Chapter 1

Introduction
The kingdom of fungi

Fungi form a large and tremendously variable group of eukaryotic microorganisms. Historically, they have long been classified in the kingdom of plants, until it was realized that the life style of plants and fungi is completely different. Plants are phototrophs, utilizing energy from light, while fungi are chemotrophs, obtaining energy from the degradation of organic materials. Likewise, autotrophic plants synthesize their organic components from atmospheric carbon dioxide, while fungi are heterotrophs and need organic material as their carbon source. Consequently, fungi play an important role in the recycling of carbon and nitrogen, and they often have a saprophytic or symbiotic life style. Fungi can be opportunistic or obligate pathogens, causing many diseases in plants and animals. In agriculture, fungi are one of the main causes of crop destruction, either by decreasing yields or by the production of toxins in crops rendering them poisonous (Carlile and Watkinson, 1994). On the other hand, many mutualistic relationships exist between fungi and other organisms. Fungi forming arbuscular mycorrhiza are obligate symbionts of plants and belong to the order Glomales. They obtain their carbon and energy sources directly from the host plant, via the colonisation of plant roots. Almost all plant species are colonized by mycorrhizal fungi and benefit from the increased potential of taking up minerals and water, which has a large impact on crop yield (Smith and Read, 1997). Another example of an intimate symbiotic relationship of fungi and plants exists in the lichens, with lichen-forming fungi acquiring their carbon source from extracellularly located green algae or cyanobacteria (Honegger, 2001). Some insects, such as tropical ant species, make use of the ability of fungi to degrade plant material by growing them in “gardens” in their nests (Silva et al., 2003). In human culture, several yeast and fungal species have a prominent role in the production of food. 

Agaricus bisporus (common mushroom) and other edible mushroom species are grown on large scale for human consumption. The yeast 

Saccharomyces cerevisiae (Baker’s yeast) is used in bread baking and brewing, and 

Aspergillus oryzae has been used for the production of Japanese soy sauce and miso (fermented soy beans) since a long time (Hesseltine, 1983). Examples of more recent inventions of applications for fungi are the production of antibiotics by e.g. Penicillium species, and the use of 

Fusarium venenatum as a producer of food with a high protein content (Quorn) used as an alternative for meat.

Most fungal species described to date can be classified into two phyla: the Basidiomycetes and the Ascomycetes. Basidiomycetes comprise for example the mushroom fungi and the rusts, which are primarily obligate plant pathogens. 

Ustilago maydis, the causal agent of corn smut disease in maize, and Cryptococcus neoformans,
an opportunistic human pathogen, are examples of Basidiomycetes of which the genomic sequences are available. The Ascomycetes can be divided over three subphyla: Taphrinomycotina, Saccharomycotina and Pezizomycotina. The Taphrinomycotina form a diverse group of yeast-like and filamentous fungi, of which *Schizosaccharomyces pombe* (also called fission yeast) is an often used model species. The subphylum Saccharomycotina consists of the “true yeasts”, including *S. cerevisiae* and the human pathogen *Candida albicans*. The largest subphylum is formed by the Pezizomycotina comprising mainly filamentous fungi producing fruiting bodies. Among the nine classes of the Pezizomycotina are the Sordariomycetes, including *Neurospora crassa* and *Magnaporthe grisea*, and the Eurotiomycetes which include *Histoplasma capsulatum* and the aspergilli.

**The aspergilli**

The filamentous fungi of the genus *Aspergillus* belong to the Ascomycetes. Up to now, 274 different species have been described (NCBI website, June 2007). The genus was first catalogued in 1729 by the Italian biologist and priest Pietro Antonio Micheli (1679-1753), who named it after the aspergillum, a brush used for the sprinkling of holy water (Micheli, 1729). The asexual spores or conidiospores of aspergilli form long rows diverting from the conidiophore which, observed under a microscope, resemble the hairs of a brush (Fig. 1). Aspergilli occur worldwide, mostly as saprophytic, soil borne fungi. Some species, such as *Aspergillus fumigatus* and *Aspergillus flavus* are opportunistic pathogens for humans and other higher animals. Most human systemic infections are found among immunocompromised patients. Other species, such as *Aspergillus niger*, are well known for their usage in food production and modern biotechnology.

