129:1412S-1417S
- Coutinho PM, Henrissat B (1999) Carbohydrate-active enzymes: an integrated database approach. In: Gilbert HJ, Davies GJ, Henrissat B, Svensson B (eds) *Recent Advances in Carbohydrate Bioengineering*. The Royal Society of Chemistry, Cambridge, pp. 3-12
- Crabeel M, de Rijke M, Seneca S, Heimberg H, Pfeiffer I, Matisova A (1995) Further definition of the sequence and position requirements of the arginine control element that mediates repression and induction by arginine in *Saccharomyces cerevisiae*. *Yeast* 11:1367-1380
- Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res.* 14:1188-1190
- 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
- Da Silva NA, Bailey JE (1991) Influence of plasmid origin and promoter strength in fermentations of recombinant yeast. *Biotechnology and Bioengineering* 37:318-324
- Dalbøge H, Heldt-Hansen HP (1994) A novel method for efficient expression cloning of fungal enzyme genes. *Mol Gen Genet* 243:253-260
- Damveld RA, Vankuyk PA, Arentshorst M, Klis FM, van den Hondel CA, Ram AF (2005) Expression of *agsA*, one of five 1,3-alpha-d-glucan synthase-encoding genes in *Aspergillus niger*, is induced in response to cell wall stress. *Fungal Genet Biol* 42:165-177
- Damveld RA, Arentshorst M, VanKuyk PA, Klis FM, van den Hondel CA, Ram AF (2005) Characterisation of CwpA, a putative glycosylphosphatidylinositol-anchored cell wall mannoprotein in the filamentous fungus *Aspergillus niger*. *Fungal. Genet. Biol.* 42:873-885
- Darwen CWE, John P (1989) Localisation of the enzymes of fructan metabolism in vacuoles isolated by a mechanical method from tubers of Jerusalem artichoke (*Helianthus tuberosus* L.) *Plant Physiol* 89:658-663
- Davies RW (1994) Heterologous gene expression and protein secretion in *Aspergillus niger*. *Prog Ind Microbiol.* 29:527-560. Review
- Davies GJ, Wilson KS, Henrissat B (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* 321:557-559
- De Groot PW, Hellingwerf KJ and Klis FM (2003) Genome-wide identification of fungal GPI proteins. *Yeast* 20:781-796
- de Ruiter-Jacobs YM, Broekhuijsen M, Unkles SE, Campbell EI, Kinghorn JR, Contreras R, Pouwels PH, van den Hondel CA (1989) A gene transfer system based on the homologous *pyrG* gene and efficient expression of bacterial genes in *Aspergillus oryzae*. *Curr Genet* 16:159-163
- De Vries RP (2003) Regulation of *Aspergillus* genes encoding plant cell wall polysaccharide-degrading enzymes; relevance for industrial production. *Appl Microbiol Biotechnol.* 61:10-20
- de Vries RP, Frisvad JC, van de Vondervoort PJ, Burgers K, Kuijpers AF, Samson RA, Visser J (2005a) *Aspergillus vadensis*, a new species of the group of black Aspergilli. *Antonie Van Leeuwenhoek.* 87(3):195-203
- de Vries, R. P., van Grieken, C., van Kuyk, P. A. & Wösten, H. A. B. (2005b). The value of genome sequences in the rapid identification of novel genes encoding specific plant cell wall degrading enzymes. *Curr Genom* 6, 157-187
- De Vries RP, Visser J (2001) *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. *Microbiol Mol Biol Rev* 65:497-522
- den Herder IF, Rosell AM, van Zuijlen CM, Punt PJ, van den Hondel CA (1992) Cloning and expression of a member of the *Aspergillus niger* gene family encoding alpha-galactosidase. *Mol Gen Genet* 233:404-410
- Diep CQ, Peng G, Bewley M, Pilaury V, Ropson I, Hopper JE (2006) Intragenic suppression of Gal3C interaction with Gal80 in the *Saccharomyces cerevisiae* GAL gene switch. *Genetics.* 172(1):77-87
- Dowzer CE & Kelly JM (1991) Analysis of the *creA* gene, a regulator of carbon catabolite repression in *Aspergillus nidulans*. *Mol Cell Biol.* 11:5701-5709
- Drysdale MR, Kolze SE, Kelly JM (1993) The *Aspergillus niger* carbon catabolite repressor encoding gene, *creA*. *Gene* 130:241-245
- Duarte JC, Costa-Ferreira M (1994) Aspergilli and lignocellulosics: enzymology and biotechnological applications. *FEMS Microbiol. Rev.* 13:377-386
- Durbin R, Eddy S (1998) Biological Sequence analysis: Probabilistic Models of Proteins and Nucleic acids. A tutorial introduction to hidden Markov models and other probabilistic modeling approaches in computational sequence analysis. Cambridge University Press, Cambridge

- Eddy SR (1998) Profile hidden Markov models. *Bioinformatics* 14:755-763
- 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
- El Alami M, Feller A, Pierard A, Dubois E (2000) Characterization of a tripartite nuclear localization sequence in the regulatory protein Lys14 of *Saccharomyces cerevisiae*. *Curr Genet* 38:78-86
- Espeso EA, Tilburn J, Sanchez-Pulido L, Brown CV, Valencia A, Arst HN Jr, Penalva MA (1997) Specific DNA recognition by the *Aspergillus nidulans* three zinc finger transcription factor PacC. *J Mol Biol*. 274(4):466-480
- Fakhoury AM, Woloshuk CP (1999) Amy1, the alpha-amylase gene of *Aspergillus flavus*: Involvement in aflatoxin biosynthesis in maize kernels. *Phytopathology* 89:908-914
- Felenbok B, Flippin M, Nikolaev I (2001) Ethanol catabolism in *Aspergillus nidulans*: a model system for studying gene regulation. *Prog Nucleic Acid Res Mol Biol*. 69:149-204. Review
- Fillinger S, Felenbok B (1996) A newly identified gene cluster in *Aspergillus nidulans* comprises five novel genes localized in the alc region that are controlled both by the specific transactivator AlcR and the general carbon-catabolite repressor CreA. *Mol Microbiol*. 20:475-88
- Flippin M, Kocialkowska J, Felenbok B (2002) Characteristics of physiological inducers of the ethanol utilization (alc) pathway in *Aspergillus nidulans*. *Biochem J*. 364:25-31
- Flynn PJ, Reece RJ (1999) Activation of transcription by metabolic intermediates of the pyrimidine biosynthetic pathway. *Mol Cell Biol* 19:882-888
- Fontaine T, Simenel C, Dubreucq G, Adam O, Delepierre M, Lemoine J, Vorgias CE, Diaquin M, Latgé JP (2000) Molecular organization of the alkali-insoluble fraction of *Aspergillus fumigatus* cell wall. *J Biol Chem* 275: 41528
- Fowler T, Berka RM, Ward M (1990) Regulation of the *glaA* gene of *Aspergillus niger*. *Curr Genet* 18:537-545
- Franck A (2002) Technological functionality of inulin and oligofructose. *Br. J. Nutr.* 87 Suppl 2:S287-291
- Frost GM, Moss DA (1987) Production of enzymes by fermentation. In: Rehm HJ, Reed G (eds) *Biotechnology*, vol 7a. VCH, Weinheim, pp 65-102
- Galagan JE, Calvo SE, Cuomo C, Ma LJ, Wortman JR, Batzoglou S, Lee SI, Basturkmen M, Spevak CC, Clutterbuck J, Kapitonov V, Jurka J, Scazzocchio C, Farman M, Butler J, Purcell S, Harris S, Braus GH, Drahí O, Busch S, d'Enfert C, Bouchier C, Goldman GH, Bell-Pedersen D, Griffiths-Jones S, Doonan JH, Yu J, Vienken K, Pain A, Freitag M, Selker EU, Archer DB, Penalva MA, Oakley BR, Momany M, Tanaka T, Kumagai T, Asai K, Machida M, Niernan WC, Denning DW, Caddick M, Hynes M, Paoletti M, Fischer R, Miller B, Dyer P, Sachs MS, Osmani SA, Birren BW (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438:1105-1115
- Gangwani L, Mikrut M, Galcheva-Gargova Z, Davis RJ (1998) Interaction of ZPR1 with translation elongation factor-1alpha in proliferating cells. *J Cell Biol* 143:1471-1484
- Gascon, S. and J. O. Lampen. 1968. Purification of the internal invertase of yeast. *J. Biol. Chem.* 243:1567-1572
- Gascon 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
- Geber A, Williamson PR, Rex JH, Sweeney EC, Bennett JE (1992) Cloning and characterization of a *Candida albicans* maltase gene involved in sucrose utilization. *J Bacteriol* 174:6992-6996
- Gellissen G, Melber K, Janowicz ZA, Dahlems UM, Weydemann U, Piontek M, Strasser AW, Hollenberg CP (1992) Heterologous protein production in yeast. *Antonie Van Leeuwenhoek*. 62(1-2):79-93. Review
- Geysens S, Pakula T, Uusitalo J, Dewerte I, Penttilä M, Contreras R (2005) Cloning and characterization of the glucosidase II alpha subunit gene of *Trichoderma reesei*: a frameshift mutation results in the aberrant glycosylation profile of the hypercellulolytic strain Rut-C30. *Appl Environ Microbiol* 71:2910-2924
- Gielkens MM, Dekkers E, Visser J, de Graaff LH (1999) Two cellobiohydrolase-encoding genes from *Aspergillus niger* require D-xylose and the xylanolytic transcriptional activator XlnR for their expression. *Appl Environ Microbiol* 65:4340-4345
- Gomez D, Cubero B, Cecchetto G, Scazzocchio C (2002) PrnA, a Zn2Cys6 activator with a unique DNA recognition mode, requires inducer for *in vivo* binding. *Mol Microbiol* 44:585-597
- Gomi K, Akeno T, Minetoki T, Ozeki K, Kumagai C, Okazaki N, Iimura Y (2000) Molecular cloning and characterization of a transcriptional activator gene, *amyR*, involved in the amylolytic gene expression in *Aspergillus oryzae*. *Biosci Biotechnol Biochem*. 64(4):816-27
- Goosen C, Yuan XL, van Munster JM, Ram AF, van der Maarel MJ, Dijkhuizen L (2007) Molecular and biochemical characterization of a novel intracellular invertase from *Aspergillus niger* with transfructosylating activity. *Eukaryot Cell* 6:674-681

