CHAPTER 6

Summary and concluding remarks

Genomics, the study of an organism's whole hereditary information, enables the elucidation of metabolic pathways, the prediction of proteins and their functions, and the gaining of deeper insight into the complex regulation within living cells. The building block of the genome is DNA (Deoxyribonucleic acid), the basic molecule for storage of information. Translation of this information via messenger molecules or mRNA (messenger Ribonucleic acids) results in the formation of proteins. Proteins perform a vast array of functions, including structural and catalytic ones. In order to gain full insight into the protein diversity of cells, one needs to know the complete DNA sequence of the genome. DNA sequencing was first established by Sanger and Coulson (1975). Over the years, technological advances in chemistry and information technology enabled scientists to sequence complete genomes of organisms. This, inevitably also has resulted in unraveling of the human genome (The IHGSC, 2001; Venter *et al.*, 2001; The IHGSC, 2004).

Genome sequencing expanded rapidly to all levels of the tree of life, where an ever increasing number of sequences are becoming publicly available (http://www.ncbi.nlm.nih.gov/Genomes). Fungi, as important agents in pathology and industry, have enjoyed the same strong focus of attention in the field of genomics. Currently, more than 40 complete genome sequences of fungi are available (Galagan *et al.*, 2005b). The Dutch company DSM has invested substantially in order to obtain the genome sequence of *Aspergillus niger* (Pel *et al.*, 2007). *A. niger* belongs to the class of the *Hyphomycetes*, under the subdivision *Deuteromycotina* (Imperfect fungi), and has enjoyed much interest from industry as a potent producer of organic acids and enzymes (Abarca *et al.*, 2004). In industry, *A. niger* is used extensively as producer of citric acid and as a source of a diversity of industrial enzymes (Olempska-Beer *et al.*, 2006; Pel *et al.*, 2007). *A. niger* also has the GRAS (generally regarded as safe) status, indicating that products and enzymes produced by this fungus are considered safe to be used in food and pharmaceutical applications. Understanding the true diversity of enzymes produced by this organism requires their identification and possible function prediction (annotation). Confirmation of enzyme activity is then done by expressing the corresponding genes in host organisms, followed by biochemical characterization of their respective activities. Under the flag of a Dutch national funding program (http://www.senternovem.nl/iopgenomics/) and in collaboration with academic and industrial partners (http://www.biopoort.net/carbnet/carbnet.html), we set forth to identify and biochemically characterize the carbohydrate active enzymes encoded in the genome of *A. niger* CBS518.33 (the strain sequenced by DSM) as well as the strain N402 (used for expression analysis and as source for targeted overexpression studies). The work described in

this thesis focused on enzymes active on sucrose, and the fructans inulin and levan (**Chapter 1**). Fructans are produced by plants, bacteria and fungi, and are used in the food industry because of their product improving and health-beneficial characteristics (Gibson *et al.*, 1995; Kaur & Gupta, 2002). Currently, a number of enzymes have been identified from *A. niger* which can hydrolyse and modify sucrose and fructans. These include the enzymes invertase (primarily hydrolysing sucrose; Boddy *et al*., 1993), endo-inulinase (endo-acting, hydrolysing inulin; Ohta *et al*., 1998), exo-inulinase (exo-acting, releasing free fructose from sucrose, inulin and levan; Arand *et al*., 2002; Moriyama *et al*., 2003) and fructosyltransferase (FTF or sucrose:sucrose 1- fructolsyltransferase (1-SST); making inulin-type short chain oligomers from sucrose; Nguyen *et al.*, 1999; L'Hocine *et al.*, 2000). These enzymes belong to glycoside hydrolase family 32 (GH32) based on the presence of conserved amino acid domains (Carbohydrate Active Enzymes database; http://www.cazy.org/; Coutinho and Henrissat, 1999; **Chapters 2 and 3**). Together with glycoside hydrolase family 68 (GH68), they belong to enzyme clan GH-J, sharing a similar three-dimensional structural fold (five-bladed β-propeller fold; Meng and Frutterer, 2003; Nagem *et al*., 2004). GH68 enzymes include inulo- and levansucrases (inulin and levan formation, respectively), as well as invertases, with currently only proteins represented from *Archaea* and *Bacteria*.

In **Chapter 2** the diversity and expression regulation characteristics of the sucrose and fructan active enzymes of *A. niger* strains CBS513.88 and N402 are described. Further chapters present detailed descriptions of a novel intracellular invertase (SucB, **Chapter 3**) and the extracellular exo-inulinase (AngInuE, **Chapter 4**) identified in the genome of *A. niger*. Focus was also placed on the putative fructan binding domain of GH32, and its possible role in catalysis and substrate recognition.