**Aspergillus niger**

*A. niger* is a common soil fungus, sometimes observed as a black mould spoiling fruit and other foods. Although this species is generally non-pathogenic, inhalation of large amounts of spores can lead to the lung disease aspergillosis (Schuster *et al.*, 2002). Oral intake of *A. niger* has been assessed as harmless by the World Health Organization, which opened the opportunity to use this versatile producer for industrial production of acids, pharmaceuticals and enzymes. Various strains of *A. niger* are applied in the large scale industrial preparation of citric acid (also known as food additive E330) and gluconic acid (E574), which serve as ingredients for the production of various foods and
drinks (Lasure and Magnuson, 2004; Mourya and Jauhri, 2000). *A. niger* is also a very efficient producer of enzymes.

**Fig. 1** *Aspergillus* species seen under a microscope. (Courtesy CBS Fungal Biodiversity Centre).

The most widely used *A. niger* enzyme is glucoamylase, which is produced by selected industrial strains overproducing this enzyme, e.g. strain *A. niger* DS03043 (van Dijck *et al.*, 2003). Glucoamylase is nowadays produced and sold by several companies under brand names like Stargen and Plusweet (from Genencor) and Fungamyl (Novozymes). These commercial preparations often combine glucoamylase with α-amylase for the release of glucose from starch granules, which can subsequently be converted into fructose by glucose isomerase yielding high fructose syrups (van der Maarel *et al.*, 2002). Also, cocktails of glucoamylase in combination with α-amylases or glucose oxidases are applied to improve dough structure in bread baking and to improve the process of brewing. Aspergilli, including *A. niger*, are also known for the production of various types of pectinases, which are used for the clarification of wine and fruit juices and production of fruit purees (Schauer and Boriss, 2004). Cellulases produced by aspergilli are in use in washing detergents to reduce pilling, and in the brewing industry. An example of a commercially produced cellulase from *A. niger* is Finizym (Novo Nordisk). *A. niger* is not only deployed in the production of its own proteins, it is also a host for the heterologous expression of proteins from various eukaryotes, including those originating from mammals, with mainly applications in the pharmaceutical industry (Yoder and Lehmbeck, 2004).

*A. niger* is one of the main host organisms for the expression and production of enzymes by the Dutch biotechnology company DSM. To improve this process and
search for novel enzyme activities, the genome was sequenced by Biomax under the authority of DSM in 2002 (Pel et al., 2007). Since that time, the genome sequences of several other Aspergillus species have also become available. This has allowed us to compare the different species with regard to gene organization as well as sequence features of specific genes of interest. In this thesis, the focus will be on new A. niger enzymes with similarity to enzymes involved in the degradation of one of its main carbon sources: starch.

**Starch and glycogen**

Starch is one of the most abundantly available carbon sources for a saprophytic fungus like *A. niger*. It is the main storage polysaccharide in plants, where it can be stored in the roots, tubers, seeds and fruits (Peters, 2006). Starch also constitutes a major component of the human diet. Rice, wheat, maize and potato are examples of plants with high starch content often used in human consumption. Products derived from starch are used for various applications in the food industry, for example as a thickener in sauces and soups or to change the texture of dairy products, but also in the production of paper and biodegradable plastics.

Starch is composed of two types of molecules: amylose and amylopectin. Amylose is an unbranched chain of α-(1,4) linked glucose residues, which generally has a degree of polymerization (DP) between 250 and 5000. Amylopectin is also mainly composed of α-(1,4)-linked glucose residues, but approximately 3-5% of these glucose moieties is additionally linked with an α-(1,6) bond, creating a branched molecule of DP 10,000-100,000 (Fig. 2). In general, starch is composed of 15-30% amylose and 70-85% amylopectin, but the exact composition depends on the source (Robyt, 1998). Starch in plants is produced from ADP-glucose formed as a result of photosynthesis (Preiss, 2004). The enzyme starch synthase couples the ADP-glucose to an existing starch molecule via an α-(1,4) glycosidic bond, thereby releasing ADP. Alternatively, a new starch molecule may be formed via a glycosylated protein acting as a primer for the elongation reaction by a starch synthase. Subsequently, α-(1,6) branches are formed in the growing glucan chains by starch branching enzymes. Starch synthesis takes place in the chloroplast or, in non-photosynthetic storage tissues, in the amyloplast.
Glycogen is the equivalent polysaccharide storage compound in bacteria, fungi and animals. Glycogen is a highly branched molecule, with between 7 and 12\% α-(1,6) linked glucose molecules. It exists in dissolved form in cells and is therefore faster accessible for degradation than starch in plants, which is normally stored in a granular form. In analogy of the synthesis route of starch, glycogen is produced from UDP-glucose via three enzymes: (1) glycogenin, a self-glycosylating enzyme acting as a starter molecule, (2) glycogen synthase which elongates the chain with α-(1,4) linked glucose residues, and (3) a branching enzyme which incorporates the α-(1,6) glycosidic bonds. The degradation of glycogen is performed by another set of specialized enzymes. Glycogen phosphorylase releases glucose-1-phosphate but cannot act near branching points. Here, a debranching enzyme first transfers a maltosyl unit from the branching point to the end of a linear α-(1,4) chain and subsequently releases glucose by hydrolysing the remaining α-(1,6) linkage (Francois and Parrou, 2001).

