INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

The Molecular Characterization of the Glucoamylase and Pyruvate Kinase Genes of *Aspergillus niger*

Rosa Eliza Zito

A Thesis

in

The Department

Of

Biology

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science at
Concordia University
Montreal, Quebec, Canada

March 2003

© Rosa Eliza Zito, 2003

National Library of Canada

Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services

Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-77944-0

ABSTRACT

The Molecular Characterization of the
Glucoamylase and Pyruvate Kinase Genes of *A. niger*

Rosa Eliza Zito

The glucoamylase gene of *Aspergillus niger*, *glaA*, is expressed at high levels in the presence of starch or maltose and repressed at the level of transcription by glucose. The promoter region of *glaA* was subjected to deletion and site-specific mutation analysis. Using autonomously replicating vectors with the *A. niger lacA* as the reporter gene, regulatory elements that control *glaA* expression in *A. niger* were mapped. Expression analysis showed that two elements that closely match the AnCF-binding consensus sequence were important for *glaA* expression. Site-directed mutagenesis of four elements that matched the CreA binding consensus showed that only two of these sites were important for regulation and that both sites were essential for *glaA* expression in maltose and glucose. Interestingly, one of the essential CreA sites overlapped the upstream portion of the single AmyR consensus element while the other site was immediately downstream from this site. The results of this study, therefore, suggested that both AnCF and AmyR were required for transcription of *glaA* and that regulation involved competition for overlapping binding sites between the transcriptional repressor CreA, and starch responsive transcription activator AmyR.

The highly expressed pyruvate kinase gene of *Aspergillus niger*, *pkiA*, encodes an enzyme in glycolysis. This study found that pyruvate kinase expression was approximately 8-fold higher under glycolytic growth conditions, as compared to gluconeogenic growth conditions. Deletion analysis of the *pkiA* promoter region found

three regions important for expression. One of the regions contains a 17 base pair sequence that is conserved in *A. niger* and *A. nidulans*. Another region contains a sequence closely matching the AnCF-binding consensus.

ACKNOWLEDGEMENTS

I wish to thank Dr. Storms and Dr. Tsang for providing me with the opportunity to work in their laboratories. The last few years gave me the occasion to gain plenty of knowledge, including a better understanding of myself. I also would like to thank Susan Sillaots, Dr. Hongshan Li, Aleks Spurmanis, Dr. Peter Ulycznyj, Nathalie Brodeur and Nadege Lory for all their help. A special thanks to my fellow lab members, Alain Bataille, Dr. Edith Munro and Yun Zheng for their technical assistance and moral support. I am grateful to Dr. Joyce and Dr. Varin for being my committee members and everyone at the Biology department that helped me during my studies at Concordia University. Most importantly, I must thank my family and Marius for their never-ending support. Home will always be a place where I am accepted as is and yet challenged to grow. I also dedicate this work to memory of Simone.

" To laugh often and much; to win the respect of intelligent people and the affection of children; to earn the appreciation of honest critics and endure the betrayal of false friends; to appreciate beauty, to find the best in others; to leave the world a little better; whether by a healthy child, a garden patch or a redeemed social condition; to know even one life has breathed easier because you have lived. This is the meaning of success."

-Ralph Waldo Emerson

TABLE OF CONTENTS

	Page
LIST OF FIGURES	x
LIST OF TABLES	xii
ABBREVIATIONS	xiii
CHAPTER 1. INTRODUCTION	1
1.1 *Aspergillus* and *Aspergillus niger*	1
1.1.1 Habitat and lifecycle	1
1.1.2 History of commercial use	4
1.1.3 Taxonomy	5
1.2 Protein production in *Aspergillus*	7
1.3 The glucoamylase gene (*glaA*) of *Aspergillus niger*	10
1.3.1 Amylases of *A. niger*	10
1.3.2 The two forms of *A. niger* glucoamylase	13
1.3.3 Structure of the *A. niger* glucoamylase gene	19
1.3.4 Regulation of the *A. niger glaA* gene	20
1.3.5 The promoter region of the *A. niger glaA* gene	22
1.3.6 Deletion analysis of the *glaA* promoter	22
1.3.7 Putative target sites for regulatory proteins in *glaA*	24

1.4 The pyruvate kinase gene of *A. niger* 27

1.5 The CCAAT – binding sequence 31

1.6 The CreA repressor 34

1.7 Regulation of *creA* 37

1.8 The transcriptional activator AmyR 39

1.9 Rationale for thesis 42

CHAPTER 2. MATERIALS AND METHODS 43

2.1 Strains and growth of *A. niger* and *E. coli* 43

2.1.1 *A. niger* and *E. coli* strains 43

2.1.2 Media and culture conditions 43

2.2 Transformation of *A. niger* and *E. coli* 44

2.3 Plasmid construction 46

2.3.1 Restriction enzymes and oligonucleotides 46

2.3.2 Isolation and manipulation of DNA 46

2.3.3 Polymerase chain reaction amplification of DNA 51

2.3.4 Starting plasmids 52

2.3.5 The *glaA* and *pkiA* deletion sets 57

2.3.6 Site–specific mutagenesis by overlap extension 60

2.3.7 CreA binding consensus mutations 63

2.3.8 AnCF-element mutants 64

2.3.9 Cloning of regulatory regions 64

2.3.10 Cloning of the *amyR* gene 65

2.4 Plasmid stability in *A. niger* 66

2.5 Enzyme assays 66

2.5.1 Secreted β-galactosidase assays 66

2.5.2 Secreted glucoamylase assays 67

CHAPTER 3. RESULTS 68

3.1 Comparison of the *glaA-lacA* chimera of ANEp2 with the native *glaA* gene of *A. niger* 68

3.2 Deletion analysis of the *glaA* promoter region 69

3.3 Analysis of putative target sites for regulatory proteins in *glaA* 72

3.3.1 Analysis of putative AnCF binding sites 72

3.3.2 Analysis of putative CreA binding sites 76

3.3.3 Electrophoretic mobility shift assays with the *glaA* promoter region 79

3.4 Expression analysis of transformants harbouring various copies and portions of the upstream regulatory region 81

3.5 The effect of AmyR on glucoamylase production 83

3.6 Deletion analysis of the *A. niger pkiA* promoter region 89

CHAPTER 4. DISCUSSION 93

4.1 Autonomously replicating shuttle vectors for use in *A. niger* 93

4.2 Deletion analysis of the *A. niger glaA* promoter region 94

4.3 Regulation of *glaA* by the AnCF-binding complex 95

4.4 Regulation of *glaA* by CreA and AmyR ____ 97

4.5 Regulation of the pyruvate kinase gene (*pkiA*) of *A. niger* ____ 102

4.6 Future studies ____ 105

WORKS CITED ____ 107

LIST OF FIGURES

Figure Number		Page
1	Life cycle of *A. niger*	3
2	Starch hydrolysis by glucoamylase	12
3	Starch hydrolysis by α-amylase and glucoamylase	15
4	Domain division in *Aspergillus niger* glucoamylase	17
5	Nucleotide sequence of the region upstream of the start codon of the *Aspergillus niger glaA* gene	26
6	Nucleotide sequence of the 5' region upstream of the start codon of the *Aspergillus niger* pyruvate kinase gene	30
7	Map of shuttle vector ANIp2	54
8	Map of shuttle vector ANEp2	56
9	Map of shuttle vector ANEp4	59
10	Site-directed mutagenesis by overlap extension	62
11	Deletion analysis of the *A. niger glaA* promoter region	71
12	Mutation analysis of putative AnCF binding sites	75
13	Mutation analysis of putative CreA binding sites	78

14	Map of the ANEpAmyR shuttle vector	86
15	Deletion analysis of the *A. niger pkiA* promoter region	92
16	The nucleotide sequence of the 5' upstream region from position –445 to –396 of the *A. niger glaA* gene	101

LIST OF TABLES

Table Number		Page
1	Homologous sequences in 5'-flanking regions of *Aspergillus* genes	23
2	Oligonucleotides used in this study	47
3	DNA sequencing results of plasmids containg wild-type and mutated *glaA* promoter regions	80
4	Expression analysis of transformants containing various copies of the URR	82
5	Glucoamylase activity of strain N593 and ANEpAmyR transformants of N593	88

ABBREVIATIONS

bp	base pair
CM	complete medium
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
EDTA	ethylene diamine tetraacetic acid
EMSA	electrophoretic mobility shift assay
g	gravity
GST	glutathione-S-transferase
kb	kilobase pair
L	litre
mg	milligram
min	minute
mL	millilitre
mM	millimolar
M	molar
MM	minimal medium
ORF	open reading frame
PCR	polymerase chain reaction

rpm	rotations per minute
μg	microgram
μL	microlitre
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

R: G/A S: C/G W: A/T Y: C/T M: A/C

1. Introduction

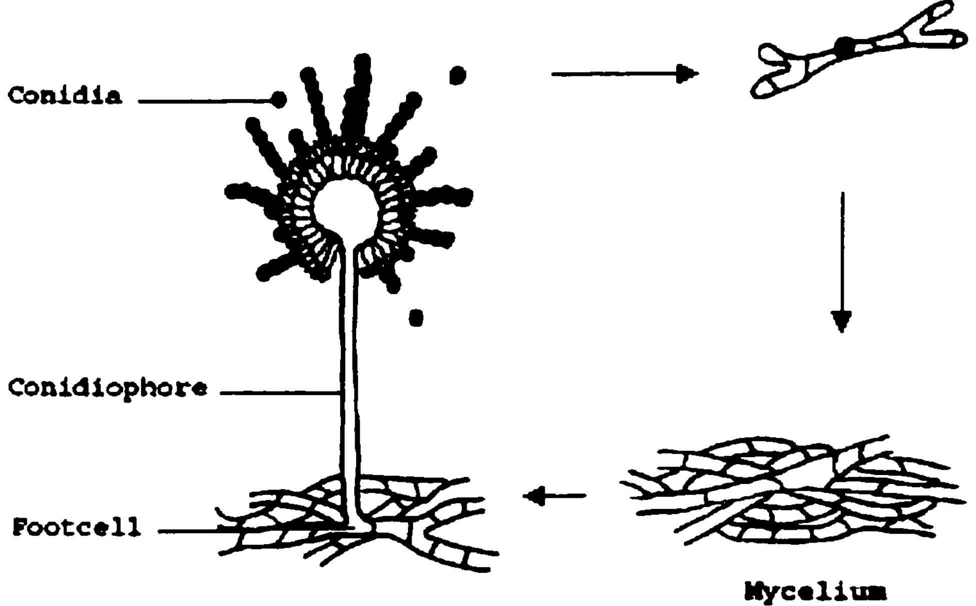
1.1 *Aspergillus* and *Aspergillus niger*

1.1.1 Habitat and lifecycle

The genus *Aspergillus* was first described and named by Micheli in 1792 for its characteristic conidiophores and sporeheads (Raper and Fennel, 1965). A member of the phylum *Ascomycetes*, the genus *Aspergillus* includes over 185 species. The most common of the species being the black *Aspergillus* referred to as *Aspergillus niger*. *Aspergillus* species are distributed worldwide and grow on a wide variety of substrates, such as soil, decaying vegetation and food. Members of the *Aspergillus* genus are also well known as biodeteriogens, causing discoloration and softening of the surface layers of wood. Approximately twenty species of the *Aspergillus* genus have been reported as causative agents of opportunistic infections in man (Kwon-Chung and Bennett, 1992). Among these, *Aspergillus fumigatus* is the most commonly isolated species, followed by *Aspergillus flavus* and *Aspergillus niger*.

The genus *Aspergillus* includes a set of fungi that are generally considered asexual, although perfect forms such as *Aspergillus nidulans* are found. As with other members of the *Aspergillus* genus, *A. niger* is a filamentous fungi but is distinctly characterized by its black conidial heads (Raper and Fennell, 1965). An obligate aerobe, *A. niger* grows readily at a temperature range of 17 °C – 42 °C and can grow over a pH range of 1.5 – 9.8. *A. niger* reproduces asexually by forming spores called conidia, which arise at the end of a reproductive mycelium called a conidiophore (Fig. 1). Conidiophores

Figure 1. Life cycle of *Aspergillus niger*. Conidiophores develop from specialized enlarged hyphal cells of the vegetative mycelium. Each spore can germinate and form a new individual. (Adapted from Raper and Fennel, 1965)


Conidia
Conidiophore
Footcell
Mycelium

develop from specialized enlarged hyphal cells of the vegetative mycelium called foot cells. Once the conidia are freed from the reproductive mycelium, each spore can germinate and form a new individual (Raper and Fennel, 1965).

1.1.2 History of commercial use

Aspergillus has a long history of being an organism of industrial importance. Although the genus was only named and described in the eighteenth century, its exploitation dates back to ancient times. For centuries, people of the orient have used *Aspergillus* species for food fermentation and in the production of soya sauce, sake and miso. Its long history of safe usage has given *Aspergillus* species such as *A. niger* the designation of GRAS (**G**enerally **R**egarded **A**s **S**afe) by the American Food and Drug Administration. Since the twentieth century discovery that *Aspergilli* could produce organic acids, *Aspergillus niger* has become widely used for the industrial production of gallic acid, citric acid and gluconic acid (Ingold, 1961) as well as enzymes such as glucoamylase, cellulase, pectinase, lipase and proteases (Bennett, 1985; Ward, 1989).

In addition to its products of fermentation, *A. niger* has other commercial uses. Due to its ease of visualization and resistance to several anti-fungal agents, *A. niger* is used to test the efficiency of preservative treatments (Jong and Ganntt, 1987). There is also interest in using *A. niger* to perform enzymatic reactions that are difficult to accomplish by strictly chemical means, such as specific additions to steroids (Jung and Gannt, 1987).

Recent events have furthered possible uses of *A. niger* in industry. DSM, an integrated international group of companies, has determined the complete DNA sequence of *A. niger* as part of the largest industrial genome project in Europe. The genome of *A. niger* is almost three times larger than that of the bakers' yeast . Ultimately, the sequence of *A. niger* enabled the identification of over 13, 000 genes encoded in a genome of approximately 34.5 million base pairs. The project was initiated two years prior to its completion when only 1-2 % of the *A. niger* genes were known. DSM is currently offering to make the results known to commercial partners in addition to academic organizations in the form of a low- barrier access program. This opens the door for genomic-based approaches to increase production of commercially important enzymes. It is estimated that only a fraction of the potential of enzymes produced by *A. niger* is currently characterized.

1.1.3 Taxonomy

The taxonomy of *Aspergillus* is primarily based on morphological features rather than physiological or biochemical features and genetic characteristics.

A. niger is both a species and a group within the genus *Aspergillus*. The taxonomy of *A. niger* is problematic. Raper and Fennell (1965) designated fifteen species as comprising the *Aspergillus niger* group, which includes all of the *Aspergilli* with black conidia. There have been suggestions to subdivide the *A. niger* group into several varieties (Al-Musallum, 1980) but currently the retention of the *A. niger* group based on black conidia seems dominant (Kusters-Van Someren *et al.*, 1990).

Classification remains difficult since molecular studies have shown that morphological classification does not reflect the genetic relatedness of strains (Kusters-Van Someren *et al.*, 1991). Therefore, a distinction between two strains may be clear at the molecular level despite a non-observable morphological difference.

Taxonomy on the basis of molecular biological techniques has proven itself to be a more reliable form of classification than one based on morphological characteristics. Kusters-van Someren *et al.* (1990) had already separated the black *Aspergilli* into *A. japonicus*, *A. carbonarius*, *A.heteromorphus*, *A. ellipticus* and a large *A. niger* aggregate. Using restriction fragment length polymorphisms (RFLPs) in ribosomal DNA, the *A. niger* aggregate was divided into two groups (Kusters-van Someren *et al.*, 1991). Phylogenic relationships between strains were determined by Kusters van Someren *et al.* (1991) by calculating the percentage of the number of hybridizing restriction fragments shared between two strains as compared with the total number of hybridizing restriction fragments present in the two strains. For instance, *A. niger* (*var. niger*) and *A. niger* (*var. awamori*) were both classified within the same group and are reported to share roughly 58% sequence identity (Kusters-van Someren *et al.*, 1991). Megnegeau *et al.* (1993), using a PCR-based technique involving the random amplification of polymorphic DNA (RAPD), further divided each group described by Kusters-van Someren *et al* (1991) into two sub-groups. Thus, according to Megnegneau *et al.* (1993), the *A. niger* aggregate can be divided into four distinct groups, types I, I', II, II', of which *A. niger*, *A. awamori*, *A. nanus*, and *A. foetidus* compose group I. The taxonomy of the black *Aspergilli* has been debatable and work in the area of DNA homology and strain relatedness is ongoing.

Hopefully, a more exhaustive use of molecular biological techniques and impending DNA sequence information may give rise to a clearer taxonomic system.

1.2 Protein production in *Aspergillus*

The utilization of fungi for protein production has been studied for years. Filamentous fungi secrete large amounts of homologous proteins and are able to grow on relatively inexpensive media. Secretion is important because it makes the product easier and cheaper to purify. The secretion of a variety of proteins, often in abundant quantities, by *Aspergillus* species such as *A. niger* has lead to a large effort to use them as hosts for expression of many heterologous proteins. Besides being effective secretors of proteins and having a long history of usage, *A. niger* is also capable of carrying out efficient post-translational modifications such as protein folding and glycosylation (Saunders *et al.*, 1989). It also has known sequences of promoters of highly expressed genes that are either constitutively expressed (*e.g. pkiA*) or regulated (*e.g. glaA*).

Heterologous protein production often requires the use of genes with regulated expression that can be tightly controlled simply by altering fermentation conditions. The choice of promoter is an important factor that affects the successful expression of heterologous proteins. For example, the production of certain heterologous proteins may be detrimental to the host. Therefore, an inducible promoter such as the promoter of the *A. niger* glucoamylase gene (*glaA*) is desirable. Although the control over gene expression offered by such glucose- repressed promoters is desirable in many instances, their use can complicate the production phase of fermentation. For protein production,

the host strain must be cultured not only under glucose- depleted conditions but also in a medium that employs a carbon source alternative to the preferred substrate, glucose.

Despite the great interest to exploit the glucoamylase expression signals for the production of heterologous proteins, regulation of the *glaA* gene is not yet completely understood. Heterologous protein production using the *glaA* promoter does not result in similar product yields obtained for the homologous fungal enzyme glucoamylase (Jeenes *et al.*, 1993; Carrez *et al.*, 1990). It is, therefore, important to further characterize the *glaA* promoter and to try to identify its mechanism of induction and repression. Other strategies developed to improve heterologous protein yields include the introduction of a large number of gene copies, the construction and use of protease deficient strains, the development of an optimal production medium and gene fusion strategies, in which the gene of interest is fused at the 3' end of a gene encoding a well secreted homologous protein (reviewed in Archer *et al.*, 1994; 1997; Verdoes *et al.*, 1994; Gouka *et al.*, 1997).

The production of heterologous gene products requires the use of plasmid vectors. Difficulty to express and secrete foreign proteins from filamentous fungi is in part due to the lack of suitable expression vectors. An expression vector must be constructed which provides the gene of interest with additional DNA sequences to allow its expression in the fungal host. These sequences typically include a strong promoter (preferably regulated), suitable restriction sites near the promoter to insert heterologous genes, a transcriptional terminator and a selectable marker. Optionally, the vector may also contain a secretion signal sequence or part of a highly expressed endogenous protein (generally from the gene which is naturally associated with the strong promoter) allowing a foreign gene to be either secreted or expressed as a fusion protein. Gene fusions have

also proved to be a powerful tool in studies of gene regulation. In this approach, a regulatory region of interest is fused to a reporter gene that can generate a product that is easily assayable.

Bi-functional plasmid vectors that enable DNA molecules to be shuttled between *E. coli* and *Aspergillus* are valuable in both the production of heterologous proteins and gene expression studies. Although a variety of shuttle vectors have been developed in model organisms such as *E. coli* and *Saccharomyces cerevisiae* (Sambrook *et al.*, 1989, Sikorski and Heiter, 1989), few are available for use with the industrial species *Aspergillus niger*.

This study describes a set of integrating and autonomously replicating shuttle vectors for use in *A. niger*. Regulation studies of the glucoamylase gene *glaA* and the pyruvate kinase gene *pkiA* of *A. niger* were carried out using shuttle vectors that contained the reporter gene *lacA*, encoding an extracellular β-galactosidase of *A. niger* (Kumar *et al.*, 1992). The plasmid vectors used for gene regulation studies also contained the selection marker *pyrG* of *A. nidulans* encoding orotidine-5'-phosphate decarboxylase (Oakley *et al.*, 1987) and the AMA1 replicator sequence (Gems *et al.*, 1991). The AMA1 (**a**utonomously **m**aintained in ***A**spergillus*) sequence is an efficient plasmid replicator and transformation enhancer in *Aspergillus*. The autonomously replicating vectors are able to persist in mycelium for an indefinite number of asexual generations without rearrangements or integration into chromosomal DNA (Alesksenko and Clutterbuck, 1995). This is important since the expression level of a reporter gene that has been integrated into a fungal chromosome is often determined by the site of integration as well as by any regulatory sequences on the expression vector.

1.3 The glucoamylase gene (*glaA*) of *Aspergillus niger*

1.3.1 Amylases of *A. niger*

A number of amylases have been characterized from strains of *Aspergillus niger* (Vihinen and Mantsala, 1989). About forty-five years ago, it was found that amylases from black mold such as *A. niger* could digest raw starch more readily than those from the yellow mold *A. oryzae*. In comparison to other fungi, *A. niger* could completely hydrolyze glutinous rice starch and glycogen to glucose without leaving any limit dextrins. The reason was later attributed to the ability of *A. niger* to secrete very high levels of the enzyme glucoamylase [α-D- (1→ 4)-glucanglucohydrolase, EC 3.2.1.3] (Ueda, S., 1981). Although starch hydrolysis in cultures of *A. niger* also results from activity of the hydrolytic enzymes α-amylase (EC 3.2.1.1) and α-glucosidase (EC 3.2.1.20), glucoamylase is the predominant secreted protein in many strains. It is typical that zero or very low levels of α-glucosidase activity are detected in starch -grown culture supernatants of *A. niger* (McCleary *et al.*, 1989). The α-amylase from *A. niger* cleaves α-1, 4 oligosaccharide links to give α-dextrins and predominantly maltose. It alone has only weak activity against raw starch but acts synergistically with glucoamylase for efficient starch degradation. Glucoamylase is an extracellular glycoprotein that produces β-D-glucose by hydrolysis of α -1, 4- glucosidic bonds, and of α -1, 6- bonds (with much less efficiency) from the non-reducing ends of starch and related oligo- and polysaccharides (Pazur and Ando, 1959) (Fig.2). The production of glucose as the end product of this reaction clearly

Fig. 2. Starch hydrolysis by glucoamylase. The glucoamylase from *A. niger* cleaves α-1, 4-glucosidic bonds and α-1, 6-bonds, although with less efficiency, from the non-reducing ends of starch to produce glucose.

