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.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

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

UMI®

Interactions Between the CREA Protein and the *glaA* Promoter Region of *Aspergillus niger*

Catherine Au

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 2002

© Catherine Au, 2002

National Library of Canada

Acquisitions and Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada

Acquisitions et services bibliographiques

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.

0-612-68403-2

Canada

ABSTRACT

Interactions Between the CREA Protein

and the *glaA* Promoter Region of *Aspergillus niger*

Catherine Au

I constructed a genomic DNA library of *Aspergillus niger*, and determined that it contained 50 000 independent clones of an average size of 14 kb. This library contained approximately twenty genome equivalents. I also constructed an expression vector using the plasmid pGEX-2T to express the fusion protein GST:CREA. This construct contained the full coding sequence of the *creA* gene of *A. niger*. The GST affinity tag is located as an amino-terminal fusion with respect to CREA. Induction of the fusion protein was performed under various conditions, and it was concluded that a construct containing a carboxy-terminal affinity tag would ameliorate the yield of full length protein, along with the expression of the construct in the Rosetta © strain of cells produced by Novagen. The protein was purified over a Glutathione Sepharose 4B column from Amersham Pharmacia Biotech, and used for electrophoretic mobility shift assays with oligonucleotides containing the CREA consensus binding sequence, and mutated segments of the *glaA* promoter of *A. niger*. Shifted bands were observed in the presence of the purified protein, and competition between labeled and unlabeled DNA indicated clearly that the protein was binding to the CREA consensus binding sequence.

ACKNOWLEDGEMENTS

I wish to thank my very tolerant supervisors Dr. Reg Storms and Dr. Adrian Tsang for providing me with the opportunity to perform research in their laboratories. It is only with their support that this project was completed, and I wish to express my gratitude for their forbearance when progress stalled. I also wish to thank my colleagues for the invaluable theoretical and technical assistance that helped me troubleshoot my experiments. Rosa Zito deserves special thanks for providing the mutated *glaA* promoter sequences. Thank you to Susan Sillaots, Amalia Martinez, Dr. Hongshen Li, Aleks Spurmanis, Dr. Peter Ulycznyj, and Edith Munro for answering all my odd questions. For moral support, I wish to thank Nathalie Brodeur, Lita Glowacki, Jill Heather Flegg, Arthur Lo, Chris Taillefer, Wade Colborne and Kirk Duplessis. All of you helped encourage me when the research and the thesis seemed interminable. As well, thank you to my parents and my sister for understanding my odd schedule. Finally, thank you to everyone in the Biology Department who offered assistance and made my life just a little bit easier.

CONTENTS

LIST OF FIGURES ix

LIST OF TABLES xii

ABBREVIATIONS xiii

CHAPTER 1. INTRODUCTION 1

1.1 *Aspergillus* and *Aspergillus niger* 1

1.1.1 Life cycle, growth and habitat 1

1.1.2 Industrial uses 5

1.2 Protein Production 6

1.2.1 Bacterial hosts for heterologous protein production 6

1.2.2 Benefits of fungal hosts 6

1.2.3 Current research 7

1.3 Carbon catabolite repression 8

1.3.1 Carbon catabolite repression in *Sacchromyces cerevisiae* 8

1.3.2 Carbon catabolite repression in *Aspergillus* 12

1.3.2.1 Proteins involved in carbon catabolite repression 12

1.4 Control of the glucoamylase gene in *Aspergillus* 13

1.4.1 Glucoamylase and its promoter region 13

1.4.2 Proteins involved in the regulation of the *glaA* promoter 14

1.4.2.1 ANCF 14

1.4.2.2 *creA* and its gene product CREA 17

1.5 *glaA* promoter region constructs 28

1.5.1 Rational behind development of constructs 28

1.5.2 Mutated consensus sequences in the promoter region 28

1.6 Rational for thesis 33

1.6.1 Outline of experimental ideas 35

CHAPTER 2. MATERIALS AND METHODS 36

2.1 Growth conditions and manipulation of *E. coli* and *A. niger* 36

2.1.1 Growth and propagation of *E. coli* 36

2.1.2 Chemically competent *E. coli* cells 36

2.1.3 Transformation of *E. coli* 37

2.1.4 Growth and propagation of *A. niger* 37

2.2 Genomic library 38

2.2.1 Preparation of pBluescript vector 38

2.2.2 Preparation of genomic DNA 38

2.2.3 Ligation 41

2.2.4 Transformation of Sure2 Ultracompetent cells 41

2.2.5 Characterization of library 42

2.2.6 Harvesting library 42

2.3 Creation of GST:CREA fusion constructs 43

2.3.1 PCR amplification of *creA* DNA sequence for inserts 43

2.3.2 Preparation of pGEX-2T vector 46

2.3.3 Preparation of *creA* insert DNA 46

2.3.4 Ligation of insert DNA into vector 46

2.3.5 Transformation into XL1-Blue competent cells 49

2.3.6 Identification and characterization of constructs 49

2.4 Induction, purification, and confirmation of presence of GST:CREA protein 49

2.4.1 Growth conditions for induction and preparation of bacterial lysate 49

2.4.2 Processing and purification of GST:CREA 50

2.5 Manipulation of *glaA* construct DNA 51

2.5.1 PCR amplification of a 227 bp region of the *glaA* promoter region 51

2.5.2 Clean up of PCR reactions 52

2.6 Electrophoretic mobility shift assays (EMSA) 52

2.6.1 20mer oligonucleotides used to optimize reaction conditions 52

2.6.2 Labeling of DNA fragments with [ã^{32}]P-ATP 53

2.6.3 Preparation of acrylamide gels and reaction mixes 53

2.6.4 Running conditions for gel shift assays 54

2.6.5 Fixing, drying, and exposure of gels to X-ray film 54

CHAPTER 3. RESULTS 56

3.1 Genomic Library 56

3.1.1 Genomic DNA inserts 56

3.1.2 pBluescript vector 56

3.1.3 Characterization of library 56

3.2 Fusion constructs 63

3.2.1 *creA* insert DNA 63

3.2.2 pGEX-2T vector 63

3.2.3 Characterization of transformants 69

3.3 GST:CREA purification 69

3.3.1 Analysis of the purification process 69

3.4 Amplified 227 bp fragments of mutated *glaA* promoter region constructs 82

3.4.1 Confirmation of successful PCR amplification, clean up and quantification 82

3.5 Electrophoretic mobility shift assays 82

3.5.1 Optimization of reaction conditions 82

3.5.2 EMSA with mutated *glaA* promoter constructs and purified GST:CREA protein 85

CHAPTER 4. DISCUSSION 95

4.1 The *A. niger* genomic library 95

4.2 Expression of the GST:CREA fusion protein of *A. niger* 95

4.3 EMSA with GST:CREA 98

4.3.1 EMSA with 20mer oligonucleotides 98

4.3.2 EMSA with mutated *glaA* promoter regions 99

4.4 Future avenues of inquiry 102

4.4.1 GST:CREA fusion protein 102

4.4.2 CREA binding to mutated *glaA* promoter sequences 103

LITERATURE CITED 104

LIST OF FIGURES

Figure 1	Life cycle of *A. niger* p. 4
Figure 2	Diagrammatic representation of Mig1p repression in *S. cerevisiae* p. 10
Figure 3	The 5' promoter region of the *A. niger glaA* gene p. 16
Figure 4	Sequence of the *creA* gene of *A. niger* including 5' and 3' untranslated sequences p. 20
Figure 5	Aligned coding sequence and translated amino acid sequence of the *creA* gene of *A. niger* p. 22
Figure 6	Amino acid sequence of zinc finger regions Finger 1 and Finger 2 of the CREA protein of *A. niger* p. 30
Figure 7	5' promoter region of the *glaA* gene of *A. niger* used in *glaA-lacA* fusion vector with CREA binding sites and HAP binding sites indicated p. 32
Figure 8	A map of the pBluescript SK- plasmid p. 40
Figure 9	Visual representation of PCR amplified *creA* sequence with engineered flanking *Bam*HI sites p. 45
Figure 10	Plasmid map of the GST-fusion expression vector pGEX-2T p. 48

Figure 11	1% agarose gel in 1X TBE containing samples from sucrose gradient fractions p. 58
Figure 12	Purified genomic DNA fragments of approximately 11 kb in size p. 60
Figure 13	Plasmid map of a genomic library clone p. 65
Figure 14	1% agarose gel containing digested DNA of 10 independent genomic clones p. 67
Figure 15	Selection of constructs containing insert DNA p. 71
Figure 16	Plasmid map of the pGCreA construct containing insert DNA oriented in the correct direction p. 73
Figure 17	Plasmid map of the pGCreA construct containing insert DNA oriented in the wrong direction p. 75
Figure 18	A 1% agarose gel of restriction digests with *Xho*I and *Mlu*I p. 77
Figure 19	Coomassie blue stained 12% acrylamide gel of samples from each step of purification p. 79
Figure 20	Western blot with anti-GST IgG, of samples from each step of purification p. 81
Figure 21	Three hour exposure of EMSA performed with 20mer DNA fragments and crude bacterial lysate p. 84

Figure 22	2.5 hour exposure of EMSA performed with 20mer DNA fragments and both purified GST:CREA protein and crude bacterial lysate p. 87
Figure 23	Five hour exposure of EMSA with D10 DNA (wild-type sequence *glaA* promoter region) and purified GST:CREA p. 89
Figure 24	Three hour exposure of EMSA with purified GST:CREA protein and all of the versions of the 227 bp mutated *glaA* promoter region p. 91
Figure 25	Three hour exposure of EMSA performed with purified protein and various mutant versions of the 227 bp mutated *glaA* promoter region p. 94

LIST OF TABLES

Table 1 Mutated *glaA* promoter constructs p. 34

Table 2 Reaction mixes for EMSA p. 55

Table 3 Confirmation that pBluescript backbone was properly prepared for ligation with genomic DNA fragments p. 61

Table 4 Analysis of genomic library p. 62

Table 5 Restriction digest of genomic library clones with *Xho*I and *Bsu*151 (*Cla* I) p. 68

ABBREVIATIONS

ATP	adenosine triphosphate
bp	basepair
ddH_2O	distilled, deionized water
dNTP	deoxyribonucleoside triphosphate
DTT	dithiothreitol
EDTA	ethylene diamine tetraacetic acid
EMSA	electrophoretic mobility shift assay
fmol	femtomole
GRAS	generally recognized as safe
GST	glutathione-s-transferase
HEPES	N-2-Hydroxyethylpiperazine-N'-2-ethanesulfonic acid
IPTG	isopropyl-beta-D-thiogalactopyranoside
kDa	kilodalton
i g	microgram
mg	milligram
i l	microlitre
ml	millilitre
mM	millimolar
M	molar
mRNA	messenger ribonucleic acid
ng	nanogram
nmol	nanomole

ORF	open reading frame
PCR	polymerase chain reaction
pmol	picomole
poly(dI-dC)	poly-deoxyinosine-deoxycytidine
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

1. Introduction

1.1 *Aspergillus* and *Aspergillus niger*

1.1.1 Life cycle, growth, and habitat.

The phylum *Ascomycetes* and the phylum *Deuteromycetes* form the largest and most diverse group of fungi (Carlile and Watkinson, 1994). I mention both these phyla together since the taxonomy of their constituent species is somewhat confusing. *Ascomycetes* are defined by their production of haploid sexual spores within an eponymous ascus (Worral, 1999). *Deuteromycetes* is considered to be a form division of *Ascomycetes*, and its members are classified by their lack of a sexual reproductive cycle (Worral, 1999).

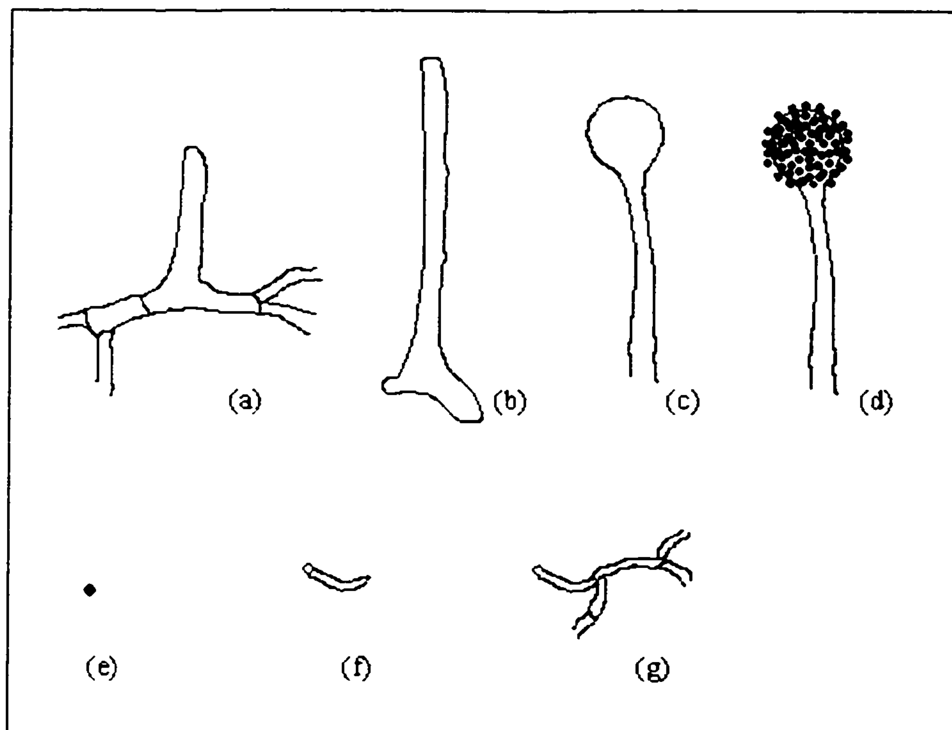
The genus *Aspergillus*, first categorized by Micheli in 1792, is one of the more commonly known members of the phylum *Ascomycetes* (Raper and Fennell, 1973). In 1945, A Manual of the *Aspergilli* was published by Thom and Raper. This was the first attempt to systematically categorize for identification purposes, the varied species of *Aspergilli*. Since then many more species have been identified, and the knowledge regarding the importance of the *Aspergilli* in our environment, health, and industry has grown astronomically (Raper and Fennell, 1973; Worrall, 1999; Stahl and Tudzynski, 1992). One of these industrially important species is *Aspergillus niger*, prized for being a non-citrus, and therefore an economically favourable, source of citric acid (Ingold, 1961). In nature, *Aspergilli* are saprophytic fungi, living in the soil and deriving nourishment

from various organic matter, which results in the decomposition and the return of essential nutrients to the soil (Webster, 1970; Ingold, 1960). *Aspergilli* are also responsible for various infections of the human and animal populations. Potentially lethal acute invasive pulmonary aspergillosis, *Aspergillus* asthma, and allergic bronchopulmonary aspergillosis are all possible consequences of inhalation of the fungal conidia (Tomee and van der Werf, 2001).

