CHAPTER 2

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 J van den Hondel,
Lubbert Dijkhuizen and Arthur F J Ram

# These authors contributed equally to this work.

Microbiology (2006), 152, 3061–3073

Abstract

As a soil fungus, *Aspergillus niger* can metabolize a wide variety of carbon sources, employing sets of enzymes able to degrade plant-derived polysaccharides. In this study the genome sequence of *A. niger* strain CBS 513.88 was surveyed, to analyse 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 genes *suc1* (invertase activity; designated *sucA*), *inuE* (exo-inulinase activity) and *inuA/inuB* (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-hydrolysing enzymes (*inuA and inuE*) indicated that they are similarly regulated and all strongly induced on sucrose and inulin. Analysis of a Δ*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 Δ*creA* background, indicating that fructose did not act as an inducer. Evidence is provided that sucrose, or a sucrose-derived intermediate, but not fructose, acts as an inducer for the expression of inulinolytic genes in *A. niger*.

1. 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 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 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). *Aspergillus 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 hydrolyses inulin internally to produce mainly inulotriose and -tetraose (Akimoto *et al*., 1999), whereas exo-inulinase hydrolyses 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 (b-fructofuranosidase, EC 3.2.1.26), encoded by the *suc1* gene (Boddy *et al*., 1993), hydrolyses the β-2,1- glycosidic bond in sucrose to produce fructose and glucose

(L'Hocine *et al*., 2000). A specific b-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 in the 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. The coding sequence for one of them appears to be a pseudogene (inuQ), while the other two genes encode intracellular invertase-like proteins that 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.

2. Methods

2.1 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 CBS 513.88 (a natural derivative of strain NRRL 3122). 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 KH_2PO_4, 70 mM $NaNO_3$, 2 mM $MgSO_4$, 76 nM $ZnSO_4$, 178 nM H_3BO_3, 25 nM $MnCl_2$, 18 nM $FeSO_4$, 7.1 nM $CoCl_2$, 6.4 nM $CuSO_4$, 6.2 nM Na_2MoO_4 and 174 nM EDTA. Erlenmeyer flasks of 300 ml were inoculated with 2 x 10^6 spores ml^{-1} and incubated at 30 ^oC in a rotary shaker at 300 r.p.m. for 21 h. Each flask contained 100 ml MM (pH 6.5) supplemented with 0.1% (w/v) Casamino acids and 2% (w/v) carbon

source. Glucose, sucrose (BDH Chemicals), xylose, fructose and maltose (Sigma-Aldrich), inulin (Sensus Frutafit, Cosun) and starch (Windmill Starch, Avebe) 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 ^{o}C on a rotary shaker at 300 r.p.m. 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 MM containing 1% (w/v) of various carbon sources and incubated at 30 ^{o}C with agitation. Mycelial samples were taken at different time points by harvesting over a Miracloth filter followed by freezing in liquid nitrogen. The samples were stored at -80 ^{o}C prior to the isolation of total RNA. Conidiospores were obtained by harvesting spores from a plate of complete medium (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 ^{o}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 XL-1 Blue (Stratagene). Transformation of XL-1 Blue was performed by the heat shock protocol as described by Inoue *et al*. (1990).

2.2 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 became recently available (Pel *et al.*, 2007). Accession numbers of currently described members of families GH32 and GH68 were selected from the Carbohydrate-Active Enzymes server at 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 http://www.ncbi.nlm.nih.gov/entrez/ and 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 at http:// hmmer.wustl.edu/. Subsequently the *A. niger* genome was searched using the HMM profiles and the Wise2 package from 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 neighbour-joining method using 1000 replicates.

2.3 Norrthern analysis

Total RNA was isolated by grinding frozen (-80 ^{o}C) mycelium in liquid nitrogen with a pestle and mortar. Powdered mycelium (200 mg) was extracted with 1 ml TRIzol Reagent (Invitrogen) in accordance with the supplier's instructions. For Northern analysis, 5 mg total RNA was incubated with 3.3 ml 6 M glyoxal, 10 ml DMSO and 2 ml 0.1 M phosphate buffer (pH 7) in a total volume of 20 ml for 1 h at 50 ^{o}C to denature RNA. RNA electrophoresis was performed in a SEA-2000 electrophoresis apparatus (Elchrom Scientific) at 10 ^{o}C. The RNA samples were separated on 1.5% (w/v) agarose gel using 0.01 M phosphate buffer (pH 5) and transferred to Hybond-N filters (Amersham) by capillary blotting. Filters were prehybridized at 65 ^{o}C for 2 h in a solution of 0.9 M NaCl, 90 mM trisodium citrate, 1.0% (w/v) Ficoll, 1.0% (w/v) polyvinylpyrrolidone, 1.0% (w/v) bovine serum albumin, 10 mM EDTA, 0.5% (w/v) SDS and 100 mg ml^{-1} single-stranded herring sperm DNA. Hybridizations were performed at 42 ^{o}C for 18 h in a solution of 50% (v/v) formamide, 10% (w/v) dextran sulphate, 0.9 M NaCl, 90 mM trisodium 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 mg ml^{-1} single-stranded herring sperm DNA. Blots were washed twice in high-stringency washing buffer (30 mM NaCl, 3 mM trisodium citrate and 0.5%, w/v, SDS) for 20 min at 65 ^{o}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 (for primer sequences see supplementary Table S1). 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 *Eco*RI digestion of the pGEMT-easy vector containing the gene encoding the inulinolytic enzyme. Fragments were purified from gel and [a-32P]dCTP-labelled probes were synthesized using the Rediprime II DNA labelling system (Amersham Pharmacia Biotech) according to the instructions of the manufacturer.