α(1-4)linkage

α(1-6)linkage

non-reducing ends

differentiates glucoamylase from α – and β – amylase (EC 3.2.1.2) (Fig.3). The 1-6 branch points are not substrates for either α – and β – amylase. After amylase-catalyzed hydrolysis, highly branched cores, called limit dextrins, remain. Limit dextrins can be further degraded only after debranching enzymes have catalyzed hydrolysis of the α -1, 6- linkages at branch points.

1.3.2 The two forms of *A. niger* glucoamylase

A. niger secretes two glycosylated forms of glucoamylase, G1 and G2, with molecular masses of 71 and 61 kDa, respectively (Pazur *et al.*, 1980; Svensson *et al.*, 1983). Both forms are encoded by a single gene (Boel *et al.*, 1984) but G1 shows increased activity towards insoluble starch. The two forms of *A. niger* glucoamylases share a common NH_2 – terminal catalytic domain followed by a heavily glycosylated serine/threonine rich linker (Fig.4), but differ in the COOH-terminal region (Svensson *et al.*, 1982). The enzyme G1 digests starch approximately eighty times faster than G2 (Southall *et al.*, 1999), which lacks about one-hundred amino acids in the region of the C-terminus also reported as the starch – binding domain (Takahashi *et al.* 1985). Much remains unanswered about the precise role of the starch-binding domain, although clearly its main function is to attach to granular starch and increase the local concentration of the substrate at the active site.

Figure 3. Starch hydrolysis by α-amylase and glucoamylase. Glucoamylase is an exo-enzyme cleaving both α (1-4) and α (1-6) linked glucose units and can convert starch to glucose. α -amylase is an endo-enzyme (hydrolyzes inside a saccharide polymer) only capable of cleaving α (1-4) linked units, in contrast to glucoamylase.

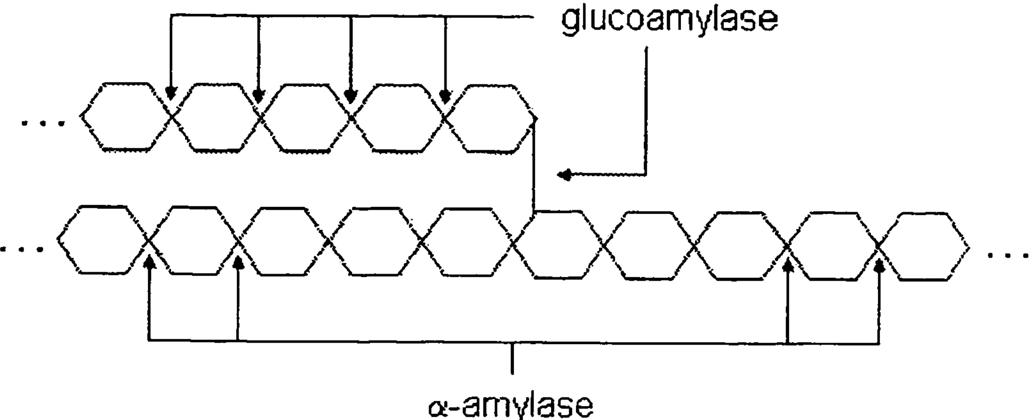
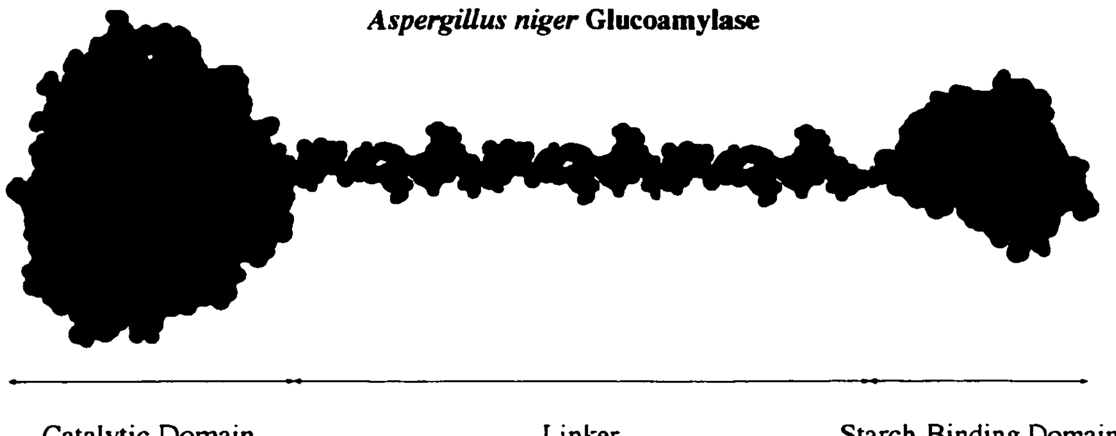

glucoamylase
α-amylase

Figure 4. Domain division in *Aspergillus niger* glucoamylase. A catalytic domain that degrades oligosaccharides from the non-reducing end, releasing glucose. A starch-binding domain that binds the enzyme to raw starch and to the cell wall. A glycosylated linker connecting the two domains. (From glucoamylase www page by Pedro M. Coutinho)


Aspergillus niger **Glucoamylase**
Catalytic Domain
~ 60 Å
Linker
~100 Å
Starch-Binding Domain
~ 30 Å

Although two forms of glucoamylase are suggested to be due to differential splicing of intervening sequences in the mRNA (Boel *et al.* 1984), the occurrence of the two forms of the enzyme has also been reported to be due to post-translational modifications of the protein (Svensson *et al.*, 1986; Dubey *et al.*, 2000). By characterizing G2 with respect to amino acid sequence and carbohydrate composition from the highly glycosylated segment, Svensson *et al.* (1986) suggest that the G2 form is generated by limited proteolysis of the larger G1 form. Dubey *et al.* (2000) identified a glycosylated 125-kDa starch-hydrolyzing enzyme as a major protein secreted by a mutant strain of *A. niger*. Antibodies raised against the 125-kDa protein were reactive to the *A. niger* 53-kDa α-amylase, the 71-kDa G1, the 61-kDa G2 and a 10-kDa non-enzymatic peptide. Antibodies raised against the 71-kDa and 61-kDa enzymes also showed reactivity towards the 125-kDa enzyme but not towards the 53-kDa α-amylase. Furthermore, the N-terminal portion of the 125-kDa enzyme was found to be identical to that of the G1 (71kDa) and G2 (61kDa) forms of *A. niger* and *A. awamori*. Dubey *et al.*, (2000), therefore, suggest that the 71-kDa glucoamylase is the N-terminal product of a proteolytically processed 125-kDa precursor enzyme, and that the 53-kDa α-amylase is the C-terminal fragment. The identification of a 10-kDa non-enzymatic protein in some culture filtrates by antibody cross-reactivity further suggests that *in vivo* post-translational proteolytic processing of the 71-kDa G1 form accounts for the occurrence of a 61-kDa glucoamylase in *A. niger*. It may, therefore, be possible that the single glucoamylase gene described by Boel *et al.*, (1984) is part of a larger gene encoding a precursor amylase, the proteolytic processing of which results in both the production of glucoamylase and α-amylase in *A. niger*.

1.3.3 Structure of the *A. niger* glucoamylase gene

The glucoamylase gene of *A. niger* was identified by screening a genomic library with a glucoamylase-specific cDNA (Boel *et al.*, 1984). A 2.5-kb *EcoRI* - *EcoRV* fragment was determined to contain the whole glucoamylase gene. Southern blot analysis identified that the *A. niger* genome contains only one copy of the glucoamylase gene.

Two initiation points of *A. niger glaA* transcription were identified by Boel *et al.* (1984) by comparing the mRNA sequence with the corresponding gene sequence. A typical TATA box is found 33 bp upstream of the first transcription initiation site. The 3'end of the gene, just before the poly (A) addition site, has five bases in common with the model sequence TTTTCACTGC of Benoist *et al.* (1980) for 3'-terminal regions of eukaryotic genes. Unlike mRNAs from higher eukaryotes, the 3' untranslated region of the *A. niger* glucoamylase mRNA does not contain the poly (A) addition signal AATAA. Boel *et al.* (1984) identified five intervening sequences in the coding region of the *glaA* gene and imply that the 169 bp intron is responsible for the production of two glucoamylase enzymes from only one gene by differential mRNA splicing. The other four introns are all very short, ranging from 55 to 75 bp, none of which have been found to persist in a translatable mRNA.

The presumptive protein product of G2 mRNA (derived from G1 mRNA differential splicing) would be similar in size to the G2 form of glucoamylase described by Svensson *et al.*, (1986). However, isolated forms of G2 have a different COOH-terminal sequence than the predicted G2 form that would result from the splicing out of

the 169 bp intervening sequence described by Boel *et al.*, (1984). Although the studies presented by Svensson *et al.*, (1986) and Dubey *et al.*, (2000) are in contradiction to what is suggested by Boel *et al.*, (1984), the possibility that one occurs cannot exclude the other. *A. niger* may produce multiple forms of G2 by either proteolysis of the G1 protein or differential splicing of glucoamylase mRNA, despite the ability to only isolate the former type of G2 protein.

1.3.4 Regulation of the *A. niger glaA* gene

Glucoamylase production is modulated by growth on different carbon sources (Nunberg *et al.*, 1984; Barton *et al.*, 1972). Expression of *glaA* is induced by growth on starch or maltose, intermediate in the presence of glucose and very low or undetectable when xylose is supplied as the sole carbon source (Fowler *et al.*, 1990).

Fowler *et al.* (1990) found that *glaA* mRNA levels of cultures grown in different carbon sources correlated well with glucoamylase protein levels. It has been suggested that transcription of the *glaA* gene is regulated by a specific *trans*-acting protein that limits its transcription (Verdces *et al.*, 1994). This hypothesis would explain the observed maximum production level of glucoamylase and the inability to significantly increase glucoamylase production by the introduction of additional gene copies (Verdoes *et al.*, 1993). It was also shown by Verdoes *et al.* (1994) that the titration effect occurs in the presence of either glucose or maltose, implying that gene expression in both carbon sources may be regulated by the same activator. Furthermore, the titration effect is

greater in glucose grown cultures, indicating that there is a greater amount of the activator in the presence of maltose.

The notion that *glaA* is regulated by a complicated pathway was further emphasized by studying glucoamylase production in pair-wise combinations of carbon sources. Fowler *et al.* (1990) found that starch or maltose in the presence of glucose can override glucose repression. It was also observed that xylose-glucose and xylose-maltose combinations resulted in a level of glucoamylase expression similar to that observed when maltose or glucose was used as the sole carbon source. However, a xylose-starch combination prevents induction of glucoamylase. Fowler *et al.* (1990) suggest that xylose ,therefore, may fully repress the production of some other starch- hydrolyzing enzyme, the product of which may act to induce *glaA* synthesis. Although Fowler *et al.* (1990) showed that xylose had no effect on the level of glucoamylase production in maltose and glucose grown cultures, Verdoes *et al.* (1994) demonstrated that the combination of xylose and a second carbon source resulted in a level of expression intermediate to that of cultures grown in xylose or the other carbon source alone. Xylose repression was also shown by Verdoes *et al.*, (1994) to be unaffected by multiple gene copies. These results lead Verdoes *et al.* (1994) to imply that xylose repression is not the result of the regulation of expression or inactivation of the titratable activator protein shown to regulate *glaA* expression in maltose and glucose.

1.3.5 The promoter region of the *A. niger glaA* gene

A comparison of the 5'- flanking region of the *A. niger glaA* gene with that of the *A. oryzae glaA* gene and the *A. oryzae amyB* gene (Wirsel *et al.*, 1989) indicated two homologous regions, in each of the three genes, which are located at almost the same distance from the translational start codon (Table 1). These regions were designated as Region I and Region II (Hata *et al.*, 1992). Region I of *A. oryzae* has 13 nucleotides identical to the upstream region of the *A. niger glaA* gene followed by eight nucleotides identical with the upstream region of the *A. oryzae amyB* gene. In Region II there is a stretch of eight identical nucleotides found in all three genes at almost the same distance from the start codon (Table 1).

1.3.6 Deletion analysis of the *glaA* promoter

Deletion analysis has been performed by several investigators in an attempt to identify regulatory elements that control *glaA* expression. Fowler *et al.* (1990) introduced plasmids with 5'- deletions of the *glaA* promoter into an *A. niger* strain whose own *glaA* gene had been disrupted. By growing the transformants on different carbon sources, Fowler *et al.* (1990) identified two regions important for *glaA* expression. A region between –562 and -318 was identified as being responsible for high-level expression of *glaA* and was called Element I. A second region designated Element II, comprising the first 224 bp upstream of the translation start point, was identified as the minimum region of the promoter required to initiate the start of transcription.

Table 1.
Homologous sequences in 5'-flanking regions of *Aspergillus* genes

Strain	Gene	Position[a]	Sequence[b]
Region I			
A. niger	*glaA*	-454 to -413	ATCGACCGGGGGACGGCGAATCCCCGGGAATT
			** * ************* * ****
A. oryzae	*glaA*	-365 to -334	TTCAGACAGGGGACGGCGAATTCACGGGCGAA
			********
A. oryzae	*amyB*	-328 to -297	CGGCCTTTTCTGCAACGCTGATCACGGGCAGC
Region II			
A. niger	*glaA*	-242 to -210	GCGAGATGGTCTCTGCAGAATTCCAAGCTCGA
			* * ******** * * *
A. oryzae	*glaA*	-231 to -200	ACTAGTCTTGGCTTGCAGAATGACTTCCGGAA
			** *** * ******** * *
A. oryzae	*amyB*	-214 to -183	CCCGGTATTGTCCTGCAGAATGCAATTTAAAC

[a] Numbers of each sequence counting from the ATG initiation codon

to the first and the last nucleotide base

[b] Identical nucleotide bases are separated by an asterisk (*)

Another study conducted by Verdoes *et al.* (1994) presented somewhat differing results. They found that the region responsible for high-level expression extended to sequences between –815 and -517 bp, since a deletion up to position –517 resulted in an expression level of about 5% that of the full length promoter. The deletion of sequences up to position –389 resulted in a complete loss of expression, leading Verdoes et *al.* (1994) to conclude that regulatory proteins interact with at least one target sequence within 517 bp upstream of the translation start codon.

Hata *et al.* (1992) performed deletion analysis of the *glaA* promoter of *A. oryzae* and found that a deletion up to position –427 did not affect high-level expression. However, the sequence from –427 to –332, containing Region I, was found to be required for both high-level expression and maltose induction.

1.3.7 Putative target sites for regulatory proteins in *glaA*

The conserved Region I, indicated by Hata *et al.* (1992), is located within the –517 and –389 sequence described by Verdoes *et al.* (1994) to be important for *A. niger glaA* expression. This region contains two 5'- SYGGGRG –3' putative CreA target sequences (Kulmberg *et al.*, 1993). There are also three other sites within the *A. niger glaA* promoter that match the CreA binding consensus sequence at positions –810, -321 and –18. Refer to Figure 5 for the sequence of the 5' region upstream of the start codon of the *A . niger glaA* gene.

In *A. nidulans*, CreA is the major regulatory protein mediating carbon repression (Kulmberg *et al.*, 1993) and will be further discussed in a following section. In the

Figure 5. Nucleotide sequence of the region upstream of the start codon of the *Aspergillus niger glaA* gene. Four sites matching the CreA binding consensus sequence are underlined. Three sites that closely match the AnCF binding consensus sequence are underlined. A single *MluI* site is also indicated. Numbers denote distance in base pairs from the ATG initiation codon of the *glaA* gene.

-655 *TGA*TGCCATT GGCGGAGGGG TCCGGACGGT CAGGAACTTA GCCTTATGAG ATGAATGATG

-595 GACGTGTCTG GCCTCGGAAA AGGATATATG GGGATCATAA TAGTACTAGC CATATTAATG

*Mlu*I site
-535 AAGGGCATAT ACC*ACGCGTT* GGACCTGCGT TATAGCTTCC CGTTAGTTAT AGTACCATCG

AnCF site1 CreA site1 CreA site2
-475 TTATACC**AGC CAATCAAG**TC ACCACGCACG ACCGGGGACG GCGAATCCCC
GGGAATTGAA

AnCF site2
-415 AGAAATTGCA TCCCAGGCCA GTGAGGCC**AG CGATTGGCC**A CCTCTCCAAG
CACACAGGGC

CreA site3 AnCF
site3
-355 CATTCTGCAG CGCTGGTGGA TTCATCGCAA TTTC**CCCCGG** CCCGGCCCGA
CACCGCTATA

-295 GGCTGGTTCT CCCACACCAT CGGAGATTCG TCGCCTAATG TCTCGTCCGT
TCACAAGCTG

-235 AAGAGCTTGA AGTGGCGAGA TGTCTCTGCA GGAATTCAAG CTAGATGCTA
AGCGATATTG

-175 CATGGCAATA TGTGTTGATG CATGTGCTTC TTCCTTCAGC TTCCCCTCGT
GCAGATGAGG

TATA box ↓
-115 TTTGGCTATA AATTGAAGTG GTTGGTCGGG GTTCCGTGAG GGGCTGAAGT
GCTTCCTCCC

↓ CreA site4
start
-55 TTTTAGACGC AACTGAGAGC CTGAGCTTCA TCCCCAGCAT CATTACACCT CAGCAATG

A. niger glaA promoter region, the two CreA sites that are located within Region I overlap with a putative target site for AmyR, a specific activator protein. This resembles the competition model presented by Kulmburg *et al.*, (1993), in which AlcR and CreA compete for the same *alc* gene promoter region.

The *glaA* promoter of *A. niger* also contains three sites that closely match the AnCF binding site consensus 5'- RRCCAATMRCR- 3' (Brakhage *et al.*, 1990). This sequence with the invariant CCAAT core is found in several fungal promoters and has been found to be important in the transcriptional enhancement of several catabolic enzymes (Littlejohn and Hynes 1992; Nagata *et al.* 1993; Kato *et al.* 1997) and penicillin biosynthesis genes (Litzka *et al.*, 1996 and Then Bergh *et al.*, 1996). The proteins that bind to CCAAT sequences will be further discussed in a subsequent section.

1.4 The pyruvate kinase gene of *A. niger*

Pyruvate kinase (ATP: pyruvate phosphotransferase, EC 2.7.1.40) is a key regulatory enzyme in glycolysis. It catalyzes the irreversible conversion of phosphoenolpyruvate (PEP) into pyruvate, generating ATP. Pyruvate kinases from a wide range of organisms have been studied. Several fungal pyruvate kinases have been isolated including those of *Neurospora crassa* (Kapoor and Tronsgaard 1972), *Mucor rouxii* (Terenzi *et al.* 1971), *A. nidulans* (Uitzetter, 1982; de Graff *et al.*, 1988) and most recently *A. niger* (de Graff *et al.*, 1992). Nearly all characterized eukaryotic pyruvate kinases are activated by fructose 1,6- bisphosphate (FBP) and tightly regulated to avoid

the simultaneous operation of glycolytic and gluconeogenic metabolic pathways (Uitzetter, 1982).

The pyruvate kinase gene (*pkiA*) is regarded as a highly expressed gene (de Graff *et al.*, 1992). Pyruvate kinase activity in *A. nidulans* has been shown to depend on the carbon source used for growth (Payton and Roberts 1976; Uitzetter 1982; de Graff *et al.*, 1988). It is reported by Uitzetter *et al.* (1982) that the use of a glycolytic carbon source leads to a 20-fold increase in enzyme activity as compared to the level found when the organism is grown on a gluconeogenic carbon source. However, de Graff *et al.*, (1988) only found a 7-fold decrease in pyruvate kinase activity under gluconeogenic growth conditions compared to glycolytic growth conditions.

The pyruvate kinase protein sequence of *A. niger* is highly homologous with that of *A. nidulans* (de Graff *et al.*, 1992). However, despite the fact that the coding regions of the two genes only differ by 33 of their 526 amino acids, the promoter regions of the two genes share little similarity. The conserved region in the *A. niger* and *A. nidulans* *pkiA* promoter is a 17 bp sequence 5'- CRCGGGGAACCGRCCGA- 3'. It is found at position –282 to –266 in *A. niger* (Fig.6) and at position –245 to –229 in *A. nidulans* (de Graff and Visser 1988). It is speculated by de Graff *et al.* (1992) that this element might be involved in the transcriptional regulation of *pkiA*, either as an activating or repressing element analogous to the upstream activating and repressing sequences found in the yeast pyruvate kinase gene (Nishizawa *et al.* 1989; McNeil *et al.*, 1990). However, it is not known whether this conserved sequence is involved in the regulation of *pkiA*. Unlike the pyruvate kinase gene of *S. cerevisiae* (PYK1), the *pkiA* genes of *A. nidulans* or *A. niger* have not previously undergone deletion analysis of their 5'- regions to map possible *cis*-

Figure 6. Nucleotide sequence of the 5' region upstream of the start codon of the *Aspergillus niger* pyruvate kinase gene. A putative TATA box and CCAAT-binding sites are underlined. The region that is conserved in the *A. niger* and the *A. nidulans pkiA* promoter is highlighted in gray. Numbers denote distance in base pairs from the ATG initiation codon of the *pkiA* gene.

```
-1041 AAATGGAAGA GAAAACCTCC GAGTACTTAC TTAGGGCCCT GTCTACTGGC
 -991 CAGAGTCTCG TCCTCTATTC ACTATGATAA ATTACCCACT GGACAAAAAA
 -941 ATAAAATAAA ATAAAAATAA AAAGGGAGAC AGCTTCTCCA TCAACTGGCA
 -891 ACTGGGTCCG TCCGAGCAGA GCAAAATTCA GCCTTATGGG TTCCGATGGA
 -841 GTCAGGGAAA TAGTTCTTGC GAAGGGCATT GGGCTTTTTT GCGAGGAGAA
 -791 AATTCAGCAC CGACAAAGCA TCCGAAATCC GACCTCGCTA GGAGAGAATG
 -741 GATCCGCGAC GATGTGGGGT CAACTGGACA GAGTGAGAGG GTATCATGTG
 -691 GTCCTGCCAG ATACTTCGCA GAATGTTGTG TGGGTGTCTG ATTGTGGCTT
 -641 GGGCGTGAAT TGCTTTTGGT CTTCCCAACC AATTATTATT GCATGCGGCG
 -591 TATGAATGCC TGAGATGCGC GGAGGGAAGG TGCCTGAGGA TGTAGTGGAC
 -541 AAATGCTGCT GATCGCTGGG CGGAAACCCT TGGCTGACCA GTGAAAAGAG
 -491 CGGACGGAGG CAGCAGGTGT ATCTACGATC AAAGAATAGT AGCAAAGCAG
 -441 TGAAAGGTGG ATCACCCAGC AAATAATTGA GTTTTGATAC CCAGCGATAG
 -391 TGCCGGGGGG GAGAAAAAGT CATTAATAAT GGGAATTATG TAGGCGATGG
 -341 GAAGTGTGAT TGTAACTACT CCGTAGCTGG AGGCACAACT AACAAGCCAG
 -291 CTCTCAACCC GCGGGGAACC GACCGACAGA TAAAAAAAAG CGTCCCAAAG
 -241 CAGGAATCCC ACCAAAAAGG GCCGATCCAG CCAATCACCG CCGCCAACAT
 -191 TTTTCCTTCC CGGGCACCCC TCCTCTAGTC CACCATCTCT CTCTTCTCTC
 -141 GCTCACCGGC CCCGTCTTTT CCTTCCCTAT TATCTCTCCC TCTTCTCCTC
  -91 CCTTCTCTCC CTCCATTCTT TCTCCCATCT TCATCACTCC CTTCTCTTCT
  -41 GTCTTCCCCC CCGGTTCAGT AGAGATCAAT CATCCGTCAA GATG
```

acting regulatory elements. The promoter region of *A. niger pkiA* also contains a possible TATA box at position –114 and two putative CCAAT-binding sites at position –211 and -613. In addition, the *A. niger* promoter region contains an extensive CT-enriched region (CT block) 190 bp preceding the translational start site, which is common to highly expressed genes in yeast and other fungi (Fig.6). A comprehensive study on the regulation of *pkiA* and the identification of its regulatory elements and possible transcription factors remains to be published.