A. niger is an *Ascomycetes* that is known to reproduce solely in an asexual manner (Webster, 1970; Raper and Fennell, 1972). At present, two varieties of this species are recognized: *A. niger* var. *awamorii*, also referred to as *A. awamorii*; and *A. niger* var. *macrosporus* (NCBI Taxonomy Browser). The asexual cycle is characterized by two stages: mycelia, which are composed of intertwined septate hyphae; and conidia, which are uninucleate single-celled entities (Ingold, 1961; Webster, 1970; Raper and Fennell, 1973). Easily carried by air currents, conidia are derived from hyphal cells (Ingold, 1961; Webster, 1970; Raper and Fennell, 1973). Upon translocation to suitable conditions, the conidia will germinate to form hyphae, which will in turn anastomose to create the mycelial mass (Raper and Fennell, 1973). Figure 1 illustrates this cycle.

A. niger is characterized by vivid black conidial heads (Raper and Fennell, 1973). This colour is dependent upon the level of copper present in the media, and the spores can range from yellow to brown to black as the copper concentration increases from 1 i g to 20 i g per liter of culture volume (Mulder, 1938). A copper concentration greater than or

Figure 1. Life cycle of *Aspergillus niger*. (a) Foot cell beginning to extend in direction perpendicular to hyphal growth. (b) Development of conidiophore. (c) Development of vesicle by swelling of terminal portion of the conidiophore. (d) Conidia develop on fruiting head and are dispersed by air currents. (e) A conidium. (f) Conidium germinates and produces hypha. (g) Hyphae grow and anastamose, forming mycelial mass. (Adapted from Thom and Raper, 1945)


(a)
(b)
(c)
(d)
(e)
(f)
(g)

equal to 20 i g per liter of media results in black coloured conidia.

1.1.2 Industrial uses.

A. niger has a long history of use in the food industry. Used to ferment soybean mash in the production of soya sauce and miso for hundreds of years, most *A. niger* derived food products have been given the designation of GRAS (Generally Recognized As Safe) by the U.S. Food and Drug Administration, which allows their use without further safety testing (http://www.cfsan.fda.gov/~rdb/opa-gras.html). In its lyophilized form, it is used as a nutritional additive in animal feed and human dietary supplements.

The industrial production of organic acids uses *A. niger* as an economically viable producer of citric acid as well as gallic and gluconic acids (Ingold, 1961; Lockwood 1975). It has been used for the mass synthesis of industrial enzymes such as amylase, glucoamylase, cellulase, pectinase, glucose oxidase, and catalase for over three decades (Underkofler, 1976; Kirimura *et al.*, 1999). For a more current example, the Lyven company of France produces pectinases from *A. niger* that are subsequently used in the clarification of fruit juice as well as á-amylases used in bread making to increase loaf volume (http://www.lyven.com/products.htm).

Aspergilli can be grown on both solid and liquid media. Due to its ability to utilize virtually anything organic as food, the fungi can be grown on such varied substrates as sawdust, processed beet pulp mash, newspaper, or garbage. It can survive a wide range

of temperatures. and requires little maintenance to sustain viability. This makes it a desirable organism to work with in industry and bioresearch.

1.2 Protein production

With the advent of the biotechnology industry and cloning, *Aspergilli* have been found to have another novel use, heterologous protein production.

1.2.1 Bacterial hosts for heterologous protein production.

When producing recombinant proteins, the original host of choice was *Escherichia coli*. With a generation time of approximately 20 minutes. and the wide availability of selectable markers to identify transformants. *E. coli* possessed many benefits. Problems, however, were encountered when expression of eukaryotic and especially mammalian proteins was desired. Lacking organelles. prokaryotes are unable to properly process eukaryotic proteins. Post-translational modifications such as trimming, refolding and glycosylation do not occur. resulting in partially functional. or more frequently, non-functional protein products. As well. eukaryotic mRNA sequences contain codons requiring tRNAs that are represented at low levels in *E. coli*. This interferes with the length of translated proteins due to stalling of the ribosomes.

1.2.2 Benefits of fungal hosts.

Using a eukaryotic host such as the *Aspergillus* fungi solved these problems. Though a eukaryote, filamentous fungi still possess some of the same advantages as bacterial hosts: they can be grown in large batches, require little hands on maintenance, and have simple

nutritional requirements. In fact, fungal hosts are more hardy and resilient than bacterial hosts due to their ability to utilize almost anything of an organic nature as sustenance, and their tolerance of a wider range of temperatures.

Aspergilli make especially good hosts for two reasons. First, secretion of heterologous proteins can be engineered. This allows harvest of the desired protein from the media and dispenses with the bother of having to crush the mycelia to extract protein. Second, certain species of *Aspergillus*, such as *A. niger*, secrete proteins at a very high level under the influence of certain promoter sequences: some proteins can be secreted at levels of over 30 grams per liter of culture as compared to 30 mg per liter of culture for bacteria (Ward, 1989; Dunn-Coleman *et al.*, 1991). This high level of production is mediated in part by very strong promoter sequences, such as those involved in the production of inducible enzymes.

1.2.3 Current research

Currently, identification of promoter sequences that mediate high levels of expression are still in progress. Researchers are investigating a number of areas, such as transcriptional regulation, and the search for novel promoter sequences.

The regulation of protein production in filamentous fungi is also a field of interest. Eukaryotic mechanisms of gene repression and induction can be studied in these organisms, and insight into the molecular mechanisms by which transcriptional control is effected can be elucidated from the incipient data. Various data can be derived regarding

DNA-protein interactions, such as the specific DNA consensus binding sequences for different repressors and transcriptional activators. Comparisons between *in vivo* and *in vitro* gene regulation are another facet of this research. This information is not just of use to scientists, it may one day lead to advances in medicine by providing the means to fight fungal infections or turn off cancer genes.

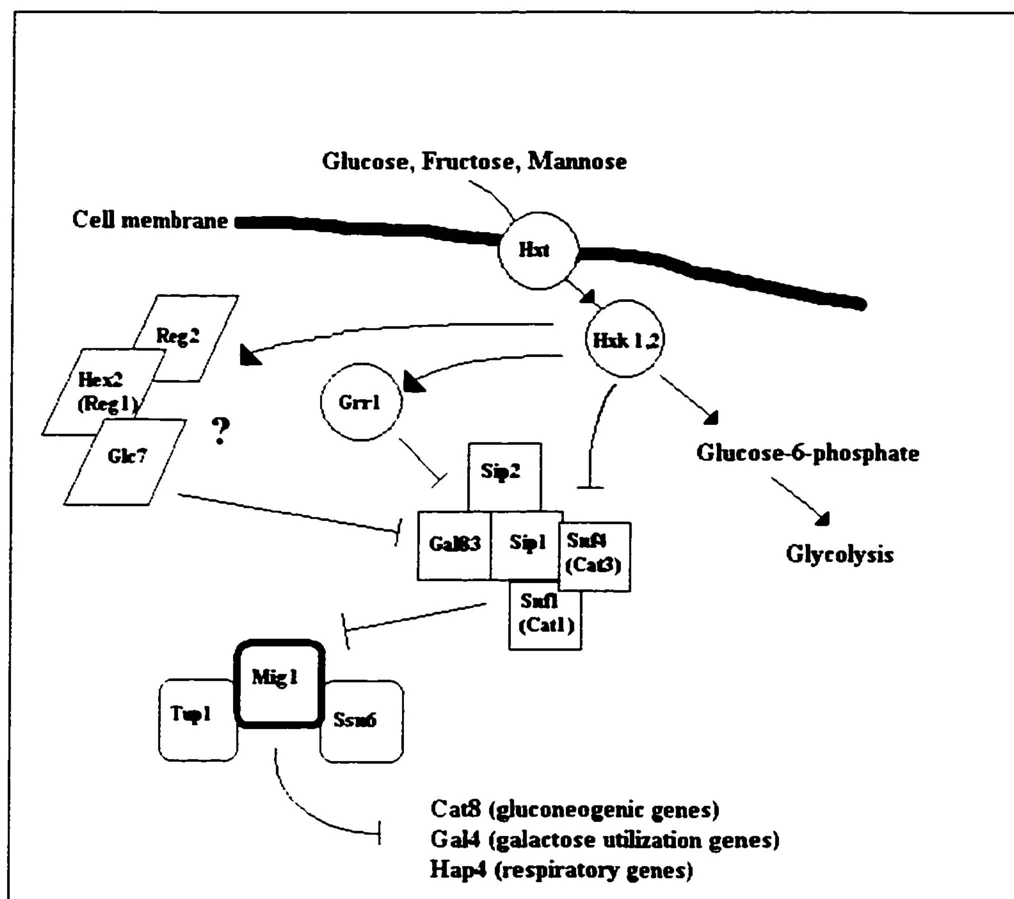
1.3 Carbon catabolite repression

1.3.1 Carbon catabolite repression in *Saccharomyces cerevisiae.*

The term carbon catabolite repression refers to the ability of an organism to suppress production of proteins required for the utilization of less favourable substrates when provided with a preferred substrate. This process has been closely observed in yeast. Preferential substrates are simple fermentable sugars such as glucose or fructose. When these sugars are present, production of various enzymes such as those involved in gluconeogenesis and the catabolism of other carbon sources, are repressed (Hohmann and Mager, 1997).

Figure 2 provides a visual summary of the different enzymes involved in carbon catabolite repression in the presence of glucose or other preferential carbon sources. Mig1p complex directly represses genes involved in gluconeogenesis, galactose utilization, and respiration (Treitel and Carlson, 1995). The Mig1p complex is not repressed by the Snf1p complex when a preferable carbon source such as glucose is

Figure 2. (Figure adapted from Hohmann and Mager, 1997) Diagrammatic representation of Mig1p repression in *S. cerevisiae*. Mig1p repression is mediated by a number of interconnected pathways.


Glucose, Fructose, Mannose
Cell membrane
Hxt
Hxk 1,2
Reg2
Hex2
(Reg1)
Glc7
?
Grr1
Sip2
Gal83
Sip1
Snf4
(Cat3)
Snf1
(Cat1)
Glucose-6-phosphate
Glycolysis
Mig1
Tup1
Ssn6
Cat8 (gluconeogenic genes)
Gal4 (galactose utilization genes)
Hap4 (respiratory genes)

present in the growth media (Celenza and Carlson, 1986). The Snf1p complex is able to repress the Mig1p complex when activated by its positive regulator Snf4p (Celenza and Carlson, 1989). Negative control of the Snf1p complex is enacted via a number of upstream mediators such as Hxk1p, Hxk2p, Grr1p, and the Hex2p complex (Gancedo, 1992; Celenza and Carlson, 1984).

Mig1p is coded for by the *Mig1* gene, a 1.5 kb coding sequence which produces a 504 aa protein (Nehlin and Ronne, 1990). To act as a functional repressor, the protein must contain two C_2H_2 zinc-finger regions at the amino terminal end of the protein, and a 24aa effector domain at the carboxy-terminal end (Nehlin and Ronne, 1990; Östling, 1998). Proteins with this highly conserved zinc-finger region are evolutionarily related DNA-binding proteins (El-Baradi and Pieler, 1991). The zinc-finger region of Mig1p bears strong similarity to that of two other DNA binding proteins: Egr, a mammalian DNA binding finger protein induced in the early growth response, and the Wilms' tumour protein (Herschman, 1991; Haber *et al.*, 1992). All of these proteins are known to bind to similar DNA sequences -- referred to as GC boxes -- in the promoter regions of the genes they control (Nehlin and Ronne, 1990; Mercado *et al.*, 1991; Herschman, 1991; Haber *et al.*, 1992). Mig1p binds to the consensus sequence 5'-SYGGRG-3' (S is G or C; Y is T or C; R is G or A), with each finger interacting with a base triplet (Flick and Johnston, 1992; Nardelli *et al.*, 1991). An AT rich sequence flanking the 5' end of the GC box is required for Mig1p binding, probably due to the AT rich segment bending more easily to accommodate the protein (Lundin *et al.*, 1993).

1.3.2 Carbon catabolite repression in *Aspergillus*

Much less is known regarding the mechanisms of carbon catabolite repression in the *Aspergilli*. As was seen in yeast, the mechanism of carbon catabolite repression in *Aspergilli* regulates specialized degredative enzymes, gluconeogenic enzymes, and repiratory enzymes as well (Ruijter and Visser, 1997). Using mainly *A. nidulans* as the model organism, the proline utilization cluster, arabinase repression pathway, ethanol regulon, glucoamylase repression pathway, and xylanase repression pathway have all been examined in varying levels of detail. While studying the mechanisms of AREA (a wide domain activator protein involved in nitrogen metabolism) in *A. nidulans*, researchers accidentally discovered that there was also a protein that globally mediated carbon catabolite metabolism – they named it CREA (Arst and Cove, 1973; Hynes and Kelly, 1977).

1.3.2.1 Proteins involved in carbon catabolite repression

The CRE protein is a global regulator of carbon catabolite repression in various fungi. Known as CREA in *A. nidulans* and *A. niger*, and CRE1 in *Sclerotinia sclerotiorum*, these proteins all contain two C_2H_2 zinc-finger domains close to their amino-termini (Dowzer and Kelly, 1991; Drysdale *et al.*, 1993; Vautard *et al.*, 1999). Proving functional identity, CRE1 of *S. sclerotiorum* was shown to complement CREA of *A. nidulans*, when the *cre1* gene was transformed into a *creA*-nonfunctional strain of *A. nidulans*, resulting in a return to a wild type phenotype (Vautard *et al.*, 1999).