Database mining and transcriptional analysis of genes encoding inulin-modifying enzymes of *A. niger*

The diversity and transcriptional regulation of GH32 enzymes in *A. niger* were explored in **Chapter 2**. Sequences of known family GH32 proteins, as well as the structurally related family GH68 proteins, were used to create database searching profiles (HMM profiles) to identify any additional members present in the genome sequence of the *A. niger* strain CBS513.88. Using this profile, a total of five GH32 proteins were identified based on the presence of a full length open reading frame, and the conserved family GH32 domains (Yuan

et al., 2006; **Chapter 2, Fig. 2**). No family GH68 members could be identified in the genome sequence of *A. niger*.

By means of multiple sequence alignments, combined in a neighbour-joining tree, it was shown that these five GH32 members represent endo-inulinase (AngInuA), exo-inulinase (AngInuE) and filamentous fungal invertase (AngSucAp, AngSucBp, AngSucCp) proteins (Yuan *et al.*, 2006; **Chapter 2, Fig. 1**). From this analysis, three striking features came to light. Firstly, no specific FTF-type enzyme could be identified in the genome sequence, although such enzyme could be purified from another *A. niger* strain (L'Hocine *et al.*, 2000). However, a single exo-inulinase with FTF activity, AngInuEp, was identified. This exo-inulinase possessed complete sequence identity with an FTF from *A. foetidus* (Rehm *et al.*, 1998), and high sequence similarities with the exo-inulinases from *A. niger* 12 and *Aspergillus awamori* (Arand *et al.*, 2002; Moriyama *et al.*, 2003). Secondly, only a single endo-inulinase (AngInuAp) could be identified in the genome sequence, compared to two (InuA and InuB) identified in *A. niger* 12 (Ohta *et al.*, 1998). The endo-inulinase of *A.niger* CBS513.88 and N402 displayed a higher deduced amino acid sequence similarity with InuA than with InuB, and was thus accordingly named InuA (**Chapter 2**). Thirdly, two novel putative intracellular invertases, AngSucB and AngSucC (**Chapters 2 and 3**) were identified. These proteins lack an N-terminal signal sequence, and are thus very likely localized intracellularly. Phylogenetic analysis showed that the abovementioned family GH32 enzymes are conserved in different *ascomycete* fungi, indicating their importance in fructan metabolism (Yuan *et al.*, 2006; **Chapter 2**, **Table 1 and Fig. S1; Chapter 3, Fig. 1 and Table 2**).

Expression analysis of the genes encoding GH32 enzymes from *A.niger* N402 indicated that *AngInuE*, *AngSucA* and *AngInuA* were only expressed when sucrose or inulin was used as carbon source. *AngSucB* was expressed (at a low level) on all carbon sources tested, whereas no expression of *AngSucC* could be detected on any of the carbon sources used (Yuan *et al.*, 2006; **Chapter 2**, **Figs**. **3, 4 and 5**). Transferring *A. niger* from medium containing xylose as sole carbon source to one containing only inulin, sucrose or maltose, confirmed that all abovementioned GH32 enzymes are only up-regulated when sucrose and inulin, but not maltose, were used as carbon sources (**Chapter 2**, **Fig. 4**). Even in the presence of increasing concentrations of fructose, no expression could be observed from any of the identified genes encoding GH32 enzymes. This observation fits well with an earlier one made for the invertase of *A. nidulans* (Vainstein & Peberdy, 1991), indicating similar regulation of fructan modifying enzymes in the two species. The expression of genes encoding these

proteins probably required additional factors not expressed in the presence of only xylose or fructose as carbon sources. Expression analysis was also carried out in an *A. niger* strain hampered in catabolite repression (*A. niger* Δ*creA*). The results obtained showed that expression of *AngInuE* and *AngSucB*, but not *AngSucA* and *AngInuA*, was directly controlled by catabolite repression (**Chapter 2**, **Fig. 3**). These findings lead to the proposal that sucrose, as main inducing molecule, could be transported over the plasma membrane. Once inside, the sucrose molecule, or a derivative thereof, could act as an inducer for the expression of genes encoding GH32 enzymes (**Chapter 2**).