1.5 The CCAAT – binding sequence

The CCAAT sequence is one of the most common *cis*-elements present in eukaryotic promoter regions. It is found in the 5' region of approximately 30% of eukaryotic genes (Bucher, P., 1990). CCAAT sequences are commonly found 50 to 200 bp upstream of the transcriptional start site and can be present in either orientation. A protein complex binding to CCAAT sequences was first identified in *Saccharomyces cerevisiae* and designated the Hap complex (Pinkham and Guarante, 1985). The Hap complex is a heteromultimer composed of four subunits, Hap2p / Hap3p / Hap4p / Hap5p (Hahn and Guarente 1988; Forsburg and Guarente 1989; McNabb *et al.* 1995). Homologues to *the Saccharomyces cerevisiae* Hap complex have been identified and characterized in *A. nidulans* and *A. oryzae* (Brakhage *et al.*, 1999; Tanaka *et al.*, 2000). Proteins binding to CCAAT sequences in *A. nidulans* have been designated AnCF (Van Heeswijck *et al.*,1991), AnCP (Nagata *et al.*, 1993) and PENR1 (Litzka *et al.*, 1996).

The Hap complex of *Saccharomyces cerevisiae* has a role in the activation of a large number of genes involved in oxidative phosphorylation in response to growth on non-fermentable carbon sources (Pinkham and Guarente, 1985). However, homologues of the Hap complex in other organisms have been found to regulate genes not involved in respiration. The mammalian CCAAT-binding complex, NF-Y/CBF, regulates genes such as the human α-globin gene, the mouse class I major histocompatibility gene and the human heat shock protein 70 gene (Chodosh *et al.*, 1988). In filamentous fungi, CCAAT sequences modulate expression of genes such as the *A. nidulans* acetamidase gene (*amds*) (Van Heeswijck *et al.*, 1991), the *A. oryzae* Taka-amylase A gene (*taa*) (Tanaka *et al.*, 2000), the *A. nidulans* penicillin biosynthesis genes (*ipnA and aat*A) (Litzka *et al.*, 1996) and the *Neurospora crassa* NADP-specific glutamate dehydrogenase gene (*am*) (Chen *et al.*, 1994).

Genes with significant homology to the *S. cerevisiae* subunits: Hap2p, Hap3p and Hap5p have been cloned from *A. nidulans* (Papagiannopoulos *et al.*, 1996 and Steidl *et al.*, 1999) and designated HapB, HapC and HapE respectively. Recently, Tanaka *et al.* (2001) have cloned the *A. oryzae* homologues for the three subunits, AoHapB / AoHapC / AoHapE, which are functionally interchangeable with the corresponding subunits in *A. nidulans* (Tanaka *et al.*, 2000).

Activator proteins consist of at least two distinct and separable domains: a DNA binding domain and a transcriptional activation domain. The Hap complex distributes these functions among different subunits. Three subunits were found in *S. cerevisiae* and mammalian CCAAT – binding complexes to be necessary for DNA-binding and subunit association (Olesen *et al.*, 1990; Maity *et al.*, 1992; Xing *et al.*, 1993; Sinha *et al.*, 1996;

Kim *et al.*, 1996). The fourth subunit, Hap4p, of the yeast Hap complex is not necessary for DNA binding but interacts directly with the Hap2p / Hap3p/ Hap5p complex and contains a transcriptional activation domain in its C-terminal region. No homologues of Hap4p have been found except in the yeast *Kluyvermyces lactis*, although a gene encoding a novel transcriptional activator, which interacted with the Hap complex was isolated from *A. nidulans* and designated *hapX* (Tanaka *et al.*, 2002).

The CCAAT sequence has been found not to be essential for the inducibility of genes such as the *A.oryzae* taka-amylase gene (*taa*), the *N.crassa* NADP-specific glutamate dehydrogenase (*am*) gene and the *A. nidulans* acetamidase (*amds*) gene. However, it has been shown to be important in determining the level of gene expression (Kato *et al.* 1996; Frederick and Kinsey 1990; Littlejohn and Hynes 1992). CCAAT–binding proteins of *A. nidulans* can bind to CCAAT sequences of a number of different genes and regulate their expression (Kato *et al.*, 1996). It is, therefore, postulated that CCAAT-binding complexes act as global transcriptional activators. Kato *et al.* (1996) propose that proteins other than AnCP interact with AnCP-binding promoters, since one *cis*-regulatory sequence can potentially bind to multiple transcription factors in eukaryotic cells (Johnson and McKnight, 1989). Later, Steidl *et al.* (1999) demonstrated that expression of the *A. nidulans amds* gene by the pathway specific regulator AmdR is dependent on AnCF function and suggests that AnCF facilitates AmdR binding to DNA *in vivo.*

The observation that the binding of AnCF is a prerequisite for a pathway-specific transcription factor, like AmdR, was further investigated by Narendja *et al.* (1999). The AnCF complex and the CCAAT sequence were found to be necessary for the formation

of a nucleosome-free, DNase I-hypersensitive region in the 5' region of the *amds* gene (Narendja *et al.* 1999). DNase I –hypersensitive sites often coincide with binding sites for transcription factors (Almer *et al.*, 1986; Moreira *et al.*, 1998). Deletion or mutation to the CCAAT sequence in the promoter region of the *amds* gene or deletion of one of the *hapB*, *hapC* and *hapE* genes resulted in not only a reduction of *amds* expression but also a distinct alteration of the gene's chromatin structure. It remains to be seen if the role of CCAAT sequences presented by Narendja *et al.* (1999), in establishing an open chromatin structure necessary for full transcriptional activation in the *amds* promoter, applies to other systems. Furthermore, studies on the interaction of CCAAT-binding complexes with pathway-specific regulators are still to be performed.

1.6 The CreA Repressor

Filamentous fungi are characterized by a versatile metabolism enabling them to grow on a variety of carbon sources. The ability to adjust carbon catabolism to existing conditions usually involves carbon repression. Carbon repression is a global regulatory mechanism in which the presence of glucose or other rapidly metabolizable carbon sources represses expression of enzymes required for the utilization of alternate carbon sources.

Genetic studies of the model organism *A. nidulans* established *creA* as the major regulatory gene controlling carbon catabolite repression (Arst and Macdonald, 1975; Arst and Bailey, 1977). The *creA* gene has been cloned and sequenced from *A. nidulans* (Dowzer and Kelly, 1989, 1991) and later from *A. niger* (Drysdale *et al.*, 1993). The *A. nidulans* CreA is a DNA binding protein of 415 amino acids that has two zinc fingers of

the Cys_2His_2 class, frequent S(T)PXX motifs and an alanine-rich region indicative of a DNA-binding protein (Dowzer and Kelly, 1991). The *A. nidulans* CreA resembles the Mig1 protein of *S. cerevisiae*, a DNA-binding protein involved in carbon catabolite repression (Nehlin *et al*., 1990), although the similarity between these two fungal repressors is restricted to the DNA-binding domain.

The *A. niger creA* gene contains an intron-less ORF of 1281bp encoding a protein of 427 amino acids (Drysdale *et al*., 1993). Its alignment with the *A. nidulans creA* shows 75% sequence identity at the nucleotide level and 90% sequence similarity at the amino acid level (Drysdale *et al*., 1993). The two proteins are functionally equivalent, since the *A. niger* gene is able to complement a *creA* mutation of *A. nidulans* (Drysdale *et al*., 1993). The *A. niger* CreA also contains two zinc-fingers which have 84% overall identity with the zinc-finger region of the *S. cerevisiae* protein Mig1p.

CreA binds the consensus sequence 5'- SYGGRG- 3' (Cubero and Scazzocchio, 1994). It belongs to the group of zinc finger proteins of the Zif268 type (Dowzer and Kelly, 1991). Crystallographic studies of a Zif268-DNA complex, by Pavletich and Pabo (1991), were used to build a model for the interaction of CreA with its DNA target (Kulmburg *et al*., 1993). It was proposed by Kulmburg *et al.* (1993) that the binding of the fingers to DNA is antiparallel, that is the amino terminis of the zinc finger regions bind the 3' end of cognate DNA targets. The first finger of CreA recognizes the triplet 5'-GRG-3', while the second finger recognizes the triplet 5'-SYG-3'. The two CreA zinc fingers bind in the major groove of the DNA in a similar fashion to Zif268, with each finger contacting a triplet in the guanine-rich strand (Pavletich and Pabo, 1991).

Many CreA-binding consensus sites can be found in promoter regions of *Aspergillus* genes. DNase I -footprinting experiments using a GST-CreA fusion protein have shown protection of a number of CreA-binding consensus sites (Kulmburg *et al.*, 1993; Cubero *et al.*, 1994). However, not all possible consensus sequences are functional. Studies of the *A. nidulans* ethanol regulon revealed that only two of the seven putative CreA consensus sites in the *alcA* promoter are functional (Panozzo *et al.*, 1998) and in the *alcR* promoter, four of the nine putative CreA sites are functional (Mathieu *et al.*, 2000). It also was found that CreA can bind to non-consensus sites in the *A. nidulans ipnA* promoter (Espeso *et al.*, 1994). Additional elements are therefore thought to be necessary for the specificity of CreA recognition. Suggestions include a possible CreA dependence on a flanking AT-rich region similar to the site selection of Mig1p shown by Lundin *et al.* (1993). Although the *prnB* promoter of *A. nidulans* has AT-rich regions upstream of the consensus site (Cubero and Scazzocchio, 1994), there are no AT-rich regions upstream of functional CreA sites in the promoter regions of *alcA* and *alcR* (Panozzo *et al.*, 1998; Mathieu *et al.*, 2000) or in Mig1p sites in *GAL1* and *GAL4*.

Interestingly, a characteristic of functional CreA binding sites is their frequent organization in *A. nidulans* as pairs of sites. Pairing of functional CreA targets is observed in the promoter region of *alcR* and *alcA* (Panozzo *et al.*, 1998; Mathieu *et al.*, 2000), *prn B* (Cubero and Scazzocchio, 1994) and *creA* itself (Strauss *et al.*, 1999). Mig1p sites are also often organized as repeats (Nehlin *et al.*, 1991), however, the *xlnA* promoter of *A. nidulans* only has one functional CreA binding site. Therefore, it is not known whether it is necessary for functional CreA binding sites to be present as pairs.

Catabolic pathways regulated by the general transcriptional repressor CreA are also often subject to pathway-specific induction. The interplay between the two regulatory systems is characterized in the *A. nidulans* ethanol utilization pathway (reviewed in Felenbok, 2001). Two enzymes, alcohol dehydrogenase and aldehyde dehydrogenase are encoded by the structural genes *alcA* and *aldA*. The two genes are subject to pathway-specific induction mediated by the activator AlcR and to carbon catabolite repression mediated by CreA (Mathieu and Felenbok, 1994). Control of ethanol catabolism in *A. nidulans* involves the direct interaction of two regulatory systems. This interaction is characterized by direct competition between the two regulators AlcR and CreA for the same sites in promoter regions of responsive genes (Mathieu and Felenbok, 1994; Panozzo *et al.*, 1998). Another characteristic of the *alc* system is the strong repression of the *alc* genes under carbon catabolite repressing conditions. This results from a "double-lock mechanism" in which there is a direct repression of the activator gene *alcR* by CreA as well as an independent repression of the structural *alcA* gene (reviewed in Felenbok, 2001). These features fine-tune the utilization of carbon sources and may be present in other catabolic systems.

1.7 Regulation of *creA*

Several regulatory events affect the function of the carbon catabolite repressor CreA. Regulation of CreA is complex involving at least three mechanisms (Strauss *et al.*, 1999). First, the addition of monosaccharides, irrespective of whether they exert carbon catabolite repression or derepression, causes a rapid increase of *creA* mRNA. The

upregulation of *creA* transcription appears to be dependent on monosaccharide uptake involving the uncharacterized *creB* gene (Arst *et al.*, 1981). Second, prolonged incubation with repressing carbon sources leads to transcriptional autorepression of *creA*, whereas incubation on derepressing carbon sources leads to a continuous slow accumulation of *creA* mRNA. Downregulation of *creA* transcription is mediated by two closely spaced CreA binding sites in the *creA* promoter (Shroff *et al.* 1996; Strauss *et al.* 1999). Third, despite elevated concentrations of *creA* mRNA under carbon derepressed conditions, an increase in *creA* mRNA levels does not correlate with an increase in CreA functional protein (Strauss *et al.*, 1999). This finding could result from either protein inactivation by covalent modification and/or protein degradation under derepressing conditions. Strauss *et al.* (1999) suggest that either CreA itself or a factor activating CreA on glucose could be a target for proteasome-mediated degradation, and that a transfer from derepressing to repressing conditions would require *de novo* protein synthesis.

It is also possible that CreA is post-transcriptionally regulated at the level of nuclear import/export and by phosphorylation/dephosphorylation, as in the case of Mig1p. In the presence of glucose, the protein kinase Snf1p is inactivated resulting in Mig1p dephosphorylation and nuclear import (DeVit *et al.*, 1997). A similar regulation of CreA function may exist since it has been shown that the CreA homolog in *Sclerontinia sclerotiorum*, (an ascomycetes), Cre1, is imported into the nucleus of *A. nidulans* in the presence of glucose (Vautard *et al.*, 1999). Repression by Mig1p in *S. cerivisiae* requires the recruitment of corepressors Snn6p and Tup1p (Keleher *et al.*, 1992) to glucose-repressible promoters and the protein complex is the actual

transcriptional repressor (Cooper *et al.*, 1994; Redd *et al.*, 1997). However, it is not yet known whether CreA interacts with any other proteins.

1.8 The transcriptional activator AmyR

Different *Aspergillus* species appear to share similar regulatory machinery for expression of amylase genes. Two transcription factors, AnCP and CreA have already been shown to be involved in amylotic gene expression in the well-characterized model fungus *A. nidulans*. However, in contrast to the regulatory mechanisms underlying high-level expression and glucose repression there is little known on specific induction of amylase genes. Promoter regions of amylase genes induced by starch/maltose share a conserved sequence Region IIIa, that has been shown to be involved in the induction of the *agdA* gene encoding alpha-glucosidase in *A. oryzae* (Minetoki *et al.*, 1996) and the induction of the *A. oryzae* alpha-amylase gene *amyB* in *A. nidulans* (Kanemori *et al.*, 1999). It was observed that the introduction of multiple copies of Region IIIa into the *agdA* promoter lead to a decrease in the production of taka-amylase and glucoamylase, two analogues coded for by the *amyB* and *glaA* genes, respectively, suggesting that a common positive factor regulates the expression of these genes (Minetoki *et al.*, 1998). A titration effect was also found by the introduction of multiple copies of the upstream region of the *A. niger glaA* gene (Verdoes *et al.*, 1994).

A starch-responsive element (SRE) was shown to be required for inducible expression of the Taka-amylase A gene (*taaG2*) of *A. oryzae* in *A. nidulans* (Tani *et al.*, 2000). Petersen *et al.* (1999) and Gomi *et al.* (2000) separately cloned a regulatory gene

for amylase synthesis from *A. oryzae* designated *amyR*. The *amyR* gene from *A. niger* was cloned in 1999 by Udagawa (unpublished) and recently, Tani *et al.* (2001) cloned and sequenced the *A. nidulans amyR* gene.

AmyR belongs to the $Zn(II)_2Cys_6$ binuclear cluster family that includes a number of other fungal transcription factors (Todd *et al.*, 1997). The *A. nidulans* AmyR shows 72% amino acid identity to the *A. oryzae* AmyR and 75% overall amino acid identity to the *A. niger* AmyR. The cys-6 zinc cluster DNA-binding motif at the NH_2-terminus is 100% conserved among the *Aspergillus* species, suggesting that all AmyRs bind to identical DNA sequences. Petersen *et al.* (1999) suggest that AmyR binds two types of sequences, $CGGN_8CGG$ and CGGAAATTTAA. The latter coincides with the SRE responsible for starch induction of the *A. nidulans taaG2* gene (Tani *et al.*, 2000) and it is believed that the SRE binder (SREB) and the *A. nidulans* AmyR are identical to each other.

The CGGAAATTTAA site is found in the promoter region of the α-amylase gene, the glucoamylase A gene and the glucoamylase B gene of *A. oryzae*. However, the $CGGN_8CGG$ site is present in most starch-induced promoters from both *A. niger* and *A. oryzae*. The *A. niger glaA* gene has two $CGGN_8CGG$ sites at approximately 800 bp and 400 bp upstream of the translational start site. Verdoes *et al.* (1994) have found, by deletion analysis of the *A. niger glaA* promoter, that deletion of the most upstream *amyR* site results in a six-fold reduction in expression and that deletion of the next *amyR* site reduces expression below the limits of detection.

Most proteins of the $Zn(II)_2Cys_6$ binuclear cluster family bind as homodimers to their targets, for example, the *S. cerevisiae* GAL4 and PPR1 proteins (Reece *et al.*, 1993).

However, some $Zn(II)_2Cys_6$ proteins, such as the *A. nidulans* AlcR, bind as a monomer to their targets (Nikolaev *et al.*, 1999). It was suggested that AmyR can bind as a monomer to a single CGG site comprising the CGGAAATTTAA motif in *A. oryzae* and *A. nidulans* (Tani *et al.*, 2001). However, *in vivo* and *in vitro* studies of the $CGGN_8CGG$ site in the *agdA* promoter with one or both of the CGG triplets altered showed that mutations disrupting either of the CGG triplets caused a drastic decrease in the gene's inducibility (Tani *et al.*, 2001). The authors concluded that although AmyR can bind to a single CGG triplet site, the directly repeated CGG triplets are required for high affinity binding. Tani *et al.* (2001) hypothesized that the direct CGG triplets allow two AmyR molecules to interact with each other and stabilize the DNA-protein interaction.

Expression of the *amyR* gene in *A. nidulans* is regulated by growth on various carbon sources (Tani *et al.*, 2001). Expression levels of the *amyR* gene are highest when grown on starch, followed by maltose, glucose and glycerol. The promoter region of the *A. nidulans amyR* gene has five putative CreA binding sites. In a *creA* mutant strain, *amyR* mRNA levels on glucose, maltose and starch are almost identical. Transcription levels on glucose and maltose were also more abundant in the *creA* mutant strain, indicating that the *amyR* gene is negatively controlled by CreA.

1.9 Rationale for thesis

Great interest has been developed in exploiting the inducible promoter of the *A. niger* glucoamylase gene (*glaA*). However, regulation of *glaA* is not fully understood. Heterologous protein production using the glucose-repressed *glaA* promoter has usually lead to less than favorable yields. It has, therefore, become important to further understand the mechanism of *glaA* induction and repression. This study intended to identify regulatory elements that control *glaA* expression and to also begin mapping possible cis-acting regulatory elements of the *A. niger* pyruvate kinase gene (*pkiA*).

This was done by subjecting promoter regions to deletion analysis and/or site-directed mutagenesis. Attention was focused on putative binding sites for the CreA repressor, the AnCF binding-complex and the transcriptional activator AmyR. These proteins have been shown to play a significant role in amylotic gene expression in the well-characterized model fungus *A. nidulans*. Nonetheless, the involvement of CreA, the AnCF binding complex or AmyR in regulating genes such as the *glaA* gene of *A. niger* has not been reported. Furthermore, despite the identification of putative target sites for regulatory proteins in the 5'region of *glaA* by other investigators, a study testing the functionality of the sites has not been published. This study is an attempt to map *cis*-acting elements of *glaA* and to further understand the regulatory mechanisms underlying high-level expression, glucose-repression and specific induction of the *A. niger glaA* gene.

2. Materials and Methods

2.1 Strains and growth of *A. niger* and *E. coli*

2.1.1 *A. niger* and *E. coli* strains

The *A. niger* strain N593 (*cspA1 pyrG6*) used in this study was obtained from C.J. Bos (Wageningen Agricultural University, Netherlands). The strain was derived from strain N402, a mutant of strain ATTC 9029 (Bos *et al.*, 1988) which carries the UV induced *cspA1* mutation conferring the short conidiophore phenotype. The *E. coli* strain DH5α (Woodcock *et al.*, 1989) was the host for maintenance and propagation of plasmid DNAs.

2.1.2 Media and culture conditions

E. coli were grown in double-strength YT medium consisting of 1.6% Bacto-tryptone, 1.0% yeast extract (Bacto) and 0.5% NaCl, with 2% agar added for solid media. *E. coli* plates and liquid cultures were incubated for 12-14 hours at 37 °C. The medium was supplemented with ampicillin (100 μg/mL) for the selection of plasmid transformants.

A. niger conidia can germinate in defined minimal medium (MM) or complete medium (CM) (Kafer *et al.*, 1977), plus appropriate supplements for auxotrophic strains. To obtain fresh conidia for inoculation of liquid cultures, conidial stock was streaked onto a suitable plate and incubated at 30 °C for six days. Conidia were harvested in 10

mL of a 0.09 M NaCl, 0.001% Tween 80 solution by rubbing the surface of the conidiating culture with a sterile spreader to dislodge the spores and stored at 4 °C. A hemacytometer was then used to determine conidial concentration. *A. niger* mycelia were grown from conidia at a concentration of 10^6 per mL of minimal medium containing various carbon sources and supplemented with 10 mM uridine. The different minimal medias consisted of either 15% (w/v) maltose, 15% (w/v) glucose or 5% (w/v) xylose (Fowler *et al.*, 1990). Liquid media cultures of 5 milliliters or greater were grown for six days at 30°C in an air incubator shaking at 200 rpm. Liquid mycelial cultures of 200 µL were grown in 96 well microtiter plates at the standard conidial concentration (10^6 conidia/mL culture volume) and incubated at 30 °C without shaking for three days. Mycelial growth was monitored using a microplate reader as decribed by Langvad (1999). Mycelia used for genomic DNA isolation or protoplast generation were grown from freshly harvested conidia in complete medium that was supplemented when required. Mycelial cultures were subsequently grown overnight with shaking at 150 rpm, at a temperature of 30 °C.