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	*Aspergillus niger* CBS 513.88	This study
AngSuc1p	Invertase	S33920	*Aspergillus niger* B60	Boddy *et al.* (1993)
AsySftAp	Fructosyltransferase	CAB89083	*Aspergillus sydowii* IAM 2544	Heyer & Wendenburg (2001)
AjaFopAp	Fructosyltransferase	BAB67771	*Aspergillus japonicus* ATCC 20611	Yanai *et al.* (2001)
AngSucBp	Putative invertase	DQ233219	*Aspergillus niger* CBS 513.88	This study
AngSucCp	Putative invertase	DQ233220	*Aspergillus niger* CBS 513.88	This study
AngInuEp	Exo-inulinase	DQ233222	*Aspergillus niger* CBS 513.88	This study
Afo1-Sstp	Fructosyltransferase	CAA04131	*Aspergillus foetidus* NRRL 337	Rehm *et al.* (1998)
Ang12InuEp	Exo-inulinase	BAD01476	*Aspergillus niger* 12	Moriyama *et al.* (2003)
Aawinu1p	Exo-inulinase	CAC44220	*Aspergillus awamori* var. 2250	Arand *et al.* (2002)
PspInuDp	Exo-inulinase	BAC16218	*Penicillium* sp. TN-88	Moriyama *et al.* (2002)
AngInuAp	Endo-inulinase	DQ233221	*Aspergillus niger* CBS 513.88	This study
AfiInu2p	Endo-inulinase	CAA07345	*Aspergillus ficuum* ATCC 16882	Uhm *et al.* (1998)
Ang12InuAp	Endo-inulinase	BAA33797	*Aspergillus niger* 12	Ohta *et al.* (1998)
Ang12InuBp	Endo-inulinase	BAA33798	*Aspergillus niger* 12	Ohta *et al.* (1998)
PpuInuAp	Endo-inulinase	BAA12321	*Penicillium purpurogenum*	Onodera *et al.* (1996)
PspInuCp	Endo-inulinase	BAB19132	*Penicillium* sp. TN-88	Akimoto *et al.* (2000)
KmaInu1p	Inulinase	CAA40488	*Kluyveromyces marxianus* ATCC 12424	Laloux *et al.* (1991)
SceSuc2p	Invertase	P00724	*Saccharomyces cerevisiae*	Taussig & Carlson (1983)
PjaInv1p	Invertase	CAA73208	*Pichia jadinii* NRRL-Y1084	Chávez *et al.* (1998)
PanInv1p	Invertase	CAA56684	*Pichia anomala* CBS5759	Pérez *et al.* (1996)
SpoInv1p	Invertase	BAA25684	*Schizosaccharomyces pombe* TP4	Tanaka *et al.* (1998)

2.4 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 59 flanking DNA and 0.9 kb of 39 flanking DNA was amplified by PCR using primers CreAP1f and CreAP2r, CreAP3f and CreAP4r (supplementary

Table S1), 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 pBlueScriptII SK to obtain plasmids pF5 and pF3. Subsequently, pF3 was digested with *Bam*HI and *Kpn*I, and the fragment obtained 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 the *Aspergillus oryzae pyrG* gene, obtained from plasmid pAO4-13 (de Ruiter-Jacobs *et al.*, 1989), resulting in the *creA* disruption plasmid pXY1.1. Plasmid pXY1.1 was linearized with *Not*I and transformed into 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 analysed by PCR to identify possible Δ*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 analyse the presence of the wild-type *creA* gene were CreAP5f and CreAP7r, CreAP8f and CreAP6r (supplementary Table S1). 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 Δ*creA* strain in the remainder of the paper.

For complementation of the Δ*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 Δ*creA* strain XY1.1 to generate XY1.1-CreA.

2.5 Nucleotide accession numbers.

The *A. niger* CBS513.88 DNA sequences encoding family GH32 members, including 1000 bp upand downstream of the ORF, 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), DQ233222 (*inuE*) and DQ233223 (inuQ).

3. Results

3.1 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 of bacterial, fungal and plant origin (Coutinho & Henrissat 1999; Pons *et al*., 1998). The two families are structurally similar (clan GH-J), sharing a similar fivefold b-propeller fold (Meng & Futterer, 2003; Nagem *et al*., 2004). Protein sequences from members of families GH32 and GH68 were extracted from the GenBank/GenPept and Swiss-Prot 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 returned 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.

A neighbour-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 subgroups 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 four 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).