**Enzymes acting on starch**

Degradation of starch by heterotrophic organisms is performed by the concerted action of different enzymes, including α-amylases, α-glucosidases and glucoamylases. Alpha-amylases hydrolyse the internal α-(1,4)-glycosidic bonds in starch, glycogen and maltooligosaccharides (α-(1,4) linked glucose oligomers) producing shorter maltooligosaccharides and maltose (4-O-α-D glucopyranosyl-D-glucose) (EC 3.2.1.1). A special type of amylases, the β-amylases, specifically release maltose from the reducing end of a substrate and are only found in bacteria and in plants. Glucoamylases
release β-glucose from the non-reducing end of maltooligosaccharides (EC 3.2.1.3), while α-glucosidase releases α-glucose from the non-reducing end (EC 3.2.1.20).

The enzymes performing these hydrolysis reactions on starch and related substrates are classified in different families according to the classification of glycoside hydrolase (GH) enzymes, which is based on sequence similarity of the proteins (see http://www.cazy.org) (Henrissat, 1991; Henrissat and Bairoch, 1993). Glucoamylases are grouped in family GH15, and most α-glucosidases group in family GH31, together with a few enzymes with different substrate specificities such as α-xylosidase and α-(1,3) glucosidase. The β-amylases belong to family GH14, while the majority of the α-amylases belong to family GH13. Some bacterial and archaeal enzymes with α-amylase activity have no apparent sequence similarity to GH13 enzymes and are placed in their own family GH57 (Henrissat and Bairoch, 1996).

**Family GH13 enzymes**

The GH13 family is the largest sequence based glycoside hydrolase family. It contains more than 30 different enzymatic specificities, and proteins originating from all kingdoms of life. Most enzymes in this family either hydrolyse or form α-(1,4) or α-(1,6) glycosidic bonds. Additionally, some enzymes act on sucrose (β-D-fructofuranosyl α-D-glucopyranoside, with a β-(1,2) glycosidic bond) or trehalose (α-D-glucopyranosyl-α-D-glucopyranoside, with an α-(1,1) glycosidic bond). To improve the predictability of the enzymatic activity of family GH13 members, the family was recently divided into 35 subfamilies, which were assigned based on the phylogenetic analysis of a large number of protein sequences (Stam et al., 2006). Most of these subfamilies contain only one enzymatic activity, while several enzyme specificities, notably α-amylases, are divided over a number of subfamilies. For example, extracellular fungal α-amylases are grouped in family GH13_1, while α-amylases from animals belong to subfamily GH13_24, and cyclodextrin glucanotransferases cluster in subfamily GH13_2 (Table 1). In general, family GH13 enzymes can be recognised by the presence of four highly conserved amino acid regions, first identified by Nakajima et al. (Nakajima et al., 1986) which contain most of the amino acids present in the active site (MacGregor et al., 2001). Within these regions, 7 amino acids important for catalytic activity and binding and positioning of the substrate are generally conserved (Table 1). These include the two catalytic residues Asp227 and Glu251 (numbering according to acid amylase from *A. niger*, unless indicated otherwise). The third completely conserved amino acid is Asp318, which plays an important role in substrate binding and transition state stabilization. Mutation of any of these three residues in family GH13 enzymes resulted in a dramatic decrease of the enzymatic activity or rendered the enzymes completely
Table 1 The four conserved regions in several family GH13, GH70 and GH77 enzymes. The numbering is based on A. niger acid amylase. The three residues essential for catalytic activity in family GH13, and equivalent residues in families GH70 and GH77, are indicated bold underlined. Other conserved residues are indicated in bold.