2.2 Transformation of *A. niger* and *E. coli*

Transformation experiments with *A. niger* required the production of protoplasts. The protocol used to generate protoplasts from young hyphae was a modification of that of Debets and Bos (1986). Approximately 4×10^8 freshly harvested conidia of the *A. niger* strain N593 (*cspA1 pyrG6*) were used to inoculate 400 mL of supplemented CM. The liquid mediim culture was incubated overnight (a maximum of 16 hr) at 30 °C and

150 rpm in an air incubator shaker. The mycelia were harvested by filtration through Miracloth (Calbiochem) and blot-dried. Mycelia then were added to a flask of Mutanase (InterSpex Product, Inc.) at a concentration of 10 mg/mL in lytic buffer (50 mM maleic acid, 0.6 M ammonium sulfate, pH 5.5). Cell wall digestion was allowed to proceed for 1-2 hours at 30 °C with shaking at 150 rpm. The digested mycelia then were filtered through sterile Miracloth and kept on ice. The protoplasts (1×10^7 to 4×10^7 per gram of mycelia) were harvested by centrifugation at 2 500 × g for 10 minutes at 4 °C. Protoplasts were suspended in 10 mL of 0.7 M KCl and washed two more times in 10 mL of S/C (1 M sorbitol, 50 mM $CaCl_2$). After the three washes, the protoplasts were suspended in S/C at a concentration of 1×10^8 protoplasts /mL and used for transformation as described by Wernars *et al.*, (1997). Protoplast suspensions of 200 µL were mixed with 1 to 10 µg of transforming DNA (in a maximum volume of 10 µL) and 50 µL of PEG buffer [25% polyethyleneglycol 8000 (w/v), 50 mM $CaCl_2$, 10 mM TrisHCl pH 7.5) followed by a 20 minute incubation on ice. After that 2 mL PEG buffer was added and the suspension was left at room temperature for 5 minutes. Following the addition of 4 mL of S/C, the transformed protoplasts were harvested by centrifugation at 2500 × g for 10 minutes at room temperature and suspended in 500 µL S/C. Aliquots of transformed cells were then added to 4 mL molten agar overlayers (45 °C) and plated onto MM plates with KCl (44.7 g per liter). Transformants harboring plasmids expressing *A. niger* β-galactosidase were plated onto plates that also contained the chromogenic indicator 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) at a final concentration of 40 µg/mL. Transformants appeared within 3-5 days. Independent

transformants were subsequently streaked onto fresh selection plates, allowed to grow for 6 days and harvested.

E. coli transformations were routinely carried out by heat-shock of $CaCl_2$ treated competent cells (Hanahan, 1983).

2.3 Plasmid construction

2.3.1 Restriction enzymes and oligonucleotides

Restriction endonuleases and DNA modifying enzymes were obtained from MBI Fermentas or New England Biolabs. Oligonucleotides used in this study (Table 2) were purchased from Biocorp Montreal.

2.3.2 Isolation and manipulation of DNA

Small scale preparations of plasmid DNA were performed by using either the Wizard© Miniprep Kit (Promega) or the QIAprep® (Spin) Miniprep Kit (Qiagen Inc.). Methods for DNA manipulation, such as restriction endonuclease digestions and ligations were done as described in Sambrook *et al.* (1989). DNA sequencing was performed by the Centre for Structural and Functional Genomics at Concordia University using the Beckman CEQ 2000XL DNA Analysis System.

Table 2. Oligonucleotides used in this study

Oligo designation	Oligo description	Oligo sequence
PGlaPr5'	*glaA* promoter sense primer the *NotI* and *AatII* sites are underlined, identical to -657 to -639	5'-AGAATGCGGCCGCATGACGTCCCACATCTGATGCCATTGG-3'
PGlaPr3'	*glaA* promoter antisense primer, the *FseI*, *XbaI* and *NheI* sites are underlined, mutations included in the oligo are italicized, identical to -24 to -1	5'-AGAATGGCCGGCCATTCTAGATGCT*AGC*GTGTAATGATGCTGGGG-3'
PGLA5'-477	*glaA* sense primer, the *MluI* site is underlined, identical to -477 to -458	5'-ATAACGCGTTTATACCAGCCAATCAAGTC-3'
PGLA5'-448	*glaA* sense primer, the *MluI* site is underlined identical to -448 to -429	5'-ATAACGCGTGACCGGGGACGGCGAATGCC-3'
PGLA5'-368	*glaA* sense primer, the *MluI* site is underlined, identical to -368 to -350	5'-ATAACGCGTGCACACAGGGCCATTCTGCAGC-3'
PGLA5'-251	*glaA* sense primer, the *MluI* site is underlined, identical to -251 to -232	5'-ATAACGCGTCCGTTCACAAGCTGAAGAGC-3'
PGLA5'-142	*glaA* sense primer, the *MluI* site is underlined, identical to -142 to -123	5'-ATAACGCGTCCGTTCACAAGCTGAAGAGC-3'
PGLA5'-142 *MluI*	*glaA* sense primer, the *MluI* and *AgeI* sites are underlined, identical to -142 to -123	5'-ATAACGCGTATACCGGTCAGCTTCCCCTCGTGCAGA-3'
PGLA5'-88	*glaA* sense primer, the *MluI* site is underlined, identical to -88 to -69	5'-ATAACGCGTGGTTCCGTGAGGGGCTGAAG-3'
PGLA3'AgeI	*glaA* antisense primer, an *AgeI* site is underlined, identical to -159 to -136	5'-CTGACCGG*T*GAAGCACATGCATC-3'

PGLA5'UMlul	*glaA* sense primer, identical to -652 to -634	5'-ATGCCATTGGCGGAGGGGTC-3'
PCREA5'1a	*glaA* sense primer, CreA1mutations included in the oligo are italicized, identical to -452 to -433	5'-ACGCACGAC*ATAC*GACGGCG-3'
PCREA3'1b	*glaA* antisense primer, CreA1 mutations included in the oligo are italicized, identical to -451 to -434	5'-TCGCCGTC*GTAT*GTCGTGCG-3'
PCREA5'2a	*glaA* sense primer, CreA2 mutations included in the oligo are italicized, identical to -436 to -414	5'-GGCGAATC*ATAC*GGAATTGAAAG-3'
PCREA3'2b	*glaA* antisense primer, CreA2 mutations included in the oligo are italicized, identical to -437 to -415	5'TTTCAATTCC*GTAT*GATTCGCCG-3'
PCREA5'1/2A	*glaA* sense primer, CreA1&2 mutations in the oligo are italicized, identical to -448 to -414	5'-ACGAC*ATAC*GACGGCGAATC*ATAC*GGAATTGAAAG-3'
PCREA3'1/2M	*glaA* antisense primer, CreA1&2 mutations in the oligo are italicized, identical to -451 to -419	5'-AATTCC*GTAT*GATTCGCCGTC*GTAT*GTCGTGCG-3'
PCREA5'3a	*glaA* sense primer, CreA3 mutations included in the oligo are italicized, identical to -331 to -319	5'-TCGCAATTTCC*ATAC*GCCCGG-3'
PCREA3'3b	*glaA* antisense primer, CreA3 mutations included in the oligo are italicized, identical to -329 to-319	5'-CCGGGC*GTAT*GGAAATTGCG-3'
PCREA3'4b	*glaA* antisense primer, *Nhe*I site is underlined, CreA4 mutations included in the oligo are italicized, identical to -30 to +1	5'-GTAAT<u>GCTAGC</u>GTGTAATGATGC*GTAT*GATGAAG-3'
PANCF5'1a	*glaA* sense primer, ANCF1 mutations included in the oligo are italicized, identical to -476 to -453	5'-GTTATACCAGC*ATACG*AAGTCACC-3'

PANCF3'1b	*glaA* antisense primer, ANCF1 mutations included in the oligo are italicized, identical to -476 to -453	5'-GGTGACTT*CG*T*AT*GCTGGTATAAC-3'
PANCF5'2a	*glaA* sense primer, ANCF2 mutations included in the oligo are italicized, identical to -390 to -370	5'-GCCAGCG*CA*T*ACG*CACCTCTC-3'
PANCF3'2b	*glaA* antisense primer, ANCF2 mutations included in the oligo are italicized, identical to -391 to-371	5'-AGAGGTG*CGT*A*TG*CGCTGGCC-3'
PANCF5'3a	*glaA* sense primer, ANCF3 mutations included in the oligo are italicized, identical to -307 to -287	5'-GACACCGCT*CAAA*CCTGGTTC-3'
PANCF3'3b	*glaA* sense primer, ANCF3 mutations included in the oligo are italicized, identical to -307 to-287	5'-GAACCAG*GTT*T*G*AGCGGTGTC-3'
PGLA5'-454-413	oligo with the identical sequence of the *glaA* promoter -454 to -413, with an underlined *Mlu*I compatible end	5'-CGCGTCCACGCACGACCGGGGACGGCGAATCCCCGGGAATTGAAAGA-3'
PGLA3'-454-143	the reverse complement of oligo PGLA5'-454-413, with an underlined *Age*I compatible end	5'-CCGGTCTTTCAATTCCCGGGGATTCGCCGTCCCCGGTCGTGCGTGGA-3'
PPKIA5'-1061	*pkiA* sense primer, an *Mlu*I site is underlined, identical to -1061 to -1041	5'-CGCGACGCGTAAATGGAAGAGAAAACCTCCG-3'
PPKIA3'-23NheI	*pkiA* sense primer, an *Nhe*I site is underlined, identical to -23 to -9	5'-CCTAGCTAGCTGATTGATCTCTACTGAACCG-3'
PPKIA5'-921	*pkiA* sense primer, an *Mlu*I site is underlined, identical to -921 to -903	5'-CGCGACGCGTAAAGGGAGACAGCTTCTCC-3'
PPKIA5'-841	*pkiA* sense primer, an *Mlu*I site is underlined, identical to -841 to -822	5'-CGCGACGCGTGTCAGGGAAATAGTTCTTGC-3'

PPKIA5'-741	*pkiA* sense primer, an *Mlu*I site is underlined, identical to -741 to -725	5'-CGCGACGCGTGATCCGCGACGATGTGG-3'
PPKIA5'-641	*pkiA* sense primer, the *Mlu*I site is underlined, identical to -641 to -623	5'-CGCGACGCGTGGGCGTGAATTGCTTTTGG-3'
PPKIA5'-541	*pkiA* sense primer, the *Mlu*I site is underlined, identical to -541 to -522	5'-CGCGACGCGTAAATGCTGCTGATCGCTGGG-3'
PPKIA5'-441	*pkiA* sense primer, the *Mlu*I site is underlined, identical to -441 to -423	5'-CGCGACGCGTTGAAAGGTGGATCACCCAG-3'
PPKIA5'-351	*pkiA* sense primer, the *Mlu*I site is underlined, identical to -351 to -334	5'-CGCGACGCGTTAGGCGATGGGAAGTGTG-3'
PPKIA5'-251	*pkiA* sense primer, the *Mlu*I site is underlined, identical to -251 to -234	5'-CGCGACGCGTCGTCCCAAAGCAGGAATC-3'
PPKIA5'-135	*pkiA* sense primer, the *Mlu*I site is underlined, identical to -135 to -121	5'-CGCGACGCGTCGGCCCCGTCTTTTC-3'
P5'AMYR	*amyR* sense primer, the *Nhe*I site is underlined, identical to +1 to +18	5'-CCTAGCTAGCATGGACTCTCATCCTTCCCCT-3'
P3'AMYR	*amyR* antisense primer, the *Fse*I site is underlined, identical to +1825 to +1845	5'-TCTTAGGCCGGCCTCAGGATTCTCCGGGACAGAA-3'

A. niger genomic DNA was isolated using a modification of the method described by Oza and Kafer (1990). Approximately 4×10^8 conidia were used to inoculate 400 mL CM medium and grown overnight at 30 °C and 150 rpm. The mycelial culture was harvested through Miracloth, rinsed with distilled water and blotted dried with Whatman® filter paper. The mycelial mass then was frozen in liquid nitrogen and ground to a fine powder. For each gram (wet weight) of mycelia ground, 5 mL of a solution of 10 mM Tris-HCl, 100 mM EDTA, pH 8.0 was added. This was followed by the addition of 20% Sarkosyl (0.5 mL/g mycelial wet weight) and 0.2 mL of RNaseA (10 mg/mL in 50 mM NaOA, pH 4.8). After an incubation at 65 °C for 30 minutes, the extract was centrifuged at room temperature (30 minutes at 23 000 × g). The supernatant was decanted into a new centrifuge tube and extracted 3 to 4 times with an equal volume of phenol:chloroform and two times with chloroform. The genomic DNA was precipitated with 0.6 volumes of isopropanol and pelleted by centrifugation (26 000 × g). The pellet was washed with 70% ethanol, air dried and dissolved in water.

2.3.3 Polymerase chain reaction amplification of DNA

PCR amplifications were done by the method described by Saiki *et al.* (1988) using either the Techne® Techgene Thermal Cycler or the Eppendorf® Mastercycler. PCR reactions were carried out in 50 µL volumes with a final dNTP concentration of 0.2 mM and 0.5 µM of each primer. The template DNA amount was 10 ng for plasmid DNA and 100 ng for genomic DNA. When using *Taq* DNA polymerase (1 U/ 50 µL reaction), each PCR reaction contained 5 µL of 10x PCR buffer [100 mM Tris-HCl (pH 8.8), 500

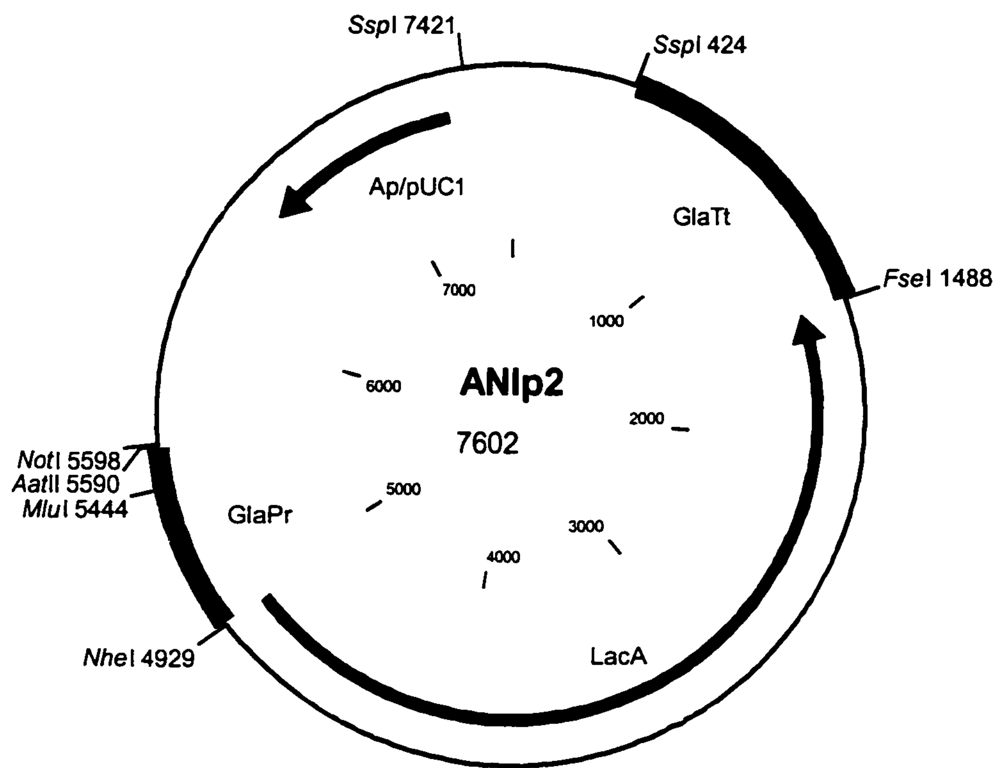
mM KCL, 0.8% Nonidet P40] and a $MgCl_2$ concentration of 1.5 mM. PCR reactions using *Pfu* DNA polymerase (1U/ 50 μL reaction) contained 5 μL of 10x PCR buffer with $MgSO_4$ [200 mM Tris-HCl (pH 8.8), 100 mM $(NH_4)_2SO_4$, 100 mM KCl, 1% Triton X-100, 1 mg/mL BSA, 20 mM $MgSO_4$]. The initial denaturation was performed over an interval of 3 minutes at 95 °C. The denaturation step was programmed at 95 °C for 30 seconds followed by primer annealing at 40 °C for 30 seconds and primer extension at 72°C for 1 minute. Depending on the amount of template DNA in the reaction mix and the expected yield of the PCR product, 25-35 cycles were usually performed. After the last cycle, the samples were usually incubated at 72 °C for 5 minutes and held at 4 °C until removed.

2.3.4 Starting plasmids

Previously constructed plasmids ANIp2, ANEp2 and ANEp4 (Dr.Hongshan Li, Concordia University) were used to construct the plasmids used in this study. ANIp2 (Fig.7) contains a 1071bp *glaA* transcription termination and poly (A) signal cassette (Gla Tt) and the *glaA* promoter cassette (Gla Pr) which includes bp –661 to bp-1 of the *glaA* gene cloned into pUC 18 (Yanisch-Perron *et al.* 1985).

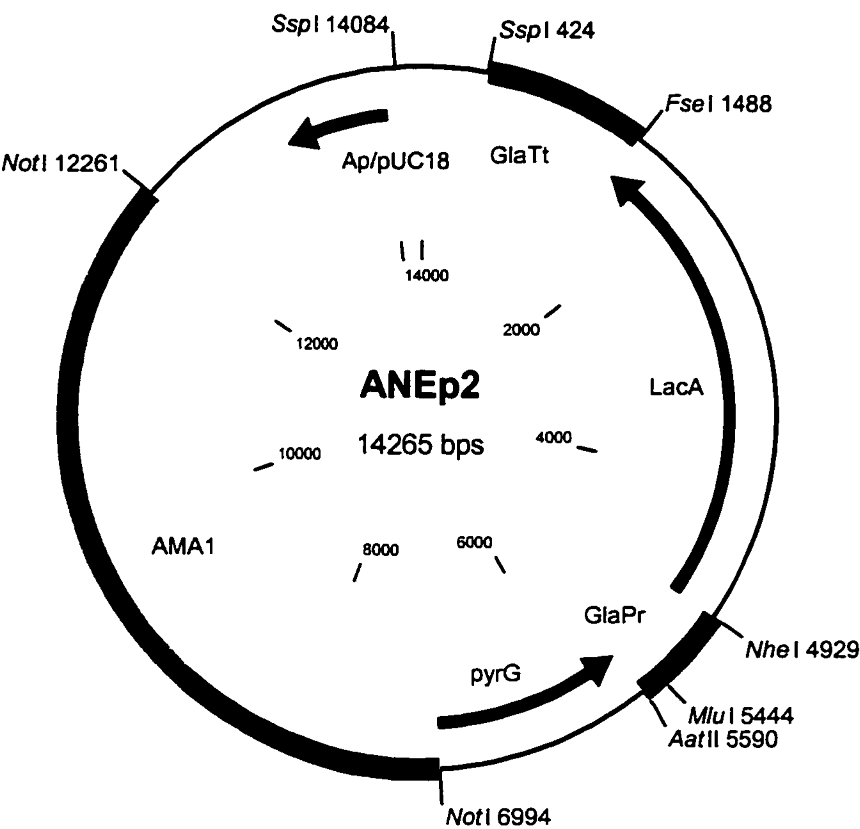
ANEp2 (Fig.8) is an autonomously replicating plasmid containing most of the previously described AMA1 replicator sequence (Gems *et al.* 1991), the *A. nidulans pyrG* gene (Oakley *et al.*, 1987) as a selectable marker and the *A.niger lacA* reporter gene cassette. The modular cassette DNAs for GlaTt, GlaPr, *pyrG* and *lacA* were prepared by PCR

Figure 7. Map of shuttle vector ANIp2. The various modules are indicated in the figure as follows, puC18 sequences (pUC), the ampicillin resistance β-lactamase gene (Ap), the glucoamylase transcription terminator (GlaTt), the glucoamylase promoter region (GlaPr), and the β-galactosidase coding region (*lacA*). The plasmid map also indicates restriction endonuclease sites.



SspI 7421
SspI 424
Ap/pUC1
GlaTt
FseI 1488
7000
1000
6000
ANIp2
7602
2000
NotI 5598
AatII 5590
MluI 5444
GlaPr
5000
3000
4000
NheI 4929
LacA

Figure 8. Map of shuttle vector ANEp2. The various modules are indicated in the figure as follows, pUC18 sequences (pUC), ampicillin resistance β-lactamase gene (Ap), the glucoamylase transcription terminator (GlaTt), the glucoamylase promoter region (GlaPr), the β-galactosidase coding region (*lacA*), the selectable marker (*pyrG*) and the sequence conferring autonomous replication in *Aspergillus* (AMA1). The plasmid map also indicates the location of several restriction endonuclease sites.



SspI 14084
SspI 424
FseI 1488
NotI 12261
Ap/pUC18
GlaTt
14000
12000
2000
ANEp2
LacA
14265 bps
10000
4000
AMA1
8000
6000
GlaPr
NheI 4929
pyrG
MluI 5444
AatII 5590
NotI 6994

amplification using oligonucleotides designed using GenBank entries KO2465, Z30918, M19132 and AR048736 respectively.

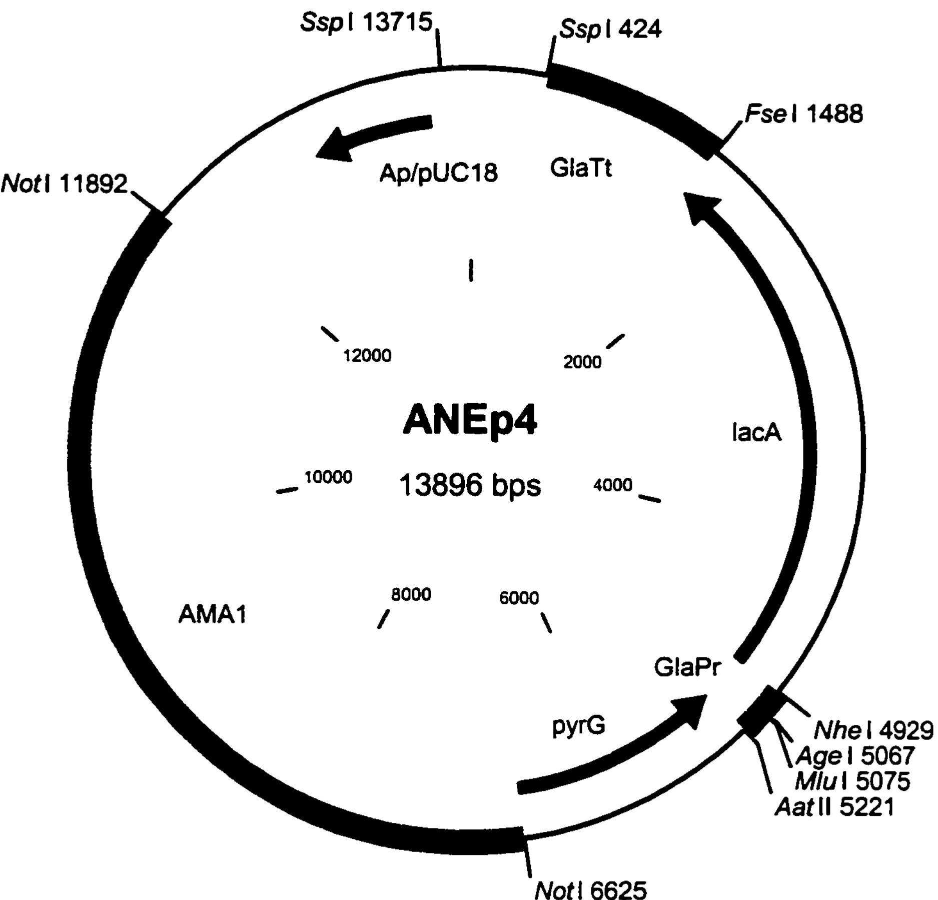
The enhancer assay vector ANEp4 (Fig. 9) was constructed by amplifying the full length *glaA* promoter cassette of ANIp2 with primers that resulted in a fragment that contained the glucoamylase promoter deleted of all functional elements except the TATA box. One of the primers introduced an *Age*I restriction site to allow directional cloning of upstream regulatory sequences (URSs) designed with *Mlu*I and *Age*I restriction sites at the 5' and 3' ends respectively.