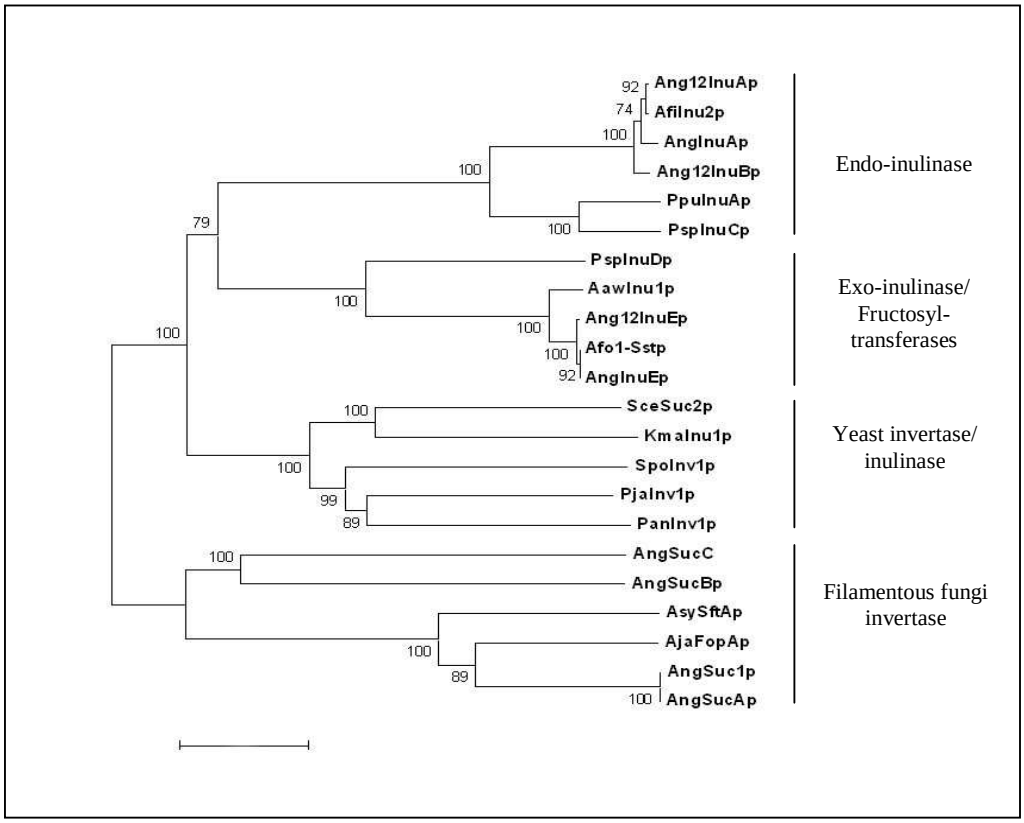

Figure. 1. Neighbour-joining tree of functionally characterized GH32 family members from filamentous fungi and yeast species. GH32 proteins identified in the genome of *A. niger* CBS513.88 are shown in bold. The proteins, their main activities, accession numbers, and the source (organisms) of the protein 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. The bar indicates 10% amino acid sequence difference.

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 blocks I, II and III by 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).

It has been previously reported that *A. niger* strain 12 contains two genes encoding 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* strains CBS513.88 and N402 was confirmed by Southern blot analysis (data not shown). Sequence comparison of the endo-inulinases of strains CBS513.88 and 12 revealed that the enzyme from strain CBS513.88 displayed higher similarity to InuA (9 amino acid differences) than to InuB (15 amino acid differences) of 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. The SucB and SucC protein sequences were analysed 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 an 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 result 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, evalue of 4×10^{-32}). SucC also displays the highest identity to the *A. niger* SucA protein (28% identity, 42% positives, evalue of 2×10^{-43}). Pairwise comparison indicated that SucB and SucC have higher identity 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).