<table>
<thead>
<tr>
<th>Source</th>
<th>Protein</th>
<th>GH family</th>
<th>Region I</th>
<th>Region II</th>
<th>Region III</th>
<th>Region IV</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Aspergillus niger</em></td>
<td>Acid amylase</td>
<td>13_1</td>
<td>YLMVDVPNH</td>
<td>DGLRIDSVLE</td>
<td>YCVQEVDN</td>
<td>NFIENHD</td>
</tr>
<tr>
<td><em>Aspergillus niger</em></td>
<td>α-glucano-transferase A</td>
<td>13_1</td>
<td>YLMMDTVINN</td>
<td>DGLRIDAAKH</td>
<td>FMTQEVLQ</td>
<td>SFSENHD</td>
</tr>
<tr>
<td><em>Aspergillus niger</em></td>
<td>α-glucan synthase A</td>
<td>13_22</td>
<td>YVIMDTLATS</td>
<td>DGDFDKAVQ</td>
<td>FLGHESTS</td>
<td>YGVSNQD</td>
</tr>
<tr>
<td><em>Aspergillus niger</em></td>
<td>AmyD</td>
<td>13_5</td>
<td>GIYMDAVLNH</td>
<td>SGGLIDAVKH</td>
<td>FVGEFNYK</td>
<td>TFVANHD</td>
</tr>
<tr>
<td><em>Histoplasma capsulatum</em></td>
<td>Amy1p</td>
<td>13_5</td>
<td>RIIIMTVLNH</td>
<td>SGGLIDAAKH</td>
<td>LIVAEFKN</td>
<td>TFVAMHD</td>
</tr>
<tr>
<td><em>Bacillus licheniformis</em></td>
<td>α-amylase</td>
<td>13_5</td>
<td>NVYMDDVINH</td>
<td>DGFLRDAVKH</td>
<td>FTVAEKWS</td>
<td>TFVANHD</td>
</tr>
<tr>
<td><em>Bacillus circulans</em> 251</td>
<td>CGTase</td>
<td>13_2</td>
<td>KVIIDFAPNH</td>
<td>DGIRKDAVKH</td>
<td>FTFGEMLQ</td>
<td>TFDINHD</td>
</tr>
<tr>
<td><em>Aspergillus oryzae</em></td>
<td>Branching enzyme</td>
<td>13_8</td>
<td>VLLLDVHSH</td>
<td>DGFRDGVTS</td>
<td>ITVAVEVS</td>
<td>AYAESHD</td>
</tr>
<tr>
<td><em>Anaerobranca horikoshii</em></td>
<td>Pullulanase</td>
<td>13_14</td>
<td>KVIKDDVYMNH</td>
<td>DGFRDDMAL</td>
<td>ILYGEFWQ</td>
<td>IYVCHD</td>
</tr>
<tr>
<td><em>Propioni-bacterium freudenreichii</em></td>
<td>Trehalose synthase</td>
<td>13_16</td>
<td>RIIIDFVMNH</td>
<td>DGGRDAVPPY</td>
<td>ILLCSEFQ</td>
<td>TFLRNHD</td>
</tr>
<tr>
<td><em>Lactobacillus reuteri</em> 121</td>
<td>Glucansucrase (GTFA)</td>
<td>70</td>
<td>QVMADRVPDQ</td>
<td>DSRVADAPDN</td>
<td>INILEQNN</td>
<td>SFVRASHD</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>Amylolamulse</td>
<td>77</td>
<td>GLYRDLEAVGV</td>
<td>GALRIDVMS</td>
<td>MVIGEDLG</td>
<td>AIIAKHD</td>
</tr>
</tbody>
</table>
inactive (Klein et al., 1992; Sogaard et al., 1993). The other conserved residues, as displayed in Table 1, often play a more indirect role in catalysis, for example in binding and positioning of the substrate (Leemhuis et al., 2002; Van der Veen et al., 2000).

The reaction mechanism of family GH13 enzymes is described as a double replacement mechanism (McCarter and Withers, 1994; Uitdehaag et al., 1999). The substrate is bound in the active site with its non-reducing end in the minus subsites. In enzymes with glucanotransferase activity, this implies that the part of the donor substrate that is transferred is bound in the minus subsites, while the acceptor substrate is bound in the plus subsites (Davies et al., 1997). The actual reaction proceeds in two steps: first, the proton donor (Glu251) protonates the oxygen atom in the scissile bond, thereby cleaving the substrate. The remaining part of the substrate forms a covalent intermediate via a β-glycosidic linkage with the nucleophile (Asp227), and the leaving group departs from the plus subsite. The existence of a covalent intermediate has been demonstrated for different enzymes, including an α-amylase (Tao et al., 1989) and a CGTase (Mosi et al., 1997; Uitdehaag et al., 1999). In the second step, an acceptor molecule is deprotonated by Glu251, which is now acting as a general base, and attacks the β-glycosidic linkage, thereby forming a new α-glycosidic linkage. The acceptor molecule may be water, in the case of hydrolysis, or another saccharide molecule in case of a transglycosylation reaction (Fig. 3).