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

Analysis of the *A. niger* genome sequence revealed the presence of *sucA*, encoding the previously characterized extracellular invertase, and two novel putative invertase genes, *sucB* and *sucC*. Database analysis revealed the presence of SucB orthologues in various other fungal species. In silico analysis indicated that these putative invertases all lack any recognizable signal peptide sequences. Phylogenetically these putative intracellular invertases of fungal origin cluster in a new distinct group (**Chapter 3, Fig. 1**). Although all 8 conserved domains characteristic of GH32 are present, the SucB subfamily showed interesting sequence variations (**Chapter 3; Table 2**). Intracellular invertases have been identified in fungi before (Gascon & Lampen, 1968; Maramatsu & Nakakuki, 1995; Nguyen *et al*., 1999), but very little is known about their biochemical characteristics and functions in the intracellular environment. The putative intracellular invertase (SucB) from *A. niger* was heterologously expressed in *Escherichia coli* and in the yeast *Saccharomyces cerevisiae* (**Chapter 3**). Purified SucB from *E. coli* expression was biochemically characterized. SucB displayed typical characteristics of an invertase: it hydrolysed sucrose and short inulin-type oligosaccharides, but not high molecular weight inulin or levan. Apart from hydrolysis, SucB also has transfructosylating properties, producing 1-kestose and nystose from sucrose and 1-kestose*,* respectively. Transfructosylation activity was directly proportional to the substrate concentration (between 20 – 50 % using a sucrose concentration range of 2.5 mM to 1 M) (**Chapter 3**, **Fig. 3**). SucB also displayed typical Michaelis Menten-type kinetics with substrate inhibition on sucrose, with apparent *Km, Ki,* and *Vmax* values of 2.0 (±0.2) mM, 268.1 (±18.1) mM and 6.6 (±0.2) µmol min^{-1} mg^{-1} of protein (total activity), respectively.

Both the *Km* and *Vmax* values are substantially lower than those reported for other fungal and bacterial invertases (Gascon & Lampen, 1968; Boddy *et al*., 1993; Reddy *et al*., 1996; Wallis *et al*., 1997; Liebl *et al*., 1998; L'Hocine *et al*., 2000). SucB also displayed a narrow pH activity range (above pH 4, below pH 7; **Chapter 3, Fig. 2A**) compared to that observed for the extracellular invertase Suc1 of *A. niger* (above pH 3, below pH10) (Boddy *et al*., 1993; Wallis *et al*., 1997). At sucrose concentrations up to 400 mM, transfructosylation (FTF) activity contributed approximately 20 to 30% to total activity. At higher sucrose concentrations, FTF increased to up to 50% of total activity. In contrast, the purified extracellular A. niger invertase did not display any detectable FTF activity, not even at a sucrose concentration as high as 2.2 M (L'Hocine *et al*., 2000). Compared to wild-type *A. niger*, disruption of *sucB* (*A. niger* Δ*sucB*) brought about an earlier onset of sporulation on solid media, independent of the carbon source. However, in liquid media no differences between the growth of wild-type *A. niger* and *A. niger* Δ*sucB* could be observed.

Taking the above observations together, it is proposed that SucB, as part of a novel group of fungal intracellular invertases (Goosen *et al*., 2007; **Chapter 3**, **Fig. 1**), functions by hydrolysing imported sucrose, kestose or nystose, producing free glucose and fructose for metabolism. By performing transfructosylation, SucB could also have a function in energy storage, regulation of osmolarity, or synthesis of the inducer of other fructan and sucrose modifying enzymes. Furthermore, disruption of the *sucB* gene showed that this enzyme does not play an essential role in the metabolism of sucrose, but could be (in) directly involved in the sporulation of *A. niger*.

Exo-inulinase of A*spergillus niger* N402: a hydrolytic enzyme with significant transfructosylating activity

Genome analysis had identified a single exo-inulinase (*AnginuE*) in *A. niger* CBS513.88 (Yuan *et al*., 2006; Pel *et al*., 2007; **Chapter 2**). Phylogenetic analysis showed that AngInuE is identical to the FTF of *A. foetidus*, a glycosylated, dimeric enzyme lacking inulin and levan hydrolytic activity (Rehm *et al*., 1998), and 99% and 91% identical to the monomeric exo-inulinases of *A.niger* 12 (InuE; Moriyama *et al*., 2003) and *A. awamori* (Inu1; Arand *et al*., 2002) (**Chapter 2**). As typical exo-inulinases, both InuE and Inu1 hydrolyse inulin, but lack any detectable FTF activity (Moriyama *et al*., 2003; Kulminskaya *et al*., 2003). However, Inu1, opposed to InuE, is able to hydrolyze levan as well.