2.3.5 The *glaA* and *pkiA* deletion sets

A *glaA* promoter deletion set had been previously constructed (Amalia Martinez-Perez, Concordia University) by cloning six different promoter fragments into the ANEp2 backbone prepared by *Mlu*I and *Nhe*I digestion and phosphatase treatment. Promoter fragments were generated by PCR with the downstream oligonucleotide PGlaPr3' in pair-wise combination with six upstream oligonucleotides having 20 nucleotide 3' ends identical to various portions of the *glaA* coding strand: PGla5'-477, PGla5'-448, PGla5'-368, PGla5'-251, PGla5'-142 and PGla5'-88 (Table 2).

The *pkiA* promoter deletions were constructed by cloning ten different promoter fragments into the ANEp2 backbone that had been digested by the restriction endonucleases *Mlu*I and *Nhe*I. Promoter fragments were generated by PCR with the common downstream oligonucleotide pkiA3'-23NheI (Table2), which contains four filler nucleotides followed by an *Nhe*I site at the 5'end. The ten upstream oligonucleotides: pkiA5'-1041, pkiA5'-921, pkiA5'-841, pkiA5'-741, pkiA5'-641, pkiA5'-541, pkiA-441,

Figure 9. Map of shuttle vector ANEp4. The various modules are indicated in the figure as follows, pUC18 sequences (pUC), ampicillin resistance β-lactamase gene (Ap), the glucoamylase transcription terminator (GlaTt), the glucoamylase promoter region (GlaPr), the β-galactosidase coding region (*lacA*), the selectable marker (*pyrG*) and the sequence conferring autonomous replication in *Aspergillus* (AMA1). The plasmid map also indicates the location of several restriction endonuclease sites.



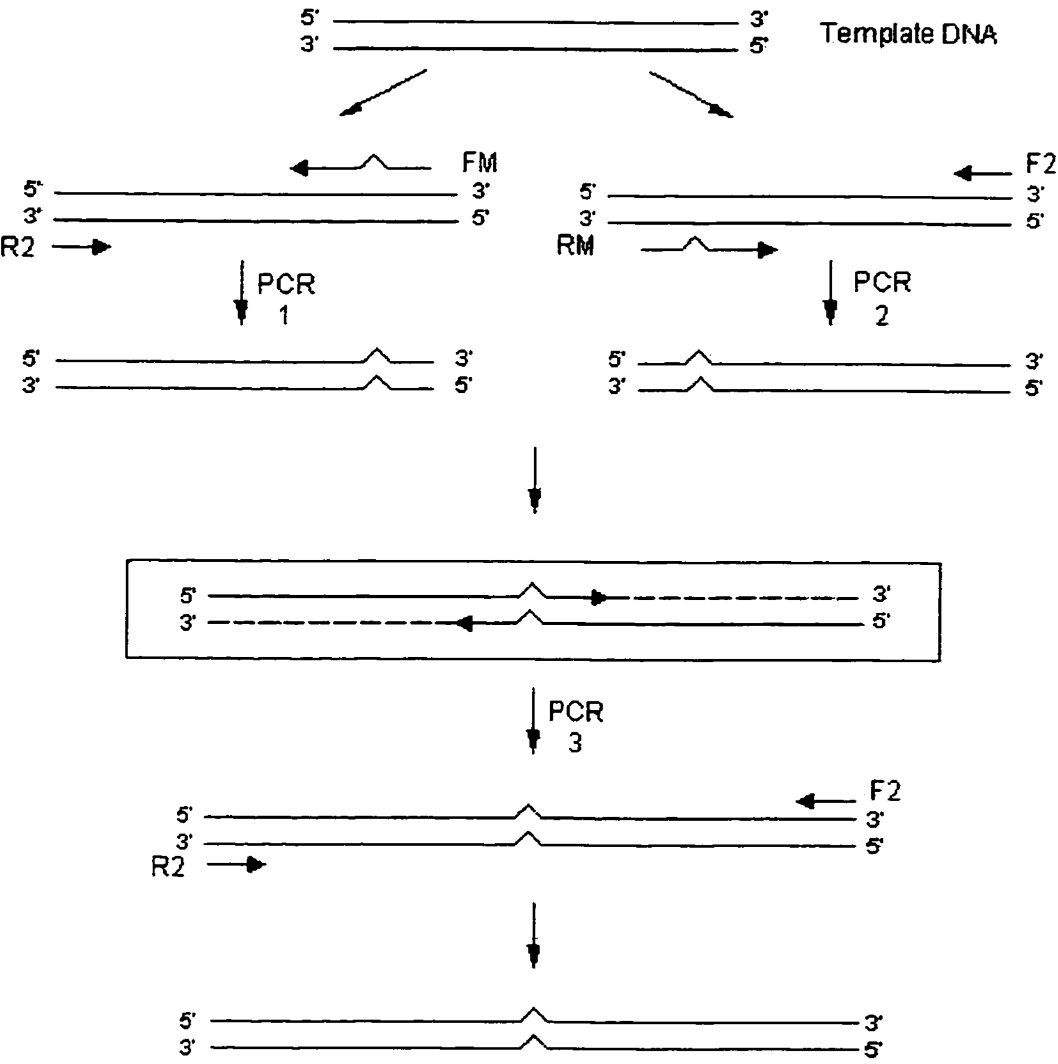
Ssp I 13715
Ssp I 424
Fse I 1488
Not I 11892
Ap/pUC18
GlaTt
12000
2000
ANEp4
lacA
10000
13896 bps
4000
AMA1
8000
6000
GlaPr
pyrG
Nhe I 4929
Age I 5067
Mlu I 5075
Aat II 5221
Not I 6625

pkiA5'-351, pkiA5'-251 and pkiA5'-135 (Table 1) have four filler nucleotides at the 5'end followed by a *Mlu*I site and a stretch of nucleotides identical to various portions of the *pkiA* coding strand. The template was genomic DNA from the *A . niger* strain N593. The ten PCR products were digested with *Mlu*I and *Nhe*I and individually ligated with the ANEp2 backbone prepared by digestion with the same endonucleases. Plasmids containing the promoter fragments were identified by restriction endonuclease mapping.

2.3.6 Site–specific mutagenesis by overlap extension

Site-specific mutations were introduced into the *glaA* promoter by overlap extension (Higuchi *et al.*, 1988; Ho *et al.*, 1989). Glucoamylase promoter derivatives were generated having base substitutions in regions closely matching the ANCF binding consensus (5'-RRCCAATMRCR-3') and the CreA binding consensus (5'-SYGGRG-3'). Four primers are used to introduce a site-specific mutation by overlap extension (Fig.10). This method requires two mutagenic primers, two flanking oligonucleotides and three PCR reactions to generate a mutation. One pair of primers is used to amplify the DNA that contains the mutation and upstream sequences, whereas the second pair of primers is used to amplify the DNA that contains the mutation with downstream sequences. The two sets of primers are used in two separate amplification reactions to amplify overlapping DNA fragments. The mutations are located in the region of overlap and are in both sets of amplified fragments. The overlapping fragments are mixed, denatured and annealed to generate heteroduplexes that are extended to produce a fragment that is the sum of the two overlapping products. A third PCR reaction using

Figure 10. Site-directed mutagenesis by overlap extension. Two fragments of the target sequence are amplified in two separate PCR reactions, 1 and 2. PCR 1 uses the forward primer (FM) containing the introduced mutation and the reverse primer containing the wild-type sequence. PCR 2 uses the reverse primer (RM) containing the introduced mutation and the forward primer containing the wild-type sequence. The boxed portion of the figure shows how the overlapping fragments, resulting from the first and second rounds of amplification, anneal and generate heteroduplexes that are extended to produce a fragment that is the sum of the two overlapping products. PCR 3 using primers R2 and F2 results in full-length mutant DNA (Modified from Sambrook *et al.*, 2001)


5'
3'
3'
5'
Template DNA
FM
5'
3'
3'
5'
R2
F2
5'
3'
3'
5'
RM
PCR
1
PCR
2
5'
3'
3'
5'
5'
3'
3'
5'
5'
3'
3'
5'
PCR
3
F2
5'
3'
3'
5'
R2
5'
3'
3'
5'

two primers that bind to the extremities of the two initial fragments results in the enrichment of a full-length DNA containing the mutations. Wild-type and mutagenized promoter regions were all generated using ANIp1 as template DNA. The two flanking oligonucleotides that matched the wild-type sequence GlauMluI (5'-ATGCCATTGGCGGAGGGGTC-3') and GlaAgeI (5'-CTGACCGGTGAAGCACATGCATC-3') included *Mlu*I and *Age*I restriction sites in the 5' region of the primers to facilitate cloning of the mutated segment of DNA.

2.3.7 CreA binding consensus mutations

Four sites matching the CreA binding consensus were identified in the *glaA* promoter. The sites CreA1 5'-CCGGGG-3', CreA2 5'-CCGGGG-3', CreA3 5'-CCGGGG-3' and CreA4 5'-CTGGGG-3' (Fig.5) were changed to (where lower-case letters indicate introduced mutations) 5'-CatacG –3', 5'-CgtatG-3', 5'-CgtatG-3' and 5'-CgtatG-3' respectively by using overlapping mutagenic primers (Table 2) in the method described above. Promoter regions containing altered CreA binding consensus sites were digested with *Mlu*I and *Age*I and cloned into the backbone ANEp4 (also digested by *Mlu*I *and Age*I) to generate eight DNA plasmids: pCreA-1 (CreA1 site mutated), pCreA-2 (CreA2 site mutated), pCreA-3 (CreA3 site mutated), pCreA-4 (CreA4 site mutated), pCreA-1&2 (CreA1 and CreA2 sites mutated), pCreA-1&3 (CreA1 and CreA3 sites mutated), pCreA-2&3 (CreA2 and CreA3 sites mutated), pCreA-1&2&3 (CreA1, CreA2 and CreA3 sites mutated).

2.3.8 AnCF-element mutants

The *glaA* promoter region contains three sites that closely match the AnCF binding consensus sequence. The sites (Fig.4) AnCF-1 (5'-AGCCAATCAAG-3'), AnCF-2 (5'-GGCCAATCGCT-3'), and AnCF-3 (5'-AGCCTATAGCG-3') were changed to 5'-AGCatAcgAAG-3', 5'-GcgtAtgCGCT-3' and 5'-cGCtcAaAcCt-3' respectively. Promoter regions with altered AnCF consensus binding sites were digested with *Mlu*I and *Age*I and cloned into ANEp4 as described above. The four DNA plasmids constructed were designated pAnCF-1 (upstream AnCF element mutated) pAnCF-2 (middle AnCF element mutated), pAnCF-3 (downstream AnCF element mutated) and pAnCF-1&2 (upstream and middle AnCF elements mutated).

Plasmids harbouring *glaA* promoter regions with CreA and AnCF binding consensus mutations were identified by *Nhe*I/*Sac*I digestion and verified by DNA sequencing analysis.

2.3.9 Cloning of regulatory regions

One to five copies of the upstream regulatory region (URR) between base pairs –477 and –368, designated URR1-ANEp4, URR2- ANEp4, URR3-ANEp4, URR4-ANEp4, URR5-ANEp4, were previously cloned into the ANEp4 *Mlu*I site (A. Martinez-Perez, Concordia University). A pair of oligonucleotides (with 5'end *MluI* sites) of which one was identical to the sense strand nucleotides –477 to -456 and the other identical to the anti-sense strand nucleotides –387 to -368, were used to amplify the

upstream regulatory region of ANEp2. The region was then cut with *Mlu*I and cloned into the *Mlu*I site of ANEp4.

One copy of the region between base pairs –454 and -413 of the upstream regulatory region was cloned into ANEp4. Equal molar concentrations of two complementary oligonucleotides representing the region were annealed following the phosphorylation of the 5'end of the oligonucleotides by T4 polynucleotide kinase. The annealed pair of oligonucleotides (flanked by *Mlu*I and *Age*I compatible ends) was directionally cloned into the *Mlu*I and *Age*I digested ANEp4 backbone. Plasmids containing the region between –454 and –413 were identified by PCR and verified by DNA sequencing analysis.

2.3.10 Cloning of the *amyR* gene

The *A. niger amyR* gene was PCR amplified from genomic DNA (strain N593) using oligonucleotides designed from the Genbank entry AF155808.

The primer 5'- CCTAGCTAGCATGGACTCTCATCCTTCCCT-3' contains four filler nucleotides at the 5'end followed by an *Nhe*I site. The other primer 5'-TCTTAGGCCGGCCTCAGGATTCTCCGGGACAGAA-3' contains five filler nucleotides at the 5'end followed by an *Fse*I site. The PCR product (1849 bp) was cut with *Nhe*I and *Fse*I and ligated into pPkiA-1041 (an ANEp2 plasmid that had the *glaA* promoter replaced by the full length *pkiA* promoter), which was also cut with the same restriction endonucleases. Correct constructs were identified by *Bgl*II and *Nhe*I, *Fse*I digestion.

2.4 Plasmid Stability in *A. niger*

Plasmid stability was assessed by following plasmid loss during nonselective growth. Conidia from pyrG^{+} transformants were transferred to MM plates with uridine and allowed to germinate and grow until an even lawn of conidia appeared. Conidia were harvested and counted. Approximately 200 conidia were spread onto nonselective MM plates containing uridine. Conidia from colonies then were streaked onto selective plates. Plasmid stability was estimated by the proportion of colonies that retained the plasmid during nonselective growth.

2.5 Enzyme assays

Two independent constructions of each plasmid were individually transformed into the *A. niger* strain N593. Three transformants from each transformation experiment were randomly selected and grown in duplicate cultures. Media from the cultures were collected by filtration through Miracloth. Cultures grown in flasks with marbles also required centrifugation (8000 X g at 4 °C for 20 min.) of the supernatant fluid to remove fragmented mycelia.

2.5.1. Secreted β-galactosidase assays

Secreted β-galactosidase assays were performed by adding 2 μL of the medium described above to 100 μL of reaction buffer [50mM sodium acetate (pH 4.1), 2 mg/mL

of ortho-nitrophenyl-β-D-galactosidase (ONPG)] and incubating for 10 minutes at 37 °C. Assays were stopped by adding nine volumes of 0.1 M sodium carbonate. Units of β-galactosidase activity per mL of culture (where one unit of enzyme hydrolyses 1 μmole of ONPG to ortho-nitrophenol per minute at 37 °C) were calculated as follows $(\Delta A_{410}/min.)(Vr)(D)/(Ve)(E_{410})$ where: $\Delta A_{410}/min.$ = change in absorbance per minute, Vr = assay volume, D = dilution factor, Ve = enzyme volume and E_{410}= μmolar extinction coefficient of ortho-nitrophenol (3.5×10^3 $\mu M^{-1}cm^{-1}$).

2.5.2 Secreted glucoamylase assays

An 0.2 mL sample of medium was mixed with 1.8 mL of 2% soluble starch (Sigma Chemical) in 0.1 M citrate buffer (pH 4.4) and incubated at 60 °C for 2 hours. The reaction was stopped by adding 10 μL of 0.5 M EDTA and boiling the reaction mixture at 100 °C for 15 minutes. The solution was assayed for glucose by the Sigma glucose kit, a procedure that is essentially that of Raabo and Terkildsen (1960) with a minor change in the quantity of chromogen to increase sensitivity. The assay required the addition of the sample to a mixture containing glucose oxidase, peroxidase and o-dianisidine. The reaction was allowed to proceed to completion in approximately 30 minutes at 37 °C. The final color intensity is proportional to the glucose concentration. Glucoamylase activity was measured as glucose forming activity (GFA, U/μL) from soluble starch. One unit of GFA is defined as 1-μmol glucose released/minute at 60°C (Mackenzie *et al.*, 2000).

3. Results

3.1 Comparison of the *glaA-lacA* chimera of ANEp2 with the native *glaA* gene of *A.niger*

Transcription initiation sites of the *glaA-lacA* chimera of ANEp2 were mapped by primer extension by Dr. Hongshan Li (Concordia University). It was determined that transcription initiates from three major sites at positions –73, -68, –54 and several minor sites. In a similar study, Nunberg *et al.* (1984) had mapped the 5' end of the *A. awamori* glucoamylase mRNA to several locations within a region –73 to –52 bp from the translational start site.

Primer extension and Northern analysis were also used to examine *glaA-lacA* expression levels in media containing different carbon sources. Fowler *et al.* (1990) had previously reported that expression of *glaA* is induced by growth on maltose, intermediate in the presence of glucose and undetectable when xylose is supplied as the sole carbon source. In this study, it was first established that *glaA-lacA* expression is specific to transformants harbouring the *glaA-lacA* chimera of ANEp2. Second, the results showed that the level of *glaA-lacA* mRNA is about 15-fold higher in the presence of maltose as compared to glucose. Third, expression of *glaA-lacA*, in the presence of xylose, was undetectable by either Northern blot analysis or primer extension. Therefore, it can be concluded that the expression of the *glaA-lacA* chimera of ANEp2 correlates well with previously established characteristics of the native *Aspergillus glaA* gene.

3.2 Deletion analysis of the *glaA* promoter region

Glucoamylase expression is modulated by growth on different carbon sources. It had been hypothesized by Verdoes *et al.* (1994) that the *glaA* gene is regulated by a complicated pathway, involving at least a specific activator protein and a repressor protein such as CreA.

In an attempt to identify the regulatory elements that control *glaA* expression, a *glaA* promoter deletion set was constructed. Six plasmids: Gla-477 → Gla-88, each containing a different promoter fragment, were used to direct the expression of the *lacA* reporter gene in medium containing various carbon sources. The ability of the deletion set to express β-galactosidase was measured and compared to the parent plasmid that harbours a 657 bp 5'-flanking region of *glaA* (Fig.11). The results indicate that two regions, one between positions –477 and –448 and the other between positions –448 and –368, contain sequences important for *glaA* expression. Deleting the region between position –477 and –448 reduced β-galactosidase expression, in maltose, from 3.7 units to 0.9 units. A further deletion up to position –368 greatly reduced expression of the reporter gene to 0.02 units. The deletion of sequences upstream of position –251 resulted in a complete loss of expression in both maltose and glucose. Although deleting sequences upstream of position –448 resulted in a lower level of *lacA* expression in maltose and glucose than the parent plasmid, transformants harbouring plasmids with sequences upstream of position –448 deleted still had an increased expression level for maltose grown cells over the level found for glucose grown cells. This suggests that sequences upstream of position –448,

Fig. 11. Deletion analysis of the *A. niger glaA* promoter region. Gla-657 is ANEp2 and harbours the upstream regulatory region out to bp –657. Deletion derivatives of ANEp2 (Gla-477 through Gla-88), each containing a varying amount of deleted upstream information, were used to direct the expression of the *lacA* reporter gene in media containing various carbon sources. A vertical line at the left end of all promoter regions indicates a unique *Mlu*I site. For each promoter class, three independently constructed plasmids were transformed into the *A. niger* strain N593 as described in Materials and Methods. Two independent transformants of each independently constructed plasmid were grown in duplicate in 5 mL liquid cultures at 30°C for six days. Units of β-galactosidase per mL of culture are indicated on the right. Values next to units of activity indicate standard deviation.

Promoter region	Units of activity		
	maltose	glucose	xylose
Gla-657 (TATA, ATG)	4.3 +/- 1.0	1.0 +/- 0.5	0.00
Gla-477 (ATG)	3.7 +/- 0.6	1.1 +/- 0.4	0.00
Gla-448 (ATG)	0.9 +/- 0.3	0.05 +/- 0.3	0.00
Gla-368 (ATG)	0.02 +/- 0.03	0.02 +/- 0.1	0.00
Gla-251 (ATG)	0 .00	0 .00	0 .00
Gla-142 (ATG)	0 .00	0 .00	0 .00
Gla-88 (ATG)	0 .00	0 .00	0 .00
Gla-142+AgeI (ATG)	0 .00	0 .00	0 .00

although required for high-level expression, do not play an important role in maltose induction and/or glucose repression.

3.3 Analysis of putative target sites for regulatory proteins in *glaA*

Analysis of all putative target sites for regulatory proteins in *glaA* involved the use of the UAS/enhancer test plasmid ANEp4. ANEp4 is a vector that harbours the TATA element and the transcription initiation region but lacks upstream sequences between bps –519 and –142 that were shown by deletion analysis to be required for expression of the *lacA* reporter gene.

Deletion analysis of the *glaA* promoter region suggests that sequences downstream of position –477 include all putative target sites for regulatory proteins in *glaA*. This study focused on this region by identifying sites matching AnCF and CreA binding consensus sequences and mutating those sites by overlap extension using PCR.

3.3.1 Analysis of putative AnCF binding sites

Three sites that closely match the AnCF binding consensus sequence (5'-RRCCAATMRCR-3') were identified within the region downstream of position –477. The three sites designated as AnCF-1, AnCF-2 and AnCF-3 are located at positions –468 to –458 on the template strand, -387 to –377 on the nontemplate strand and –312 to –292 on the nontemplate strand respectively. To determine if these sites played a role in *glaA* expression, β-galactosidase expression levels directed by mutant promoter regions were

compared to expression levels generated by the wild-type *glaA* promoter. β-galactosidase expression levels of transformants harbouring plasmids containing the mutant and wild type forms of the *glaA* promoter regions are shown in Fig.12. Altering the most upstream AnCF binding consensus site resulted in an approximate 65% reduction in expression of the reporter gene when maltose was used as the sole carbon source. Altering the middle AnCF-binding consensus site reduced expression in maltose by about 60%. However, altering the AnCF-3 site resulted in an expression level similar to that of the wild type promoter. Although changing either AnCF-1 or AnCF-2 resulted in a reduced level of reporter gene expression in transformants grown in either maltose or glucose, both types of mutations were still expressed at high levels in maltose as compared to glucose or xylose. Simultaneous mutations in both the most upstream and middle AnCF-binding consensus site reduced the expression level of the reporter gene in maltose by at least 90% relative to the amount expressed when the sites are individuallyaltered. Changing any of the AnCF-binding consensus sites also did not alter gene expression in xylose suggesting that none of the sites are required for xylose repression of *glaA*. Therefore, these results show that changing either AnCF-1 or AnCF-2 leads to a decreased level of gene expression and reveals that the integrity of at least one site is necessary for maltose induction. It can also be inferred that simultaneously changing AnCF-1&3 or AnCF-2&3 would result in an expression profile similar to what is produced by an AnCF-1 or AnCF-2 mutation respectively.