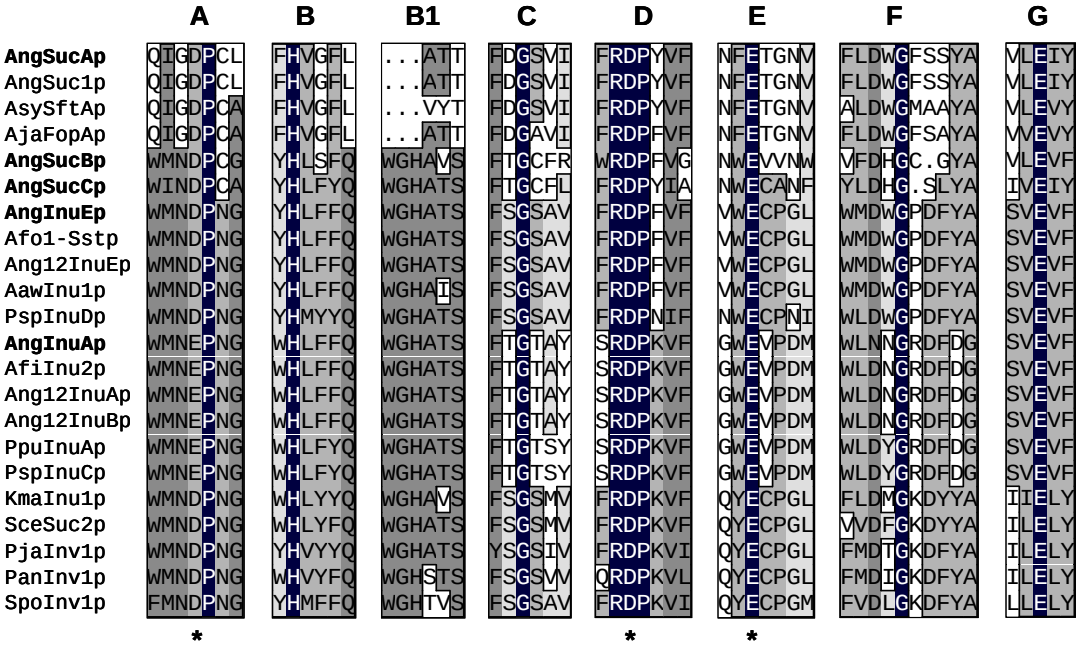

Figure. 2. Conserved domain alignment of family GH32 enzymes from filamentous fungi and yeast species. Conserved residues are shaded by different intensities based on homology level in the alignment. Black, 100% homology; dark grey, ≥75% homology; light grey, ≥50% homology; very light grey, ≥33% homology.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

3.2 Identification of glycoside hydrolase family 32 members in other fungal genomes

The genomes of the ascomycete fungi *Aspergillus nidulans*, *Aspergillus fumigatus*, *Neurospora crassa*, *Gibberella zeae* and *Magnaporthe grisea* and the basidiomycete fungus *Ustilago maydis* were analysed for the presence of GH32 family members. The results are summarized in Table 2 and a phylogenetic tree is available as supplementary Fig. S1.

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-hydrolysing 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) has high identity to the *A. niger* SucB protein (Table 2). As in *A. niger*, the protein is predicted to be intracellular in *A. nidulans*, as an N-terminal signal sequence is not present. Orthologues 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 neighbour joining tree (Fig. S1). *N. crassa* contains only a single GH32 family member. This protein displays the highest sequence identity with β-fructofuranosidases of bacterial origin (*Bacillus megaterium* FruA; e-value 3×10^{-72}) (Chiou *et al*., 2002) and groups in the fungal tree together with the yeast-like invertases (Fig. S1). *G. zeae* (anamorph *Fusarium graminearum*) contains five GH32 family members. FG08415.1 is the putative orthologue 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 *Saccharomyces cerevisiae* Suc2 protein (e-value 3×10^{-82}) and clusters together with yeast-like invertases. FG03288.1 encodes a protein with high identity to the SucC protein of *A. niger*. The protein is also predicted to be intracellularly localized.

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

Fungus	ORF number	Closest *A. niger*	homolog	Closest functionally
A. nidulans	AN3837.2	SucB; 1e-144	AngSucA; 1e-33	Intracellula
	AN5012.3*1	InuE; 4e-166	AngInuE; 4e-166	Extracellular
A. fumigatus	Afu2g01240	InuE; 2e-70	PanInv1; 3e-79	Extracellular
	Afu5g00530	InuA; 4e-183	PpuInuC; 0.0	Extracellular
	Afu5g00480	InuE; 3e-169	PpuInuD; 0.0	Extracellular
	Afu6g05000	SucB; 1e-182	AngSucA; 2e-36	Intracellular
N. crassa	NCU04265.2	InuE; 3e-41	BmeFruA; 3e-72	Extracellular
G. zeae	FG02067.1	InuA; 5e-05	PmaCft; 2e-09	Intracellular
	FG02339.1	InuE; 7e-59	ScSuc2; 3e-82	Extracellular
	FG03288.1	SucC; e-136	AngSucA;1e-50	Intracellular
	FG06451.1	InuE; 4e-61	BsuSacC; 1e-72	Extracellular
	FG08415.2*1	SucA; 0.0	AngSucA; 0.0	Extracellular
M. grisae	MG02507.4	InuE; 1e-55	KmaInu1; 3e-83	Extracellular
	MG05785.4	SucA; 2e-49	AngSucA; 2e-49	Extracellular
	MG07837.4	SucB; 8e-36	AsySft1; 7e-30	Extracellular
	MG10748.4	SucA; 1e-08	AngSucA; 1e-08	Truncated protein; Domain A, B1, B, C, D, F are missing
	MG10767.5*1	SucA; 2e-14	AngSucA; 2e-14	Extracellular
U. maydis	UM01945.1	InuE; 1e-66	SpoInv1; 2e-90	Extracellular
	UM03605.1	SucB; 2e-89	AngSucA; 3e-37	Intracellular

*BLASTP searches were performed with the full-length protein sequences.

†Localization was predicted using the 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. ‡Annotation improved manually.

FG06451.1 shows the highest identity (e-value 1 x 10^{-72}) to the *Bacillus subtilus* 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 2 x 10^{-09}) to a cycloinulooligosaccharide fructanotransferase from *Bacillus macerans* (Kim & Choi, 2001). *M. grisea* contains five GH32 enzymes, four

of which cluster in a separate branch different from the other fungal proteins (Fig. S1). 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, Fig. S1), whereas the remaining four form a distinct group with high sequence identities to each other (e-value 1 x 10^{-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, *U. 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 most closely 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 to be unique to *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 (Fig. S1).