- Gouka RJ, Hessing JG, Stam H, Musters W, van den Hondel CA. (1995) A novel strategy for the isolation of defined *pyrG* mutants and the development of a site-specific integration system for *Aspergillus awamori*. *Curr Genet.* 27(6):536-540.
- Grassin C, Fauguenbergue P (1999) Enzymes, fruit juice processing. In: Flickinger MC, Drew SW (eds) *Encyclopedia of bioprocess technology: fermentation, biocatalysis, and bioseparation*. Wiley & Sons, Inc, New York
- Gronenborn AM (2005) The DNA-binding domain of GATA transcription factors - a prototypical type IV Cys2-Cys2 Zinc finger. In zinc finger proteins: from atomic contact to cellular function. Iuchi S and Kuldell N (eds). Landes Bioscience, Georgetown, TX. pp. 26-30
- Grootwassink JWD, Hewitt GM (1983) Inducible and constitutive formation of β -fructofuranosidase (inulase) in batch and continuous cultures of the yeast *Kluyveromyces fragilis*. *J Gen Microbiol* 129:31-41
- Grün CH, Hochstenbach F, Humbel BM, Verkleij AJ, Sietsma JH, Klis FM, Kamerling JP, Vliegenthart JF (2005) The structure of cell wall alpha-glucan from fission yeast. *Glycobiology* 15:245-257
- Guasconi V, Yahi H, Ait-Si-Ali S (2002) Transcription factors. *Atlas Genet Cytogenet Oncol Haematol*. URL: <http://www.infobiogen.fr/services/chromcancer/Deep/TranscripFactorsID20043.html>
- Gupta AK and Kaur N (1997) Fructan storing plants - a potential source of high fructose syrups. *J Sci Ind Res* 56:447-452
- Gupta AK and Kaur N (eds) (2000) Fructan metabolism in Jerusalem artichoke and chicory; in Carbohydrate reserves in plants - synthesis and regulation (The Netherlands: Elsevier Science) pp 223-248
- Hannig G, Makrides SC (1998) Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends Biotechnol*, 16:54-60
- Hartley JL, Temple GF, Brasch MA (2000) DNA cloning using in vitro site-specific recombination. *Genome Res.* 10:1788-1795
- Hasper AA, Trindade LM, van der Veen D, van Ooyen AJ, de Graaff LH (2004) Functional analysis of the transcriptional activator XlnR from *Aspergillus niger*. *Microbiology* 150:1367-1375
- Hata Y, Ishida H, Kojima Y, Ishikawa E, Kawato A, Suginami K, Imayasu S (1997) Comparison of two glucoamylases produced by *Aspergillus oryzae* in solid-state (koji) and in submerged culture. *J. Ferment. Bioeng.* 84:532-537
- Hata Y, Kitamoto K, Gomi K, Kumagai C, Tamura G (1992) Functional elements of the promoter region of the *Aspergillus oryzae glaA* gene encoding glucoamylase. *Curr Genet.* 22(2):85-91
- Hata Y, Tsuchiya K, Kitamoto K, Gomi K, Kumagai C, Tamura G, Hara S (1991) Nucleotide sequence and expression of the glucoamylase-encoding gene (*glaA*) from *Aspergillus oryzae*. *Gene* 108:145-150
- Hawksworth DL (2001) The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycol. Res.* 105:1422-1432
- Hayashida S, Kuroda K, Ohta K, Kuhara S, Fukuda K and Sakaki Y (1989) Molecular cloning of the glucoamylase I gene of *Aspergillus awamori* var. *kawachi* for localization of the raw-starch-affinity site. *Agric. Biol. Chem.* 53:923-929
- Hendry G (1993) Evolutionary origins and natural functions of fructans: A climatological, biogeographic and mechanistic appraisal. *New Phytol.* 123:3-14
- Hendry GAF, Wallace RK (1993) The origin, distribution and evolutionary significance of fructans. In *Science and technology of fructans*. pp 119-139. Edited by Suzuki, M. & Chatterton, N. J. Boca Raton, FL: CRC Press
- Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J.* 280:309-316
- Henrissat B, Bairoch A (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J.* 293:781-788
- Hesse SJ, Ruijter GJ, Dijkema C, Visser J (2002) Intracellular pH homeostasis in the filamentous fungus *Aspergillus niger*. *Eur. J. Biochem.* 269:3485-3494
- Heyer AG, Wendenburg R (2001) Gene cloning and functional characterization by heterologous expression of the fructosyltransferase of *Aspergillus sydowii* IAM 2544. *Appl Environ Microbiol* 67:363-370
- Hirayama M, Sumi N, Hidaka H (2006) Purification and characterization of a fructooligosaccharide-producing beta-fructofuranosidase from *Aspergillus niger* ATCC 20611. *Agric. Biol. Chem.* 53:667-673
- Hochstenbach F, Klis FM, Van Den EH, van DE, Peters PJ, Klausner RD (1998) Identification of a putative alpha-glucan synthase essential for cell wall construction and morphogenesis in fission yeast. *Proc Natl Acad Sci USA* 95:9161-9166

- Holz C, Prinz B, Bolotina N, Sievert V, Büsow K, Simon B, Stahl U, Lang C (2003) Establishing the yeast *Saccharomyces cerevisiae* as a system for expression of human proteins on a proteomescale. *J Struct Funct Genomics*, 4:97-108
- Horisberger M, Lewis BA, Smith F (1972) Structure of a (1, 3)-a-D-glucan (pseudonigeran) of *Aspergillus niger* NNRL 326 cell wall. *Carbohydr Res* 23:183-188
- Hoshiko S, Makabe O, Nojiri C, Katsumata K, Satoh E, Nagaoka K (1987) Molecular cloning and characterization of the *Streptomyces hygroscopicus* alpha-amylase gene. *J. Bacteriol.* 169:1029-1036
- Hull EP, Green PM, Arst HN, Jr., Scazzocchio C (1989) Cloning and physical characterization of the L-proline catabolism gene cluster of *Aspergillus nidulans*. *Mol Microbiol* 3:553-559
- Inoue H, Nojima H, Okayama H (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 96:23-28
- Ito T, Tani S, Itoh T, Tsukagoshi N, Kato M, Kobayashi T (2004) Mode of AmyR binding to the CCGN8AGG sequence in the *Aspergillus oryzae* taaG2 promoter. *Biosci Biotechnol Biochem* 68:1906-1911
- Jespersen HM, MacGregor EA, Sierks MR, Svensson B (1991) Comparison of the domain-level organization of starch hydrolases and related enzymes. *Biochem. J.* 280:51-55. Review
- Johnston IR (1965) The composition of the cell wall of *Aspergillus niger*. *Biochem J* 96:651-658
- Johnston M (1987) A model fungal gene regulatory mechanism: the GAL genes of *Saccharomyces cerevisiae*. *Microbiol Rev.* 51(4):458-476
- Johnston IR (1965) The composition of the cell wall of *Aspergillus niger*. *Biochem. J.* 96:651-658
- Kaneko A, Sudo S, Sakamoto Y, Tamura G, Ishikawa T, Ohba T (1996) Molecular-cloning and determination of the nucleotide-sequence of a gene encoding an acid-stable alpha-amylase from *Aspergillus kawachii*. *J. Ferment. Bioeng.* 81:292-298
- Kang HK, Lee JH, Kim D, Day DF, Robyt JF, Park KH, Moon TW (2004) Cloning and expression of *Lipomyces starkeyi* alpha-amylase in *Escherichia coli* and determination of some of its properties. *FEMS Microbiol. Lett.* 233:53-64
- Kaper T, van der Maarel MJ, Euverink GJ, Dijkhuizen L (2004) Exploring and exploiting starch modifying amylomaltases from thermophiles. *Biochem. Soc. Trans.* 32:279-282
- Kaplan H, Hutkins RW (2003) Metabolism of fructooligosaccharides by *Lactobacillus paracasei* 1195. *Appl Environ Microbiol* 69, 2217-2222
- Karnaoukhova L, Taylor M, Neiditch B (2003) Construction of cDNA libraries by recombination using the CloneMiner™ cDNA Library Construction Kit. *FOCUS* 25.2:320-324
- Kato M, Kato Y, Kin T, Kokubun A, Maeda H, Maeyama N, Maruyama J, Nagasaki H, Nakajima T, Oda K, Okada K, Paulsen I, Sakamoto K, Sawano T, Takahashi M, Takase K, Terabayashi Y, Wortman JR, Yamada O, Yamagata Y, Anazawa H, Hata Y, Koide Y, Komori T, Koyama Y, Minetoki T, Suharnan S, Tanaka A, Isono K, Kuhara S, Ogasawara N, Kikuchi H (2005) Genome sequencing and analysis of *Aspergillus oryzae*. *Nature* 438:1157-1161
- Kato M, Sekine K, Tsukagoshi N (1996) Sequence-specific binding sites in the Taka-amylase A G2 promoter for the CreA repressor mediating carbon catabolite repression. *Biosci Biotechnol Biochem.* 60:1776-1779
- Kato N, Murakoshi Y, Kato M, Kobayashi T, Tsukagoshi N (2002a) Isomaltose formed by alpha-glucosidases triggers amylase induction in *Aspergillus nidulans*. *Curr Genet* 42:43-50
- Kato N, Suyama S, Shirokane M, Kato M, Kobayashi T, Tsukagoshi N (2002b) Novel alpha-glucosidase from *Aspergillus nidulans* with strong transglycosylation activity. *Appl Environ Microbiol* 68:1250-1256
- Kaur N, Gupta AK (2002) Applications of inulin and oligofructose in health and nutrition. *J Biosci* 27:703-714
- Keller NP, Hohn TM. (1997) Metabolic pathway gene clusters in filamentous fungi. *Fungal. Genet. Biol.* 21:17-29
- Kelly R, Kwon-Chung KJ (1992) A zinc finger protein from *Candida albicans* is involved in sucrose utilization. *J Bacteriol* 174:222-232
- Kelly JM, Hynes MJ (1985) Transformation of *Aspergillus niger* by the *amdS* gene of *Aspergillus nidulans*. *EMBO J.* 4:475-479
- Kim HY, Choi YJ (2001) Molecular characterization of cyclo-inulooligosaccharide fructanotransferase from *Bacillus macerans*. *Appl Environ Microbiol* 67:995-1000
- Kimura A, Takata M, Sakai O, Matsui H, Takai N, Takayanagi T, Nishimura I, Uozumi T, Chiba S (1992) Complete amino acid sequence of crystalline alpha-glucosidase from *Aspergillus niger*. *Biosci. Biotechnol. Biochem.* 56:1368-1370
- Kirkpatrick CR, Schimmel P (1995) Detection of leucine-independent DNA site occupancy of the yeast Leu3p transcriptional activator *in vivo*. *Mol Cell Biol* 15:4021-4030
- Kitamoto HK, Ohmomo S, Iimura Y (1998) Isolation and nucleotide sequence of the gene encoding phosphoenolpyruvate carboxykinase from *Kluyveromyces lactis*. *Yeast* 14:963-967