Figure 12. Mutation analysis of putative AnCF binding sites. Derivatives of ANEp4 containing a variety of wild-type (filled boxes) and altered (Xs) AnCF binding consensus sites were constructed by inserting fragments into *Mlu*I and *Age*I digested ANEp4 backbone. A vertical line at the left end of the promoter regions represents a unique *Mlu*I site. For each promoter class, three independently constructed plasmids were transformed into the *A. niger* strain N593 as described in Materials and Methods. Two independent transformants of each independently constructed plasmid were grown in liquid mycelial cultures of 200 μL MM cultures in 96 well plates for 72 hours. Units of β-galactosidase per mL of culture are indicated on the right. Values next to units of activity indicate standard deviation.

Promoter region **Units of activity**

ATG

*Age*I site

TATA

Construct	maltose	glucose	xylose
wild-type	5.6 +/- 1.3	1.1 +/- 0.5	0. 00
ANCF-1	2.0 +/- 0.6	0.15 +/- 0.03	0. 00
ANCF-2	2.2 +/- 0.9	0.16 +/- 0.1	0. 00
ANCF-3	4.1 +/- 1.5	1.0 +/- 0.4	0. 00
ANCF-1&2	0.11 +/- 0.5	0.11 +/- 0.1	0. 00
ANCF-2&3	2.2 +/- 0.9	0.16 +/- 0.1	0. 00
ANCF-1&3	2.0 +/- 0.6	0.16 +/- 0.03	0. 00
ANCF-1,2&3	0.11 +/- 0.5	0.11 +/- 0.1	0. 00

3.3.2 Analysis of putative CreA binding sites

Four sites matching the CreA binding consensus site (5'-SYGGRG-3') were identified within the region downstream of position –477 of the *glaA* promoter. They were designated CreA1 at position –444 to –439 of the template strand, CreA2 at position –329 to –324 of the non-template strand, CreA3 at position –321 to –316 of the non-template strand, and CreA4 at position –324 to –319 of the non-template strand. The regulatory role of these sites was assessed by changing them individually and in various combinations. Beta-galactosidase expression levels of transformants harbouring plasmids with various CreA mutations were compared to expression levels generated by the wild type *glaA* promoter (Fig.13). Surprisingly, altering either the CreA1 or CreA2 site individually resulted in a complete loss of gene expression when either maltose or glucose was the carbon source. However, altering either of the two more downstream sites, CreA3 and CreA4, did not alter gene expression. Promoter regions with multiple CreA consensus sites changed also resulted in a loss of gene expression if one of the altered sites was either CreA1 or CreA2. Since sites CreA1 and CreA2 matched the CreA repressor binding consensus sequence, it was expected that changing these sites would alleviate glucose repression. However, the results evidently indicate that a transcription factor necessary for *glaA* expression interacts with these two sites.

Figure 13. Mutation analysis of putative CreA binding sites. Derivatives of ANEp4 containing a variety of wild-type (blue triangles) and altered (yellow triangles) CreA binding consensus sites were constructed by inserting fragments into *Mlu*I and *Age*I digested ANEp4 backbone. A vertical line at the left end of the promoter regions represents a unique *Mlu*I site. For each promoter class, three independently constructed plasmids were transformed into the *A. niger* strain N593 as described in Materials and Methods. Two independent transformants of each independently constructed plasmid were grown in liquid mycelial cultures of 200 μL MM cultures in 96 well plates for 72 hours. Units of β-galactosidase per mL of culture are indicated on the right. Values next to units of activity indicate standard deviation.

Promoter region	Units of activity		
	maltose	glucose	xylose
wild-type (ATG, *Age*I site)	5.6 +/- 1.3	1.1 +/- 0.5	0.00
Cre-1 (ATG)	0.00	0.00	0.00
Cre-2 (ATG)	0.00	0.00	0.00
Cre-3 (ATG)	5.9 +/- 1.0	1.4 +/- 0.4	0.00
Cre-4 (ATG)	5.2 +/- 1.9	0.9 +/- 0.7	0.00
Cre-1&2 (ATG)	0.00	0.00	0.00
Cre-1&3 (ATG)	0.00	0.00	0.00
Cre-2&3 (ATG)	0.00	0.00	0.00
Cre1,2&3 (ATG)	0.00	0.00	0.00

3.3.3 Electrophoretic mobility shift assays with the *glaA* promoter region

Electrophoretic mobility shift assays done by Catherine Au (MSc thesis, 2002) used 227 bp DNA fragments generated by PCR amplification of plasmids produced in this study. The DNA fragments generated were the result of the amplification of the *glaA* region between –501 and –274 of plasmids (described above) containing the wild-type *glaA* sequence and CreA mutations in various combinations.

Electrophoretic mobility shift assays using purified GST:CreA protein and fragments confirmed to have the correct sequence (Table 3) resulted in the production of an intense shifted band when a promoter fragment contained either the wild-type sequence or a CreA3 site mutation. One faint band was visible when the promoter fragment containing a CreA1 site mutation was used. However, no band was observed for fragments containing a CreA2 site mutation, a CreA1 and CreA2 double mutation or a CreA1/CreA2/CreA3 triple mutation. The DNA fragment C2/3-2a was found to contain a point mutation outside of any CreA consensus site (at position –240) in addition to a CreA2 site and CreA3 site mutation. It produced, like the other double and triple CreA mutants, no visible shifted band. These results suggest that CreA sites 1 and 2 are important for CreA binding.

Table 3. DNA sequencing results of plasmids containing wild-type and mutated *glaA* promoter regions

Plasmid designation	Description of sequence inserted into ANEp4	Divergence from the expected sequence
D9	wild-type *glaA* promoter region (-519 to -142)	none
D10	wild-type *glaA* promoter region (-519 to -142)	none
oligo6	wild-type *glaA* promoter region (-454 to -413)	none
oligo8	wild-type *glaA* promoter region (-454 to -413)	none
C1a	CreA1 mutation	1 point mutation[b] at -453
C1c	CreA1 mutation	none
C1f	CreA1 mutation	none
C2a[a]	CreA2 mutation	no mutation present
C2b	CreA2 mutation	none
C2e[f]	CreA2 mutation	1 point mutation at -414
C3e	CreA3 mutation	none
C3g	CreA3 mutation	none
C4a	CreA4 mutation	none
C4c[e]	CreA4 mutation	1 point mutation at -67
C1/2e1	CreA1 &CreA2 double mutation	none
C1/2(3)	CreA1 &CreA2 double mutation	none
C1/2(6)	CreA1 &CreA2 double mutation	none
C2/3-2a[f]	CreA2 & CreA3 double mutation	1 point mutation at -240
C2/3-2b[d]	CreA2 & CreA3 double mutation	-
C1/3a[f]	CreA1 &CreA3 double mutation	1 point mutation at -453
C1/3c[c]	CreA1 &CreA3 double mutation	-
C1/3d	CreA1 &CreA3 double mutation	none
C1/2e4[f]	CreA1&CreA2&CreA3 triple mutation	1 point mutation at - 336
C1/2g4	CreA1&CreA2&CreA3 triple mutation	none
H1-3g	AnCF-1 mutation	none
H1-3e	AnCF-1 mutation	1 point mutation at -452
H2-4e	AnCF-2 mutation	none
H2-3e	AnCF-2 mutation	also contains a CreA3 mutation
H1/2-2[d]	AncF-1&2 mutation	-
H1/2-2[d]	AncF-1&2 mutation	-
H3b	AnCF-3 mutation	none
H3-c5[e]	AnCF-3 mutation	1 point mutation at -374
H3-a4[e]	AnCF-3 mutation	I point mutation at -371

[a] not used to calculate expression level

[b] all unforeseen point mutations lie outside of CreA & AnCF consensus binding sites

[c] poor sequencing result but transformants harbouring the plasmid behave identically to those containing C1/3d and are unable to express the reporter gene

[d] poor sequencing result but at least two other independent constructs of its type show identical expression levels

[e] expression identical to wild-type promoter despite addition point mutation

[f] reporter gene is still not expressed despite the additional point mutation

3.4 Expression analysis of transformants harbouring various copies and portions of the upstream regulatory region

The region between positions –477 and –368 was first implicated by deletion analysis to be important for gene expression. It contains two putative CreA-binding sites (CreA1 & CreA2) and two putative AnCF-binding sites (AnCF-1 & AnCF-2) which upon site directed mutagenesis, were found to play a significant role in expression. One to five copies of the upstream regulatory region (URR) were cloned into the ANEp4 *Mlu*I site. Beta-galactosidase expression levels of transformants containing the ANEp4 derivative plasmids showed that this region is able to activate transcription (Table 4). Furthermore, increasing the copy number of the upstream regulatory region from zero to three leads to an increase in expression levels. The expression level increase from undetectable to 0.17 units to 1.25 units to 2.5 units as the copy number of upstream regulatory regions increased from zero to three is suggestive of cooperative binding of a transcription factor(s) to this region when multiple copies of the URR are present.

Transcription activation activity of a smaller region of the upstream regulatory region between position –454 and –413 was also examined. This region, containing a smaller portion of the URR, harboured the two putative CreA-binding sites (CreA1 & CreA2) and not the two surrounding putative AnCF elements (AnCF-1 & AnCF-2). Having shown by site-directed mutagenesis that the integrity of both CreA1 and CreA2 are necessary for gene expression, it was of interest to determine if this region was also sufficient to drive expression of the reporter gene. However, cloning one copy of the

Table 4
Expression analysis of transformants containing various copies of the URR

Number of URR [a] cloned into ANEp4	β-galactosidase expression [b]
0	0. 00
1	0.17 (+/- 0.08)
2	1.25 (+/- 0.4)
3	2.5 (+/- 0.5)
4	2.1 (+/- 0.7)
5	1.8 (+/- 0.2)

[a] The *glaA* upstream regulatory region between position -477 and -368 (URR) was amplified and cloned into the unique *Mlu*I site of ANep4 in various copies as indicated the Materials and Methods.

[b] β-galactosidase activity was determined as described in Materials and Methods. Transformants were grown in medium containing maltose as the sole carbon source

region between position –454 and –413 did not result in expression of the reporter gene in medium containing maltose, glucose or xylose as the carbon source. This result indicates that although the integrity of this region is necessary, it alone is not sufficient for gene expression.

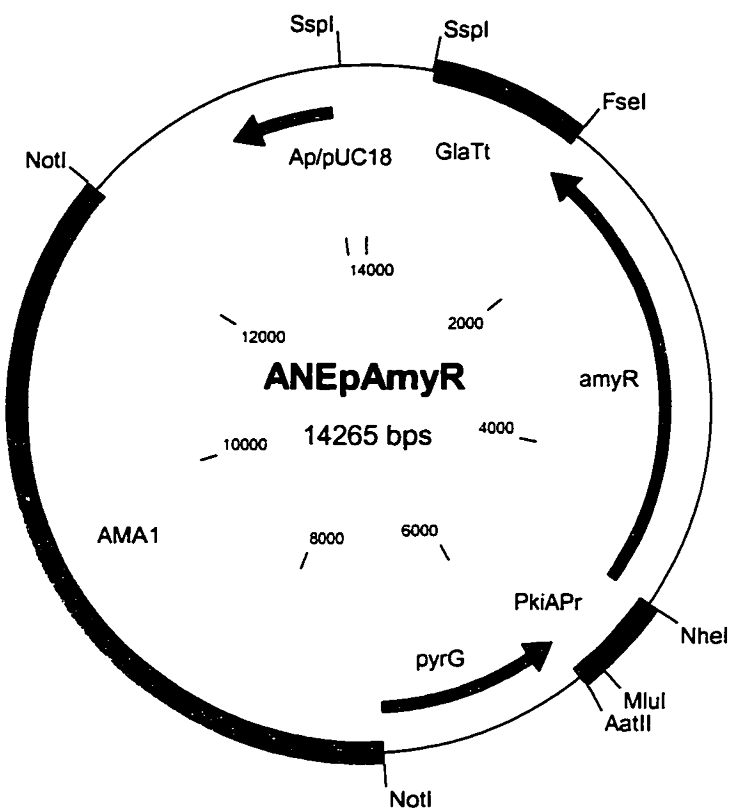
3.5 The effect of AmyR on glucoamylase production

The ANEpAmyR plasmid (Fig.14) was constructed by PCR amplification of the *A. niger amyR* gene from genomic DNA and the cloning of it downstream of the full length *pkiA* promoter of pPkiA-1041 (refer to Materials and Methods for more details). The pyruvate kinase gene (*pkiA*) is a constitutively expressed gene that demonstrates a twenty fold increase in enzyme activity when grown on a glycolytic carbon source such as glucose as compared to the level found when a gluconeogenic carbon source is used (Uitzetter *et al.*, 1982). The *amyR* gene is differentially expressed by growth on various carbon sources (Tani *et al.*, 2001). The promoter region of the *A. nidulans amyR* gene has five putative CreA binding sites. Similar to the *glaA* gene, the *amyR* gene is highly expressed by growth on starch or maltose and repressed when grown on glucose as the sole carbon source.

It was, therefore, of interest to investigate if direct repression of the activator gene *amyR* by CreA contributes to the glucose repression demonstrated by the *glaA* gene. In order to do so, the ANEpAmyR plasmid was transformed into the *A. niger* strain N593 (*cspA1 pyrG6*). Transformants harbouring the plasmid were grown in MM media

containing either glucose or xylose as the sole carbon source and assayed for secreted glucoamylase activity (as described in Materials and Methods). Two independent

Figure 14. Map of the ANEpAmyR shuttle vector. The various modules are indicated in the figure as follows, pUC18 sequences (pUC), ampicillin resistance β-lactamase gene (Ap), the glucoamylase transcription terminator (GlaTt), the pyruvate kinase promoter region (PkiAPr), the *A. niger* amyR gene, the selectable marker (*pyrG*) and the sequence conferring autonomous replication in *Aspergillus* (AMA1). The plasmid map also indicates the location of several restriction endonuclease sites.


SspI
SspI
FseI
NotI
Ap/pUC18
GlaTt
14000
12000
2000
ANEpAmyR
amyR
14265 bps
10000
4000
AMA1
8000
6000
PkiAPr
NheI
pyrG
MluI
AatII
NotI

plasmid constructions were transformed into the strain of which three randomly selected transformants were grown in duplicate. Transformants were first grown in medium having glucose as the only carbon source. Under these conditions, the *pkiA* promoter should drive expression of the usually repressed *amyR* gene. Secreted glucoamylase activity was measured as glucose forming activity (GFA) per µL of culture supernatant from soluble starch. It was, therefore, necessary to measure the amount of glucose in each sample of culture supernatant used to perform the assay and subtract it from the amount of glucose measured following the reaction described in Materials and Methods. Assaying strain N593 (*cspA1pyrG6*) following growth in glucose resulted in an average GFA of 243.0 µmol glucose/min and a glucoamylase activity (GA) of 1.22 gfa/µL. Transformants harbouring ANEpAmyR had an average GFA of 538.5 µmol glucose/min and glucoamylase activity of 2.69 gfa/µL (Table 5). Despite the greater than 2-fold glucoamylase activity of the transformants, it was a concern that the presence of glucose in the culture supernatant produced a background level too high to get a clear indication of the plasmid's effect on glucoamylase production. Transformants were, therefore, grown in medium containing xylose as the sole carbon source and assayed for glucoamylase activity. The average glucoamylase activity of three transformants from each independent plasmid construction is shown in Table 5. Glucoamylase activity is undetectable in culture supernatants from strain N593 grown in xylose. Transformants harbouring ANEpAmyR had an average glucose forming activity of 44.4 µmol glucose/min and a glucoamylase activity of 0.22 gfa/µL.

While developing and conducting the reproducibility tests of the glucoamylase assay (results not shown) it was first noticed that one transformant designated 1.2

Table 5.

Glucoamylase activity of strain N593[a] and ANEpAmyR transformants of N593[b]

Sample	Carbon source of medium	GFA[c] (μmol glucose/min)	GA[d](gfa/μL)
Strain N593	glucose	243.0 (+/-103.5)	1.29 (+/-0.5)
ANEpAmyR transformants	glucose	538.5 (+/-77.0)	2.69 (+/-0.4)
Strain N593	xylose	0. 00	0. 00
ANEpAmyR transformants[e]	xylose	44.4 (+/- 8.7)	0.22 (+/- 0.04)

[a] The *A. niger* strain N593 (*cspA1 pyrG6*)

[b] Two independent constructions of the ANEpAmyR plasmid were transformed into strain N593 of which three randomly selected transformants were grown in duplicate.

[c] One unit of glucose forming activity (GFA) is defined as 1μmol glucose released/minute at 60°C

[d] Glucoamylase activity(GA) is measured as gfa/μL

[e] Excluding transformant 1.2 that has an expression level significantly higher than all other transformants tested [GFA of 129.2 (+/- 14.7) and GA of 0.65 (+/- 0.07)]

consistently demonstrated higher glucoamylase activity when grown in either glucose or xylose. Plasmid stability tests were, therefore, done to determine if the difference in activity between the one transformant and all others assayed was due to plasmid integration into the chromosomal DNA. However, all transformants behaved similar to autonomously replicating plasmids. Under nonselective growth conditions, all transformants were unstable and could be readily cured of their plasmids. More tests such as determining the plasmid copy number of each transformant would have to be done.

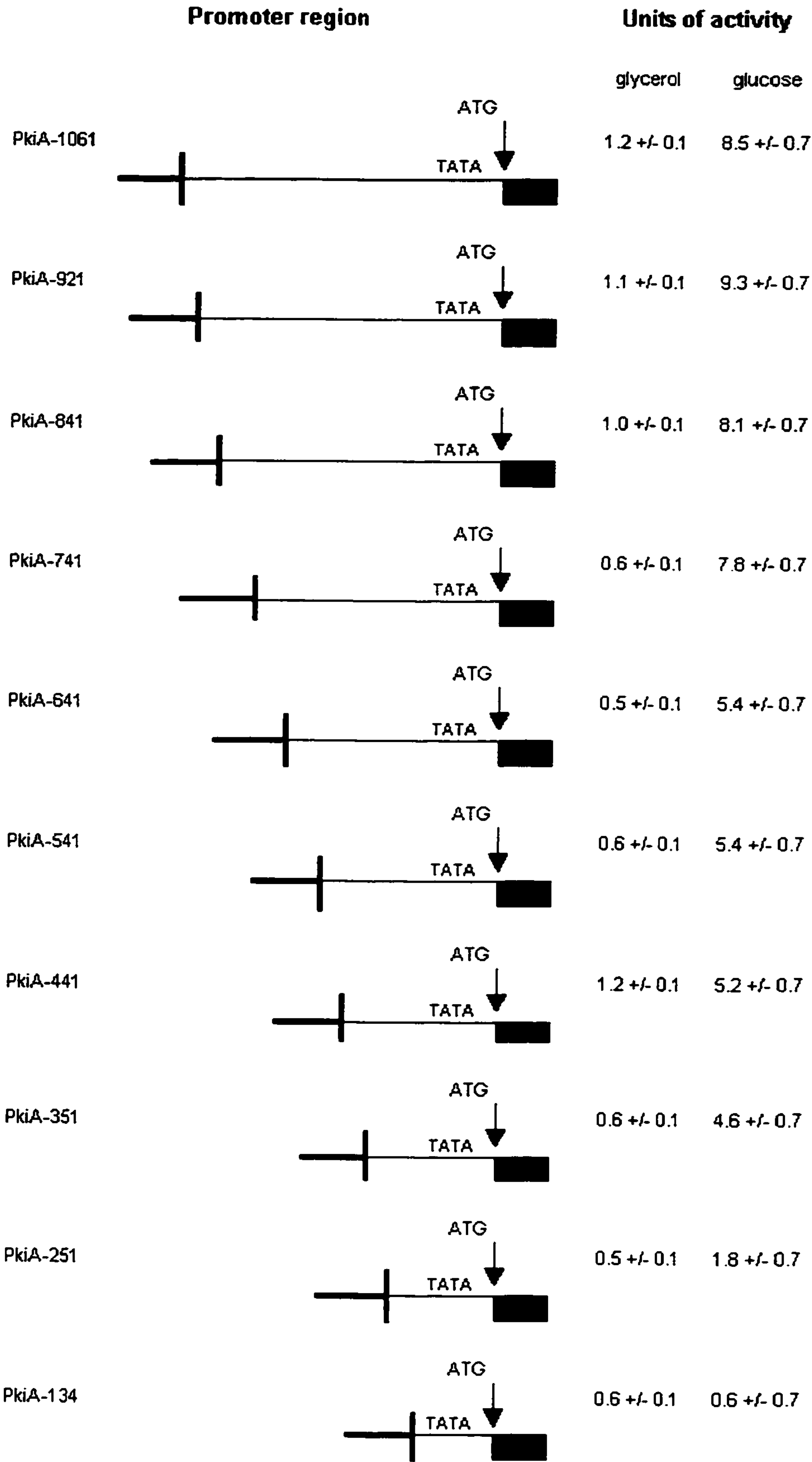
3.6 Deletion analysis of the *A. niger pkiA* promoter region

The pyruvate kinase gene (*pkiA*) is considered to be a highly expressed gene that is induced by glucose (de Graff *et al.*, 1988; Uitzetter 1982). Although there is a concurrence that pyruvate kinase activity in *Aspergillus* is lower under gluconeogenic growth conditions as compared to glycolytic conditions, reported differences in the level of activity produced under the two conditions vary. The level of pyruvate kinase activity in *A. nidulans* under glycolytic conditions has been reported by some investigators as being as much as 20-fold higher than under gluconeogenic conditions (Uitzetter *et al.*, 1982) and by others as only 7-fold higher (de Graff *et al.*, 1985).

This study, therefore, attempted to determine the expression level of the *A. niger* pyruvate kinase gene under gluconeogenic and glycolytic conditions and identify the regulatory elements that control *pkiA* expression. In order to do so, a set of ten promoter deletion plasmids (each containing a different promoter fragment) were used to direct

expression of the *lacA* reporter gene in medium containing glucose or glycerol as the sole carbon source (Fig. 15). The results demonstrated that *pkiA* expression is about 8-fold higher when glucose was the sole carbon source. Deletion analysis showed that the promoter region downstream of position –741 included all the information necessary to direct full levels of normally regulated expression. The deletion of sequences upstream of position –251, containing the 17 bp conserved sequence 5'-CRCGGGGAACCGRCCGA-3', (at position –282 to –266) resulted in a significant decrease in expression in glucose but did not alter expression when glycerol was the carbon source.