3.3 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 of the Northern blot (not shown). The detection of two different sized mRNAs for the *sucB* gene suggests two different mRNA start sites, or two different polyadenylation 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 that 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 all four genes, *inuE*, *sucA*, *inuA* and *sucB*, was 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 *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 in Fig. 4 also indicate that, although the different genes are co-regulated and all induced on sucrose or inulin, there were slight differences in expression pattern. 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 may be due to differences in mRNA stability, or to expression also being 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 binding sites for PacC (5'-GCCARG-3') (see for review Penalva & Arst, 2002) and CreA (5'-SYGGRG-3') (see for review Ruijter & Visser, 1997) in the promoters of most of these genes might support the latter explanation.

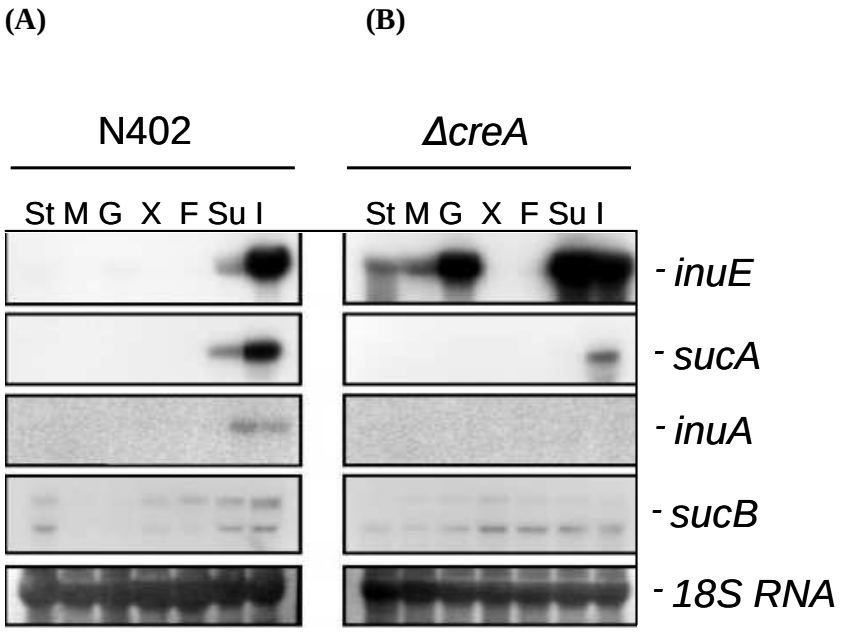

Figure. 3. Expression analysis of *A. niger* inulinolytic genes in wild-type N402 (A) and the Δ*creA* strain (B). Total RNA was extractd from mycelia grown for 21 h at 30 ^{o}C in MM containing 2% (w/v) of different carbon sources. RNA (10 mg) was separated by agarose gel electrophoresis, blotted, and hybridized with ^{32}P-labelled 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.

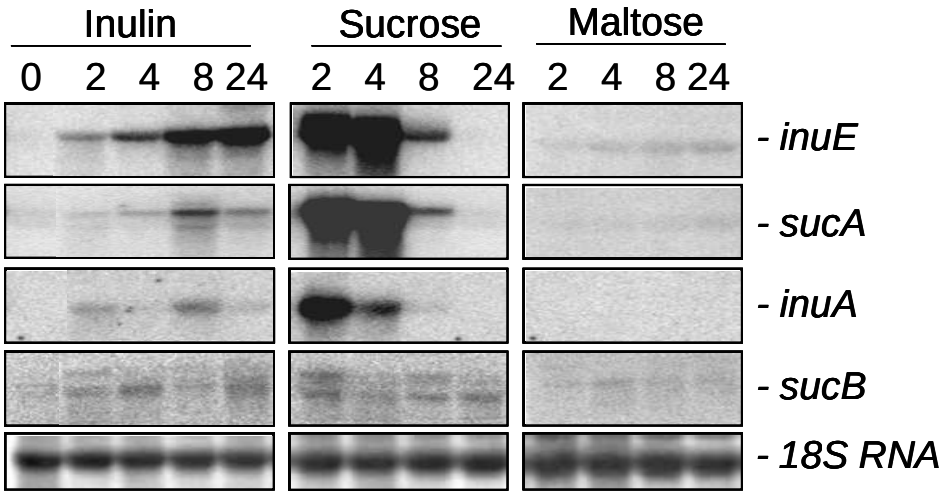

Figure. 4. Northern blot analysis of inulinolytic genes in *A. niger* wild-type strain after transfer from xylose to different carbon sources. *A. niger* was pregrown in MM containing 2% (w/v) xylose and 0, 1% (w/v) Casamino acids at 30 ^{o}C for 18 h. Total RNA was prepared from mycelia grown for 2, 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 mg total RNA was loaded on the gel. Ethidium bromide staining of 18S rRNA was used as loading control.