- Kohlhaw GB (2003) Leucine biosynthesis in fungi: entering metabolism through the back door. *Microbiol Mol Biol Rev* 67:1-15
- Kolar M, Punt PJ, van den Hondel CA, Schwab H (1988) Transformation of *Penicillium chrysogenum* using dominant selection markers and expression of an *Escherichia coli lacZ* fusion gene. *Gene* 62:127-134
- Korman DR, Bayliss FT, Barnett CC, Carmona CL, Kodama KH, Royer TJ, Thompson SA, Ward M, Wilson LJ, Berka RM (1990) Cloning, characterization, and expression of two alpha-amylase genes from *Aspergillus niger* var. *awamori*. *Curr Genet* 17:203-212
- Kralj S, van Geel-Schutten GH, van der Maarel MJ, Dijkhuizen L (2004) Biochemical and molecular characterization of *Lactobacillus reuteri* 121 reuteransucrase. *Microbiology* 150:2099-2112
- Kralj S, van Geel-Schutten GH, Dondorff MM, Kirsanovs S, van der Maarel MJ, Dijkhuizen L (2004) Glucan synthesis in the genus *Lactobacillus*: isolation and characterization of glucansucrase genes, enzymes and glucan products from six different strains. *Microbiology* 150:3681-3690
- Krappmann S, Braus GH (2005) Nitrogen metabolism of *Aspergillus* and its role in pathogenicity. *Med Mycol* 43 Suppl 1:S31-40. Review
- 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
- Kulmburg P, Prangé T, Mathieu M, Sequeval D, Scazzocchio C, Felenbok B (1991) Correct intron splicing generates a new type of a putative zinc-binding domain in a transcriptional activator of *Aspergillus nidulans*. *FEBS Lett* 280(1):11-16
- Kulmburg P, Sequeval D, Lenouvel F, Mathieu M, Felenbok B (1992) Identification of the promoter region involved in the autoregulation of the transcriptional activator ALCR in *Aspergillus nidulans*. *Mol Cell Biol* 12:1932-1939
- Kulminskaya AA, Arand M, Eneyskaya EV, Ivanen DR, Shabalin KA, Shishlyannikov SM, Saveliev AN, Korneeva OS, Neustroev KN (2003) Biochemical characterization of *Aspergillus awamori* exoinulinase: substrate binding characteristics and regioselectivity of hydrolysis. *Biochim Biophys Acta* 1650, 22-29
- Kumar S, Tamura K, Nei M (1993) MEGA: Molecular Evolutionary Analysis, version 1.01., The Pennsylvania State University, University Park, PA
- Kuriki T, Imanaka T (1999) The concept of the alpha-amylase family: structural similarity and common catalytic mechanism. *J. Biosci. Bioeng.* 87:557-565
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685
- Laloux O, Cassart JP, Delcour J, Van Beeumen J, Vandenhaute J (1991) Cloning and sequence analysis of the gene encoding invertase from the yeast *Kluyveromyces marxianus* var. *marxianus* ATCC 12424. *FEBS Lett* 289, 64-68
- Laughon A, Gesteland RF (1982) Isolation and preliminary characterization of the GAL4 gene, a positive regulator of transcription in yeast. *Proc Natl Acad Sci U S A* 79(22):6827-6831
- Lawrence CE, Altschul SF, Boguski MS, Liu JS, Neuwald AF, Wootton JC (1993) Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science* 262:208-214
- Le Gal-Coeffet MF, Jacks AJ, Sorimachi K, Williamson MP, Williamson G, Archer DB (1995) Expression in *Aspergillus niger* of the starch-binding domain of glucoamylase. Comparison with the proteolytically produced starch-binding domain. *Eur. J. Biochem.* 233:561-567
- Leandro MJ, Goncalves P, Spencer-Martins I (2006) Two glucose/xylose transporter genes from the yeast *Candida intermedia*: first molecular characterization of a yeast xylose-H⁺ symporter. *Biochem J* 395:543-549
- Leemhuis H, Wehmeier UF, Dijkhuizen L (2004) Single amino acid mutations interchange the reaction specificities of Cyclodextrin Glycosyltransferase and the acarbose-modifying enzyme Acarviosyl Transferase. *Biochemistry* 43:13204-13213
- L'Hocine L, Wang W, Jiang B, Xu S (2000) Purification and partial characterization of fructosyltransferase and invertase from *Aspergillus niger* A50023. *J Biotechnol* 81:73-84
- Liebl W, Brem D, Gottschlich A. (1998) Analysis of the gene for beta-fructosidase (invertase, inulinase) of the hyperthermophilic bacterium *Thermotoga maritima*, and characterisation of the enzyme expressed in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 50:55-64
- Lockington R, Scazzocchio C, Sequeval D, Mathieu M, Felenbok B (1987) Regulation of *alcR*, the positive regulatory gene of the ethanol utilization regulon of *Aspergillus nidulans*. *Mol Microbiol.* 1:275-81
- Ma J, Ptashne M (1987) Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* 48(5):847-853
- MacGregor EA, Janecek S, Svensson B (2001) Relationship of sequence and structure to specificity in the alpha-amylase family of enzymes. *Biochim Biophys Acta* 1546:1-20

- Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, Kusumoto K, Arima T, Akita O, Kashiwagi Y, Abe K, Gomi K, Horiuchi H, Kitamoto K, Kobayashi T, Takeuchi M, Denning DW, Galagan JE, Nierman WC, Yu J, Archer DB, Bennett JW, Bhatnagar D, Cleveland TE, Fedorova ND, Gotoh O, Horikawa H, Hosoyama A, Ichinomiya M, Igarashi R, Iwashita K, Juvvadi PR, Kato M, Kato Y, Kin T, Kokubun A, Maeda H, Maeyama N, Maruyama J, Nagasaki H, Nakajima T, Oda K, Okada K, Paulsen I, Sakamoto K, Sawano T, Takahashi M, Takase K, Terabayashi Y, Wortman JR, Yamada O, Yamagata Y, Anazawa H, Hata Y, Koide Y, Komori T, Koyama Y, Minetoki T, Suharnan S, Tanaka A, Isono K, Kuhara S, Ogasawara N, Kikuchi H (2005) Genome sequencing and analysis of *Aspergillus oryzae*. *Nature* 438:1157-1161
- Marchler-Bauer A, Bryant SH (2004) CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* 32:W327-W331
- Marion CL, Rappleye CA, Engle JT, Goldman WE (2006) An alpha-(1,4)-amylase is essential for alpha-(1,3)-glucan production and virulence in *Histoplasma capsulatum*. *Mol Microbiol* 62:970-983
- Marmorstein R, Carey M, Ptashne M, Harrison SC (1992) DNA recognition by GAL4: structure of a protein-DNA complex. *Nature*. 356(6368):408-414
- Martens-Uzunova ES, Zandleven JS, Benen JA, Awad H, Kools HJ, Beldman G, Voragen AG, Van den Berg JA, Schaap PJ (2006) A new group of exo-acting family 28 glycoside hydrolases of *Aspergillus niger* that are involved in pectin degradation. *Biochem J*. 400:43-52
- Martin I, Debarbouille M, Ferrari E, Klier A, Rapoport G (1987) Characterization of the levanase gene of *Bacillus subtilis* which shows homology to yeast invertases. *Mol Gen Genet* 208, 177-184
- Mathieu M, Fillinger S, Felenbok B (2000) In vivo studies of upstream regulatory cis-acting elements of the alcR gene encoding the transactivator of the ethanol regulon in *Aspergillus nidulans*. *Mol Microbiol*. 36(1):123-131
- Mathieu M, Nikolaev I, Sczacchio C, Felenbok B (2005) Patterns of nucleosomal organization in the alc regulon of *Aspergillus nidulans*: roles of the AlcR transcriptional activator and the CreA global repressor. *Mol Microbiol*. 56(2):535-548
- Matsubara T, Ben Ammar Y, Anindyawati T, Yamamoto S, Ito K, Iizuka M, Minamiura N (2004) Molecular cloning and determination of the nucleotide sequence of raw starch digesting alpha-amylase from *Aspergillus awamori* KT-11. *J. Biochem. Mol. Biol.* 37:429-438
- Mattey M, Allan A (1990) Glycogen accumulation in *Aspergillus niger*. *Biochem Soc Trans.* 18:1020-1021
- May T, Mackie RI, Fahey C, Cremin JC, Garleb KA (1994) Effect of fiber source on short-chain fatty acid production and on the growth and toxin production by *Clostridium difficile*. *Scand. J. Gastroenterol.* 29:916-922
- Meeuwssen PJ, Vincken JP, Beldman G, Voragen AG. (2000) A universal assay for screening expression libraries for carbohydrases. *J Biosci Bioeng.* 89(1):107-109
- Megnagneau B, Debets F, Hoekstra RF (1993) Genetic variability and relatedness in the complex group of black *Aspergilli* based on random amplification of polymorphic DNA. *Curr Genet.* 23(4):323-329
- Meng G, Futterer K (2003) Structural framework of fructosyl transfer in *Bacillus subtilis* levansucrase. *Nat Struct Biol* 10, 935-941
- Merika M, Orkin SH (1993) DNA-binding specificity of GATA family transcription factors. *Mol Cell Biol.* 13(7):3999-4010
- Mewes HW, Albermann K, Heumann K, Liebl S, Pfeiffer F (1997) MIPS: a database for protein sequences, homology data and yeast genome information. *Nucleic Acids Res* 25:28-30
- Meyer V, Arentshorst M, El-Ghezal A, Drews AC, Kooistra R, van den Hondel CA, Ram AF (2007) Highly efficient gene targeting in the *Aspergillus niger* kusA mutant. *J. Biotechnol.* 128:770-775
- Miller J, McLachlan AD, Klug A (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* 4(6):1609-1614
- Minetoki T, Gomi K, Kitamoto K, Kumagai C, Tamura G (1995) Nucleotide sequence and expression of alpha-glucosidase-encoding gene (*agdA*) from *Aspergillus oryzae*. *Biosci. Biotechnol. Biochem.* 59 (8):1516-1521
- Morita T, Tanaka N, Hosomi A, Giga-Hama Y, Takegawa K (2006) An alpha-amylase homologue, aah3, encodes a GPI-anchored membrane protein required for cell wall integrity and morphogenesis in *Schizosaccharomyces pombe*. *Biosci Biotechnol Biochem* 70:1454-1463
- Moriyama S, Tanaka H, Uwataki M, Muguruma M, Ohta K (2003) Molecular cloning and characterization of an exoinulinase gene from *Aspergillus niger* strain 12 and its expression in *Pichia pastoris*. *J Biosci Bioeng* 96:324-331
- Moriyama S, Akimoto H, Suetsugu N, Kawasaki S, Nakamura T, Ohta K (2002) Purification and properties of an extracellular exoinulinase from *Penicillium* sp. strain TN-88 and sequence analysis of the encoding gene. *Biosci. Biotechnol. Biochem.* 66, 1887-1896
- Moriyama S, Muguruma M, Ohta K (2006) Quantitative expression analysis of inulinase gene cluster of *Penicillium* sp. strain TN-88. *J Biosci Bioeng* 101:277-279

- Morkeberg R, Carlsen M, Nielsen J (1995) Induction and repression of alpha-amylase production in batch and continuous cultures of *Aspergillus oryzae*. Microbiology. 141:2449-2454
- Mouyna I, Fontaine T, Vai M, Monod M, Fonzi WA, Diaquin M, Popolo L, Hartland RP, Latgé JP (2000) Glycosylphosphatidylinositol-anchored glucanotransferases play an active role in the biosynthesis of the fungal cell wall. J. Biol. Chem. 275:14882-14889
- Mouyna I, Morelle W, Vai M, Monod M, L  ch  ne B, Fontaine T, Beauvais A, Sarfati J, Pr  vost MC, Henry C, Lat  g   JP (2005) Deletion of GEL2 encoding for a beta(1-3)glucanotransferase affects morphogenesis and virulence in *Aspergillus fumigatus*. Mol. Microbiol. 56:1675-1688
- Muramatsu M, Nakakuki T (1995) Enzymatic synthesis of novel fructosyl and oligofructosyl trehaloses by *Aspergillus sydowii* beta-fructofuranosidase. Biosci. Biotechnol. Biochem. 59:208-212
- Myers AM, Morell MK, James MG, Ball SG (2000) Recent progress towards understanding biosynthesis of the amylopectin crystal. Plant Physiol. 122:989-997
- Nagem RA, Rojas AL, Golubev AM, Korneeva OS, Eneyskaya EV, Kulminkaya AA, Neustroev KN, Polikarpov I (2004) Crystal structure of exo-inulinase from *Aspergillus awamori*: The enzyme fold and structural determinants of substrate recognition. J Mol Biol 344:471-480
- Nakajima T, Imanaka T, Aiba S (1986) Comparison of amino acid sequences of eleven different alpha-amylases. Appl Microbiol Biotechnol 23:355-360
- Nakamura A, Nishimura I, Yokoyama A, Lee DG, Hidaka M, Masaki H, Kimura A, Chiba S, Uozumi T (1997) Cloning and sequencing of an alpha-glucosidase gene from *Aspergillus niger* and its expression in *A. nidulans*. J Biotechnol 53:75-84
- Nakamura T, Maeda Y, Tanoue N, Makita T, Kato M, Kobayashi T (2006) Expression profile of amyolytic genes in *Aspergillus nidulans*. Biosci Biotechnol Biochem 70:2363-2370
- Nakamura A, Haga K, Yamane K (1993) Three histidine residues in the active center of cyclodextrin glucanotransferase from alkalophilic *Bacillus sp.* 1011: effects of the replacement on pH dependence and transition-state stabilization. Biochemistry 32:6624-6631
- Nakao M, Nakayama T, Kakudo A, Inohara M, Harada M, Omura F, Shibano Y (1994) Structure and expression of a gene coding for thermostable alpha-glucosidase with a broad substrate specificity from *Bacillus sp.* SAM1606. Eur J Biochem 220:293-300
- Narendja F, Goller SP, Wolschek M, Strauss J (2002) Nitrate and the GATA factor AreA are necessary for *in vivo* binding of NirA, the pathway-specific transcriptional activator of *Aspergillus nidulans*. Mol Microbiol. 44(2):573-583
- Naumoff DG (2001) beta-fructosidase superfamily: homology with some alpha-L-arabinases and beta-D-xylosidases. Proteins 42:66-76
- Nguyen QD, Mattes F, Hoschke A, Rezessy-szab   J, Bhat MK (1999) Production, purification and identification of fructooligosaccharides produced by beta-fructofuranosidase from *Aspergillus niger* IMI 303386. Biotechnol. Lett. 21:183-186
- Ni BF, Needleman RB (1990) Identification of the upstream activating sequence of MAL and the binding sites for the MAL63 activator of *Saccharomyces cerevisiae*. Mol Cell Biol 10:3797-3800
- Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, Berriman M, Abe K, Archer DB, Bermejo C, Bennett J, Bowyer P, Chen D, Collins M, Coulson R, Davies R, Dyer PS, Farman M, Fedorova N, Fedorova N, Feldblyum TV, Fischer R, Fosker N, Fraser A, Garcia JL, Garcia MJ, Goble A, Goldman GH, Gomi K, Griffith-Jones S, Gwilliam R, Haas B, Haas H, Harris D, Horiuchi H, Huang J, Humphray S, Jimenez J, Keller N, Khouri H, Kitamoto K, Kobayashi T, Konzack S, Kulkarni R, Kumagai T, Lafon A, Lat  g   JP, Li W, Lord A, Lu C, Majoros WH, May GS, Miller BL, Mohamoud Y, Molina M, Monod M, Mouyna I, Mulligan S, Murphy L, O'Neil S, Paulsen I, Penalva MA, Perte   M, Price C, Pritchard BL, Quail MA, Rabinowitsch E, Rawlins N, Rajandream MA, Reichard U, Renauld H, Robson GD, Rodriguez de CS, Rodriguez-Pena JM, Ronning CM, Rutter S, Salzberg SL, Sanchez M, Sanchez-Ferrero JC, Saunders D, Seeger K, Squares R, Squares S, Takeuchi M, Tek  ia F, Turner G, Vazquez de Aldana CR, Weidman J, White O, Woodward J, Yu JH, Fraser C, Galagan JE, Asai K, Machida M, Hall N, Barrell B, Denning DW (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature 438:1151-1156
- Nikolaev I, Lenouvel F, Felenbok B (1999) Unique DNA binding specificity of the binuclear zinc AlcR activator of the ethanol utilization pathway in *Aspergillus nidulans*. J Biol Chem. 274(14):9795-9802
- Niness KR (1999) Inulin and Oligofructose: What Are They? J. Nutr. 129:1402S-1406S
- Nunberg JH, Meade JH, Cole G, Lawyer FC, McCabe P, Schweickart V, Tal R, Wittman VP, Flatgaard JE, Innis MA (1984) Molecular cloning and characterization of the glucoamylase gene of *Aspergillus awamori*. Mol. Cell. Biol. 4 (11):2306-2315
- Ohara O, Temple G (2001) Directional cDNA library construction assisted by the *in vitro* recombination reaction. Nucleic Acids Research, 29: E22
- Ohta K, Akimoto H, Matsuda S, Toshimitsu D, Nakamura T (1998) Molecular cloning and sequence analysis of two endoinulinase genes from *Aspergillus niger*. Biosci Biotechnol Biochem 62:1731-1738