Fig. 15. Deletion analysis of the *A. niger pkiA* promoter region. PkiA1061 harbours a 1061 bp region of the *A. niger pkiA* gene from the start codon to bp -1061. Deletion derivatives of PkiA-1061 each contain a varying amount of deleted upstream information to direct the expression of the *lacA* reporter gene in medium containing various carbon sources. A vertical line at the left end of all promoter regions indicates a unique *Mlu*I site. For each promoter class, three independently constructed plasmids were transformed into the *A. niger* strain N593 as described in Materials and Methods. Two independent transformants of each independently constructed plasmid were grown in duplicate in 5 mL liquid cultures at 30°C for six days. Units of β-galactosidase per mL of culture are indicated on the right. Values next to units of activity indicate standard deviation.


Promoter region
Units of activity
glycerol
glucose
PkiA-1061
ATG
TATA
1.2 +/- 0.1
8.5 +/- 0.7
PkiA-921
ATG
TATA
1.1 +/- 0.1
9.3 +/- 0.7
PkiA-841
ATG
TATA
1.0 +/- 0.1
8.1 +/- 0.7
PkiA-741
ATG
TATA
0.6 +/- 0.1
7.8 +/- 0.7
PkiA-641
ATG
TATA
0.5 +/- 0.1
5.4 +/- 0.7
PkiA-541
ATG
TATA
0.6 +/- 0.1
5.4 +/- 0.7
PkiA-441
ATG
TATA
1.2 +/- 0.1
5.2 +/- 0.7
PkiA-351
ATG
TATA
0.6 +/- 0.1
4.6 +/- 0.7
PkiA-251
ATG
TATA
0.5 +/- 0.1
1.8 +/- 0.7
PkiA-134
ATG
TATA
0.6 +/- 0.1
0.6 +/- 0.7

4. Discussion

4.1 Autonomously replicating shuttle vectors for use in *A. niger*

This study intended to map possible *cis*-acting regulatory elements that control glucoamylase and pyruvate kinase gene expression in *A. niger*. Deletion analysis and site-directed mutagenesis of the *glaA* promoter region tested the functionality of putative target sites for regulatory proteins such as the CreA repressor, the AnCF binding complex, and the transcriptional activator AmyR. Unlike the *glaA* gene, a study identifying regulatory elements and possible transcription factors involved in *pkiA* expression in *Aspergillus* has not been published. This study involved deletion analysis of the *A. niger pkiA* promoter region to begin identifying activating and/or repressing elements analogous to those found in the yeast pyruvate kinase gene (Nishizawa *et al.*, 1989, McNeil *et al.*, 1990).

This study of the *A. niger* glucoamylase and the pyruvate kinase genes used autonomously replicating shuttle vectors. Their ability to replicate autonomously greatly simplifies the analysis of gene expression levels since context effects associated with chromosome integration can be avoided. As there are few plasmid vectors available for use with the industrial species *A. niger*, the expression vectors used in this study are not only able to facilitate studies of gene regulation but may also prove useful in the production of heterologous proteins in *A. niger*.

4.2 Deletion analysis of the *A. niger glaA* promoter region

Deletion analysis of the *glaA* promoter region found that sequences downstream of position –477 are important for *glaA* expression. The results indicate that the region between positions –477 and –448 is required for high-level expression, whereas the region between positions –448 and –368 is important for maltose induction and/or glucose repression. Although similar to the findings of other investigators, this study further narrows the region between –562 and –318 reported by Fowler *et al.* (1990) as being responsible for high-level expression of *glaA* by 215 base pairs. The results of this study also differ from what was published by Verdoes *et al.* (1994). Verdoes *et al.* (1994) found that a deletion up to position –517 resulted in an expression level of about 5% that of the full length promoter and concluded that the region responsible for high-level expression extended to sequences between positions –815 and –517.

Hata *et al.* (1992) conducted a deletion analysis of the *glaA* promoter region of *A. oryzae* and found that the region between –427 and -332, containing the conserved Region I, was required for both high–level expression and maltose induction. The conserved Region I indicated by Hata *et al.* (1992) corresponds, with regard to sequence similarity, to a region between positions –454 and –413 of the *A. niger glaA* gene. This region contains in part the sequence –448 to –368 found in this study to be important for maltose induction and glucose repression. This region contains two putative CreA repressor target sites that overlap with a putative target site for the transcriptional activator protein AmyR.

4.3 Regulation of *glaA* by the AnCF-binding complex

The CCAAT sequence is a *cis*-element commonly found in eukaryotic promoter regions. A multimeric protein complex that binds to CCAAT sequences has been identified and characterized in several eukaryotes such as *S. cerevisiae* (Pinkham and Guarante, 1985) and *mammals* (Chodosh *et al.*, 1988). In filamentous fungi, the CCAAT binding complex is designated AnCF and has been found to modulate gene expression in *A. nidulans* and *A. oryzae* (Brakhage *et al.*, 1999; Tanaka *et al.*, 2000). CCAAT sequences in filamentous fungi are not essential for gene inducibility but are important for determining the level of gene expression (Kato *et al.*, 1996). It has been postulated that the CCAAT-binding complex acts as a global transcriptional activator (Kato *et al.*, 1996) and that the binding of the multimeric protein complex is a prerequisite for pathway-specific transcription factors (Narendja *et al.*, 1999).

The *A. niger* promoter region contains three sites that closely match the 5'-RRCCAATMRCR-3' AnCF binding consensus sequence. Deleting the region upstream of position –477, just before the most upstream CCAAT binding site designated AnCF-1, resulted in an expression level similar to that of the full-length promoter. However, deletion of the *glaA* promoter region containing the AnCF-1 site significantly reduced gene expression in maltose without relieving glucose repression. Deletion of the *glaA* region upstream of position –368 harboring AnCF-1, the middle AnCF site and the conserved Region I dramatically reduced expression.

Cloning one copy of the upstream regulatory region between positions –477 and –368, containing the conserved Region I flanked by two putative AnCF-binding sites,

resulted in a low level of transcription. However, cloning one copy of the region between positions –454 and – 413 corresponding to Region I in *A.niger* did not result in the expression of the reporter gene. Although deletion analysis and site-directed mutagenesis showed that the integrity of the region between –454 and – 413 is necessary for *glaA* expression in *A. niger*, the region alone is not sufficient for gene expression. The inability of the smaller portion of the upstream regulatory region, containing Region I without the two surrounding AnCF-binding consensus sites, to activate transcription further emphasizes the importance of the two upstream AnCF sites for *glaA* expression. The lower expression level of plasmids harbouring the upstream regulatory region between positions –477 and – 368 as compared to plasmids containing the intact *glaA* promoter region downstream of position –477 also demonstrates that the region downstream of position –368 plays a role in *A. niger glaA* expression. The conserved Region II, indicated by Hata *et al.*, (1992) corresponds to the region between position –242 and –210 of the *A. niger glaA* gene. This region contains a stretch of eight identical nucleotides that are also found in the *A. oryzae glaA* and *amyB* genes at approximately the same distance from the translational start codon (Table 2). Perhaps these eight nucleotides constitute an element important for *glaA* expression.

My expression analysis of promoter derivatives with site-directed mutations in putative CCAAT-binding sites revealed that the two upstream AnCF sites were important for *glaA* expression, whereas the third site did not seem important under the conditions tested. Mutagenesis of either the first or middle putative AnCF-binding site resulted in a reduced expression level without affecting carbon source-dependent regulation. Simultaneous mutagenesis of the two upstream sites resulted in greatly reduced

expression in maltose and glucose and a loss of induction by maltose demonstrating that the integrity of at least one of the sites is necessary for maltose induction.

4.4 Regulation of *glaA* by CreA and AmyR

CreA is the major regulatory protein controlling carbon catabolite repression in *A. nidulans* (Arst and Bailey, 1977). The *A. niger creA* gene has been cloned in *A. niger* (Drysdale *et al.*, 1993) but a study establishing its role in *A. niger* gene expression has yet to be published. CreA binds the consensus sequence 5'-SYGGRG-3' (Cubero and Scazzocchio, 1994). It has been found that although many CreA binding sites are found in promoter regions of *Aspergillus* genes, few are functional (reviewed in Felenbok, 2001). Interest lies in determining elements that are necessary for the specificity of CreA recognition and identifying characteristics of functional CreA binding sites. Genes found to be regulated by CreA often have functional CreA binding sites organized as pairs. Furthermore, genes regulated by the general transcriptional repressor CreA are also often subject to pathway-specific induction. The interplay between CreA and pathway specific systems has been described in the *A. nidulans* ethanol utilization pathway (reviewed in Felenbok, 2001). The interaction is characterized by the direct competition of the activator protein AlcR and CreA for the same sites in promoters of ethanol utilization genes. The system also displays a "double-lock mechanism" in which the activator gene is also repressed by CreA.

The *A. niger glaA* promoter region contains a pair of CreA-binding consensus sites separated by 9 bp. One of these, the downstream site, overlaps by 6 bp an AmyR-

binding consensus site found in the *glaA* promoter. The region upstream of position –368 of *glaA* between –469 and –376 contains these two CreA sites, an AmyR site and the two AnCF-binding sites shown to be critical for high levels of expression. Deleting this region (Gla-368) dramatically reduced expression. Site-directed mutagenesis of either of the two putative CreA-binding sites contained in Region I resulted in a loss of gene expression in every condition tested, whereas mutagenesis of the two downstream CreA elements did not alter gene expression. These results were surprising. Assuming that these sites bound the CreA repressor, it was expected that changing these sites would alleviate glucose repression.

Electrophoretic mobility shift assays conducted by Catherine Au (MSc thesis, 2002) suggested that these two upstream CreA consensus sites play a role in CreA binding. However, the *in vivo* results of this study evidently indicate that a transcription factor necessary for *glaA* expression interacts with these sites. Attention, was therefore, focused on the conserved Region I. Visual inspection of the sequence revealed an AmyR binding sequence of the $CGGN_8CGG$ type. The *amyR* gene has been cloned in *A. niger* (Udagawa, unpublished). Expression of the *amyR* gene in *A. nidulans* is similar to that of *glaA*. It is induced by growth on starch or maltose and repressed by CreA when grown in glucose (Tani *et al.*, 2001).

One of the CGG triplets of the $CGGN_8CGG$ site is located two base pairs downstream of the CreA1 site, whereas the second CGG triplet is located within the CreA2 site (Fig. 16). Mutagenesis of either the CreA1 site just upstream the first CGG triplet or of the CGG triplet contained in the CreA2 site results in a complete loss of gene expression. *In vivo* and *in vitro* studies done by Tani *et al.* (2001) of the $CGGN_8CGG$

site in the *A. nidulans agdA* promoter, with both or either one of the CGG triplets altered, showed that mutations disrupting either of the CGG triplets cause a drastic decrease in the gene's inducibility. It is hypothesized that directly repeated CGG triplets are required for high affinity binding of AmyR by allowing two AmyR molecules to interact with each other and stabilize the DNA-protein interaction (Tani *et al.*, 2001). Altering the CreA2 site disrupts the downstream CGG triplet whereas changing the CreA1 site alters the sequence just adjacent to the upstream CGG triplet of the AmyR binding consensus sequence ($CGGN_8CGG$). It is, therefore, possible that the sequence just adjacent to one of the CGG triplets is also required for the binding of one of the AmyR molecules to the $CGGN_8CGG$ site of the *glaA* promoter region.

Figure 16. The nucleotide sequence of the 5' upstream region from position –445 to –396 of the *A. niger glaA* gene. The site matching the AmyR binding consensus sequence (CGGN_8CGG) is underlined. The CGG triplets are underlined twice. The two sites matching the CreA binding consensus sequence (5'-SYGGRG-3) in this region are highlighted in grey.

CreA1 site CreA2 site

-445 A[illegible]ACG GCGAAT[illegible] [illegible]GAATTGAA

AmyR site

-415 AGAAATTGCA TCCTAGGCCA

Besides matching the CreA binding consensus sequence and having the characteristic organization of a pair of sites associated with other functional CreA sites, the only evidence that CreA1 and CreA2 are functional CreA binding sites comes from the results of the electrophoretic mobility shift assays of *glaA*. Taking into consideration the results of both the *in vitro* and *in vivo* studies of *glaA*, it can be hypothesized that the glucoamylase gene of *A. niger* is regulated by direct competition between the pathway-specific activator protein AmyR and the general transcriptional repressor CreA in a system similar to the interplay between two regulatory proteins characterized in the *A. nidulans* ethanol utilization pathway (reviewed in Felenbok, 2001). The strong repression of the *glaA* gene of *A. niger* under carbon catabolite repressing conditions may also result from the "double-lock mechanism" in which CreA directly represses the activator *amyR* gene and its target *glaA*. Supporting this double lock model I found that directing expression of the *amyR* gene by the constitutively expressed *pkiA* promoter resulted in significantly increased glucoamylase expression by both glucose and xylose grown cells. It can, therefore, be suggested that the lack of glucoamylase expression when xylose is supplied as the sole carbon source may in part be due to the absence of the activator protein AmyR.

4.5 Regulation of the pyruvate kinase gene (*pkiA*) of *A. niger*

Expresssion studies with the full-length promoter of the pyruvate kinase gene (*pkiA*) found that growth with the glycolytic carbon source glucose leads to an

approximate eight-fold increase in reporter gene expression as compared to the level found when the organism is grown on the gluconeogenic carbon source glycerol.

Deletion analysis of the pyruvate kinase promoter region found three regions that are important for *pkiA* expression in glucose. Deleting regions between position –741 and –641, -351 and –251 and between –251 and –134 resulted in decreased expression levels in glucose. Two putative CCAAT-binding sites are found within the *pkiA* promoter region. Deleting the region between position –641 and –541, containing the most upstream putative CCAAT-binding site, did not result in a lower expression level than observed when the region upstream of position –641 was deleted. However, when the region between position –251 and –134, harboring the downstream CCAAT-binding consensus site was deleted the level of expression was reduced significantly. Both putative CCAAT-binding sites contain the CCAAT core sequence that is required for the multimeric protein complex to bind (Mantovani, 1998). Perhaps the downstream CCAAT site is functionally more important because it (5'-AGCCAATCACC-3') is flanked by nucleotides that are a closer match to the 11 bp AnCF binding consensus sequence (5'-RRCCAATMRCR-3') (Brakhage, 1998) than the upstream CCAAT site (5'-AACCAATTATT-3'). Deletion analysis of the region upstream of the *A. niger pkiA* coding region revealed that the downstream AnCF consensus sequence is the only possible AnCF-binding site within the *pkiA* promoter region.

Deleting the region upstream of position –251, containing a 17 bp sequence that is conserved in the otherwise non-homologous *A. niger* and *A. nidulans pkiA* promoters, resulted in a significantly reduced amount of reporter gene expression in glucose. This conserved element is a likely candidate element required for the recruitment of

transcription factors involved in the expression of *pkiA* and may be analogous to the upstream activating sequences found in the yeast pyruvate kinase gene.

Nishizawa *et al.* (1989) identified two upstream activating sequences (UAS) and an upstream repressible sequence (URS) in the promoter region of the yeast pyruvate kinase gene (*pyk1*). The most upstream positive regulatory element, designated UAS_{PYK1} (5'-ACCCAGACATCGGGCTTCC-3'), covers the region between –653 and –634 of the pyruvate kinase gene. It is an essential element, functional in either orientation and does not show strict position dependency. The downstream positive regulatory element, UAS_{PYK2}, is located between positions –811 and –714 of the *pyk1* gene. UAS_{PYK2} is a *cis*-acting element required for full transcriptional activation of the *pyk1* promoter and is able to activate transcription if the sequence containing the URS has been removed. The URS of the *pyk1* gene, located between positions –468 and –344, represses transcription when cells are grown in a gluconeogenic carbon source and to a lesser extent in glucose grown cells.

In this study, reporter gene expression of glycerol grown transformants harboring *pkiA* promoter derivatives did not reveal a repressor element. It is still possible that the *pkiA* promoter region of *A. niger* contains a repression sequence comparable to the upstream repressible sequence of the yeast pyruvate kinase gene (Nishizawa *et al.*, 1989). For example, I could have missed a negative acting regulatory element if it was located within one of the deleted fragments that also contained a regulatory element that positively impacts expression. Mapping a *cis*-acting negative regulatory element within the *pkiA* promoter region may, therefore, only be possible by performing smaller internal deletions in the promoter region rather than the series of increasingly large 5' deletions

used here. Repression of the *pkiA* gene may also not result from the direct binding of a repressor protein to a *cis*-acting regulatory element but by some other mechanism such as the inactivation, sequestering or regulated expression of a transcriptional activator.

4.6 Future studies

This study only began investigating the regulation of the *A. niger pkiA* gene. Additional experiments should commence by performing site-directed mutagenesis of the putative regulatory elements identified by deletion analysis.

Future studies to characterize the glucoamylase gene of *A. niger* should include additional electrophoretic mobility shift assays using purified CreA protein, AmyR protein and crude extract from both wild-type and variously mutated strains. Site-directed mutagenesis of only the CCG triplets of the AmyR binding sequence should also be done. Expression studies using the various plasmids constructed in this study should also include the use of *creA* and *amyR* mutated strains. Finally, to identify mutations that affect glucoamlylase production, a close inspection of the glucoamylase gene of *A. niger* strains selected for industrial use (due to their elevated glucoamylase production) can be made. This, however, would entail the use of strains that are often unavailable to academic researchers. Nonetheless, industrial companies are cooperating with academic institutions. For example, there is now a low-barrier access program to the complete DNA sequence of *A. niger*.

Insight into the molecular mechanisms responsible for the repression and induction of *glaA* and a better comprehension of the interplay between different

regulatory systems in *A. niger* has been difficult to obtain using conventional methods. Genomic-based approaches hopefully will help elucidate how *Aspergillus* fine-tunes gene expression in response to the carbon source availability.

With the determination of the complete sequence of the *A. niger* genome and the use of new technologies such as microarray based transcription profiling , a further understanding of *Aspergillus* gene regulation will evidently arise.

Works Cited

Aleksenko, A.Y. and Clutterbuck A.J. 1995. Recombinational stability of replicating plasmids in *Aspergillus nidulans* during transformation, vegetative growth and sexual reproduction. *Curr. Genet.* 28(1): 87-93

Almer, A. and Horz, W. 1986. Nuclease hypersensitive regions with adjacent positioned nucleosomes mark the gene boundaries of the PHO5/PHO3 locus in yeast. *EMBO J.* 5(10): 2681-7

Al-Musallum, A. 1980. Revision of the black *Aspergillus* species. Thesis, Rijksuniversiteit, Utrecht

Archer, D.B., Jeenes, D.J. and Mackenzie, D.A. 1994. Strategies for improving heterologous protein production from filamentous fungi. *Antoine Van Leeuwenhoek* 65(3):245-50

Arst, H.N. Jr. and MacDonald, D.W. 1975. A gene cluster in *Aspergillus nidulans* with an internally located *cis*-acting regulatory region. *Nature* 254(5495): 26-31

Arst, H.N. Jr., Jones, S.A. and Bailey, C.R. 1981. A method for the selection of deletion mutations in the L-proline catabolism gene cluster of *Aspergillus nidulans*. *Genet. Res.* 38(2): 171-95

Au, C. 2002. Interactions between the CreA protein and the *glaA* promoter region of *Aspergillus niger*. M.Sc. Thesis, Concordia University

Barton, L.L., Georgi, C.E. and Lineback, D.R. 1972. Effect of maltose on glucoamylase formation by *Aspergillus niger*. *J. Bacteriol.* 111(3): 771-7

Bennett, J.W. 1985. Taxonomy of fungi and biology of the *Aspergilli*. In A.L. Demain and N.A. Solomon, (eds.), Biology of industrial microorganisms. Benjamin Cummings Publishing Co., Menlo Park, CA.

Benoist, C., O'Hare K., Breathnach, R. and Chambon, P. 1980. The ovalbumin gene-sequence of putative control regions. *Nucleic Acids Res.* 8(1): 127-42

Boel, E., Hansen, M.T., Hjort, I. and Fill, N.P. 1984. Two different types of intervening sequences in the glucoamylase gene from *Aspergillus niger*. *EMBO J.* 3: 1581-1585

Boel, E., Hjort, I., Svensson, B., Norris, F., Norris K.E. and Fill, N.P. 1984. Glucoamylase G1 and G2 from *Aspergillus niger* are synthesized from two different but closely related mRNAs. *EMBO J.* 3: 1097-1102

Bos, C.J., Debets, A.J., Swart, K., Huybers, A., Kobus, G. and Slakhorst, S.M. 1988. Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in *Aspergillus niger*. *Curr. Genet.* 14(5): 437-43

Brakhage, A.A., Andrianopoulos, A., Kato, M., Steidl, S., Davis, M.A., Tsukagoshi, N. and Hynes, M.J. 1999. HAP-like CCAAT-binding complexes in filamentous fungi: implications for biotechnology. *Fungal Genetics and Biology* 27: 243-252

Bucher, P. 1990. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol.* 212(4): 563-78

Carrez, D., Janssens W., Degrave P., van den Hondel, C.A., Kinghorn, J.R., Fiers, W. and Contreras, R. 1990. Heterologous gene expression by filamentous fungi: secretion of human interleukin-6 by *Aspergillus nidulans*. *Gene* 94: 147-154

Chen, H. and Kinsey, J.A. 1994. Sequential gel mobility shift scanning of 5' upstream sequences of the *Neurospora crassa am* (*GDH*) gene. *Mol.Gen.Genet.* 242(4): 399-403

Chodosh L.A., Baldwin A.S., Carthew R.W. and Sharp P.A. 1988. Human CCAAT-binding proteins have heterologous subunits. *Cell.* 53(1): 11-24

Cooper, J.P., Roth, S.Y. and Simpson, R.T. 1994. The global transcriptional regulators, SSN6 and TUP1, play distinct roles in the establishment of a repressive chromatin structure. *Genes Dev.* 8(12): 1400-10

Coutinho, P.M. http://www.public.iastate.edu/~pedro/glase/glase.html

Cubero, B. and Scazzocchio, C. 1994. Two different, adjacent and divergent zinc finger binding sites are necessary for CREA-mediated carbon catabolite repression in the proline gene cluster of *Aspergillus nidulans*. *EMBO J.* 13(2): 407-415

Debets, A.J.M. and Bos ,C.J. 1986. Isolation of small protoplasts from *Aspergillus*. *Fungal Genetics Newsletter* 3324

de Graaff, L., van den Broeck, H. and Visser, J. 1992. Isolation and characterization of the *Aspergillus niger* pyruvate kinase gene. *Curr.Genet.* 22(1): 21-7

DeVit, M.J. and Johnston, M. 1999. The nuclear exportin Msn5 is required for nuclear export of the Mig1 glucose repressor of *Saccharomyces cerevisiae*. *Curr. Biol.* 9(21):1231-41

Dowzer, C.E.A.and Kelly, J.M. 1991. Analysis of the *creA* gene, a regulator of carbon catabolite repression in *Aspergillus nidulans*. *Mol. Cell. Biol.* 11(11): 5701-5709

Drysdale, M.R., Kolze, S.E. and Kelly, J.M. 1993. The *Aspergillus niger* carbon catabolite repressor encoding gene, *creA*. *Gene* 130: 241-245