The 3D structure of the enzymes in family GH13 is characterized by the presence of several domains (Jespersen et al., 1991). The main part of the proteins, domain A, is a (β/α)8 barrel or TIM-barrel, a symmetrical fold composed of 8 β-strands surrounded by 8 α-helices. The active site is located in the loops at the C-termini of several of the β-strands. In addition to domain A, all family GH13 enzymes contain a loop protruding between the third β-strand and the third α-helix of the (β/α)8 barrel called domain B (Janecek et al., 1997). Domain B is thought to be involved in the formation of the substrate binding cleft, and its composition and 3D structure can be highly variable between the different GH13 subfamilies. Studies on highly homologous barley α-amylases with different B-domains have shown the importance of this domain in substrate specificity and stability at low pH (Rodenburg et al., 1994; Juge et al., 1995).

A variety of other domains may be present in family GH13 enzymes, depending on their enzymatic reaction and substrate specificities. For example, α-amylases (EC 3.2.1.1) often contain an additional C-terminal domain C which has an antiparallel β-sandwich fold (Jespersen et al., 1991). Cyclodextrin glucanotransferases contain two extra domains following the C-domain: domain D has an unknown function, while domain E is a starch binding domain (SBD) (Penninga et al., 1996). Such a SBD, which binds strongly to granular starch and cyclodextrins, belongs in Carbohydrate Binding Module family 20 (CBM20) (Knegtel et al., 1995). Some α-amylases and most glucoamylases
from family GH15 also contain a domain from family CBM20 (see e.g. Iefuji et al. (1996) and Svensson et al. (1986)).

Family GH13 enzymes share their (β/α)_8 barrel structure and their reaction mechanism with the families GH70 and GH77. Together, these three families form the α-amylase superfamily, or clan GH-H. The members of family GH70 are glucansucrases, large bacterial enzymes which use sucrose to form long α-glucan polymers with different
types of glycosidic linkages (EC 2.4.1.5 and 2.4.1.140) (Monchois et al., 1999). The
\((\beta/\alpha)\) barrel of GH70 enzymes is thought to be circularly permuted compared to GH13
enzymes, and several additional domains with unknown function are present
(MacGregor et al., 1996; van Hijum et al., 2006). Family GH77 contains amylomaltase
enzymes from bacteria and archaea, including many thermostable examples with
industrial applications (Kaper et al., 2004). These intracellular enzymes have
disproportionating activity and transfer \(\alpha\)-(1,4) linked glucans to an acceptor substrate,
which may be starch, maltooligosaccharides or glucose (EC 2.1.4.25). The enzymes in
the families GH13, GH70 and GH77 share their catalytic machinery, implicating that
the amino acids involved in catalysis are largely conserved. However, there are also
several differences within the four conserved regions of these families. For example,
His143 is almost completely conserved in family GH13, and several studies have shown
that mutation of this residue can have a strong effect on catalytic activity or reaction
specificity of the enzyme (Chang et al., 2003; Nakamura et al., 1993). The equivalent of
this His is, however, replaced by a conserved Glu in glucansucrase enzymes, while the
equivalent residue is not conserved at all in amylomaltases of family GH77 (Table 1).

Function and regulation of family GH13 enzymes in fungi
Filamentous fungi often live on decaying plant material. Consequently, they need to
produce and secrete a full range of enzymes to degrade plant polysaccharides such as
pectin, cellulose and starch. Indeed many fungi possess powerful enzyme mixtures to
hydrolyse a variety of plant polysaccharides and form small oligosaccharides, which can
be taken up via specialised transporters to serve as carbon and energy sources. To
degrade starch, fungi produce extracellular \(\alpha\)-amylases, \(\alpha\)-glucosidases and
glucoamylases. The extracellular fungal \(\alpha\)-amylases identified thus far show
considerable mutual similarity and have been classified in subfamily GH13_1. In
A. niger, the enzymes from this subfamily include acid amylase (Boel et al., 1990) and
the almost identical proteins AmyA/AmyB identified \(A. \ niger \ var \ \ awamori\) (Korman et
al., 1990). These two identical enzymes are, however, not present in the genome
sequence of \(A. \ niger\) strain ATCC 1015, which harbours only one orthologous protein
with 72% identity to AmyA/B. In addition, a protein identical to \(A. \ oryzae\) TAKA-
amylase has been identified in an industrial \(A. \ niger\) strain (Vujicic-Zagar and Dijkstra,
2006). Other previously identified starch-acting enzymes in \(A. \ niger\) include
glucoamylase GlaA (Boel et al., 1984) and \(\alpha\)-glucosidase AglA (Nakamura et al.,
1997), which has been renamed AgdA to prevent confusion with \(\alpha\)-galactosidases (den
Alpha-glucan acting enzymes in Aspergillus niger

Herder et al., 1992). Thus far, no enzymes from family GH70 or GH77 have been identified in Eukaryotes including fungi.