- Omichinski JG, Clore GM, Schaad O, Felsenfeld G, Trainor C, Appella E, Stahl SJ, Gronenborn AM (1993) NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science* 261:438-446
- Onodera S, Murakami T, Ito H, Mori H, Matsui H, Honma M, Chiba S, Shiomi N (1996) Molecular cloning and nucleotide sequences of cDNA and gene encoding endo-inulinase from *Penicillium purpurogenum*. *Biosci Biotechnol Biochem* 60:1780-1785
- Orlean P (1997) Biogenesis of yeast wall and surface components. In: Pringle J, Broach JR, Jones EW (eds) *Cell Cycle and Cell Biology*. Cold Spring Harbor Laboratory Press, New York, pp. 229-362
- Ozimek LK, Kralj S, van der Maarel MJ, Dijkhuizen L (2006) The levansucrase and inulosucrase enzymes of *Lactobacillus reuteri* 121 catalyse processive and non-processive transglycosylation reactions. *Microbiology*. 152:1187-1196
- Paldi T, Levy I, Shoseyov O (2003) Glucoamylase starch-binding domain of *Aspergillus niger* B1: molecular cloning and functional characterization. *Biochem. J.* 372:905-910
- Pandey A, Soccol CR, Selvakumar P, Soccol VT, Krieger N, Fontana JD (1999) Recent developments in microbial inulinases. Its production, properties, and industrial applications. *Appl Biochem Biotechnol* 81, 35-52
- Panozzo C, Capuano V, Fillinger S, Felenbok B (1997) The zinc binuclear cluster activator AlcR is able to bind to single sites but requires multiple repeated sites for synergistic activation of the alcA gene in *Aspergillus nidulans*. *J Biol Chem*. 272(36):22859-22865
- Parenticová L, Benen JAE, Samson RA, Visser J (1997) Evaluation of RFLP analysis of the classification of selected black aspergilli. *Mycol. Res.* 101:810-814
- Parenticova L, Benen JA, Kester HC, Visser J (1998) *pgaE* encodes a fourth member of the endopolygalacturonase gene family from *Aspergillus niger*. *Eur. J. Biochem*. 251:72-80
- Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, Schaap PJ, Turner G, de Vries RP, Albang R, Albermann K, Andersen MR, Bendtsen JD, Benen JA, van den Berg M, Breestraat S, Caddick MX, Contreras R, Cornell M, Coutinho PM, Danchin EG, Debets AJ, Dekker P, van Dijk PW, van Dijk A, Dijkhuizen L, Driessen AJ, d'Enfert C, Geysens S, Goosen C, Groot GS, de Groot PW, Guillemette T, Henrissat B, Herweijer M, van den Hombergh JP, van den Hondel CA, van der Heijden RT, van der Kaaij RM, Klis FM, Kools HJ, Kubicek CP, van Kuyk PA, Lauber J, Lu X, van der Maarel MJ, Meulenberg R, Menke H, Mortimer MA, Nielsen J, Oliver SG, Olsthoorn M, Pal K, van Peij NN, Ram AF, Rinas U, Roubos JA, Sagt CM, Schmoll M, Sun J, Ussery D, Varga J, Vervecken W, van de Vondervoort PJ, Wedler H, Wosten HA, Zeng AP, van Ooyen AJ, Visser J, Stam H (2007) Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat Biotechnol*. 25(2):221-231.
- Penalva MA, Arst HN Jr (2002) Regulation of gene expression by ambient pH in filamentous fungi and yeasts. *Microbiol Mol Biol Rev*. 66(3):426-446. Review
- Penalva MA, Arst HN Jr. (2004) Recent advances in the characterization of ambient pH regulation of gene expression in filamentous fungi and yeasts. *Annu Rev Microbiol*. 58:425-451. Review
- Penalva MA, Arst H (2002). Regulation of gene expression by ambient pH in filamentous fungi and yeasts. *Microbiol Mol Biol Rev* 66, 426-446
- Perez JA, Rodriguez J, Rodriguez L, Ruiz T (1996) Cloning and sequence analysis of the invertase gene INV1 from the yeast *Pichia anomala*. *Curr Genet* 29, 234-240
- Peters D (2006) Carbohydrates for fermentation. *Biotechnol J* 1:806-814
- Petersen KL, Lehmebeck J, Christensen T (1999) A new transcriptional activator for amylase genes in *Aspergillus*. *Mol Gen Genet* 262:668-676
- Pilauri V, Bewley M, Diep C, Hopper J (2005) Gal80 dimerization and the yeast GAL gene switch. *Genetics* 169:1903-1914
- Pilon-Smiths EA, Ebskamp MJ, Paul MJ, Jeuken MJ, Weisbeek PJ, Smeeckens SC (1995) Evaluation of safety on inulin and oligofructose as dietary fiber. *Plant Physiol* 107:125-130
- Pina C, Goncalves P, Prista C, Loureiro-Dias MC (2004) Ffz1, a new transporter specific for fructose from *Zygosaccharomyces bailii*. *Microbiology* 150:2429-2433
- Pollock CJ, Chatterton NJ (1988) Fructans. In *The Biochemistry of Plants* (Stumpf, P.K. and Conn, E.E., eds), Vol. 14, Carbohydrates (Preiss, J., ed.). London: Academic Press, pp. 109-140
- Pons T, Naumoff DG, Martinez-Fleites CM, Hernández L (2004). Three acidic residues are at the active site of a beta-propeller architecture in glycoside hydrolase families 32, 43, 62 and 68. *Proteins* 54, 424-432
- Pons T, Olmea O, China G, Beldarrain A, Marquez G, Acosta N, Rodriguez L, Valencia A (1998) Structure model for family 32 of glycosyl-hydrolase enzymes. *Proteins* 33, 383-395
- Popolo L, Vai M, Gatti E, Porello S, Bonfante P, Balestrini R, Alberghina L (1993) Physiological analysis of mutants indicates involvement of the *Saccharomyces cerevisiae* GPI-anchored protein gp115 in morphogenesis and cell separation. *J. Bacteriol.* 175:1879-1885