3.4 Expression analysis of the inulinolytic genes in a Δ*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. 3(a) 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 a Δ*creA* mutant of *A.*

niger strain N402. A Δ*creA* null mutant was generated and verified as described in Methods. The phenotype of the Δ*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 Δ*creA* strain with the wild-type *creA* gene fully complemented the reduced growth and reduced conidiation phenotype of the Δ*creA* strain (data not shown).

The expression analysis of the *inuE* gene in the Δ*creA* strain showed that the expression of this gene was higher than in the wild-type strain, indicating that it is under control of carbon metabolite repression. Expression of *inuE* was found in the Δ*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 was undetectable in both the wild-type and the Δ*creA* strain. Apparently, the expression of the *inuE* gene requires an activator or inducer molecule and is not expressed under derepressing (Δ*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 inulinolytic enzymes showed a different expression pattern in the Δ*creA* strain. *sucA* and *inuA*, both expressed specifically on inulin and sucrose in the wild-type strain, were expressed at a lower level (*sucA*), or not detectable (*inuA*) in the Δ*creA* strain (Fig. 3b). Expression of *sucA* and *inuA* was not detected after growth on xylose and fructose, similar to what was observed for *inuE*. In contrast to the expression profile of the *inuE* gene, no expression was detected of *sucA* and *inuA* 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 (six and five sites, respectively), the expression of both genes does not seem to be directly controlled by CreA. Disruption of catabolite repression in the Δ*creA* strain could lead to the inactivation of complex pathways, which might bring about decreased or total loss of expression. The expression pattern of *sucB* 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 Δ*creA* background showed two remarkable features: (i) expression of s*ucB* in the Δ*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 Δ*creA* background favours the transcription of the smaller-sized mRNA molecule. The 1 kb promoter region of *sucB* contains seven putative CreA-binding sites. Expression of the *sucC* gene was also not detected in the Δ*creA* mutant (data not shown). Clearly, the different responses of *inuE*, *sucB* and *sucA/inuA* in a Δ*creA* mutant background suggest the involvement of additional factors other than the presence of an inducer and repression via CreA. These might include environmental factors such as pH, nitrogen availability and temperature.

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

Fructose has been shown to induce expression of inulinases in the yeast *K. fragilis* (Grootwassink & Hewitt, 1983). The expression analysis of the inulinolytic genes in the wild-type *A. niger* strain and the Δ*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 pregrown 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 mM, 50 mM, 5 mM, 500 nM, 50 nM, 5 nM, and no carbon sources. As shown in Fig. 5(a), 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 in 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 50 mM sucrose, indicating that some form of repression at high sucrose concentrations may exist, e.g. catabolite repression by released glucose from sucrose hydrolysis. The addition of 50 mM or 5 mM of glucose, 1-kestose or 1-nystose did not trigger induction of the inulinolytic system (data not shown).

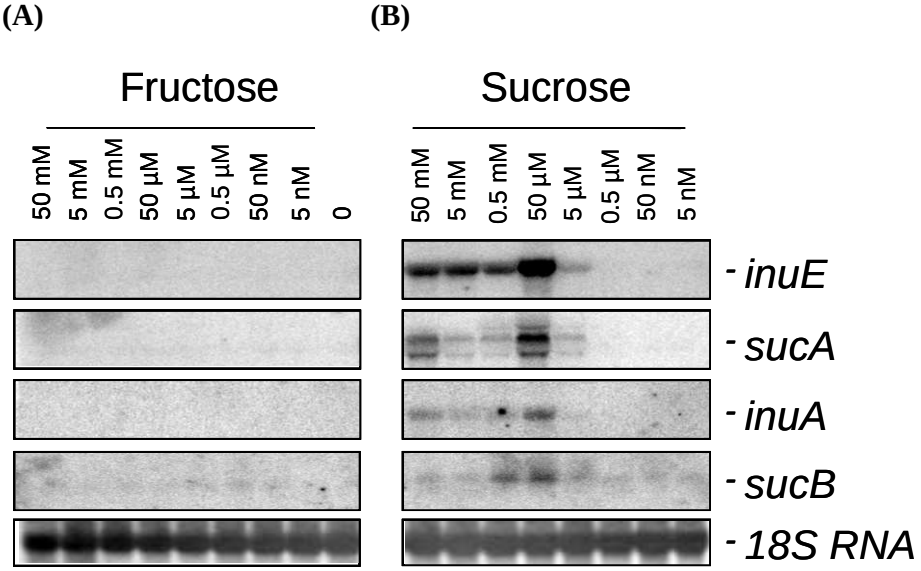

Figure. 5. Northern blot analysis of genes encoding inulinolytic enzymes of *A. niger* N402 using carbon-shift experiments. *A. niger* was pregrown in MM containing 2% (w/v) glycerol at 30 ^{o}C for 18 h. Total RNA was prepared from mycelia grown for 4 h after transfer to fresh MM containing concentrations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