1.4 Control of the glucoamylase gene in *Aspergillus*

1.4.1 Glucoamylase and its promoter region

The *glaA* gene of *A. niger* codes for glucoamylase; a secreted degradative enzyme. The enzyme is found in two forms that are produced by the differential splicing of an intron (Boel *et al.*, 1984; Nunberg *et al.*, 1984)). Tightly repressed in the presence of a xylose concentration above 20 mg· l^{-1}, maximal expression of *glaA* is observed when starch or maltose is the sole carbon source, and lower levels of expressed protein are observed in the presence of a glucose concentration above 10 mg· l^{-1} (Fowler *et al.*, 1990; Santerre Henriksen, 1999). When grown on mixed carbon source media containing maltose and xylose, or glucose and xylose, expression of *glaA* is unaffected, however, when grown on a mixed carbon source of starch and xylose, expression is reduced to minute levels, thus implying the existence of a complicated regulatory mechanism (Fowler *et al.*, 1990). Using GFP as a reporter gene, it was demonstrated that the *glaA* promoter region is not completely repressed in the presence of xylose, and that previous conclusions regarding total repression were due to lack of sensitivity of the assay used (Santerre Henriksen, 1999; Fowler *et al.*, 1990).

When the copy number of a plasmid containing a *glaA* promoter region fused to a reporter gene was increased, a relationship between promoter copy number and level of expression was demonstrated: increased copy number resulted in reduced expression (Verdoes, 1994; Verdoes *et al.*, 1994). Xylose induced repression, however, was not affected by copy number (Verdoes, 1994; Verdoes *et al.*, 1994).

The sequence between –318 and –562 with respect to the ATG was shown to be necessary for high-level expression by testing various truncations of the *glaA* promoter region fused to a reporter gene under three different growth conditions: starch, starch/xylose, and maltose/xylose (Fowler *et al.*, 1990). It was determined that basal expression levels depend on sequences within the region between –1 and –224 (Fowler *et al.*, 1990). In *A. oryzae*, the promoter region necessary for high-level expression in the presence of maltose was determined to encompass the bases –332 to –427, and the region necessary for basal expression was determined to be from –210 to –218 (Hata *et al.*, 1992). Please refer to Figure 3 for the sequence of the 5' promoter region of the *glaA* gene of *A. niger*.

1.4.2 Proteins involved in regulation of the *glaA* promoter

The *glaA* promoter in *A. niger* is subject to repression by CREA, the previously mentioned global regulator of carbon catabolite repression (Ruijter and Visser, 1997). The *creA* gene and CREA protein will be discussed in greater detail in the following sections.

1.4.2.1 ANCF

First identified and named in the early 1990's, AnCF was identified as a CCAAT-binding element that bound to a site in the promoter of *amdS* in *A. nidulans* (van Heeswijck and Hynes, 1991). Over time, it was discovered that this CCAAT-binding protein complex was related to both the HAP protein complex of yeast and to other eukaryotic CCAAT-binding proteins (Kato *et al.*, 1997; Kato *et al.*, 1998; Steidl, 1999). It soon became

Figure 3. The 5' promoter region of the *A. niger glaA* gene. Numbering counts backward from ATG of the *gla*A gene. Putative **CCAAT** (HAP binding) sites are indicated with pale gray highlighting. Putative SYGGRG (CREA binding) sites are indicated with black highlighting. [illegible] sites are indicated with dark gray highlighting. ***Restriction enzyme sites*** are indicated by bold italic lettering.

```
                                                                  HAP site CRE site
-695  ATATGAGTTC ATCCTGCAGA ATACCGCGGC GTTCCACATC TGATGCCATT G[illegible]GGG

-635  TCCGGACGGT CAGGAACTTA GCCTTATGAG ATGAATGATG GACGTGTCTG GCCTCGGAAA

                                                                        MluI site
-575  AGGATATATG GGGATCATAA TAGTACTAGC CATATTAATG AAGGGCATAT ACCACGCGTT

                                                                    HAP site
-515  GGACCTGCGT TATAGCTTCC CGTTAGTTAT AGTACCATCG TTATACCAGC CAATCAAGTC

                 CRE site        CRE site
-455  ACCACGCACG ACCGGGCACG GCGAATCCCC GGGAATTGAA AGAAATTGCA TCCTAGGCCA

                 HAP site
-395  GTGAGGCCAG CGATTGGCCA CCTCTCCAAG CACACAGGGC CATTCTGCAG CGCTGGTGGA

                   CRE site
-335  TTCATCGCAA TTTCCCCCGG CCCGGCCCGA CACCGCTATA GGCTGGTTCT CCCACACCAT

-275  CGGAGATTCG TCGCCTAATG TCTCGTCCGT TCACAAGCTG AAGAGCTTGA AGTGGCGAGA

-215  TGTCTCTGCA GGAATTCAAG CTAGATGCTA AGCGATATTG CATGGCAATA TGTGTTGATG

-155  CATGTGCTTC TTCCTTCAGC TTCCCCTCGT GCAGATGAGG TTTGGCTATA AATTGAAGTG

-95   GTTGGTCGGG GTTCCGTGAG GGGCTGAAGT GCTTCCTCCC TTTTAGACGC AACTGAGAGC

                 CRE site
-35   CTGAGCTTCA TCCCCAGCAT CATTACACCT CAGCCATG
```

evident that these HAP-like protein complexes had already been identified in other species of fungi and had been given their own names: PENR1, AnCP, CBF, HAP, or NF-Y (Kato *et al.*, 1997; Kato *et al.*, 1998; Litzka *et al.*, 1998; Brakehage *et al.*, 1999; Papagiannopoulos, 1996). It has been demonstrated in *S. cerevisiae* that the *hap* gene products are involved in activating genes in the respiratory pathways, and it has been shown in *A. nidulans* that the related protein complex enhances similar gene expression (Kato *et al.*, 1997; Bonnefoy *et al.*, 1995). This, along with other data, suggests that the HAP-like complex functions as a transcriptional activator by binding to CCAAT sequences in the promoter regions of target genes. Binding sites for this HAP-like complex have been experimentally verified in the *taa* promoter of *A. oryzae*, as well as the *taa*, *amdS*, *gatA*, *ipnA* and *acvA* promoters of *A. nidulans* (Litzka *et al.*, 1998; Kato *et al.*, 1997; Nagata *et al.*, 1992; van Heeswijck and Hynes, 1991). Although putative HAP-like protein complex binding sites have been identified in the *glaA* gene of *A. niger*, none have been confirmed experimentally. Refer to figure 3 for the locations of the putative CCAAT binding sites in the *glaA* promoter.

1.4.2.2 *creA* and its gene product CREA

The *creA* genes of both *A. nidulans* and *A. niger* have been cloned and sequenced (Dowzer and Kelly, 1991; Drysdale *et al.*, 1993). The *creA* gene of *A. nidulans* contains an intron-less ORF of 1245bp, resulting in a protein of 415aa (Dowzer and Kelly, 1991). In *A. niger*, the *creA* gene contains an intron-less 1281bp ORF, and codes for a protein of 427aa (Drysdale *et al.*, 1993). At the amino-terminal end, CREA has two zinc-finger domains of the Cys_2His_2 variety that bear great similarity to the zinc-finger domains of

Mig1p: these are followed by an Ala-rich region and S(T)PXX motifs, both features common to DNA binding proteins (Nehlin and Ronne, 1990; Dowzer and Kelly, 1991; Drysdale *et al.*, 1993). When the *creA* sequences of *A. nidulans* and *A. niger* are aligned, there is an 82% sequence similarity and 90% identity at the amino acid level (Dowzer and Kelly, 1991). Complementation was observed when an *A. nidulans* strain containing a mutated *creA* gene was transformed with a plasmid containing the *A. niger* homologue (Drysdale *et al.*, 1993). This demonstrated the functional identity of both CREA proteins. More interestingly, CREA also contains a region of similarity to Rgr1p of *S. cerevisiae* that is located between amino acids 284 and 325 (Drysdale *et al.*, 1993). The Rgr1p protein of *S. cerevisiae* is involved in carbon catabolite repression, and it has been hypothesized that CREA combines functional domains of both Rgr1p and Mig1p in one protein (Sakai *et al.*, 1990; Drysdale *et al.*, 1993). Please refer to Figure 4 for the coding sequence of the *creA* gene of *A. niger* with flanking 5' and 3' sequences, and to Figure 5 for the amino acid sequence of the *A. niger* CREA protein aligned with the coding sequence of the gene.

Various homologues of the CRE protein have been isolated in other genera of fungi. Complementation experiments have demonstrated that the CREA proteins of *Gibberella fujikuroi* and *Botrytis cineraea*, and the CRE1 proteins of *Trichoderma reesei* and *Sclerotinia sclerotiorum*, are functionally equivalent to the CREA protein of *A. nidulans*, and therefore functionally equivalent to the CREA protein of *A. niger* (Drysdale *et al.*, 1993; Strauss *et al.*, 1995; Vautard *et al.*, 1999; Tudzynski *et al.*, 2000). Thus,

Figure 4. Sequence of the *creA* gene of *A. niger* including 5' and 3' untranslated sequences (Drysdale *et al.*, 1993). Numbering begins from the ATG. Negative numbers count backwards from the ATG. The TAG stop codon and the ATG are indicated in **bold type**. The ORF is indicated by capital letters and the untranslated regions are indicated by lowercase letters. The three underlined regions indicate regions of similarity between the *creA* promoter regions of *A. niger* and *A. nidulans*. Each different style of underlining indicates a different region of similarity.

-1038 aagagatcat attcatcaga gaacacaaga tctgagagag tgtgtaagag cggtagatac
-978 ccaagattgt aaaatcatct atcaagtcaa gaaaattata aaaagaataa gaagagcatg
-918 aaataactac aaagcactgg gttccatctg actgggaaag ctgactgacc ggtttgaagc
-858 cccaagaagc aaagcagccg atctggttca agcacgttct ttttcccttt cctgcttttt
-798 ttgccccccc gatatttatc cccacacaaa gtacatagtc ttttcttttt tcgatttttt
-738 tttttaaatc cctttttttt tattcttatt tccctttcg taattttttt tatccatctt
-678 tttttcgatt ttttcaattt tctttttcct tttttttttc tttttttttt ttttttttctt
-618 cgcgttccca ttctgctctc cgattccgat aacccacccc ctctacgact cgcccctcttt
-558 tccctcccct cccccgaatt ccgtttcctt cttcttcccc tccattcctc atcttttcgc
-498 cctttccgat ttcttctctt cttatatctt cgtctcccca gatcatcttc tccaggtttt
-438 cttcctccct ccctcttttc gagaccattt gctcaacatc acccttgccc gactcgctac
-378 ttattacccc ggggtccatt tccgattccg gctcacccaa cacttatcat aactaccaca
-318 ttccgtatcc cttcaataat tgaaaggaat tcgtcgttat cagcccttgg ttgaattcac
-258 aattgaatac cgcctcgcct tggttcgtca tcagggccct gtgatcctca ctacccttgc
-198 ggtcttcctc ctatcatctc tccttgccaa ggttaggtct ggggtgctag cgggaggtga
-138 tcgactggcc gatatagcta cccctggtcg cgacgtgcat cactatcgcc tcggcttctc
-78 atcccttccg aacgtgtcct ttttaaaccc ggctctgtcc taccttttta ccctcgtcgg
-18 tcttgggaca agcttcac**AT G**CCGCCACCG GCCTCTTCAG TAGATTTCTC AAATCTGCTG
43 AACCCTCAAA ACAATTCCAC TGATTCCACC CCTTCCACTC CTGTGGACAG CTCCAAGACC
103 CCTTCTACTC CGTCCAGTAC ACAGTCGAAC TCCAACATGG CGTCGTCAGT GAGCCTACTT
163 CCGCCGCTTA TGAAGGGTGC CCGGCCTGCG ACCGAAGAAG TTCGTCAGGA TCTTCCCCGC
223 CCCTACAAGT GCCCCCTCTG CGATCGCGCA TTCCACCGTC TGGAGCACCA GACAAGACAC
283 ATCCGCACGC ACACGGGTGA AAAGCCGCAT GCCTGTCAAT TCCCCGGCTG CACTAAGCGC
343 TTCAGTCGTT CTGACGAGCT TACGCGCCAC TCGCGAATCC ACAACAACCC TAACTCGAGA
403 CGTCGAAACA AGGCTCAGCA TCTGGCTGCG GCCGCTGCAG CTGCTGCAGG GCAGGACAAC
463 GCCATGGCCA ACACCGCCAG CGCTATGATG CCTCCTCCCA GCAAGCCAAT GACCCGGTCG
523 GCCCCTGTCT CGCAGGTCGG CTCCCCCGAC ATCTCTCCCC CTCACTCCTT CTCCAACTAC
583 GCCAGTCACA TGCGATCGAA CTTGGGCCCG TATGCTCGCA AAGGCGACGA GGCATCGTCG
643 GGCATGGAAC TATACCTCCT GGCGACTGCT GCGTCGCAGG TGGAACGCGA CGAACACTTC
703 GACTTCCATG CAGGGCCACG CAATCACCAT TTGTTCAGCT CGCGCCACCA TGGAAGTGGC
763 CGTCTGCCCC TCCTTGCTGC TTACGCGATC ACACACAACA TGAGCCGGTC GCATTCACCT
823 GAAGACGACG ACGGTTATTC GCACCGTGTG AAACGTTCAA GACCCAACTC TCCCAACTCG
883 ACTGCTCCTT CCTCTCCCAC CTTCTCCCAC GACTCCCTCT CACCCACTCC CGATCACACT
943 CCCTTGGCGA CTCCCGCTCA CTCACCGCGG TTGAGACCCC TGGGCTCCAG CGATCTGCAT
1003 TTGCCGTCTA TCCGTCATTT GTCGCTCCAC CACACACCAG CTCTTGCTCC GATGGAACCC
1063 CAACCGGAAG GGCCCAACTA CTACAGTCCT TCCCAGGGTC ACCATGGTCC CAGCATCTCG
1123 GACATCATGT CTAAACCGGA CGGAACACAG CGCAAGCTGC CAGTCCCTCA GGTGCCAAAA
1183 GTCGCCGTTC AGGATATGTT AAATCCCGGT AGTGGATTTT CGTCGGTCCA TTCGTCCACC
1243 GCGAATTCCG TTGCCGGTGG CGACTTGGCA GAACGCTTC**T AG**ttcgaaca ttcttcagcc
1303 acacgttggt ttgtgtaaaa attgggttca aaaaatcagc agttctttta tgcgcgctac
1363 gaccgaatag acttgtgcat ttacaatggt tcatgggcat cattggtgtc gggtgattgg
1423 gtggtttttc ttcctcagct cttctgttgg atttatcttt cacatttatt ttccctttta
1483 ctcttttttt attgcaaggc ctctctactg atagatggac gggatatcct gtgaatttgt
1543 tatttttgtc ccctcttctg atcctccttg ctcttctctc caaatacttt tttttttac

Figure 5. Aligned coding sequence and translated amino acid sequence of the *creA* gene of *A. niger* (Drysdale *et al.*, 1993). The upper line is the DNA coding sequence for the CREA protein of *A. niger*. Underneath is the corresponding amino acid sequence. Eukaryotic specific codons are indicated by the **bold typeface.** Black highlighting indicates the region containing the zinc fingers. Underlining indicates the alanine rich sequence. **Light grey highlighting** indicates the region of similarity to Rgr1p of *S. cerevisiae*.