Expression of the starch degrading enzymes in Aspergillus species is regulated by AmyR, a positive transcriptional regulator with a GAL4-type Zn(II)$_2$Cys6 cluster. The six conserved cysteine residues in this type of regulator need two Zn atoms (or occasionally Cd atoms) to form a so called binuclear cluster, which allows binding to a specific stretch of DNA (Pan and Coleman, 1990). The specific DNA sequence for binding of AmyR is identified as 5’-CGGN$_6$(C/A)GG-3’ (Petersen et al., 1999;Ito et al., 2004). Expression of starch degrading enzymes is induced via AmyR when breakdown products of starch are present, notably maltose or isomaltose. In A. nidulans, isomaltose was found to be a stronger inducer of the AmyR system than maltose. It was suggested that maltose is converted into isomaltose by an α-glucosidase with α-(1,6) transglycosylation activity (Kato et al., 2002a; Kato et al., 2002b). The expression of starch degrading enzymes is repressed in the presence of high concentrations of an easily accessible, monosaccharide carbon source such as glucose. This process of carbon catabolite repression is mediated by the transcriptional repressor CreA (Dowzer and Kelly, 1991). Expression of the amyR gene is regulated both by CreA and by AmyR itself (Tani et al., 2001).

Apart from starch catabolism, family GH13 enzymes in fungi are also involved in the formation and breakdown of the storage compound glycogen. The presence of intracellular glycogen has been demonstrated in several fungal species (Mattey and Allan, 1990; Bahia et al., 1997; Francois and Parrou, 2001), and the intracellular enzymes needed for its production and degradation are encoded in the available genome sequences. Two of these enzymes, glycogen branching and debranching enzyme, are members of family GH13. In S. cerevisiae, glycogen storage is regulated via the phosphorylation state of the main synthesising and degrading enzymes, glycogen synthase and glycogen phosphorylase, and is dependent on the availability of the carbon source (Francois and Parrou, 2001).

The fungal cell wall

Recently, several studies have shown that different family GH13 enzymes of fungal origin may be involved in the formation of α-glucans in the fungal cell wall, rather than in starch degradation. The fungal cell wall is usually made up of chitin, β-glucan, α-glucan and covalently attached cell wall proteins. All these components form together an extensively crosslinked complex (Bowman and Free, 2006).

The most abundant cell wall polysaccharide in yeasts and fungi is β-glucan, which mainly exists as β-(1,3) glucan with occasional β-(1,6) linked branches (Lesage and
Bussey, 2006). Additionally, a β-(1,6) linked glucan which is highly branched with β-(1,3) linkages is present in S. cerevisiae and S. pombe (Klis et al., 2001), but was not identified in several other ascomycete fungi. Among the aspergilli, the cell wall glucans of A. fumigatus are the best studied, due to the pathogenic nature of this species. Apart from β-(1,3) glucan, A. fumigatus contains a linear glucan with β-(1,3) and β-(1,4) linkages (Fontaine et al., 2000). β-Glucans are formed by plasma-membrane localized β-glucan synthases, which use UDP-glucose to form the glucan intracellularly, and subsequently transport it out of the cell (Cabib and Kang, 1987). Additionally, several enzymes localized in the cell wall area are known to either hydrolyze or transglycosylate β-glucans. In S. cerevisiae, these enzymes are mainly encoded by the GAS gene family (Ragni et al., 2007), and homologous proteins have been described in S. pombe (De Groot et al., 2007), C. albicans (Hartland et al., 1991) and A. fumigatus (Mouyna et al., 2000). As β-glucan forms a vital part of the cell wall for many pathogenic species, a group of anti-fungal drugs known as echinocandins is on the market, which act through the inhibition of β-glucan synthases (Beauvais and Latgé, 2001).

Chitin, a linear chain of β-(1,4)-N-acetyl-glucosamine, is present in most yeast species as well as filamentous fungi. In S. cerevisiae it is mainly located in the scars formed after budding and constitutes only 1-2% of the cell wall dry weight (Lesage and Bussey, 2006). In filamentous fungi, chitin forms a larger component of the cell wall (10-20%) because it is responsible for maintaining the cell wall structure (Bowman and Free, 2006). It is produced by chitin synthases, integral membrane proteins that require UDP-N-acetyl-glucosamine as donor molecule.