In order to determine the true enzymatic characteristics of AngInuE, the gene encoding this enzyme was heterologously expressed in and purified from *E. coli*. Results showed that AngInuE is a monomeric enzyme of 57 kDa, which displayed pH and temperature optima characteristics close to what has been described for InuE and Inu1. However, AngInuE also displayed clear transfructosylation activity in the presence of sucrose, 1-kestose and nystose, and produced oligosaccharides of both the inulin, levan as well as neo-series inulin type (for chemical structures, see **Chapter 1;** Goosen *et al.*, submitted for publication; **Chapter 4**).

AngInuE hydrolysed inulin and levan to release free fructose (Goosen *et al.*, submitted for publication; **Chapter 4**). Strikingly, although AngInuE differed from InuE by only three amino acids (Gln199His, Ser476Gly and Ser499Thr), the two enzymes display different substrate and product specificities. Following site directed mutagenesis of the three amino acids AngInuE to their InuE counterparts, no changes in inulin or levan hydrolysis and FTF activity could be observed.

A structural feature widely present in GH32 enzymes is the C-terminal β-sandwich domain, which contains the conserved sequence motif SVEVF (GH32 domain G). Although the precise function of this domain is unkown, its presence appears to be essential for activity. Site-directed mutagenesis of Ser469 in the conserved family GH32 domain G significantly reduced activity on sucrose, inulin and levan (Goosen *et al.*, submitted for publication; **Chapter 4)**. Mutations in Ser469 also resulted in decreased transfructosylation activity, compared to wild-type AngInuE. SVEVF is highly conserved among fungal polymeric fructan hydrolysing enzymes, but not in invertases (Yuan *et al.*, 2006), and has been proposed to play a role in polymer binding (Burne *et al.*, 1992; Ohta *et al.*, 1998; Moriyama *et al.*, 2003).

Our data show that AngInuE, expressed and purified from *E. coli*, functions as a monomeric, non-glycosylated exo-inulinase, hydrolysing sucrose, inulin and levan. The enzyme is also capable of performing significant transfructosylation activity with increasing sucrose concentrations. Differences in molecular weight and biochemical characteristics between 1-SST, InuE and Inu1 may be based on differences in (1) analytic methods used, (2) *E. coli* (non-glycosylated) and *A. niger* produced enzymes, or (3) production of different isoforms of the enzyme in *A. niger*. Finally, our data show that the conserved motif SVEVF in domain G of the family GH32 β-sandwich domain clearly plays a significant role in overall enzyme activity, and may also be responsible for polymeric fructan binding.

The GH32 enzyme network of *A. niger*

Enzymes identified by *in silico* genome mining, together with their known and experimentally elucidated enzyme characteristics (Yuan *et al*., 2006; Goosen *et al*., 2007; Goosen *et al*., submitted for publication; **Chapters 2-4**) enabled us to propose a model for the sucrose and fructan modifying enzyme network of *A. niger* (Fig. 1). In the environment, *A. niger* encounters sucrose or fructan rich decaying plant material. This triggers the induction of sucrose and fructan modifying enzymes (Yuan *et al*., 2006; **Chapter 2**). Sucrose, but not inulin, most probably enters the fungal hyphe *via* a specific importer, inducing the transcription of the sucrose and fructan modifying enzymes. Hydrolytic enzymes containing signal sequences are exported via a protein secretion pathway (see Pel *et al*., 2007), enabling them to get into direct contact with extracellular sources of sucrose and fructans. The fungus *A. niger* is able to hydrolyse these carbohydrates using the extracellular invertase SucA, the exo-inulinase AngInuE and the endo-inulinase InuA. SucA releases free glucose and fructose from sucrose. InuA specifically hydrolyses inulin into smaller inulo-oligosaccharides. By doing this, more terminal non-reducing ends from these inulo-oligosaccharides become available as substrates for the exo-inulinase AngInuE. This enzyme completely hydrolyses these inulo-oligosaccharides, as well as sucrose, to free fructose and glucose. Apart from the hydrolysis activity, AngInuE produced small amounts of inulin-, levan and neo-series inulo-oligosaccharides (**Chapter 4**). These oligosaccharides might be imported into the cell, but are most likely hydrolysed again to free fructose and glucose.