1 ATGCCGCCACCGGCCTCTTCAGTAGATTTCTCAAATCTGCTGAACCCTCAAAACAATTCC
1 M P P P A S S V D F S N L L N P Q N N S

61 ACTGATTCCACCCCTTCCACTCCTGTGGACAGCTCCAAGACCCCTTCTACTCCGTCCAGT
21 T D S T P S T P V D S S K T P S T P S S

121 ACACAGTCGAACTCCAACATGGCGTCGTCAGTGAGC**CTA**CTTCCGCCGCTTATGAAGGGT
41 T Q S N S N M A S S V S **L** L P P L M K G

181 GCCCGGCCTGCGACCGAAGAAGTTCGTCAGGATCTT**CCC**CGC**CCC**TACAAGTGC**CCC**CTC
61 A R P A T E E V R Q D L P P P

241 TGCGATCGCGCATTCCACCGTCTGGAGCACCAGACA**AGA**CACATCCGCACGCACACGGGT
81 R

301 GAAAAGCCGCATGCCTGTCAATTC**CCC**GGCTGCACTAAGCGCTTCAGTCGTTCTGACGAG
101 P

361 CTTACGCGCCACTCGCGAATCCACAACAACCCTAACTCG**AGA**CGTCGAAACAAGGCTCAG
121 N N P N S **R** R R N K A Q

421 CATCTGGCTGCGGCCGCTGCAGCTGCTGCAGGGCAGGACAACGCCATGGCCAACACCGCC
141 H L A A A A A A A A G Q D N A M A N T A

481 AGCGCTATGATGCCTCCT**CCC**AGCAAGCCAATGACCCGGTCGGCCCCTGTCTCGCAGGTC
161 S A M M P P **P** S K P M T R S A P V S Q V

541 GGCTCC**CCC**GACATCTCT**CCC**CCTCACTCCTTCTCCAACTACGCCAGTCACATGCGATCG
181 G S **P** D I S **P** P H S F S N Y A S H M R S

601 AACTTGGGCCCGTATGCTCGCAAAGGCGACGAGGCATCGTCGGGCATGGAA**CTA**TACCTC
201 N L G P Y A R K G D E A S S G M E **L** Y L

661 CTGGCGACTGCTGCGTCGCAGGTGGAACGCGACGAACACTTCGACTTCCATGCAGGGCCA
221 L A T A A S Q V E R D E H F D F H A G P

721 CGCAATCACCATTTGTTCAGCTCGCGCCACCAT**GGA**AGTGGCCGTCTG**CCC**CTCCTTGCT
241 R N H H L F S S R H H **G** S G R L **P** L L A

781 GCTTACGCGATCACACACAACATGAGCCGGTCGCATTCACCTGAAGACGACGACGGTTAT
261 A Y A I T H N M S R S H S P E D D D G Y

841 TCGCACCGTGTGAAACGTTCA**AGACCC**AACTCT**CCC**AACTCGACTGCTCCTTCCTCT**CCC**
281 S H R **V K R S R P N S P N S T A P S S P**

901 ACCTTCTCCCACGACTCCCTCTCA**CCC**ACT**CCC**GATCACACT**CCC**TTGGCGACT**CCC**GCT
301 **T F S H D S L S P T P D H T P L A T P A**

961 CACTCACCGCGGTTG**AGACCCC**CTGGGCTCCAGCGATCTGCATTTGCCGTCTATCCGTCAT
321 **H S P R L R P** L G S S D L H L P S I R H

1021 TTGTCGCTCCACCACACACCAGCTCTTGCTCCGATGGAA**CCC**CAACCGGAAGGG**CCC**AAC
341 L S L H H T P A L A P M E **P** Q P E G **P** N

1081 TACTACAGTCCTTCCCAGGGTCACCATGGT**CCC**AGCATCTCGGACATCATGTCTAAACCG
361 Y Y S P S Q G H H G **P** S I S D I M S K P

1141 GAC**GGA**ACACAGCGCAAGCTGCCAGTCCCTCAGGTGCCAAAAGTCGCCGTTCAGGATATG
381 D **G** T Q R K L P V P Q V P K V A V Q D M

1201 TTAAAT**CCC**GGTAGTGGATTTTCGTCGGTCCATTCGTCCACCGCGAATTCCGTTGCCGGT
401 L N **P** G S G F S S V H S S T A N S V A G

1261 GGCGACTTGGCAGAACGCTTCTAG
421 G D L A E R F •

generalizations as to the mechanisms of CRE protein action can be deduced from research spanning many genera and species. Regulation of the *creA* gene occurs at both the transcriptional and post-transcriptional level (Strauss *et al.*, 1999). The 5' DNA sequences flanking the *creA* gene of *A. niger* and the *creA* gene of *A. nidulans* both contain poly(T) stretches, and share three regions of similarity (Drysdale *et al.*, 1993). Refer to Figure 4 for the exact locations of these regions of similarity. Down regulation of *creA* mRNA in the presence of repressing carbon sources is effected by autoregulation, and requires functional CREA binding sites in the *creA* promoter region (Strauss *et al.*, 1999). In *A. nidulans*, two closely spaced CREA binding sites located around –560 with respect to ATG, were identified and shown to be necessary for *creA* repression *in vivo* (Strauss *et al.*, 1999).

To study the functional regions of the CREA protein, a set of various *A. nidulans creA* mutants was developed. Mutants contained one of the following alterations: missense mutations in either the first or second zinc-finger; a pericentic inversion resulting in premature termination after a disrupted alanine-rich region; or premature termination either before or after the acidic region, resulting in removal of the Rgr1p-similar domain (Shroff *et al.*, 1996). All mutants display reduced growth rates and colony diameters, and mutants disrupted at or beyond amino acid 107 display an impaired ability to conidiate (Shroff *et al.*, 1996). Originally it was postulated that a deletion mutant of *creA* resulted in a leaky-lethal phenotype, however, further study revealed that null alleles are viable although they display the same reduced growth rates, colony size, and impaired ability to

conidiate as those mutants studied by Shroff *et al.* in 1996 (Dowzer and Kelly, 1991; Shroff *et al.*, 1997). Functional studies of *creA* mutants of *A. niger* have also been performed, and similar observations noted (Ruijter *et al.*, 1997). The repressing effect of the various mutants differed with respect to the mutant, the growth media, and the gene examined (Shroff *et al.*, 1996; 1997; Ruijter *et al.*, 1997). Artificial down regulation of the *creA* gene has been performed via antisense silencing, and results in 50% derepression compared to a null allele without affecting growth or colony morphology (Bautista *et al.*, 2000).

The consensus binding site for CREA in *A. nidulans*, has been refined to 5'-SYGGRG-3' through DNAse1 footprinting and EMSA using the promoter regions of *alcA*, *alcR*, *prnB*, *prnD*, *otaA*, and *ipnA* (Sophianopoulou *et al.*, 1993; Kulmburg *et al.*, 1993; Cubero and Scazzocchio, 1994; Espeso and Peñalva, 1994; Dzikowska *et al.*, 1999). The 5'-SYGGRG-3' consensus sequence has been shown to be recognized by CRE homologues in both *Aspergillus oryzae* and *Humicola grisea*, but no work has been published regarding *A. niger* to date (Kato *et al.*, 1996; Takashima *et al.*, 1998). It should be noted that this is also the consensus binding sequence recognized by Mig1p of *S. cerevisiae* (Flick and Johnston, 1992). Binding of CREA to certain non-consensus sites has been observed *in vitro* in the *ipnA* promoter, however, it should be noted that a large proportion of putative binding sites identified via *in vitro* methods are later proven to be inactive *in vivo* (Espeso and Peñalva, 1994; Mathieu *et al.*, 2000; Orejas *et al.*, 2001). As was observed in *S. cerevisiae*, most of the 5'-SYGGRG-3' binding sites shown to be functional *in vivo* are preceded by AT-rich sequences (Nehlin *et al.*, 1991; Kulmburg *et*

al., 1993; Lundin *et al.*, 1993; Cubero and Scazzocchio, 1994; Mathieu *et al.*, 2000). Even though a CREA consensus binding site is not active *in vivo*, this does not necessarily indicate that CREA does not exert repression over a gene, but rather, that the repression might be exerted in an indirect manner (Orejas *et al.*, 2001). Whether CREA acts as a dimer or a monomer is a much debated question. Certainly, there is evidence to support both points of view: CREA binding appears to be stabilized when two sites are placed closely together, which supports the dimerization theory; GST exists as a dimer in solution, and this evidence of dimerization of CREA may just be a result of the GST region of the GST:CREA fusion protein dimerizing with the GST region of another fusion protein; certain investigators believe their EMSA results indicate monomeric binding to consensus binding sites (Cubero and Scazzocchio, 1994; Espeso and Peñalva, 1994; Panozzo *et al.*, 1998; Mathieu *et al.*, 2000). Creation of a full length CREA protein with fusion domains removed will help to answer this question.

The CREA protein consists of two zinc fingers of the Cys_2His_2 type, and each finger contains an antiparallel â-ribbon comprised of two â-strands, and an á-helix (Pàrraga *et al.*, 1988; Drysdale *et al.*, 1993). Two cysteines located near the turn in the â-ribbon and two histidines in the carboxy-terminal end of the á-helix are coordinated with a zinc ion, creating the zinc finger (Pàrraga *et al.*, 1988). The á-helix region of each zinc finger nestles into the major groove of the DNA, and the second â-strand stabilizes the DNA-protein complex by interacting with the sugar phosphate backbone of the DNA (Pavletich and Pabo, 1991). When binding to the DNA, the peptide is oriented "antiparallel" to the G-rich strand of DNA; the first zinc finger interacting with the second triplet sequence

and the second zinc finger interacting with the first triplet sequence, when reading the DNA in the 5' to 3' direction (Pavletich and Pabo, 1991). Counting from the amino terminal end of the protein, the two zinc finger regions are referred to as Finger 1 and Finger 2. Finger 1 is a type 2 zinc finger region, and interacts with the sequence 5'-GCG-3' (Nardelli *et al.* 1991). Finger 2 is a type 1 zinc finger region, and interacts with the sequence 5'-GGG-3' (Nardelli *et al.* 1991). Type 1 fingers are differentiated from type 2 fingers by the nature of the third and sixth amino acid residues in their á-helix structure: type 1 fingers contain a glutamic acid at the third position and an arginine at the sixth position, whereas type 2 fingers contain a histidine at the third position but show variability at the sixth position (Nardelli *et al.* 1991). Each zinc finger fits into the major groove of the DNA helix, and spans a distance of 3 bp (Nardelli *et al.*, 1991). Hydrogen bonds are formed between amino acid residues of the zinc finger, and the DNA bases in the recognition site, however, the zinc finger does not interact with each of the three bases in the triplet (Pavletich and Pabo, 1991). The amino acid residues that interact directly with the DNA bases are the residue immediately before the á-helix as well as the third and sixth residues (Pavletich and Pabo, 1991). The residue immediately before the á-helix can interact with the third base of the recognition triplet, the third residue of the á-helix can interact with the second base of the recognition triplet, and the sixth residue of the á-helix can interact with the first base of the recognition triplet, bearing in mind that all three interactions do not occur at one triplet (Pavletich and Pabo, 1991). Binding specificity is determined by the characteristics of the amino acids at these positions (Pavletich and Pabo, 1991). Please refer to Figure 6 for a diagram indicating the amino

acid sequence of each of the zinc fingers of CREA. Structural features of the zinc finger regions are indicated on the figure.

1.5 *glaA* promoter region constructs

1.5.1 Rational behind development of constructs.

To study the important functional regions of the *glaA* promoter, a set of constructs were created, containing the region from -146 to -516 in relation to the ATG. This region was identified as being the most important for high-level expression (unpublished data). Each member of this set had point mutations in one or more of the putative DNA consensus sequences for CREA or ANCF (Hap-like protein) binding. The sequence of this region with the location of CREA consensus binding sites and ANCF consensus binding sites along with sequences of their engineered mutations, are shown in Figure 7.

These mutated promoter regions are fused to the *lacA* reporter gene in an expression vector, and then transformed into *A. niger* protoplasts. The activity of the transformants is determined through a â-galactosidase assay of the media that was previously cleared of mycelia. In this manner, the important sites for promoter regulation can be determined.

1.5.2 Mutated consensus sequences in the promoter region.

Each of the specific consensus binding sites has been mutated to disrupt binding of the regulator protein. The intention is to investigate the activity of all combinations of

Figure 6. Amino acid sequence of zinc finger regions Finger 1 and Finger 2 of the CREA protein of *A. niger*. Black highlighting indicates the amino acid residues forming the á-helix. **Grey highlighting** indicates residues involved in coordination with the zinc ion. Underlining indicates residues forming a â-strand. Triangles ( ) indicate residues that form hydrogen bonds with the bases of the DNA consensus binding site.

Finger 1

P R P <u>Y K C</u> P L C D <u>R A F</u> H R I I H Q I R H I R I H I

Finger 2

E K P <u>H A C</u> Q F P G C T <u>K R F</u> S R S D F I I R H S R I H N

Figure 7. 5' promoter region of the *glaA* gene of *A. niger* used in *glaA-lacA* fusion vector with CREA binding sites and HAP binding sites indicated. Numbering counts backward from ATG of the *gla*A gene. **CCAAT** (HAP binding) sites are indicated with pale gray highlighting. SYGGRG (CREA binding) sites are indicated with black highlighting. ***Restriction enzyme sites*** are indicated by bold italic lettering. The mutated sequence for each CREA binding site is below the wild type sequence. The PCR amplified region used in electrophoretic mobility shift assays is underlined.