- Punt PJ, Oliver RP, Dingemanse MA, Pouwels PH, van den Hondel CA (1987) Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*. *Gene* 56:117-124
- Punt PJ, Strauss J, Smit R, Kinghorn JR, van den Hondel CA, Scazzocchio C (1995) The intergenic region between the divergently transcribed *niiA* and *niaD* genes of *Aspergillus nidulans* contains multiple NirA binding sites which act bidirectionally. *Mol Cell Biol* 15:5688-5699
- Punt PJ, van Biezen N, Conesa A, Albers A, Mangnus J, van den Hondel C (2002) Filamentous fungi as cell factories for heterologous protein production. *Trends Biotechnol.* 20(5):200-206
- Punt PJ, van den Hondel CA (1992) Transformation of filamentous fungi based on hygromycin B and phleomycin resistance markers. *Methods Enzymol* 216:447-457
- Raina S and Missiakas D (1997) Making and breaking disulfide bonds. *Annu. Rev. Microbiol.* 51:179-202
- Ram AF and Klis FM (2006) Identification of fungal cell wall mutants using susceptibility assays based on Calcofluor White and Congo Red. *Nature Protocols* 1:2253-2256
- Ram AF, Wolters A, Ten Hoopen R, Klis FM (1994) A new approach for isolating cell wall mutants in *Saccharomyces cerevisiae* by screening for hypersensitivity to calcofluor white. *Yeast* 10:1019-1030
- Raper KB, Fennell DI (1955) New species and varieties of *Aspergillus*. *Mycologia* 47:68-89
- Raper KB, Fennell DI (1965) The genus *Aspergillus*. Williams and Wilkins, Baltimore
- Rappleye CA, Engle JT, Goldman WE (2004) RNA interference in *Histoplasma capsulatum* demonstrates a role for alpha-(1,3)-glucan in virulence. *Mol Microbiol* 53:153-165
- Rappleye CA, Goldman WE (2006) Defining virulence genes in the dimorphic fungi. *Annu Rev Microbiol* 60:281-303
- Rappleye CA, Engle JT, Goldman WE (2004) RNA interference in *Histoplasma capsulatum* demonstrates a role for alpha-(1,3)-glucan in virulence. *Mol. Microbiol.* 53:153-165
- Ravagnani A, Gorfinkiel L, Langdon T, Diallinas G, Adjadj E, Demais S, Gorton D, Arst HN, Scazzocchio C (1997) Subtle hydrophobic interactions between the seventh residue of the zinc finger loop and the first base of an HGATAR sequence determine promoter-specific recognition by the *Aspergillus nidulans* GATA factor AreA. *EMBO J.* 16:3974-3986
- 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
- Reese AJ, Doering TL (2003) Cell wall alpha-1,3-glucan is required to anchor the *Cryptococcus neoformans* capsule. *Mol. Microbiol.* 50:1401-1409
- Rehm J, Wilimitzer L, Heyer LG (1998) Production of 1-kestose in transgenic yeast expressing a fructosyl transferase from *Aspergillus foetidus*. *J Bacteriol* 180, 1305-1310
- Ritsema T, Smeekens S (2003) Fructans: beneficial for plants and humans. *Curr Opin Plant Biol* 6, 223-230
- Roberfroid MB, Delzenne NM (1998) Dietary fructans. *Annu Rev Nutr* 18, 117-143
- Roby JF (1998) *Essentials of Carbohydrate Chemistry*. Springer, New York.
- Rubio MC, Maldonado MC (1995) Purification and characterization of invertase from *Aspergillus niger*. *Curr. Microbiol.* 31:80-83
- Ruepp A, Zollner A, Maier D, Albermann K, Hani J, Mokrejs M, Tetko I, Guldener U, Mannhaupt G, Munsterkotter M, Mewes HW (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res* 32:5539-5545
- Ruijter GJ, Visser J (1997) Carbon repression in *Aspergilli*. *FEMS Microbiol Lett* 151:103-114
- Sambrook J, Fritsch E, and Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Sangeetha PT, Ramesh MN, Prapulla SG (2004) Production of fructosyl transferase by *Aspergillus oryzae* CFR 202 in solid-state fermentation using agricultural by-products. *Appl Microbiol Biotechnol* 65, 530-537
- Sasangka P, Matsuno A, Tanaka A, Akasaka Y, Suyama S, Kano S, Miyazaki M, Akao T, Kato M, Kobayashi T, Tsukagoshi N (2002) Structural features of the glycogen branching enzyme encoding genes from *Aspergilli*. *Microbiol. Res.* 157:337-344
- Sauer J, Sigurskjold BW, Christensen U, Frandsen TP, Mirgorodskaya E, Harrison M, Roepstorff P, Svensson B (2000) Glucoamylase: structure/function relationships, and protein engineering. *Biochim Biophys Acta* 1543:275-293
- Scazzocchio C (2000) The fungal GATA factors. *Curr Opin Microbiol.* 3(2):126-131
- Schjerling P, Holmberg S (1996) Comparative amino acid sequence analysis of the C6 zinc cluster family of transcriptional regulators. *Nucleic Acids Res* 24:4599-4607
- Schuster E, Dunn-Coleman N, Frisvad JC, van Dijk PWM (2002) On the safety of *Aspergillus niger* – a review. *Appl. Microbiol. Biotechnol* 59:426-435

- Semova N, Storms R, John T, Gaudet P, Ulyczynj P, Min XJ, Sun J, Butler G, Tsang A (2006) Generation, annotation, and analysis of an extensive *Aspergillus niger* EST collection. BMC Microbiol. 6:7
- Sghir A, Chow JM, Mackie RI (1998) Continuous culture selection of *bifidobacteria* and *lactobacilli* from human faecal samples using fructooligosaccharide as selective substrate. J. Appl. Microbiol. 85:769-777
- Sharrocks AD (2000) Introduction: the regulation of eukaryotic transcription factor function, CMLS Cell. Mol. Life Sci. (57) 1147-1148
- Shibuya I, Gomi K, Iimura Y, Takahashi K, Tamura G, Hara S (1990) Molecular cloning of the glucoamylase gene of *Aspergillus shirousami* and its expression in *Aspergillus oryzae*. Agric. Biol. Chem. 54 (8):1905-1914
- Shibuya I, Tamura G, Ishikawa T, Hara S (1992) Cloning of the alpha-amylase cDNA of *Aspergillus shirousamii* and its expression in *Saccharomyces cerevisiae*. Biosci Biotechnol Biochem. 56:174-179
- Shroff RA, O'Connor SM, Hynes MJ, Lockington RA, Kelly JM (1997) Null alleles of *creA*, the regulator of carbon catabolite repression in *Aspergillus nidulans*. Fungal Genet Biol 22:28-38
- Sil AK, Alam S, Xin P, Ma L, Morgan M, Lebo CM, Woods MP, Hopper JE. (1999) The Gal3p-Gal80p-Gal4p Transcription Switch of Yeast: Gal3p Destabilizes the Gal80p-Gal4p Complex in Response to Galactose and ATP. Mol Cell Biol. 19(11):7828-7840
- Somiari R, Brzeski H, Tate R, Bieleck S, Polak J (1997) Cloning and sequencing of an *Aspergillus niger* gene coding for beta-fructofuranosidase. Biotechnol. Lett. 19:1243-1247
- Southall SM, Simpson PJ, Gilbert HJ, Williamson G, Williamson MP (1999) The starch-binding domain from glucoamylase disrupts the structure of starch. FEBS Letters 447:58-60
- Stagg CM, Feather MS (1973) The characterization of a chitin-associated D-glucan from the cell walls of *Aspergillus niger*. Biochim. Biophys. Acta. 320:64-72
- Stam MR, Danchin EG, Rancurel C, Coutinho PM, Henrissat B (2006) Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins. Protein Eng Des Sel 19:555-562
- Steup, M (1988) Starch degradation. In The Biochemistry of Plants (Stumpf, P.K. and Conn, E.E., eds), Vol. 14, Carbohydrates (Preiss, J., ed.). London: Academic Press, pp. 255-296
- Stryer, L (1999) Biochemistry, fourth edition. W.H. Freeman and Company, New York, pp. 998-999
- Suzuki M and Yagi N (1994) DNA recognition code of transcription factors in the helix-turn-helix, probe helix, hormone receptor, and zinc finger families. Proc Natl Acad Sci U S A. 91(26):12357-12361
- Suzuki M, Pollock CJ (1986) Extraction and characterization of the enzymes of fructan biosynthesis in timothy grass (*Phleum pratense* L.). Can J Bot 64:1884-1887
- Tada S, Iimura Y, Gomi K, Takahashi K, Hara S and Yoshizawa K (1989) Cloning and nucleotide sequence of the genomic Taka-amylase A gene of *Aspergillus oryzae*. Agric. Biol. Chem. 53:593-599
- Takaha T, Yanase M, Okada S, Smith SM (1993) Disproportionating enzyme (4-alpha-glucanotransferase; EC 2.4.1.25) of potato. Purification, molecular cloning, and potential role in starch metabolism. J. Biol. Chem. 268:1391-1396
- Tanaka N, Ohuchi N, Mukai Y, Osaka Y, Ohtani Y, Tabuchi M, Bhuiyan MS, Fukui H, Harashima S, Takegawa K (1998) Isolation and characterization of an invertase and its repressor genes from *Schizosaccharomyces pombe*. Biochem Biophys Res Commun 245:246-253
- Tani S, Itoh T, Kato M, Kobayashi T, Tsukagoshi N (2001) In vivo and in vitro analyses of the AmyR binding site of the *Aspergillus nidulans* *agdA* promoter; requirement of the CGG direct repeat for induction and high affinity binding of AmyR. Biosci Biotechnol Biochem 65:1568-1574
- Tani S, Katsuyama Y, Hayashi T, Suzuki H, Kato M, Gomi K, Kobayashi T, Tsukagoshi N (2001a) Characterization of the *amyR* gene encoding a transcriptional activator for the amylase genes in *Aspergillus nidulans*. Curr Genet 39:10-15
- Tannock GW (1997) Probiotic properties of *lactic-acid bacteria*: plenty of scope for fundamental R & D. Trends Biotechnol. 15:270-274
- Taussig R, Carlson M (1983) Nucleotide sequences of the yeast SUC2 gene for invertase. Nucleic Acids Res 11:1943-1954
- Teste MA, Enjalbert B, Parrou JL, Francois JM (2000) The *Saccharomyces cerevisiae* YPR184w gene encodes the glycogen debranching enzyme. FEMS Microbiol Lett 193:105-110
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673-4680
- Tilburn J, Sarkar S, Widdick DA, Espeso EA, Orejas M, Mungroo J, Peñalva MA, Arst HN Jr. (1995) The *Aspergillus* PacC zinc finger transcription factor mediates regulation of both acid- and alkaline-expressed genes by ambient pH. EMBO J. 14:779-790