The industrially importance of *A.niger* (and *A.oryzae*) is firmly nestled on their ability to secrete large quantities of proteins. These filamentous fungi also contain the highest percentage of extracellular carbohydrate active enzymes, which are extensively used in industry (Abarca *et al*, 2004; Pel *et al*., 2007). Phylogenetic analysis have shown that apart from a large number of extracellular enzymes, Aspergilli also contain a significant repertoire of putative carbohydrate-active enzymes lacking any detectable export signals (Pel *et al*., 2007). Although little is known regarding the biochemical properties of these putative intracellular enzymes, their widespread occurrence in different Aspergilli indicate that they are conserved in function. These enzymes could play an important role in the intracellular metabolism of imported di- or small oligosaccharides, in gene regulation by synthesis of specific inducer molecules, in the salvaging of cellular carbohydrates following mycelial breakdown, in osmolarity, or even in differentiation related processes, e.g. sporulation in case of SucB.

SucB, as putative intracellular invertase, is situated in the *A. niger* genome next to a putative hexose importer (An15g00310; Pel *et al.*, 2007), which might function by importing small amounts of sucrose from the extracellular environment. Once inside, SucB may hydrolyse this sucrose to make glucose and fructose available for catabolism, or perform transfructosylation resulting in the formation of small inulin-, levan- and neo-series oligosaccharides. Disruption of the *sucB* gene in *A. niger* did not cause any difference in growth in liquid media, showing that this enzyme does not play an essential role in sucrose and fructan metabolism. However, on solid media the SucB disrupted strains displayed earlier onset of sporulation compared to the wild-type *A. niger*. SucB might thus be (in)directly responsible for the onset of sporulation, either by relieving stress on the organism, by regulation of osmolarity, or intracellular storage of energy, or by removing free sucrose via transfructosylation. The presence of SucB homologues in other fungal species indicates that these proteins may all play an important, but not essential role in the intracellular metabolism of sucrose (Yuan *et al.*, 2006; Goosen *et al.*, 2007; **Chapters 2 and 3).**

A. *niger* thus employs multiple enzymes in order to facilitate sucrose and fructan hydrolysis. These enzymes are specifically induced by their substrates, or substrate derivates, as in the case of the induction of *InuA* by the inulin derivate sucrose. Finally, the precise metabolic roles of intracellular sucrose hydrolysing enzymes are still unclear. Elucidation of their specific function in overall metabolism in *A. niger* provides an interesting and challenging topic for future research.
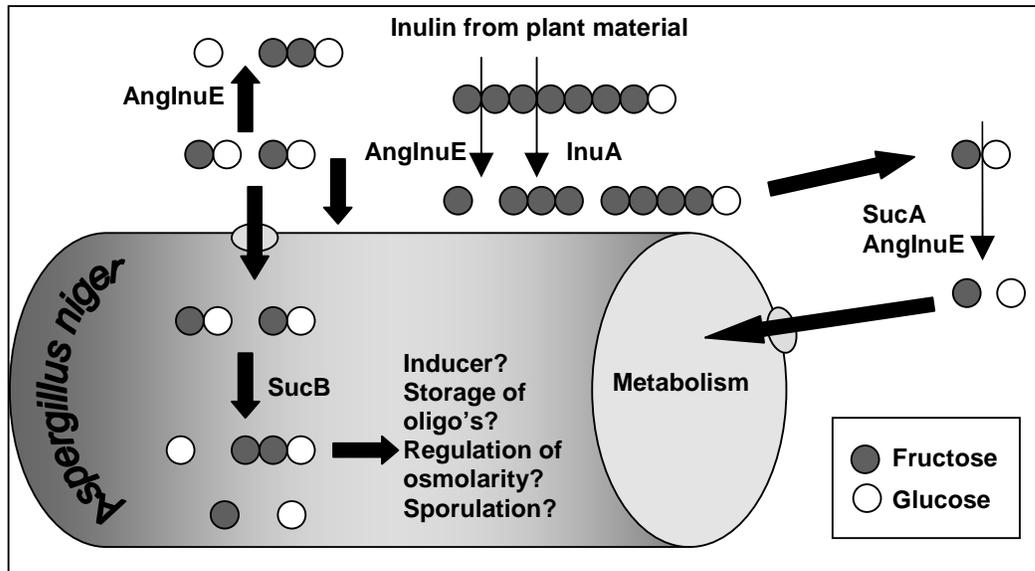
Figure 1. Schematic representation of the sucrose and fructan modifying enzymes network of *Aspergillus niger* (see text for details), depicting both extracellular (AngInuE, InuA, SucA) and intracellular (SucB) sucrose and fructan modifying enzymes.