***MluI* site**

-523 ACC***ACGCGT***T GGACCTGCGT TATAGCTTCC CGTTAGTTAT AGTACCATCG TTATACCAG**C**

HAP site 1 **CREA site 1** **CREA site 2**

-463 **CAAT**CAAGTC ACCACGCACG ACCGGGGACG GCGAATCCCC GGGAATTGAA AGAAATTGCA

CATACG CATA CG

HAP site 2

-403 TCCTAGGCCA GTGAGGCCAG CG**ATTGG**CCA CCTCTCCAAG CACACAGGGC CATTCTGCAG

CREA site 3

-343 CGCTGGTGGA TTCATCGCAA TTTCCCCCGG CCCGGCCCGA CACCGCTATA GGCTGGTTCT

CATACG

-283 CCCACACCAT CGGAGATTCG TCGCCTAATG TCTCGTCCGT TCACAAGCTG AAGAGCTTGA

-223 AGTGGCGAGA TGTCTCTGCA GGAATTCAAG CTAGATGCTA AGCGATATTG CATGGCAATA

***AgeI* site**

-163 TGTGTTGATG CATGTGCTTC ***ACCGGT***

mutations of the CREA consensus binding sites C1, C2, and C3. At the time the thesis was prepared, not all of the constructs were created. Table 1 indicates the constructs that were donated, which mutation or mutations they contain, and the sequence of each wild-type consensus sequence and the mutated sequence that replaces it.

1.6 Rationale for thesis

1.6.1 Outline of experimental ideas

There is little work on the mechanism of CREA repression in *A. niger*, although studies regarding these mechanisms in the proline utilization gene cluster and the ethanol utilization pathway have been performed in *A. nidulans*. It was my intent to investigate the binding of CREA to several putative regulatory sites in the promoter region of *glaA* of *A. niger*. Analysis of the *glaA* promoter localized the information for its transcriptional regulation to within 470 bp of the start codon, a region that contains three putative CREA binding sites. To this end, fusion constructs of glutathione-s-transferase and CREA were created within the vector pGEX-2T. The constructs were transformed into competent cells of the strain Rosetta®, and fusion proteins were induced, and the resulting bacterial lysate was purified over a Glutathione Sepharose 4B column. The purified fusion protein was used in electrophoretic mobility shift assays with radioactively labeled DNA probes of Rosa Zito's various mutants of the *glaA* promoter region. Shifting of the bands was observed to determine which of the putative sites

Table 1.
Mutated *glaA* promoter constructs

Constructs donated [a]

Construct Name	Site C1	Site C2	Site C3
C1a [c]	mutated	-	-
C1c [b]	mutated	-	-
C1f [b]	mutated	-	-
C2a [d]	-	mutated	-
C2b [b]	-	mutated	-
C2c [c]	-	mutated	-
$C3_1e$ [b]	-	-	mutated
$C3_2g$ [b]	-	-	mutated
C2/3-2a [b]	-	mutated	mutated
C2/3-2b [c]	-	mutated	mutated
C2/3-9 [e]	-	mutated	mutated
D9 [c]	-	-	-
D10 [b]	-	-	-
E1 [d]	mutated	mutated	mutated
E4 [b]	mutated	mutated	mutated

Mutated sequences

Site	Wild-type sequence	Mutated sequence
C1 (sense strand)	5'-CCGGGG-3'	5'-C**ATAC**G-3'
C2 (anti-sense strand)	5'-CCGGGG-3'	5'-C**GTAT**G-3'
C3 (anti-sense strand)	5'-CCGGGG-3'	5'-C**GTAT**G-3'

[a] The constructs were provided by Rosa Zito of Concordia University. Some were used by her for *in vivo* expression studies.
[b] These constructs were sequenced and found to possess the correct mutated CREA binding site(s).
[c] These constructs were sequenced and found to possess the correct mutated CREA binding site(s). but contained a point mutation in a region of the promoter away from the CREA binding sites.
[d] These constructs were sequenced and found not to possess the correct mutated CREA binding site(s)
[e] These constructs were not sequenced at the time the thesis was written.

bound to CREA, and how mutation of the various potential CREA sites affected CREA binding to the *glaA* regulatory sequence between –138 and –526.

Since so little is known regarding promoter regulation in *A. niger*, a genomic library was constructed so that other genes regulated by CREA, and genes important to catabolite repression could be pulled out of the library using techniques such as *in vivo* cloning. This library will also serve as a useful tool for other researchers wishing to probe for non-coding sequences or for genes and their associated 5' and 3' untranslated regions.

2. Materials and Methods

2.1 Growth conditions and manipulation of *E. coli* and *A. niger*.

2.1.1 Growth and propagation of *E. coli*.

Please refer to Table 3 for a list of *E. coli* and *A. niger* strains used. All strains were streaked from glycerol stocks kept at –80 °C onto the appropriate solid media for the strain. For mini-preps, protein induction, and competent cell preparation, liquid overnight cultures were prepared using 1 ml of the appropriate media inoculated with a sterile stick that was touched to an appropriate colony on the solid media. Both plates and overnight cultures were incubated for a minimum of 12 hours and a maximum of 16 hours at 37 °C. Liquid cultures were incubated while shaking at 250 rpm, with a culture volume to flask volume ratio of 1:10.

2.1.2 Chemically competent *E.coli* cells (adapted from protocol obtained from the web site: http://research.nwfsc.noaa.gov/protocols/rbcl2.html).

Solution composition is detailed at the referenced web site. Psi broth was inoculated with fresh overnight culture in a 1:100 dilution. Cultures were grown with shaking at 250 rpm, at 37 °C, until the OD_{550} = 0.48. The culture was incubated on ice for 15 minutes, then cells were pelleted at 5000 x g for five minutes. The supernate was discarded, and the pellet resuspended in 0.4 of the original volume with Tfb1. Incubation on ice and pelleting of cells was repeated. The supernate was discarded, and the pellet resuspended in 0.04 of the original volume with Tfb2, followed by incubation on ice for 15 minutes.

Cells were aliquoted by 100i l into microfuge tubes, quick frozen with liquid nitrogen, and stored at –80 °C.

2.1.3 Transformation of *E. coli.*

The protocol used to transform *E. coli* was adapted from the protocol provided with the Sure2 ultracompetent cells from Stratagene. Cells were thawed on ice and 50 i l aliquots were transferred into sterile Falcon™ tubes. 1 i l of provided â-mercaptoethanol was added to each tube, and the tubes were incubated on ice for 10 minutes with gentle tapping every two minutes. 1 to 10 i l of the ligation mix was added to the competent cells, and the cells were incubated on ice for another 30 minutes. The cells were subjected to heat shock at 42 °C for exactly 30 seconds, then put on ice for two minutes, followed by addition of 450 i l of NZY+ broth preheated to 42 °C. Cells were incubated for one hour at 37 °C with shaking at 250 rpm, then plated onto selective media and incubated overnight for 12-16 hours at 37 °C.

2.1.4 Growth and propagation of *A. niger*

Fresh conidia were used to inoculate liquid culture. Streaking old conidial stock onto an appropriate solid media plates, and incubating the plates at 30 °C for six days produced fresh conidia. Conidia were harvested by pooling 10 ml of a 0.09M NaCl, 0.001% Tween 80 solution onto the plate and scraping lightly with a sterile spreader. Fresh conidia were stored in this saline/Tween solution at 4 °C.

For growth of mycelial cultures in liquid media, the appropriate liquid media was inoculated with 10^6 conidia of the desired strain for every milliliter of culture volume. Flasks were incubated overnight with shaking at 200 rpm, at a temperature of 30 °C, for mycelia used to harvest genomic DNA.

2.2 Genomic Library

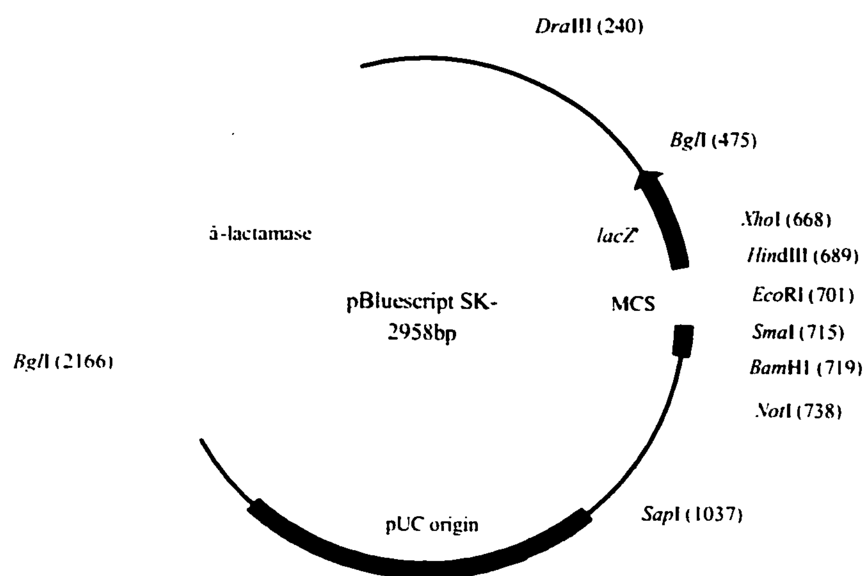
2.2.1 Preparation of pBluescript vector

The plasmid vector pBluescript from Stratagene was cut with the restriction enzyme *Bam*HI, followed by dephosphorylation of the sticky ends with calf intestinal alkaline phosphatase. The DNA was cleaned up via phenol/chloroform extraction and ethanol precipitation. For storage and later usage, the DNA was resuspended in sterile ddH_2O. Please see Figure 8 for a map of the pBluescript vector.

2.2.2 Preparation of genomic DNA.

Inserts for constructing a genomic library of *A. niger* DNA were created by partially digesting genomic DNA with the restriction enzyme *Sau*3A. Five microfuge tubes containing 20 μg of genomic DNA were set up. The first three tubes had enzyme added to a final concentration of 0.1 units/μg DNA. The fourth tube had enzyme added to a final concentration of 0.2 units/μg DNA. The fifth tube had a final enzyme concentration of 0.5 units/μg DNA. The first three tubes were incubated at 37 °C for 30 minutes, 45 minutes, and 60 minutes respectively. Tubes four and five were incubated at 37 °C for 30 minutes. The resultant fragments were ethanol precipitated and then resuspended in 21 μl sterile TE per tube. Following resuspension the fragments were pooled and then

Figure 8. A map of the pBluescript SK- plasmid. Please note the Multiple Cloning Site (MCS) located inside the *lacZ* gene.


DraIII (240)
BglI (475)
XhoI (668)
HindIII (689)
EcoRI (701)
SmaI (715)
BamHI (719)
NotI (738)
SapI (1037)
BglI (2166)
ä-lactamase
lacZ
MCS
pBluescript SK-
2958bp
pUC origin

loaded onto a 10% to 40% sucrose gradient in Beckman Ultraclear™ tubes. In a SW-28 rotor, the gradients were spun at 40 000 X g for 24 hours at 20 °C. Acceleration and deceleration were set to the slowest settings. Fractions were pumped out from the bottom of the tube using a peristaltic pump. Forty-four aliquots of 750 µl were collected and labeled. 40 µl aliquots were removed from odd numbered samples, vacuum concentrated in a Savant Speedvac™ to 20 µl, followed by the addition of 4 µl of 6X loading dye to each sample. Samples were subjected to electrophoresis on a 1% agarose gel in 1 X TBE for three hours at 60V. Samples 23-26 inclusive were determined to contain DNA fragments of around 11 kB in size. After ethanol precipitation and resuspension of these fragments, they were pooled in a total of 200 µl sterile TE. Fragments were ethanol precipitated, and resuspended in a total of 25 µl sterile ddH_2O, followed by quantification via agarose gel electrophoresis.

2.2.3 Ligation

10 ng of pBluescript digested and dephosphorylated DNA, and 33 ng of *A. niger* genomic DNA (approximate fragment size 11 kb), were combined with 4 units of T4 DNA ligase in a 5 µl reaction mix. The reaction was incubated at 21 °C for 1 hour, followed inactivation of the enzyme at 65 °C for 12 minutes.

2.2.4 Transformation of Sure2 Ultracompetent cells

The chemically competent cells were transformed with the ligation mix DNA according to the printed protocol accompanying the purchased cells. Two transformation mixes, each containing 100 µl of cells, were transformed with half of the ligation mix. After

following the transformation protocol, the transformation mixes were diluted to a total volume of 9ml with NZY^{+} broth. The mix was plated in aliquots of approximately 200 i l onto 54 LB agar with 50 i g/ml of ampicillin, IPTG, and X-gal (LBA-I/X) plates to select for transformants. The plates were incubated at 37 °C for 13 hours.

2.2.5 Characterization of library

Four plates were chosen at random, and the numbers of blue and white colonies were counted for each plate. This information was used to calculate the percentage of colonies containing insert DNA, and the total number of independent clones. Confirmation of blue/white selection for insert DNA was calculated by random selection of 10 white colonies and 4 blue colonies, extraction their DNA, followed by an *Eco*R1 digest and agarose gel electrophoresis. Digestion of the white colonies with *Xba*I and *Bsu*151 followed by agarose gel electrophoresis allowed determination of the insert size.

2.2.6 Harvesting the library

One of the 54 LBA-I/X plates was flooded with 5ml of LB, and the colonies were scraped off into the media with a sterile scraper. Media was transferred to another plate with a sterile pipette and the scraping was repeated. This was repeated for a total of 10 plates, and then the media was put on ice. The first plate was reflooded with 2 ml of media, and the scraping and transferring procedure was repeated. All 54 plates were processed in this manner resulting in 22.5 ml of collected media. To prepare for storage at −80 °C, 4.5 ml of glycerol and 3.0 ml of LB were added to the collected media and vortexed thoroughly to produce a solution containing a final concentration of 15%

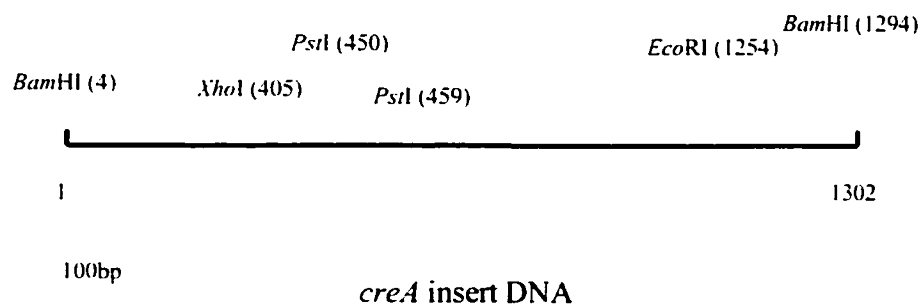
glycerol. This mixture was divided between 20 cryotubes by 1.5 ml aliquots and then stored at –80 °C.