Dubey, A.K., Suresh, C., Kavitha, R., Karanth, N.G. and Umesh-Kumar S. 2000. Evidence that the glucoamylases and alpha-amylase secreted by *Aspergillus niger* are proteolytically processed products of a precursor enzyme. *FEBS Lett.* 471(2-3): 251-5

Espeso, E.A. and Penalva, M.A. 1994. In vitro binding of the two-finger repressor CreA to several consensus and non-consensus sites at the *ipnA* upstream region is context dependent. *FEBS Letters* 342: 43-48

Felenbok, B., Flipphi, M. and Nikolaev, I. 2001. Ethanol catabolism in *Aspergillus nidulans*: a model system for studying gene regulation. *Prog. Nucleic Acid Res. Mol. Biol.* 69:149-204

Forsgurg, S.L. and Guarente, L. 1989. Identification and characterization of HAP4: A third component of the CCAAT-bound HAP2/HAp3 heteromer. *Genes Dev.* 3: 1166-1178

Fowler, T., Berka, R.M. and Ward, M. 1990. Regulation of the *glaA* gene of *Aspergillus niger*. *Curr. Genet.* 18: 537-545

Frederick, G.D. and Kinsey, J.A. 1990. Distant upstream regulatory sequences control the level of expression of the *am* (GDH) locus of *Neurospora crassa*. *Curr. Genet.* 18(1): 53-8

Gancedo, J.M. 1998. Yeast carbon catabolite repression. *Microbiol. Mol.Biol. Rev.* 62: 334-361

Gems, D., Johnstone, I.L. and Clutterbuck, A.J. 1991. An autonomously replicating plasmid transforms *Aspergillus nidulans* at high frequency. *Gene* 98: 61-67

Gomi, K., Akeno, T., Minetoki, T., Ozeki, K., Kumagai, C., Okazaki, N. and Iimura, Y. Molecular cloning and characterization of a transcriptional activator gene, *amyR*, involved in the amylolytic gene expression in *Aspergillus oryzae*. *Biosci. Biotechnol. Biochem.* 64(4): 816-27

Gonzalez, R., Gavrias, V., Gomez, D., Scazzocchio, C. and Cubero, B. 1997. The integration of nitrogen and carbon catabolite repression in *Aspergillus nidulans* requires the GATA factor AreA and an additional positive-acting element, ADA. *EMBO J.* 16: 2937-2944

Gouka, R.J., Punt, P.J. and van den Hondel, C.A. 1997. Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. *Appl. Microbiol. Biotechnol.* 47(1): 1-11

Hahn, S. and Guarente, L. 1988. Yeast HAP2 and HAP3: transcriptional activators in a heteromeric complex. *Science* 240(4850): 317-21

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166(4): 557-80

Hata, Y., Kitamoto, K., Gomi, K., Kumagai, C. and Tamura, G. 1992. Functional elements in the promoter region of the Aspergillus oryzae *glaA* gene encoding glucoamylase. *Curr. Genet.* 22: 85-91

Higuchi, R., Krummel, B. and Saiki, R.K. 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res.* 16(15): 7351-67

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease L.R. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 77(1): 51-9

Ingold, C.T. 1961. The biology of fungi. Hutchinson Educational Ltd. London, UK

Jeenes, D.J., Marczinke, B., MacKenzie, D.A. and Archer, D.B. 1993. A truncated glucoamylase gene fusion for heterologous protein secretion from *Aspergillus niger*. *FEMS Microbiol. Lett.* 107: 267-271

Jong, S.C. and Gantt, M.J. (eds.) 1987. Catalogue of fungi and yeasts, 17th edition. American Type Culture Collection, Rockville, MD.

Kafer, E. 1977. Meiotic and mitotic recombination in *Aspergillus* and its chromosomal aberrations. *Adv. Genet.* 19: 33-131

Kanemori, Y., Gomi, K., Kitamoto, K., Kumagai, C. and Tamura, G. 1999. Insertion analysis of putative functional elements in the promoter region of the *Aspergillus oryzae* Taka-amylase A gene (*amyB*) using a heterologous *Aspergillus nidulans amdS-lacZ* fusion gene system. *Biosci Biotechnol Biochem.* 63(1): 180-3

Kapoor, M. and Tronsgaard, T.M. 1972. Pyruvate kinase of *Neurospora crassa*: purification and some properties. *Can. J. Microbiol.* 18(6):805-15

Kato, M., Aoyama A., Naruse, F., Kobayashi, T. and Tsukagoshi N. 1997. An *Aspergillus nidulans* nuclear protein, AnCP, involved in enhancement of Taka-amylase A gene expression, binds to the CCAAT-containing *taaG2*, *amdS*, and *gatA* promoters. *Mol. Gen. Genet.* 254: 119-126

Kato, M., Aoyama A., Naruse, F., Tateyama Y., Hayashi, K., Miyazaki, M., Papagiannopoulos, P., Davis, M.A., Hynes, M.J., Kobayashi T. and Tsukagoshi, N. 1998. The *Aspergillus nidulans* CCAAT-binding factor AnCP/AnCF is a heteromeric protein analogous to the HAP complex of *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* 257: 404-411

Keleher, C.A., Redd, M.J., Schultz, J., Carlson, M. and Johnson, A.D. 1992. Ssn6-Tup1 is a general repressor of transcription in yeast. *Cell* 68(4): 709-19

Kim, I.S., Sinha, S., de Crombrugghe, B. and Maity, S.N. 1996. Determination of functional domains in the C subunit of the CCAAT-binding factor (CBF) necessary for formation of a CBF-DNA complex: CBF-B interacts simultaneously with both the CBF-A and CBF-C subunits to form a heterotrimeric CBF molecule. *Mol. Cell. Biol.* 16(8): 4003-13

Kozowski, M. and Stepien, P.P. 1982. Restriction analysis of mitochondrial DNA of members of the genus *Aspergillus* as an aid in taxonomy. *Jour. Gen. Microbiol.* 128: 471-476

Kulmburg, P., Mathieu, M., Dowzer, C., Kelly, J. and Felenbok, B. 1993. Specific binding sites in the *alcR* and *alcA* promoters of the ethanol regulon for the CREA repressor mediating carbon catabolite repression in *Aspergillus nidulans*. *Molecular Microbiology* 7: 847-857

Kumar, V., Ramakrishnan, S., Teeri, T.T., Knowles, J.K. and Hartley, B.S. 1993. *Saccharomyces cerevisiae* cells secreting an *Aspergillus niger* beta-galactosidase grow on whey permeate. *Biotechnology (NY)* 10: 82-85

Kusters-Van Someren, M., Samson, R.A., and Visser, J. 1990. In: Samson RA, Pitts JI (eds)_Modern concepts in *Penicillium* and *Aspergillus* classification. NATO ASI Series vol 185, Plenum Press, New York, pp 321-334

Kusters-Van Someren, M., Samson, R.A., and Visser, J. 1991. The use of RFLP analysis in the classification of the black *aspergilli*: Reinterpretation of the *Aspergillus niger* aggregate. *Curr. Gen.* 19: 21-261

Kwon-Chung, K.J., Varma, A., Edman, J.C. and Bennett, J.E. 1992. Selection of *ura5* and *ura3* mutants from the two varieties of *Cryptococcus neoformans* on 5-fluoroorotic acid medium. *J. Med. Vet. Mycol.* 30(1): 61-9

Langvad, F. 1999. A rapid and efficient method for growth measurement of filamentous fungi. *J. Microbiol. Methods* 37(1): 97-100

Littlejohn, T.G. and Hynes, M.J. 1992. Analysis of the site of action of the *amdR* product for regulation of the *amdS* gene of *Aspergillus nidulans*. *Mol. Gen. Genet.* 5(1): 81-8

Litzka, O., Then Bergh, K. and Brakhage, A.A. 1996. The *Aspergillus nidulans* penicillin-biosynthesis gene *aat (penDE)* is controlled by a CCAAT-containing DNA element. *Eur. J. Biochem.* 238(3): 675-82

Lundin, M., Nehlin, J.O. and Ronne, H. 1994. Importance of a flanking AT-rich region in target site recognition by the GC-box-binding zinc finger protein MIG1. *Mol. Cell. Biol.* 14(3): 1979-1985

MacKenzie, D.A., Jeenes, D.J., Gou, X. and Archer D.B. 2000. Molecular basis of glucoamylase overproduction by a mutagenised industrial strain of *Aspergillus niger*. *Enzyme Microb. Technol.* 26(2-4):193-200

Maity S.N., Sinha S., Ruteshouser E.C. and de Crombrugghe B. 1992. Three different polypeptides are necessary for DNA binding of the mammalian heteromeric CCAAT binding factor. *J. Biol. Chem.* 267(23): 16574-80

Mantovani, R. 1998. A survey of 178 NF-Y binding CCAAT boxes. *Nucleic Acids Res.* 26: 1135-1143

Mathieu, M. and Felenbok, B. 1994. The *Aspergillus nidulans* CREA protein mediates glucose repression of the ethanol regulon at various levels through competition with the ALCR specific transactivator. *EMBO J.* 13: 4022-4027

Mathieu, M., Fillinger, S. and Felenbok, B. 2000. *In vivo* studies of upstream regulatory cis-acting elements of the alcR gene encoding the transactivator of the ethanol regulon in *Aspergillus nidulans*. *Mol. Microbiol.* 36(1): 123-31

McCleary B.V., Gibson, T.S., Sheehan. H., Casey, A., Horgan, L.and O'Flaherty J. 1989. Purification, properties and industrial significance of transglucosidase from *Aspergillus niger*. *Carbohyd. Res.* 185: 147-62

McNabb, D.S., Xing, Y. and Guarente, L. 1995. Cloning of yeast HAP5: a novel subunit of a heterotrimeric complex required for CCAAT binding. *Genes Dev.* 9(1): 47-58

McNeil, J.B., Dykshoorn, P., Huy, J.N. and Small, S. 1990. The DNA-binding protein RAP1 is required for efficient transcriptional activation of the yeast PYK glycolytic gene. *Curr. Genet.* 18(5): 405-12

Megnegneau, B., Debets, F. and Hoekstra, R.F. 1993. Genetic variability and relatedness in the complex group of black *Aspergilli* based on random amplification of polymorphic DNA. *Curr. Genet.* 23: 323-329

Minetoki, T., Nunokawa, Y., Gomi,K., Kitamoto, K., Kumagai, C. and Tamura, G. 1996. Deletion analysis of promoter elements of the *Aspergillus oryzae agdA* gene encoding alpha-glucosidase. *Curr. Genet.* 30(5): 432-8

Moreira, J.M. and Holmberg S. 1998. Nucleosome structure of the yeast CHA1 promoter: analysis of activation-dependent chromatin remodeling of an RNA-polymerase-II-transcribed gene in TBP and RNA pol II mutants defective in vivo in response to acidic activators. *EMBO J.* 17(20): 6028-38

Nagata, O., Takashima, T., Tanaka, M. and Tsukagoshi, N. 1993. *Aspergillus nidulans* nuclear proteins bind to a CCAAT element and the adjacent upstream sequence in the promoter region of the starch-inducible Taka-amylase A gene. *Mol. Gen. Genet.* 237: 251-260

Narendja, F.M., Davis, M. A., Hynes, M.J. 1999. AnCF, the CCAAT binding complex of *Aspergillus nidulans*, is essential for the formation of a DNase I-hypersensitive site in the 5'region of the *amdS* gene. *Mol. Cell. Biol.* 19: 6523-6531

Nehlin, J.O., Carlberg, M. and Ronne, H. 1991. Control of yeast GAL genes by MIG1 repressor: a transcriptional cascade in the glucose response. *EMBO J.* 11: 3373-3377

Nikolaev, I., Coche, M.F., Lenouvel, F. and Felenbok, B. 1999. A single amino acid, outside the AlcR zinc binuclear cluster, is involved in DNA binding and in transcriptional regulation of the *alc* genes in *Aspergillus nidulans*. *Mol. Microbiol.* 31(4): 1115-24

Nishizawa, M, Araki, R, Teranishi, Y. 1989. Identification of an upstream activating sequence and an upstream repressible sequence of the pyruvate kinase gene of the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 9(2): 442-51

Nunberg, J.H., Meade, J.H., Georgette, C., Lawyer, F.C., McCabe, P., Schweickart, V., Tal, R., Wittman, V.P., Flatgaard, J.E.and Innis, M.A. 1984. Molecular cloning and characterization of the glucoamylase gene of *Aspergillus awamori*. *Mol. Cell. Biol.* 4: 2306-2315

Oakley, B.R., Rinehart, J.E., Mitchell, B.L., Oakley, C.E., Carmona, C., Gray, G.L. and May, G.S. 1987. Cloning, mapping and molecular analysis of the *pyrG* (orotidine-5'-phosphate decarboxylase) gene of *Aspergillus nidulans*. *Gene* 61: 385-399

Olesen, J.T., Fikes, J.D. and Guarente, L. 1991. The *Schizosaccharomyces pombe* homolog of *Saccharomyces ceerevisiae* HAP2 reveals selective and stringent conservation of the small essential core protein domain. *Mol.Cell. Biol.* 11: 611-619

Oza, K. and Kafer E. 1990. Cloning of the DNA repair gene, *uvsF*, by transformation of *Aspergillus nidulans*. *Genetics* 125: 341-349

Panozzo, C., Cornillot, E. and Felenbok B. 1998. The CreA repressor is the sole DNA-binding protein responsible for carbon catabolite repression of the *alcA* gene in *Aspergillus nidulans* via its binding to a couple of specific sites. *J. Biol. Chem.* 273 (11): 6367-6372

Papagiannopoulos, P., Andrianopoulos, A., Sharp, J.A., Davis, M.A., Hynes, M.J. 1996. The *hapC* gene of *Aspergillus nidulans* is involved in the expression of CCAAT-containing promoters. *Mol.Gen. Genet.* 251: 412-421

Pavletich, N.P. and Pabo, C.O. 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science.* 252(5007): 809-17

Payton, M. and Roberts, C.F. 1976. Mutants of *Aspergillus nidulans* lacking pyruvate kinase. *FEBS Lett.* 66(1): 73-6

Pazur, J.H. and Ando, T. 1959. The action of an amyloglucosidase of *Aspergillus niger* on starch and malto-oligosaccharides. *J. Biol. Chem.* 234: 1966-1970

Pazur, J.H, Tominaga Y., Forsberg S. and Simpson, D.L. 1980. Glycoenzymes: an unusual type of glycoprotein structure for a glucoamylase. *Carbohydr Res.* 84(1): 103-14

Petersen, K.L., Lehmbeck, J. and Christensen, T. 1999. A new transcriptional activator for amylase genes in *Aspergillus*. *Mol. Gen. Genet.* 262(4-5): 668-76

Pinkham, J. and Guarente, L. 1985. Cloning and molecular analysis of the HAP2 locus: A global regulator of respiratory genes in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 5: 3410-3416

Raper, K.B. and Fennel, D.I. 1965. The genus *Aspergillus*. Williams and Wilkins Co., Baltimore, MD

Reece, R.J. and Ptashne, M. 1993. Determinants of binding-site specificity among yeast C6 zinc cluster proteins. *Science* 261(5123): 909-11

Redd, M.J., Arnaud, M.B. and Johnson, A.D. 1997. A complex composed of *tup1* and *ssn6* represses transcription *in vitro*. *J. Biol. Chem.* 272(17): 11193-7

Saiki, R.K., Gelfand, D.H., Stoffelm S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239(4839): 487-91

Sambrook, J., Fritsh, E.F. and Maniatis T. 1989. Molecular cloning: A Laboratory Manual. 2nd edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Santerre Henriksen, A.L., Even, S., Muller, C., Punt, P.J., van den Hondel C.A. and Nielsen, J. 1999. Study of the glucoamylase promoter in *Aspergillus niger* using green fluorescent protein. *Microbiology*. 145 (3): 729-734

Saunders, G., Picknett, T.M., Tuite, M.F. and Ward, M. 1989. Heterologous gene expression in filamentous fungi. *Trends Boitechnol*. 7: 283-287

Shroff, R.A., Lockington, R.A. and Kelly, J.M. 1996. Analysis of mutations in the *creA* gene involved in carbon catabolite repression in *Aspergillus nidulans*. *Can. J. Microbiol*. 42(9): 950-9

Sikorski, R.S. and Hieter, P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122: 19-27

Sinha, S., Maity, S.N., Seldin, M.F. and de Crombrugghe, B. 1996. Chromosomal assignment and tissue expression of CBF-C/NFY-C, the third subunit of the mammalian CCAAT-binding factor. *Genomics*. 37(2): 260-3

Southall, S.M., Simpson, P.J., Gilbert, H.J., Williamson, G. and Williamson, M.P. 1999. The starch-binding domain from glucoamylase disrupts the structure of starch. *FEBS Lett*. 447(1): 58-60

Steidl, S., Papagiannopoulos, P., Litzka, O., Andrianopoulos, Davis, M.A., Brakhage, A.A., Hynes, M.J. 1999. AnCF, the CCAAT binding complex of Aspergillus nidulans, contains products of the *hapB, hapC, hapE* genes and is required for activation by the pathway-specific regulatory gene *amdR*. *Mol. Cell. Biol*. 19(1): 99-106

Strauss, J., Horvath, H.K., Abdallah, B.M., Kindermann, J., Mach, R.L. and Kubicek, C.P. 1999. The function of CreA, the carbon catabolite repressor of *Aspergillus nidulans*, is regulated at the transcriptional and post-transcriptional level. *Mol. Microbiol*. 32(1): 169-178

Svensson, B., Larson, K., Gunnarsson, A. 1986. Characterization of a glucoamylase G2 from *A. niger*. *Eur. J. Biochem*. 154: 497-502

Svensson, B., Larson, K., Svendsen, I. and Boel, E. 1983. The complete amino acid sequence of the glycoprotein, glucoamylase G1, from *Aspergillus niger*. *Carlsberg Res. Commun*. 48: 529-544

Svensson, B., Pedersen, T.G., Svendsen, I., Sakai, T. and Ottesen, M. 1982. Characterization of the two forms of glucoamylase from *A. niger*. *Carlsberg Res. Commun.* 47: 55-69

Tanaka A., Kato, M., Hashimoto, H., Kamei, K., Naruse, F., Papagiannopoulos, P., Davis, M.A., Hynes, M.J., Kobayashi, T. and Tsukagoshi, N. 2000. An *Aspergillus oryzae* CCAAT-binding protein, AoCP, is involved in the high-level expression of the Taka-amylase A gene. *Curr. Genet.* 37(6): 380-7

Tani, S., Kawaguchi, T., Kato, M., Kobayashi T. and Tsukagoshi, N. 2000. A novel nuclear factor, SREB, binds to a cis-acting element, SRE, required for inducible expression of the *Aspergillus oryzae* Taka-amylase A gene in *A. nidulans*. *Mol. Gen. Genet.* 263(2): 232-8

Takahashi, T., Kato, K., Ikegami, Y. and Irie, M. 1985. Different behavior towards raw starch of three forms of glucoamylase from a Rhizopus sp. *J. Biochem (Tokyo)* 98(3): 663-71

Terenzi, HF, Roselino, E. and Passeron, S. 1971. Two types of pyruvate kinase in *Mucor rouxii*. Changes in the enzymatic pattern related to aerobic development. *Eur. J. Biochem.* 18(3): 342-50

Then Bergh, K., Litzka, O. and Brakhage, A.A. 1996. Identification of a major *cis*-acting DNA element controlling the bidirectionally transcribed penicillin biosynthesis genes *acvA* (*pcbAB*) and *ipnA* (*pcbC*) of *Aspergillus nidulans*. *J Bacteriol.* 178: 3908-3916

Todd, R.B., Murphy, R.L., Martin, H.M., Sharp, J.A., Davis, M.A., Katz, M.E. and Hynes M.J. 1997. The acetate regulatory gene *facB* of *Aspergillus nidulans* encodes a Zn(II)2Cys6 transcriptional activator. *Mol. Gen. Genet.* 254(5): 495-504

Ueda, S. 1981. Fungal glucoamylases and raw starch digestion. *Trends in Biochemical Sciences* 6: 89-90

Uitzetter, J.H., Bos, C.J. and Visser, J. 1982. Growth characteristics of *Aspergillus nidulans* mutans defective in carbohydrate metabolism. *Antoine Van Leeuwenhoek.* 48(3): 219-27

van Heeswijck, R. and Hynes, M.J. 1991. The *amdR* product and a CCAAT-binding factor bind to adjacent, possibly overlapping DNA sequences in the promoter region of the *Aspergillus nidulans amdS* gene. *Nucleic Acids Res.* 19(10): 2655-60

Vautard-Mey G., Cotton P. and Fevre M. 1999. Expression and compartmentation of the glucose repressor CRE1 from the phytopathogenic fungus *Sclerotinia sclerotiorum*. *Eur. J. Biochem.* 266(1): 252-9

Verdoes, J.C., Punt, P.J., Schrickx, J.M., van Verseveld, H.W., Stouthamer, A.H. and van den Hondel, C.A.M.J.J. 1993. Glucoamylase overexpression in *Aspergillus niger*: molecular genetic analysis of strains containing multiple copies of the *glaA* gene. *Transgenic Research* 2: 84-92

Verdoes, J.C., Punt, P.J. and van den Hondel, C.A.M.J.J. 1994. The effect of multiple copies of the upstream region on expression of the *Aspergillus niger* glucoamylase-encoding gene. *Gene* 145: 179-187

Verdoes, J.C., van Diepeningen, A.D., Punt, P.J., Debets, A.J.M., Stouthhamer, A.H. and van den Hondel, C.A.M.J.J. 1994. Evaluation of the molecular and genetic approaches to generate glucoamylase overproducing strains of *Aspergillus niger*. *Journal of Biotechnology* 36: 165-175

Vihinen, M. and Mantsala, P. 1989. Microbial amylolytic enzymes. *Crit. Rev. Biochem. Mol. Biol.* 24(4):329-418

Visser, J. and Ruijter, G.J.G. 1997. Carbon repression in *Aspergilli*. *FEMS Microbiology Letters* 151: 103-114

Ward, O.P. 1989. Fermentation biotechnology. Prentice Hall, Englewood Cliffs, NJ.

Wernars, K., Goosen, T., Wennekes, B.M., Swart, K., van den Hondel, C.A. and van den Broek, H.W. 1987. Cotransformation of *Aspergillus nidulans*: a tool for replacing fungal genes. *Mol. Gen. Genet.* 209(1): 71-7

Wirsel, S., Lachmund, A., Wildhardt, G. and Ruttkowski, E. 1989. Three alpha-amylase genes of *Aspergillus oryzae* exhibit identical intron-exon organization. *Mol Microbiol.* 3(1): 3-14

Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S.S., Michael, M.Z. and Graham, M.W. 1989. Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res.* 17(9): 3469-78

Xing, Y., Fikes, J.D. and Guarente, L. 1993. Mutations in yeast HAP2/HAP3 define a hybrid CCAAT box binding domain. *EMBO J.* 12(12): 4647-55

Yanisch-Perron, C., Vierra J. and Messing J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33: 103-19