- Todd RB, Andrianopoulos A (1997) Evolution of a fungal regulatory gene family: the Zn(II)₂Cys₆ binuclear cluster DNA binding motif. *Fungal Genet Biol* 21:388-405
- Tsuji G, Kenmochi Y, Takano Y, Sweigard J, Farrall L, Furusawa I, Horino O, Kubo Y (2000) Novel fungal transcriptional activators, Cmr1p of *Colletotrichum lagenarium* and pig1p of *Magnaporthe grisea*, contain Cys₂His₂ zinc finger and Zn(II)₂Cys₆ binuclear cluster DNA-binding motifs and regulate transcription of melanin biosynthesis genes in a developmentally specific manner. *Mol Microbiol* 38:940-954
- Tsakagoshi N, Kobayashi T, Kato M (2001) Regulation of the amylolytic and (hemi-)cellulolytic genes in *Aspergilli*. *J Gen Appl Microbiol* 47:1-19
- Tsakamoto A, Kimura K, Ishii Y, Takano T, Yamane K (1988) Nucleotide sequence of the maltohexaose producing amylase gene from an alkalophilic *Bacillus* sp. #707 and structural similarity to liquefying type alpha-amylases. *Biochem Biophys Res Commun* 151:25-31
- Uhm T, Chae S, Lee D, Kim S, Cassart P, Vandenhaute J (1998) Cloning and nucleotide sequence of the endoinulinase-encoding gene, *inu2*, from *Aspergillus ficuum*. *Biotechnol Lett* 20:809-812
- Uitdehaag JC, Mosi R, Kalk KH, van der Veen BA, Dijkhuizen L, Withers SG, Dijkstra BW (1999) X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the alpha-amylase family. *Nat Struct Biol* 6:432-436
- Vainstein MH, Peberdy JF (1990) Solubilisation of a cell wall bound invertase in *Aspergillus nidulans*. *FEMS Microbiol Lett* 59, 265-269
- Vainstein MH, Peberdy JF (1991) Regulation of invertase in *Aspergillus nidulans*: effect of different carbon sources. *J Gen Microbiol* 137, 315-321
- Vallee BL, Coleman JE, and Auld DS (1991) Zinc fingers, zinc clusters, and zinc twists in DNA-binding protein domains. *Proc. Natl. Acad. Sci. USA*. Vol 89:999-1003
- van der Kaaij RM, Yuan XL, Franken A, Ram AFJ, Punt PJ, Maarel MJEC, Dijkhuizen L (2007a) Two novel, putatively cell wall-associated and glycosylphosphatidylinositol-anchored alpha-glucanotransferase enzymes of *Aspergillus niger*. *Eukaryot Cell* 6:1178-1188
- Van der Kaaij RM, Janecsek S, van der Maarel MJEC, Dijkhuizen L (2007b) Phylogenetic and biochemical characterization of a novel cluster of intracellular fungal α -amylase enzymes. *Microbiology* 153: 4003-4015
- van der Maarel MJ, van der Veen B, Uitdehaag JC, Leemhuis H, Dijkhuizen L (2002) Properties and applications of starch-converting enzymes of the α -amylase family. *Journal of Biotechnology* 94:137-155
- van der Vlugt-Bergmans CJB and van Ooyen AJJ (1999) Expression cloning in *Kluyveromyces lactis*. *Biotechnol. Technol.* 13:87-92
- van Dijk PW, Selten GC, Hempenius RA (2003) On the safety of a new generation of DSM *Aspergillus niger* enzyme production strains. *Regul. Toxicol. Pharmacol.* 38:27-35
- Van Gorcom RFM, van Hartingsveldt W, van Paridon PA, Veenstra AE, Luiten RGM, Selten GCM (1991) Cloning and expression of microbial phytase. *European Patent Application* 0.420.358
- van Hartingsveldt W, Mattern IE, van Zeijl CM, Pouwels PH, van den Hondel CA (1987) Development of a homologous transformation system for *Aspergillus niger* based on the *pyrG* gene. *Mol Gen Genet* 206:71-75
- van Hartingsveldt W, van Zeijl CM, Harteveld GM, Gouka RJ, Suykerbuyk ME, Luiten RG, van Paridon PA, Selten GC, Veenstra AE, van Gorcom RFM, van den Hondel (1993) Cloning, characterization and overexpression of the phytase-encoding gene (*phyA*) of *Aspergillus niger*. *Gene*. 127(1):87-94
- van Hartingsveldt W, Mattern IE, van Zeijl CM, Pouwels PH, van den Hondel CA (1987) Development of a homologous transformation system for *Aspergillus niger* based on the *pyrG* gene. *Mol Gen Genet* 206: 71-75
- Van Loo J, Coussement P, DeLeenheer L, Hoebregs H, Smits G (1995) On the presence of inulin and oligofructose as natural ingredients in the Western diet. *CRC Crit. Rev. Food Sci. Nutr.* 35:525-552
- van Peij NN, Gielkens MM, de Vries RP, Visser J, de Graaff LH (1998b) The transcriptional activator XlnR regulates both xylanolytic and endoglucanase gene expression in *Aspergillus niger*. *Appl Environ Microbiol* 64:3615-3619
- van Peij NN, Visser J, de Graaff LH (1998a) Isolation and analysis of *xlnR*, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. *Mol Microbiol* 27:131-142
- Vandamme EJ, Derycke DG (1983) Microbial inulinase: fermentation process, properties, and applications. *Adv. Appl. Microbiol* 29:139-176
- Vankuyk PA, Diderich JA, MacCabe AP, Hererro O, Ruijter GJ, Visser J (2004) *Aspergillus niger* *mstA* encodes a high-affinity sugar/H⁺ symporter which is regulated in response to extracellular pH. *Biochem J* 379:375-383
- Varga J, Kevei F, Vriesema A, Debets F, Kozakiewicz Z, Croft JH (1994) Mitochondrial DNA restriction fragment length polymorphisms in field isolates of the *Aspergillus niger* aggregate. *Can J Microbiol.* 40(8):612-621

- Ventura L, Gonzalez-Candelas L, Perez-Gonzalez JA, Ramon D (1995) Molecular cloning and transcriptional analysis of the *Aspergillus terreus gla1* gene encoding a glucoamylase. Appl. Environ. Microbiol. 61 (1):399-402
- Verdoes JC, Punt PJ, van den Hondel CA (1995) Molecular genetic strain improvement for the overproduction of fungal proteins by filamentous fungi. Appl. Microbiol. Biotechnol 43:195-205
- Verhaest M, Ende WV, Roy KL, De Ranter CJ, Laere AV, Rabijs A (2005) X-ray diffraction structure of a plant glycosyl hydrolase family 32 protein: fructan 1-exohydrolase IIa of *Cichorium intybus*. Plant J. 41:400-411
- Vienken K, Scherer M, Fischer R (2005) The Zn(II)2Cys6 putative *Aspergillus nidulans* transcription factor repressor of sexual development inhibits sexual development under low-carbon conditions and in submersed culture. Genetics 169:619-630
- Vijn I, Smeekens S (1999) Fructan: more than a reserve carbohydrate. Plant Physiol. 120:351-359
- Wainright M (1992) An introduction to fungal biotechnology. John Wiley & Sons, Inc., New York, N. Y.
- Wallis GLF, Hemming FW, Peberdy JF (1997) Secretion of two β -fructofuranosidases by *Aspergillus niger* growing in sucrose. Arch Biochem Biophys 345, 214-222
- Walton JD (2000) Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: a hypothesis. Fungal Genet Biol 30:167-171
- Waterhouse AL, and Chatterton NJ (1993) Glossary of Fructan Terms. In Science and Technology of Fructans, Suzuki, M and Chatterton, N.J., eds. CRC Press, Boca Raton, pp. 1-8
- Weenink XO, Punt PJ, van den Hondel CA, Ram AF (2006) A new method for screening and isolation of hypersecretion mutants in *Aspergillus niger*. Appl Microbiol Biotechnol 69:711-717
- Wernars K, Goosen T, Wennekes LM, Visser J, Bos CJ, van den Broek HW, van Gorcom RF, van den Hondel CA, Pouwels PH (1985) Gene amplification in *Aspergillus nidulans* by transformation with vectors containing the *amdS* gene. Curr. Genet. 9:361-368
- Williamson G, Belshaw NJ, Williamson MP (1992) O-glycosylation in *Aspergillus* glucoamylase. Conformation and role in binding. Biochem. J. 282:423-428
- Wilson RA, Arst HN (1998) Mutational analysis of AREA, a transcriptional activator mediating nitrogen metabolite repression in *Aspergillus nidulans* and a member of the "streetwise" GATA family of transcription factors. Microbiol Mol Biol Rev. 62(3):586-596
- Wirsel S, Lachmund A, Wildhardt G, Ruttkowski E (1989) Three alpha-amylase genes of *Aspergillus oryzae* exhibit identical intron-exon organization. Mol Microbiol. 3:3-14
- Woloshuk CP, Foutz KR, Brewer JF, Bhatnagar D, Cleveland TE, Payne GA (1994) Molecular characterization of *aflR*, a regulatory locus for aflatoxin biosynthesis. Appl Environ Microbiol 60:2408-2414
- Woloshuk CP, Yousibova GL, Rollins JA, Bhatnagar D, Payne GA (1995) Molecular characterization of the afl-1 locus in *Aspergillus flavus*. Appl Environ Microbiol 61:3019-3023
- Wolschek MF, Kubicek CP (1997) The filamentous fungus *Aspergillus niger* contains two "differentially regulated" trehalose-6-phosphate synthase-encoding genes, *tpsA* and *tpsB*. J Biol Chem 272:2729-2735
- Yamamoto M, Ko LJ, Leonard MW, Beug H, Orkin SH, Engel JD (1990) Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. Genes Dev. 4:1650-1662
- Yamazaki H, Yamazaki D, Takaya N, Takagi M, Ohta A, Horiuchi H (2007) A chitinase gene, *chiB*, involved in the autolytic process of *Aspergillus nidulans*. Curr Genet 51:89-98
- Yanai K, Nakane A, Kawate A, Hirayama M (2001) Molecular cloning and characterization of the fructooligosaccharide-producing beta-fructofuranosidase gene from *Aspergillus niger* ATCC 20611. Biosci. Biotechnol. Biochem. 65:766-773
- Yu J, Chang P, Bhatnagar D, Cleveland TE (2000) Cloning of a sugar utilization gene cluster in *Aspergillus parasiticus*. Biochim Biophys Acta 1493:211-214
- Yuan XL, Goosen C, Kools H, van der Maarel MJ, van den Hondel CA, Dijkhuizen L, Ram AF (2006) Database mining and transcriptional analysis of genes encoding inulin-modifying enzymes of *Aspergillus niger*. Microbiology 152:3061-3073
- Yuan XL, Roubos JA, van den Hondel CA, Ram AF (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
- Yun JW (1996). Fructooligosaccharides: Occurrence, preparation, and application. Enzyme Microb Technol 19:107-117