2.3 Creation of GST:CREA fusion constructs

2.3.1 PCR amplification of *creA* DNA sequence for inserts

Two 31mer oligonucleotides referred to as P5'CreABamHI (5'-CGCGGATCCATGCCGCCACCGGCCTCTTCAG-3') and P3'CreABamHI (5'-GCCGGATCCCTAGAAGCGTTCTGCCAAGTCG-3') were used to PCR amplify the *creA* gene from *A. niger* genomic DNA. This resulted in the *creA* coding sequence with flanking *Bam*HI sites on either side. The locations of the restriction sites were designed to facilitate insertion into the MCS of pGEX-2T in the same reading frame as GST. A reaction mix of 100 i l containing 20 nmol dNTPs, 30 pmol of P5'CreABamHI, 30 pmol of P3'CreABamHI, 100 ng of *A. niger* genomic DNA, and 2 units of *Taq* polymerase was combined on ice. The mixture was split into two PCR tubes and placed in a Perkin-Elmer PCR 9700 machine. The PCR amplification program consisted of three segments. The first segment of the amplification program lasted 3 minutes at 94 °C. The second segment consisted of a sequence repeated for 25 cycles: 94 °C for 30 seconds; 60 °C for 30 seconds; then 72 °C for one minute and thirty seconds. The third segment involved two steps: five minutes at 72 °C followed by a hold at 4 C until the samples were removed. The reaction mixes were cleaned up using Qiagen Qiaquick PCR spin columns. Please refer to Figure 9 for a map of the amplified *creA* insert.

Figure 9. Visual representation of PCR amplified *creA* sequence with engineered flanking *Bam*HI sites. Numbering of basepairs is in the 5' to 3' direction of the sense strand. Please note the location of the *Xho*I site in relation to the beginning of the *creA* gene.



creA insert DNA

2.3.2 Preparation of the pGEX-2T vector.

The plasmid pGEX-2T was obtained from glycerol stocks stored at −80 °C, and streaked onto LBA plates for propagation (Gulick lab). The plasmid DNA was extracted using a standard mini plasmid preparation protocol, and purified via phenol/chloroform extraction and ethanol precipitation. Enzymatic digestion was performed in a 50 i l volume containing approximately 20 i g of pGEX-2T DNA and 10 units of *Bam*HI, for one hour at 37 °C. Ten more units of enzyme were added, and another hour of incubation was performed to ensure complete digestion. This was followed by dephosphorylation with the addition of 1 unit calf intestinal alkaline phosphatase to the reaction mix. The vector was purified using a Qiaquick PCR column to remove enzyme. Please refer to Figure 10 for a map of the pGEX-2T vector.

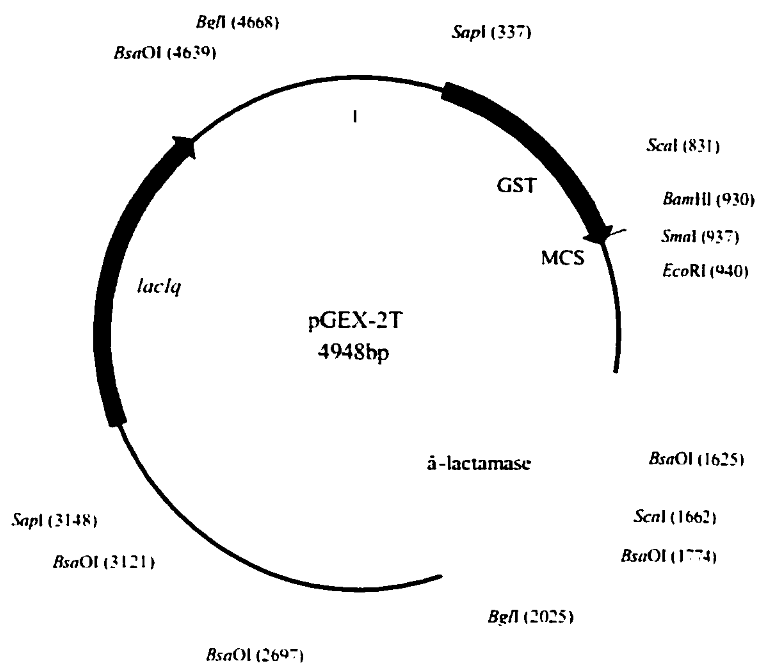
2.3.3 Preparation of *cre*A insert DNA

The insert DNA was fully digested by the restriction enzyme *Bam*HI, as per the protocol used to prepare the pGEX-2T vector. The vector was cleaned up with a Qiaquick PCR column to remove enzyme in preparation for ligation.

2.3.4 Ligation of insert DNA into vector

40 ng of vector DNA and 46 ng of insert DNA were combined in a 20 μl reaction mix with 4 units of T4 DNA ligase. The reaction mix was incubated for 1.5 hours at 16 °C followed by another 20 minutes at 21 °C. The ligase was inactivated by incubation for 20 minutes at 65 °C.

Figure 10. Plasmid map of the GST-fusion expression vector pGEX-2T. The MCS is located at the end of the coding region of GST.


BglI (4668)
BsaOI (4639)
SapI (337)
ScaI (831)
GST
BamHI (930)
SmaI (937)
MCS
EcoRI (940)
lacIq
pGEX-2T
4948bp
â-lactamase
BsaOI (1625)
ScaI (1662)
BsaOI (1774)
SapI (3148)
BsaOI (3121)
BglI (2025)
BsaOI (2697)

2.3.5 Transformation into XL1-Blue competent cells.

XL1-Blue cells were rendered competent via the rubidium chloride competent cell preparation protocol (refer to 2.1.2). For transformation, the Stratagene protocol was followed, using 50 i l of competent cells and 5 i l of the ligation reaction mix. After transformation, the 500 i l of transformed cells is split into three aliquots of 167 i l, and plated onto three LBA plates. The plates are incubated overnight at 37 °C.

2.3.6 Identification and characterization of constructs.

Twenty-four colonies at a time were selected at random for analysis. These colonies were set up as overnight cultures, their DNA to be extracted via standard mini plasmid preparation protocol. The plasmid DNA was run on an agarose gel to determine its approximate size in kilobases. Plasmids which appeared to carry insert DNA were digested with *Xho*I and *Mlu*I to confirm the correct orientation of the insert DNA (see Figures 16 and 17).

2.4 Induction, purification, and confirmation of presence of GST:CREA protein.

2.4.1 Growth conditions for induction and preparation of bacterial lysate.

Fresh transformants of the plasmid DNA in Rosetta© strain cells from Novagen were grown as overnight cultures in media containing 34 i g/ml chloramphenicol and 50 i g/ml ampicillin. The next morning, SOB media containing the appropriate antibiotics was inoculated with overnight culture in a 1:100 ratio of overnight culture to culture volume

in the flask. Cultures were grown at 37 °C with shaking at 250 rpm until the OD_{600} was approximately equal to 1.0. 100 mM IPTG was added to each flask to a final concentration of 1mM to induce GST:CREA production. Induction proceeded for another 4.5 hours. Cultures were chilled on ice for 15 minutes, followed by pelleting at 6000 x g for 10 minutes at 4 °C. The supernate was discarded and the pellet resuspended in ice cold 1X PBS which contained Proteinase Cocktail Set III from Novagen. The cell suspension was sonicated on ice at setting 3 for 125 x 6 second pulses, interspersed with a 20 second pause between pulses.

2.4.2 Processing and purification GST:CREA

Twenty percent Triton-X-100 was added to the sonicate to a final concentration of one percent. 1 M $ZnCl_2$ was added to the sonicate so that there was a 5 x molar excess of Zn^{+2} ions to GST:CREA proteins. This was calculated assuming a yield of one percent of the final protein product. The sonicate was gently rocked for 24 hours at 4 °C in order to solubilize the proteins. Spinning at 16 000 x g, at 4 °C, for 30 minutes resulted in pelleting of insoluble matter. The centrifugation step was repeated once more, and the supernate was stored at 4 °C until its application on to the Glutathione Sepharose column.

While the supernate sat at 4 °C, a column of Glutathiose Sepharose resin was prepared. Using a bed volume of 1 ml, the column was equilibrated by the addition of 30 ml of ice cold 1X PBS. At each step of the purification, a small aliquot was kept for later analysis. The bacterial supernate was slowly added to the column at a rate of 0.2 ml/minute, and allowed to flow through and drain into a bottle. Once the supernate had cleared the

column. cellular debris and non-specifically bound proteins were removed by 3 X 10 ml washes with 1X PBS. Protein was eluted off the column by the application of 3 X 1 ml applications of elution buffer (10mM reduced glutathione in 50mM Tris-HCl, pH 8.0).

Confirmation of the presence of the GST:CREA protein was carried out via SDS-PAGE followed by transfer to a nitrocellulose membrane and Western blotting with GST-specific antibodies. A protein of 74 kDa was expected. The protein was quantified by the Bradford protein microassay method.

2.5 Manipulation of *glaA* construct DNA

2.5.1 PCR amplification of a 227 bp portion of the *glaA* promoter region.

Two primers. 5'CreMutAll (5'-GCTTCCCGTTAGTACC-3') and 3'CreMutAll (5'-CGATGGTGTGGGAGAACC-3'), were designed to amplify the region between -501 and -274 with respect to the ATG of *glaA*. A 2 i l aliquot of each of the mutants was provided by Rosa Zito of Concordia University. This aliquot was diluted 1:50 with sterile ddH$_2$0 and 1 i l was used in one PCR reaction mix. The promoter region was PCR amplified by two rounds of PCR. During the first round, an appropriately buffered 100 i l reaction mix containing 20 nmol dNTPs, 30 pmol of 5'CreMutAll, 30 pmol of 3'CreMutAll. 1 i l of pre-diluted template DNA. 1.5 mM MgCl$_2$, and 2 units of *Taq* polymerase was prepared on ice. The reaction mix was split between two PCR tubes and placed in the Perkin Elmer 9700 PCR machine. A program designated as CreAllMut1,

consisting of three segments, was run. The first segment of the amplification program lasted 5 minutes at 94 °C. The second segment consisted of a sequence repeated for 27 cycles: 94 °C for 30 seconds; 47.5 °C for 30 seconds; then 72 °C for 20 seconds. The third segment involved two steps: five minutes at 72 °C followed by a hold at 4 °C until the samples were removed. The amount of product DNA was quantified, and if of sufficient quality and quantity, was used in the second round of PCR amplification. The second round of amplification was set up in the same manner as the first, except the reaction was scaled up by ten times to a total volume of 1 ml, and 200 ng of template DNA was used. These samples were aliquoted into PCR tubes and then subjected to the program CreMutAll1.

2.5.2 Clean up of PCR reactions

The PCR reactions were purified via phenol/chloroform extraction followed by ethanol precipitation. The pelleted DNA was resuspended in sterile ddH_20 and quantified via agarose gel electrophoresis with ethidium bromide.

2.6 Electrophoretic mobility shift assays (EMSA)

2.6.1 20mer oligonucleotides used to optimize reaction conditions

Two sets of complementary 20mer oligonucleotides containing either a CCGGGG binding site or a CCAAT binding site were created. The CCGGGG containing oligonucleotides were: 5'-CCGGGG (5'-GTTCCTTCCGGGGAGTGGCG-3') and 3'-

CCGGGG (5'-CGCCACTCCCCGGAAGGAAC-3'). The CCAAT containing oligonucleotides were: 5'-CCAAT (5'-ATACCAGCCAATCAAGTCAA-3') and 3'-CCAAT (5'-TTGACTTGATTGGCTGGTAT-3'). In a microfuge tube, 4 i l of Boehringer Manheim 10X buffer B (to give a final NaCl concentration of 200 mM) was mixed with 1 nmol of each complementary oligonucleotide and ddH$_2$0 to bring the reaction volume up to 20 i l. The tube was incubated at 80 ^{0}C for 5 minutes and then cooled slowly to room temperature. This provided a stock solution of double stranded DNA with a concentration of 50 pmol/ i l, which was used later in labeling reactions.

2.6.2 Labeling of DNA fragments with [ã^{32}]P-ATP.

A reaction mixture containing 2 i l of10X Buffer A, 3 pmol of DNA fragments, and 9.09 i l of sterile ddH$_2$0 was assembled on ice, followed by the addition of 9 pmol [ã^{32}]P-ATP (3000Ci/mmol) and 20 units of T4 polynucleotide kinase from MBI. After thorough mixing, the reaction mix was incubated at 37 ^{0}C for one hour. The reaction was stopped by the addition of 1 i l of 0.5M EDTA, followed by inactivation of the enzyme by incubation at 65 ^{0}C for 20 minutes. 9 i l of sterile ddH$_2$0 was added to the tube to bring the final concentration of the radioactively labeled probe to 100 fmol/i l. The probe was good for use for the next two days.

2.6.3 Preparation of acrylamide gels and reaction mixes.

6% polyacrylamide gels in 0.5X TBE were cast in pairs and left to polymerize overnight. Once polymerization has taken place, gels were stored under a thin layer of ddH$_2$0 until use. Each reaction mix contained the following components in the final volume: 25 mM

HEPES-KOH (pH 7.9), 0.01 M DTT, 0.01 M $MgCl_2$, 2.5 i g poly-dI-dC, and 5% ficoll (adapted from Kulmburg *et al.*, 1992). Varying amounts of protein, labeled probe, and unlabelled competitor DNA were added to this mixture, and the volume was made up with sterile ddH_20 to a final volume of 20 i l, 30 i l, or 40 i l, adjusted depending on the volume of protein added. Please see Table 2 for an example of a set of reaction mixes.

2.6.4 Running conditions for gel shift assays.

Before loading onto the gels, samples were mixed with 6X loading dye. The plates were set up against a water-cooled core to keep the gels and the proteins within as cool as possible. Ice and ddH_20 were used to make up the running buffer of 0.5X TBE to further cool the plates. Gels were run at a constant voltage of 150V until the first dye front was close to the bottom of the plate.

2.6.5 Fixing, drying, and exposure of gels to X-ray film.

Once the gels had been run, the smaller plate was pried loose to expose the surface of the gel. The gel surface was thinly covered with 7% acetic acid, and left to sit on the surface for 5 minutes. This served to fix the gel. The acid was poured off, and the gel was transferred to a piece of filter paper. The gels were dried under vacuum for two hours at 80 °C. The dried gel was exposed to X-ray film between 3 and 6 hours.