4 Discussion

The availability of the *A. niger* genome sequence 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 ORF. In addition to three extracellular inulinolytic activities of *A. niger*, the invertase encoded by the *suc1* gene, the exoinulinase 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 enzymic activities in *Aspergillus* spp. 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 was originally published as an enzyme from *A. niger* strain ATCC 20611, but this strain has been reclassified as an *Aspergillus japonicus* strain (ATCC culture collection database; http://www.lgcpromochem-atcc.com/). 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 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 enzymic 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, is therefore likely to encode a fructofuranosidase lacking any detectable fructosyltransferase activity under the given assay conditions. A second fungal fructosyltransferase-encoding gene was identified in *Aspergillus 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-mass 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 orthologues were not identified in the genomes 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 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 6C 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 mg total RNA was loaded on the gel. Ethidium bromide staining of 18S rRNA was used as loading control. http://mic.sgmjournals.org 3069 Genes encoding inulin-modifying enzymes of *A. niger* 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 *Aspergillus 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-termini 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 *Aspergillus 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 b-2,6 glycosidic linkages) activity. No transfructosylation activity has

been reported for the A. awamori Inu1 protein. The different enzymic 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. Biochemical analysis of the recombinant SucB protein expressed in *E. coli* indeed indicates that SucB has hydrolysing 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-kestose, respectively (**Chapter 3**; Goosen *et al*., 2007).

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 endoplasmic reticulum 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 nonclassical 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 widespread among filamentous fungi and not specific to *A. niger*.

The presence of carbohydrate-degrading enzymes that do not contain a signal sequence is 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*., 2005). The presence of these intracellular enzymes strongly suggests that fungi are capable of transporting oligosaccharides into the cell which are subsequently hydrolysed by the intracellular enzymes. Alternatively, these intracellular enzymes may also possess transglycosylation activity, e.g. for the synthesis of inducer molecules that activate transcription factors.

The genome of *A. niger* also contained an ORF which showed homology to the group of exo-inulinases. However, to assemble this ORF encoding 137 amino acids,

containing some of the conserved domains 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). 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 hydrolyse 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 s*ucA* and *inuE* on inulin and sucrose is in agreement with previous obserrvations (Wallis *et al*., 1997; Moriyama *et al*., 2003; Rehm *et al*., 1998). The expression of the single endoinulinase (*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 five putative CreA-binding motifs (SYGGRG) in their first 1000 bp. creA encodes a wide-domain regulatory protein that binds to the promoter of target genes to prevent or decrease expression if a favourable carbon source (such as glucose or fructose) is present (Dowzer & Kelly, 1991; Ruijter & Visser, 1997). No such binding sites were observed when analysing 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 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 mM 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 mM 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 analysed 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 of fructose may have resulted in starvation and the inability of the fungus to induce expression due to 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 Δ*creA* strain on fructose should lead to a high level of expression of the genes encoding the inulinolytic enzymes. However, in Fig. 3(b), we show that the inulinolytic genes are not expressed in the Δ*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 glucose, 1- kestose or 1-nystose did not induce expression of the inulinolytic genes either, we suggest a mechanism by which

sucrose is transported across the plasma membrane. Once intracellular, the sucrose molecule, or a derivative of it, 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).

Table S1. Primers used in this study

Primer name	Sequence 5’ to 3’
InuEP1f	CTGGTGGCATCGAGTGGGG
InuEP2r	GTGGCGGGGAATGGCCAT
InuAP1f	TGGATGAACGAGCCAAACG
InuAP2r	ACGGCCTCTCCTTGTCCGCCA
SucAP1f	ATGAAGCTTCAAACGGCTTC
SucAP2r	AAGATACTCACCGAACCCAA
SucBP1f	CTCCGAATGCCCTGACTGCTA
SucBP2r	GGGTCTGAAAGGGCTGGTTT
SucCP1f	GGAACGCAGCCGGTTTAT
SucCP2r	GCTAACACCAAGCGAGACGG
InuQP1f	TGTATCAGAATCCGCGGT
InuQP2r	TAGTCCGGTGCGAATACG
CreAP1f	TTGCGGCCGCCGACACCCAACAATACGGG
CreAP2r	CGGGATCCCGGTCGACGCATGTGAAGCTTGTCCCAA
CreAP3f	CGGGATCCTCAGCCACACGTTGGTTTG
CreAP4r	GGGGTACCGGGAATGGTCTGGTCTCCGT
CreAP5f	CCTTACAGCTTTACCTTAC
CreAP6r	GTAAAGTACCCCGACTGC
CreAP7r	CGACGCCATGTTGGAGTTC
CreAP8f	AGTCCTTCCCAGGGTCACCA
PAO10	TTCGCGAGACTGAATGCG
PAO9	AATGTCAATTCCAGCAGCG