Publications

- **Xiao-Lian Yuan**, Tao Hua Zhou, Da Lian Yao (2000) *Study on the effects of soil-film cover to the growth of transplanted cotton*, Journal of Anhui Agricultural University, 28
- **Xiao-Lian Yuan***, Coenie Goosen*, Harrie Kools, Marc J.E.C. van der Maarel, Cees A.M.J.J van den Hondel, Lubbert Dijkhuizen and Arthur F.J. Ram (2006) *Database mining and transcriptional regulation of genes encoding inulin modifying enzymes in Aspergillus niger*, Microbiology 152(Pt 10):3061-3073
- **Xiao-Lian Yuan**, Johannes A. Roubos, Cees A. M. J. J. van den Hondel, Arthur F. J. Ram (2008) *Identification of InuR, a new Zn(II)₂Cys₆ transcriptional activator involved in the regulation of inulinolytic enzymes in Aspergillus niger*, Molecular Genetics and Genomics, 279(1):11-26
- Coenie Goosen*, **Xiao-Lian Yuan***, Jolanda M van Munster, Arthur F J Ram, Marc J E C van der Maarel, Lubbert Dijkhuizen (2007) *Molecular and biochemical characterization of a novel intracellular invertase from Aspergillus niger with transfructosylating activity*, Eukaryotic cell, 6(4):674-681
- Rachel M. van der Kaaij, **Xiao-Lian Yuan**, Franken A, Arthur F J Ram , Peter J. Punt, Marc J.E.C van der Maarel, Lubbert Dijkhuizen (2007) *Two novel, putatively cell wall-associated and glycosylphosphatidylinositol-anchored alpha-glucanotransferase enzymes of Aspergillus niger*, Eukaryotic cell, 6(7):1178-1188
- **Xiao-Lian Yuan***, Rachel M. van der Kaaij*, Cees A.M.J.J. van den Hondel, Peter J. Punt, Marc J.E.C van der Maarel, Lubbert Dijkhuizen, Arthur F.J. Ram *Aspergillus niger genome wide analysis reveals a large alpha-glucan acting enzymes with unexpected expression profiles*, submitted.

Curriculum Vitae

Xiaolian Yuan, was born on November 25, 1974 in Feixi, Anhui Province, China. After middle school, she started studying Agronomy at Anhui Agricultural University (AAU) in 1993 and graduated with honor in 1997. From July 1997 until August 2000 she worked at AAU as an assistant lecturer, focusing on the courses *seeds physiology* and *marketing of agro-products*. During this period she was also a student co-ordinator at AAU.

In August 2000, she was granted an NFP fellowship (NUFFIC) and started studying MSc in Biotechnology at Wageningen University. From September 2001 to February 2002 she did her MSc internship at Prof. Ton Bissling's group in the department of molecular plant science of Wageningen University. Her research topic was "*Characterizing actin gene family in legume plant Medicago truncatula with respect to Rhizobium symbiosis*". She obtained her MSc degree in Biotechnology in March 2002.

In April 2002, she started a PhD project at the Institute of Biology of Leiden University (IBL), under supervision of Dr. Arthur Ram and Prof. Cees van den Hondel. The topic of the project was "*Disclosing the carbohydrate modifying enzyme network of Aspergillus niger by functional genomics*" and the results are presented in this book.

Since January 2008, she is working at Microarray Department of Amsterdam University as a molecular biologist.



Acknowledgements

In early 2002, when I was finishing my MSc in Biotechnology at Wageningen University, I wanted to further broaden my knowledge on molecular biology, especially in the field of functional genomics. An interesting PhD position was available from the fungal IOP program. I applied for this and a few months later I started my PhD project at the Institute of Biology of Leiden University (IBL), in the fungal genetics group of Prof. van den Hondel.

Now, a few years later, with a lot more knowledge and a lot more experience, the project is finished. This thesis is the final tangible result. The project itself, however, was enjoyable, of course with its difficult moments but I look back on a very nice period in which I developed myself a lot, both professionally and personally.

It would not have been possible to finish the project without many contributions, both scientifically and personally. First of all, I want to thank the fungal IOP program for financially supporting my PhD project. This was essential for the successful completion of the project.

In the first place I would like to thank my colleagues at the fungal genetics group at the Institute of Biology of Leiden University (IBL). I enjoyed the time we were working together with you all. I must say that I will miss the working environment of the group.

Mark, thank you for being a great colleague and friend. You did your best to help this fresh colleague, both with her experiments and with her personal life. Your kind assistance made the project going more smoothly. Davy, without you the lab would lose much of its character. With your active and funny attitude we always had something nice to talk about. Thank you for letting me use your computer in the last period of the project.

Jaap, I would like to thank you for helping me with project discussions and guiding me in the right direction. Trish, thank you for helping me set up those experiments. Yamira and Neuza, thank you for your contributions to the project. I have a good memory of us working together. Rolf, I enjoyed sharing an office and having nice discussions with you. Sandra, Caroline, Xavier, Robbert, Simon, Angelique, Vera, Lucienne, and Chika, I enjoyed working with you and thank you for all the fun we had together in the group.

My IOP genomics partners, Elena, Harrie, Rachel, Coenie, Joris, Andreja and Hanem, thank you for all the work we did together and for the fun we had during our evening activities after our project meetings. Elena, thank you for assisting me with the microarray

experiments for a whole week in the beginning of the project. Harrie, thank you for helping out on those bioinformatics questions. Rachel and Coenie, thank you for sharing your results with me. They have been included in this thesis.

My Chinese friends, Herman, Xinying and Wang Ping, thank you for being able to share happy and unhappy moments from each other.

Marianne and Bert, thank you for your care and bringing me a good feeling of a real family life during the past few years.

My parents, thank you for your understanding and support for me being abroad. Because of your way of raising me I took the courage to take a big step and go abroad. Due to my life and your health situation, life has been not that easy for you. I know how much you would like to have me close to you. You often admire those parents having their daughter nearby, looking after each other. Nevertheless, your wishes to have me obtain the PhD degree would never change. 我的父母，谢谢你们对我远离家乡追求学业的理解和支持。在过去的几年中，因我的生活加上你们每况愈下的身体，让你们承受了很大的痛苦。我知道你们很羡慕邻居家的父母，常有女儿相伴。我的每一个困难让你们担忧不已，我深知可怜天下父母心的道理。你们的愿望也最终促使我完成博士学位。

Finally, I would like to give my deep appreciation to my husband, Patrick. You have brought me to a completely fresh new life. Your care, wisdom, honesty all make me feel so safe to be with you. All your support has encouraged me a lot to complete the thesis.

At last of course, my lovely son Jason. Although your arrival delayed mama's PhD progress, your laughing, crying, screaming and even your naughtiness have made me learn what life is all about. You drove mama through the heavy task of completing her thesis.

Xiao-Lian Yuan



Amsterdam, December 2007

Starch and Inulin
Modifying Network of
Aspergillus niger