Table 2.
Reaction mixes for EMSA

Name of compound	DNA[a]	pGEX-2T[b]	Crude[c]	Purified[d]	C + 1 D10[e]	C + 5 D10[f]	C + 1 E1[g]	C + 5 E1[h]
	Name of Tube							
4X EMSA buffer	5 i l	5 i l	5 i l	5 i l	5 i l	5 i l	5 i l	5 i l
20% Ficoll	5 i l	5 i l	5 i l	5 i l	5 i l	5 i l	5 i l	5 i l
poly-dI-dC 2.5 i g/i l	1 i l	1 i l	1 i l	1 i l	1 i l	1 i l	1 i l	1 i l
Protein 1i g/tube	-	4 i l	0.6 i l	1.9 i l	0.6 i l	0.6 i l	0.6 i l	0.6 i l
100 fmol [ā^{32}]P - D10	1 i l	1 i l	1 i l	1 i l	1 i l	1 i l	1 i l	1 i l
1 pmol D10	-	-	-	-	1.3 i l	-	-	-
5 pmol D10	-	-	-	-	-	6.5 i l	-	-
1 pmol E1	-	-	-	-	-	-	1.4 i l	-
5 pmol E1	-	-	-	-	-	-	-	7.1 i l
ddH$_2$0	8 i l	4 i l	7.4 i l	6.1 i l	6.1 i l	0.9 i l	6 i l	0.3 i l
Total reaction volume	20 i l	20 i l	20 i l	20 i l	20 i l	20 i l	20 i l	20 i l

[a] Contains only labeled wild type (D10) DNA.
[b] Contains labeled wild type (D10) DNA and 1 i g of cell lysate from cells carrying plasmid pGEX-2T.
[c] Contains labeled wild type (D10) DNA and 1 i g of crude bacterial lysate containing GST:CREA.
[d] Contains labeled wild type DNA and 1 i g of purified GST:CREA.
[e] Contains labeled wild type DNA, 1 i g of crude lysate, and 1 pmol unlabeled wild type DNA.
[f] Contains labeled wild type DNA, 1 i g of crude lysate, and 5 pmol unlabeled wild type DNA.
[g] Contains labeled wild type DNA, 1 i g of crude lysate, and 1 pmol of unlabeled triple mutant (E1) DNA.
[h] Contains labeled wild type DNA, 1 i g of crude lysate, and 5 pmol of unlabeled triple mutant (E1) DNA.

3. Results

3.1 Genomic Library

3.1.1 Genomic DNA inserts

Fractions 23, 24, 25, and 26 from the 10% to 40% sucrose gradient were of the desired size of 10 kb to 15 kb (Figure 11). After pooling, ethanol precipitation, and resuspension in sterile ddH_2O, quantification of the purified fragments resulted in a concentration of 33 ng/i l (Figure 12) of the correct size.

3.1.2 pBluescript vector

Correct preparation of the pBluescript vector was confirmed via a set of test ligations and transformations using samples of the pBluescript vector from each stage of its preparation (Table 3)

3.1.3 Characterization of library

To calculate the number of independent clones and the percentage of transformants containing insert, four plates were selected at random from the 54 plates used to grow the library, and the numbers of blue and white colonies were counted (Table 4). The total number of independent clones was calculated to be approximately 50 000. The percentage of transformants containing insert DNA was as estimated by the proportion of white colonies which was 92%.

Figure 11. 1% agarose gel in 1X TBE containing samples from sucrose gradient fractions. Marker lanes contain a total of 0.5 ì g of DNA per lane. ë indicates the MBI Fermentas ë-DNA digested with *Hin*dIII. indicates the MBI Fermentas 1 kb ladder. Fraction numbers are indicated at the top of each lane.

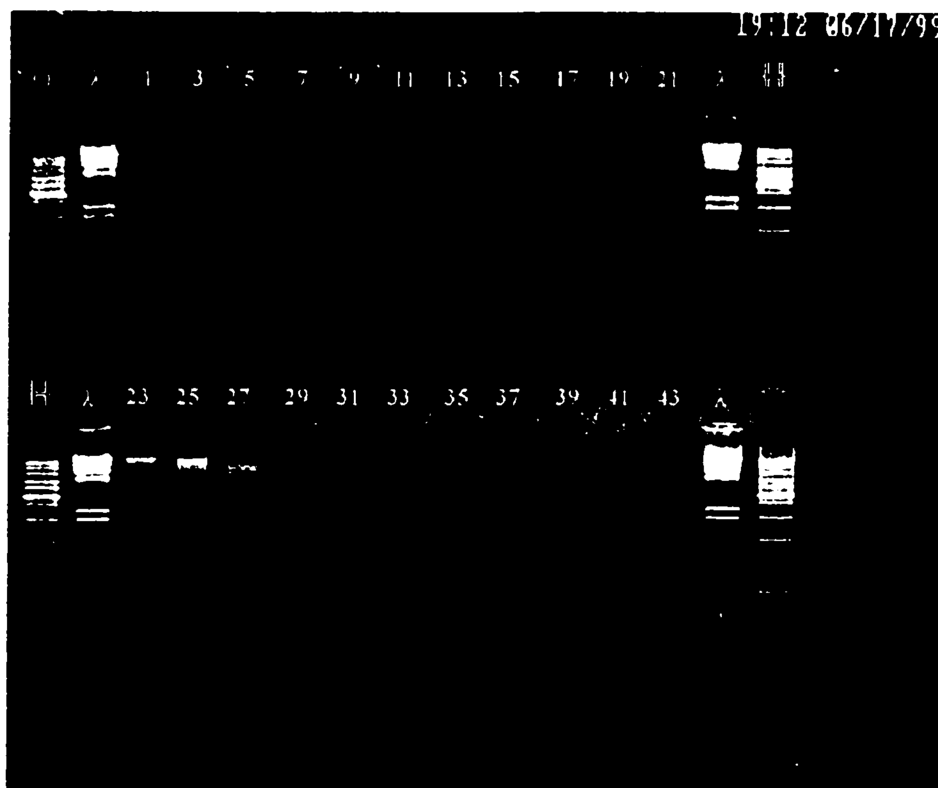


Figure 12. Purified genomic DNA fragments of approximately 11 kb in size. The last three lanes contain 1 i l, 2 i l, and 3 i l respectively, of the purified genomic DNA fragments. The first two lanes are marker lanes. Marker lanes contain a total of 0.5 i g of DNA per lane. ë indicates the MBI Fermentas ë-DNA digested with *Hin*dIII. indicates the MBI Fermentas 1 kb ladder.

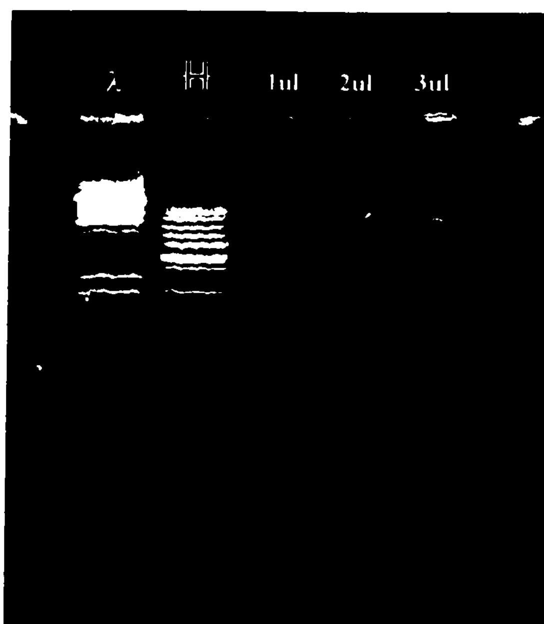

λ
1ul
2ul
3ul

Table 3.
Confirmation that pBluescript backbone was properly prepared for ligation with genomic DNA fragments

Sample	Transformation Frequency [a] (transformants/i g of DNA)
Bluescript	7.8×10^4
Cut Bluescript	1.3×10^3
Cut Bluescript with ligase	1.8×10^4
Cut, dephosphorylated Bluescript	0
Cut, dephosphorylated Bluescript with ligase	2.5×10^2
No plasmid	0

[a] Self-prepared $RbCl_2$ competent XL1-Blue cells were used for test transformations.

Table 4.
Analysis of genomic library

Plate #	Total colonies	Blue colonies	White colonies
1	817	37	780
2	757	56	701
3	785	73	712
4	1321	130	1191

Total # plates	54
Average # colonies per plate	920
Average # white colonies per plate	846
Percentage of white colonies	92%
Total # individual clones in library	50000

A map of the library construct shows the pBluescript vector with an insert of approximately 10 kb to 15 kb in the *Bam*HI site of the MCS (Figure 13). There are restriction enzyme sites for *Bsu*151 (*Cla*I) and *Xba*I on either side of the insert that are unique in the pBluescript vector, and an unknown number of sites within the inserts. Digestion with these two enzymes produces a fragment of 3 kb, which represents the pBluescript plasmid, and a number of other fragments whose sizes add up to the size of the insert DNA. Both the agarose gel of the fragments (Figure 14) and a table summarizing the results are shown (Table 5). The average insert size was calculated to be 14.7 kb.

3.2 Fusion constructs

3.2.1 *creA* insert DNA

Presence of a PCR product of the correct size was ascertained via agarose gel electrophoresis (data not shown). The concentration of the *creA* insert DNA was calculated to be 2 ng/i l.

3.2.2 pGEX-2T vector

Proper preparation of the pGEX-2T vector was confirmed with a set of test ligations and transformations (data not shown). Vector, prepared vector with ligase, and prepared vector and insert with ligase were all transformed into competent cells. The vector was deemed to be properly prepared.

Figure 13. Plasmid map of a genomic library clone. Actual clones will contain inserts of varied size. Genomic DNA insert is indicated by thick dark region. Restriction enzyme sites within the pBluescript plasmid vector are indicated. Please note the location of both the *Bsu*151 (*Cla*I) site at 684 and the *Xba*I site at 11731.

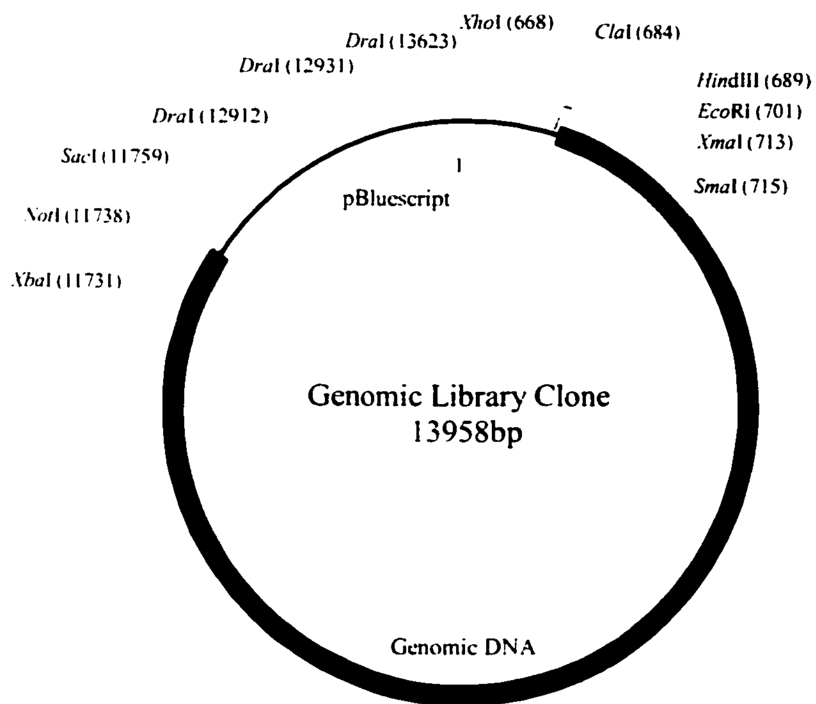

XhoI (668)
DraI (13623)
ClaI (684)
DraI (12931)
HindIII (689)
EcoRI (701)
DraI (12912)
XmaI (713)
SacI (11759)
SmaI (715)
pBluescript
NotI (11738)
XbaI (11731)
Genomic Library Clone
13958bp
Genomic DNA

Figure 14. 1% agarose gel containing digested DNA of 10 independent genomic clones. DNA was digested with *Xba*I and *Bsu*151 (*Cla*I). Sample lanes each contain 1 i g of digested DNA. Marker lanes contain a total of 0.5 i g of DNA per lane. ë indicates the MBI Fermentas ë-DNA digested with *Hin*dIII. indicates the MBI Fermentas 1 kb ladder. The location of the pBluescript band is indicated with an arrow. Partially digested bands in lanes 1, 3, 4, and 8 were ignored during the calculation of genomic DNA insert size.

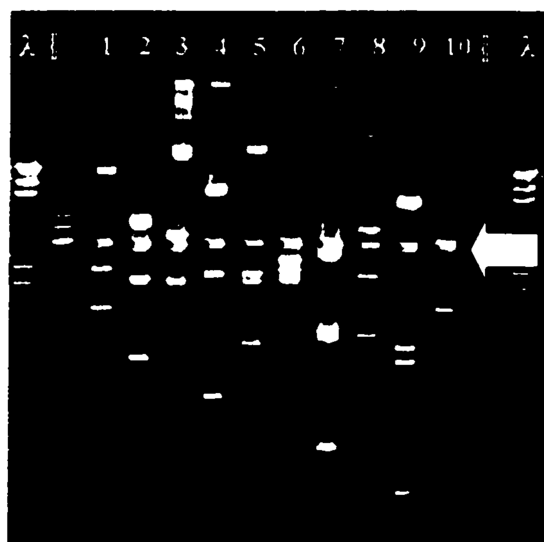

λ
1 2 3 4 5 6 7 8 9 10
λ

Table 5.
Restriction digest of genomic library clones with *Xba*I and *Bsu*151 (*Cla* I)

Clone #	DNA fragments in kb[a]					Insert size
1	1.6	2.3	20			23.9
2	1.4	2.1	3.75			7.45
3	2	3.25	28			33.25
4	0.8	2.2	8			11
5	1.4	2.1	2.2	31		36.7
6	2.2	2.4	2.5			7.1
7	0.6	1.4	1.5	2.7		6.2
8	1.4	2.2	3.5			7.1
9	0.4	1.2	1.4	6.5		9.5
10	0.6	1.6	3.1			5.3

Average DNA insert size 14.7 kb

[a] Only the length of non-vector fragments are included.