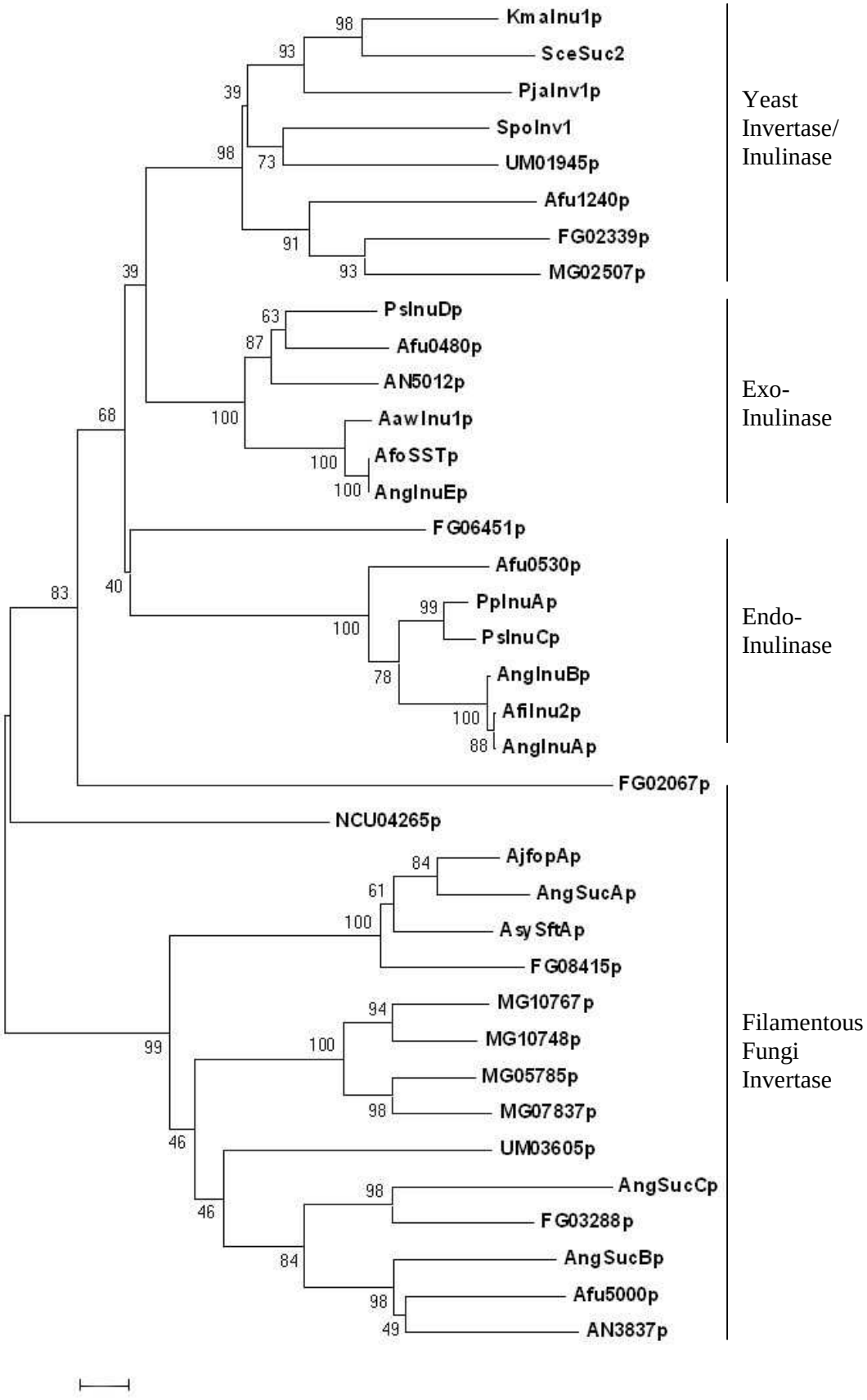
KmaInu1p
SceSuc2
PjaInv1p
SpoInv1
UM01945p
Afu1240p
FG02339p
MG02507p
PsInuDp
Afu0480p
AN5012p
AawInu1p
AfoSSTp
AngInuEp
FG06451p
Afu0530p
PpInuAp
PsInuCp
AngInuBp
AfiInu2p
AngInuAp
FG02067p
NCU04265p
AjfopAp
AngSucAp
AsySftAp
FG08415p
MG10767p
MG10748p
MG05785p
MG07837p
UM03605p
AngSucCp
FG03288p
AngSucBp
Afu5000p
AN3837p
Yeast Invertase/ Inulinase
Exo-Inulinase
Endo-Inulinase
Filamentous Fungi Invertase

Figure. S1. Neighbour-joining tree of GH32 family members identified in the genomes of *A. niger, A. nidulans, A. fumigatus*, *N. crassa*, *G. zeae*, *M. grisea* and *U. maydis*, together with functionally described GH32 family members from filamentous fungi and yeasts. If the fungal protein has a highest blastp hit with a bacterial GH32 enzyme, this enzyme was included in the tree. BmeFruA, *Bacillus megaterium* FruA (AAM19071); BsuSacC, *Bacillus subtilus* SacC (CAA29137); BmaCft, *Bacillus macerans* Cft (Q9F0I5). Proteins predicted to lack an N-terminal signal sequence were considered as intracellular enzymes and indicated by the grey background. Accession numbers of the proteins are listed in Tables 1 and 2 of the main paper. Bootstrap values are indicated at the node of each branch. The tree was created with Mega 3.1 usingdefault settings for gap and extension penalties. Bar indicates 10% amino acid sequence difference.