Cell wall mannoproteins (proteins with covalently linked mannose residues) represent a considerable part of up to 40% (w/w) of the fungal cell wall (Brul et al., 1997). Two types of cell wall localized proteins are described in yeasts and fungi: the proteins with internal repeats (PIR proteins) and those which are linked to the plasma membrane or cell wall via a glycosylphosphatidylinositol (GPI) anchor. This last category of proteins is found among all eukaryotes. The GPI-anchored proteins contain an N-terminal signal sequence for secretion, which targets them to the endoplasmic reticulum (ER) where they are glycosylated. Additionally, a specific C-terminal sequence motif, described by De Groot et al. (2003), is recognized and cleaved, and a ready-made GPI-anchor is attached to the new C-terminus (the ω-site). The complex GPI structure consists of mannose and glucosamine residues linked to an inositol - PO₄ lipid component referred to as phosphatidylinositol (PI), and is linked to the C-terminus of the protein via a phosphoethanolamine group (Fig 4). Upon secretion, the protein is bound to the plasma membrane via the fatty acid chains. Some GPI-anchored proteins are subsequently released from their GPI-anchors by the action of phospholipase C (PLC) which cleaves between the inositol - PO₄ and the lipid anchoring. The proteins
Alpha-glucan acting enzymes in *Aspergillus niger*

are then covalently linked to the cell wall β-(1,6) glucan via their modified GPI-anchor. Whether GPI-anchored proteins are specifically targeted to the plasma membrane or to the cell wall is a matter of debate. In *S. cerevisiae* GPI-anchored proteins, the presence of two basic amino acids preceding the ω-site correlated with their location on the plasma membrane, rather then in the cell wall (Caro *et al.*, 1997) and later mutational studies have confirmed this relation (Frieman and Cormack, 2003). Additionally, the presence of stretches of serine and threonine residues in GPI-anchored proteins can overrule the basic amino acids and direct a protein to the cell wall (Frieman and Cormack, 2004). Searches for putatively GPI-anchored proteins in the genomes of various fungi yielded 33 candidates in *S. pombe*, 66 in *S. cerevisiae*, 97 in *N. crassa* and 104 in *C. albicans*. Among these predicted proteins are cell wall-associated proteins with unknown functions, enzymes acting on β-glucans, and a variety of other predicted glycoside hydrolases (De Groot *et al.*, 2003). Mutation of genes encoding components of the GPI-anchor biosynthetic pathway in *N. crassa* resulted in abnormal phenotype and cell lysis, showing that GPI-anchoring of proteins is critical for proper cell wall function (Bowman *et al.*, 2006). GPI-anchors may also serve to attach certain polysaccharide cell wall components to the plasma membrane, as was demonstrated for galactomannan in *A. fumigatus* (Costachel *et al.*, 2005).

Fig. 4 Schematic representation of a GPI-anchor structure from *A. fumigatus* proteins linked to the cell membrane (Fontaine *et al.*, 2003). The amount of mannose residues and the place where the phosphate group is linked to a mannose residue may vary between different species. Man: mannose, glcNH₂: glucosamine, ino: inositol.

The PIR proteins have thus far only been described in yeast species, although searches in fungal genomes have indicated that this type of proteins is also
present in ascomycete fungi (De Groot et al., 2005). PIR proteins in \textit{S. cerevisiae} contain repeats of a sequence of 19 amino acids and are linked to the cell wall $\beta$-(1,3) glucan via a mild-alkali sensitive linkage (Toh-e A \textit{et al.}, 1993; Kapteyn \textit{et al.}, 1999).

Cell wall $\alpha$-glucan is present in most Ascomycetes and Basidiomycetes, including \textit{S. pombe}; but is absent from yeast species belonging to the Saccharomycetes. Generally, $\alpha$-glucan in fungal cell walls is of the $\alpha$-(1,3) type with a small percentage of $\alpha$-(1,4) glycosidic bonds, also called pseudonigeran (Grün \textit{et al.}, 2005; Horisberger \textit{et al.}, 1972). Additionally, an $\alpha$-glucan with alternating $\alpha$-(1,3)/(1,4) glycosidic bonds (nigeran) has been identified in \textit{A. niger} and several other Ascomycetes (Barker and Carrington, 1953; Woranovicz-Barreira \textit{et al.}, 1999; Johnston, 1965). A few cases of glucan types with mainly $\alpha$-(1,4) bonds have been reported in \textit{S. pombe} as well as \textit{A. niger} (Garcia \textit{et al.}, 2006; Kirimura \textit{et al.}, 1999), but none of these were investigated in detail. It is generally believed that the $\alpha$-glucan compound is produced by $\alpha$-glucan synthases, although this synthesizing process was never directly demonstrated. Alpha-glucan synthases are predicted transmembrane enzymes with two catalytic domains, as shown in Fig. 5. The C-terminal, intracellular part has similarity to glycogen- and starch-synthases, members of family glycosyltransferase (GT) 5 which use UDP- or ADP glucose to form an $\alpha$-glucan chain. The N-terminal, extracellular domain of $\alpha$-glucan synthases has resemblance to family GH13 enzymes and is thought to be involved in the coupling of extruded glucan chains (Hochstenbach \textit{et al.}, 1998; Grün \textit{et al.}, 2005).

In addition to $\alpha$-glucan synthases, two other types of enzymes, both family GH13 homologs, were recently shown to play a role in fungal $\alpha$-glucan formation. The first was Aah3p, a family GH13_1 protein identified in the fission yeast \textit{S. pombe} (Morita \textit{et al.}, 2006). The knockout of the corresponding gene caused an aberrant cell shape and hypersensitivity towards cell wall degrading enzymes, indicating that cell wall biogenesis was somehow affected. Subsequent expression of mutated Aah3p in which the conserved aspartate and glutamate residues in the active site were replaced by alanine failed to rescue the phenotype of the knockout. This was an indication for the importance of the enzymatic activity of Aah3p, rather than its structural properties. Expression of a tagged version of Aah3p showed that the protein was localized in the cell membrane fraction, and that it was attached via a GPI-anchor, as was already predicted by De Groot \textit{et al.} (2003).

A second $\alpha$-amylase homologue with a proposed role in cell wall formation is Amy1p from \textit{H. capsulatum}. This predicted protein was shown to have high similarity to an $\alpha$-amylase from \textit{Bacillus licheniformis}, and had no recognisable signal sequence. A functional knockout of Amy1p lost completely the ability to form cell wall $\alpha$-(1,3) glucan, which is a very relevant finding while $\alpha$-(1,3) glucan is critical for virulence in this pathogenic fungus (Rappleye \textit{et al.}, 2004).
Alpha-glucan acting enzymes in *Aspergillus niger*

Fig. 5 Schematic representation of a fungal α-glucan synthase protein, based on the model of *S. pombe* Ags1p as proposed by Hochstenbach et al. (1998).

**Scope of this thesis**

The starting point for the studies described in this thesis was the genome sequence of *A. niger* (Pel et al., 2007). The sequence was used to identify novel enzymes resembling family GH13, 15 and 31 enzymes putatively acting on starch. A selection of the predicted proteins was then studied with regard to regulation of expression, phylogeny, physiological function and biochemical characteristics.

Chapter 1 describes the phylogenetic analysis of the family GH13, 15 and 31 enzymes in *A. niger* and related aspergilli. This is combined with the transcriptional analysis of these enzymes in *A. niger* grown on different carbon sources. The comparison of gene expression in a wild type strain and in an AmyR deletion strain allowed for the identification of genes encoding enzymes involved in starch degradation. Expression of some of the genes of interest appeared not to be regulated by AmyR or the presence of maltose, indicating that the encoded proteins may be involved in other physiological processes than starch degradation.
Chapter 1 Introduction

In chapter 2 two novel GH13 enzymes from *A. niger* are expressed, purified and characterised. The enzymes, named AgtA and AgtB, act as α-glucanotransferases on maltooligosaccharide substrates and starch. They are the first described family GH13 members which are predicted to be anchored to the cell wall or cell membrane with a GPI anchor. An *A. niger* knockout strain of *agtA* has an increased sensitivity for calcofluor white, a cell wall disturbing compound.

The two novel α-glucanotransferases are studied in more biochemical detail in chapter 3. By using a synthetically blocked substrate, we have determined the specific activity and *K*ₐₚ for both enzymes. Additionally, the efficiency of both enzymes to use various acceptor substrates is compared.

In chapter 4, the phylogenetic analysis of a novel group of intracellular fungal α-amylases is presented. These enzymes are similar to *H. capsulatum* Amy1p and belong to the subfamily GH13_5. One enzyme from this novel group, AmyD from *A. niger*, is expressed in *E. coli* and biochemically characterised.

Finally, a summary of the data described in this thesis combined with a discussion of the results is included. A summary in Dutch, meant for those with little knowledge of microbiology and biochemistry, is also provided.