3.2.3 Characterization of transformants

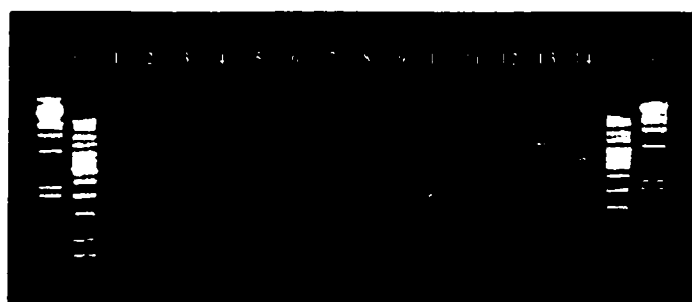
First, transformants were run on an agarose gel to confirm presence of insert DNA (Figure 15). Three insert containing constructs were selected for further analysis. According to the construct maps, there are two possible orientations that the insert DNA can assume in terms of reading frame: pGCreA-right direction (Figure 16) and pGCreA-wrong direction (Figure 17). Digestion with *Xho*I and *Mlu*I, followed by agarose gel electrophoresis, will differentiate between the two constructs. Constructs with the correct orientation of *creA* will produce two bands: 2.6 kb and 3.6 kb. Constructs with the incorrect orientation of *creA* will produce one band of 3.1 kb. As shown in Figure 18, the construct designated as pGCreA8 contains the insert in the correct direction. A second construct designated as pGCreA45 was identified in this manner. These two independent constructs were saved and used for further studies.

3.3 GST:CREA purification

3.3.1 Analysis of the purification process.

Each step of the isolation and purification process was analyzed via SDS-PAGE and Western blot. As can be seen (Figure 19), purification appears to enrich the lower molecular weight proteins in the original lysate. After Western blotting, however, it can be seen that enrichment of a few specific bands in the molecular weight range of 20kDa to 55 kDa is much greater than enrichment of the full-length protein of 74 kDa (Figure 20). A small fraction of the purified eluate contains fusion protein of the correct size.

Figure 15. Identification of constructs containing insert DNA. The pGEX-2T plasmid is approximately 5 kb. The pGCreA construct is approximately 6.2 kb. The clones containing insert DNA are retarded by the gel. Please note lanes 8, 13, and 17. The outer most lanes contain MBI Fermentas ë-DNA digested with *Hin*dIII. The lanes second from the outside contain MBI Fermentas 1 kb ladder. Marker lanes contain 0.5 ì g of DNA per lane.



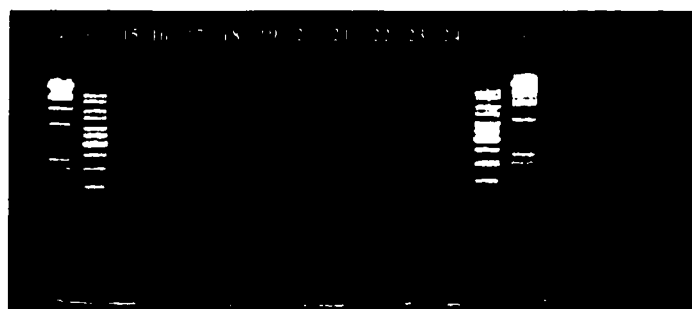
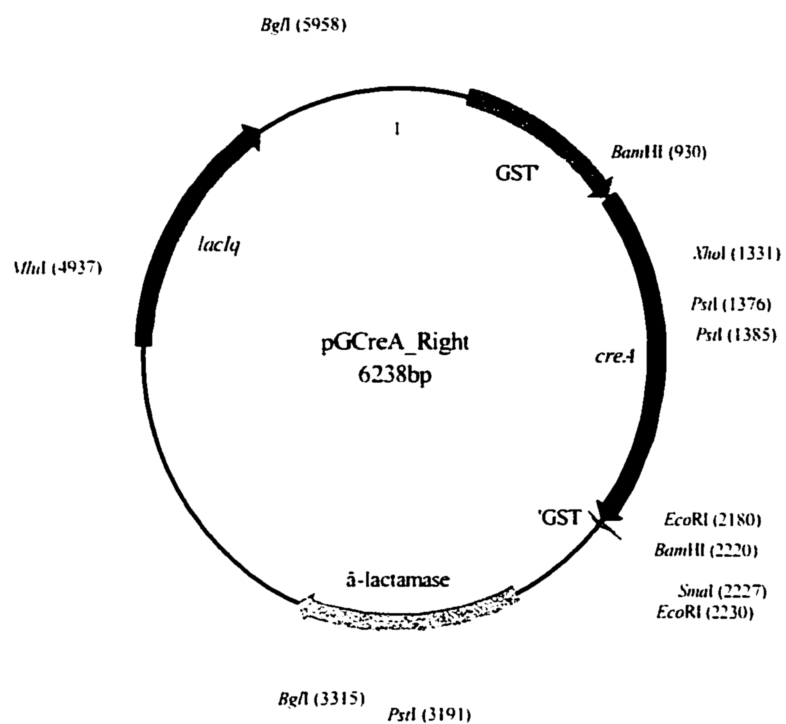
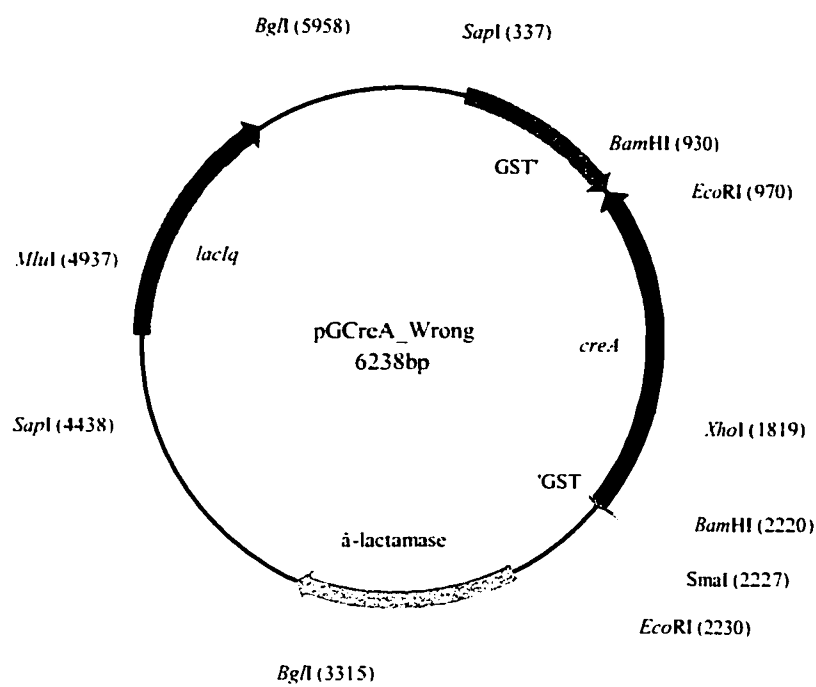


Figure 16. Plasmid map of the pGCreA construct containing insert DNA oriented in the correct direction. The correct direction refers to the *creA* gene being oriented in the same direction as the GST coding fragment. Arrow indicates the direction of transcription. Please note the locations of the following restriction sites used to determine insert orientation: *Mlu*I at 4937 and *Xho*I at 1331.



BglI (5958)
I
GST'
BamHI (930)
XhoI (1331)
PstI (1376)
PstI (1385)
creA
MluI (4937)
lacIq
pGCreA_Right
6238bp
'GST
EcoRI (2180)
BamHI (2220)
SmaI (2227)
EcoRI (2230)
â-lactamase
BglI (3315)
PstI (3191)

Figure 17. Plasmid map of the pGCreA construct containing insert DNA oriented in the wrong direction. The wrong direction refers to the *creA* gene being oriented in the opposite direction to the gene coding for GST. Arrow indicates the direction of transcription. Please note the locations of the following restriction sites used to determine insert orientation: *Mlu*I at 4937 and *Xho*I at 1819.



BglI (5958)
SapI (337)
BamHI (930)
GST'
EcoRI (970)
MluI (4937)
lacIq
pGCreA_Wrong
6238bp
creA
SapI (4438)
XhoI (1819)
'GST
BamHI (2220)
à-lactamase
SmaI (2227)
EcoRI (2230)
BglI (3315)

Figure 18. A 1% agarose gel of restriction digests with *Xho*I and *Mlu*I. Constructs 8, 13, and 17 contain insert DNA, but only construct 8 contains the insert in the correct orientation. indicates the MBI Fermentas 1 kb ladder. The marker lane contains 0.5 i g of DNA.

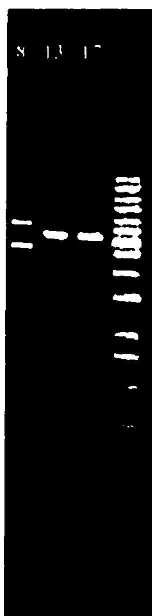


Figure 19. Coomassie blue stained 12% acrylamide gel of samples from each step of purification. Black bars on the left hand side indicate positions of molecular weight markers. From top to bottom, the bars indicate 97 kDa, 66 kDa, 45 kDa, and 30 kDa. The black bar on the right hand side indicates the position of the full length GST-CREA fusion protein. From left to right, the lanes are: (M) molecular weigh markers; (L) bacterial lysate; (S) supernate containing soluble protein fraction; (F) flow through from column; (W1) wash 1, (W2) wash 2, and (W3) wash 3; (E) pooled eluates.

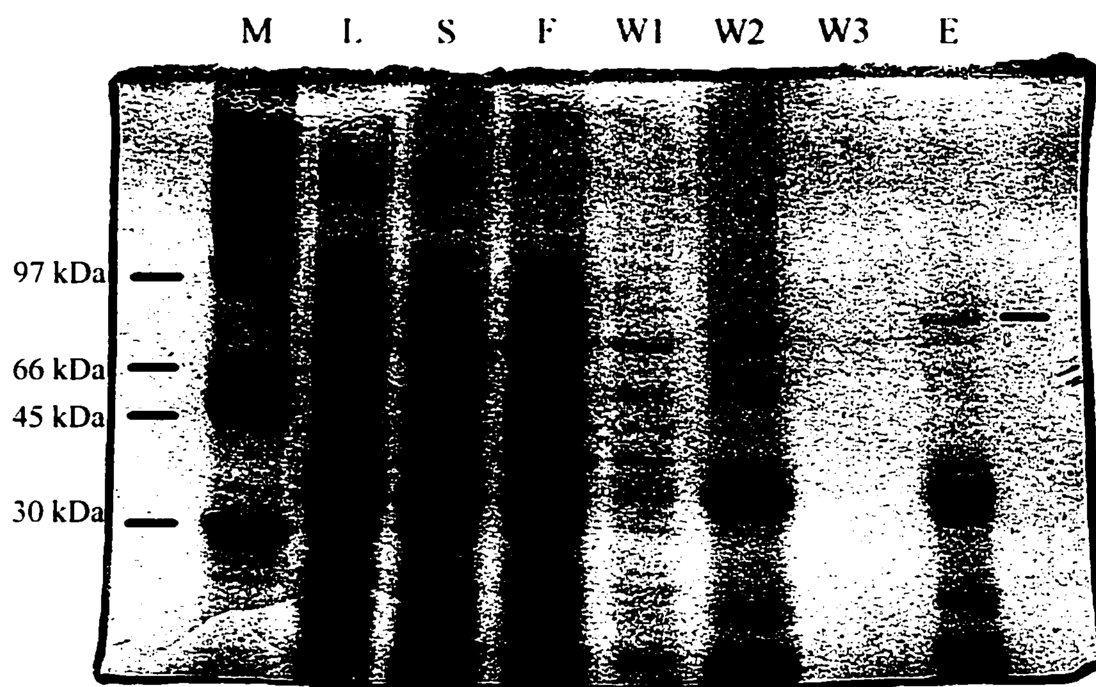
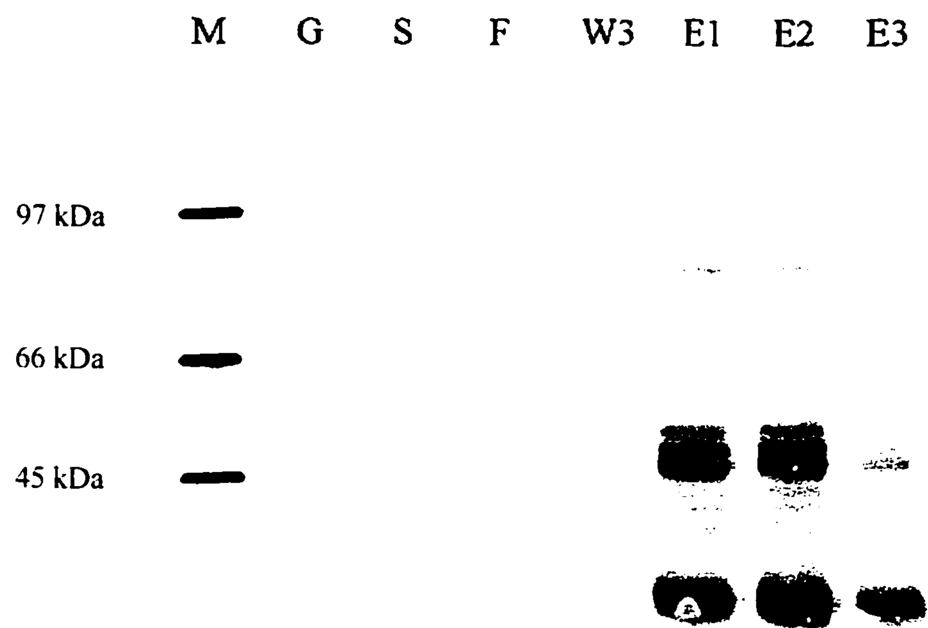

M
L
S
F
W1
W2
W3
E
97 kDa
66 kDa
45 kDa
30 kDa

Figure 20. Western blot with anti-GST IgG, of samples from each step of purification. Black bars on the left hand side indicate positions of molecular weight markers. From top to bottom, the bars indicate 97 kDa, 66 kDa, and 45 kDa. From left to right, the lanes are: (M) molecular weight markers; (G) GST; (S) supernate containing soluble protein fraction; (F) column flow through; (W3) wash 3; and (E1) eluate 1, (E2) eluate 2, and (E3) eluate 3. Please note that no band is visible in the GST lane since GST weighs 26 kDa and has run off the gel.



M
G
S
F
W3
E1
E2
E3
97 kDa
66 kDa
45 kDa

3.4 Amplification of wild type and mutant versions of the *A. niger glaA* promoter region from –244 to –501.

3.4.1 Confirmation of successful PCR amplification, cleanup and quantification.

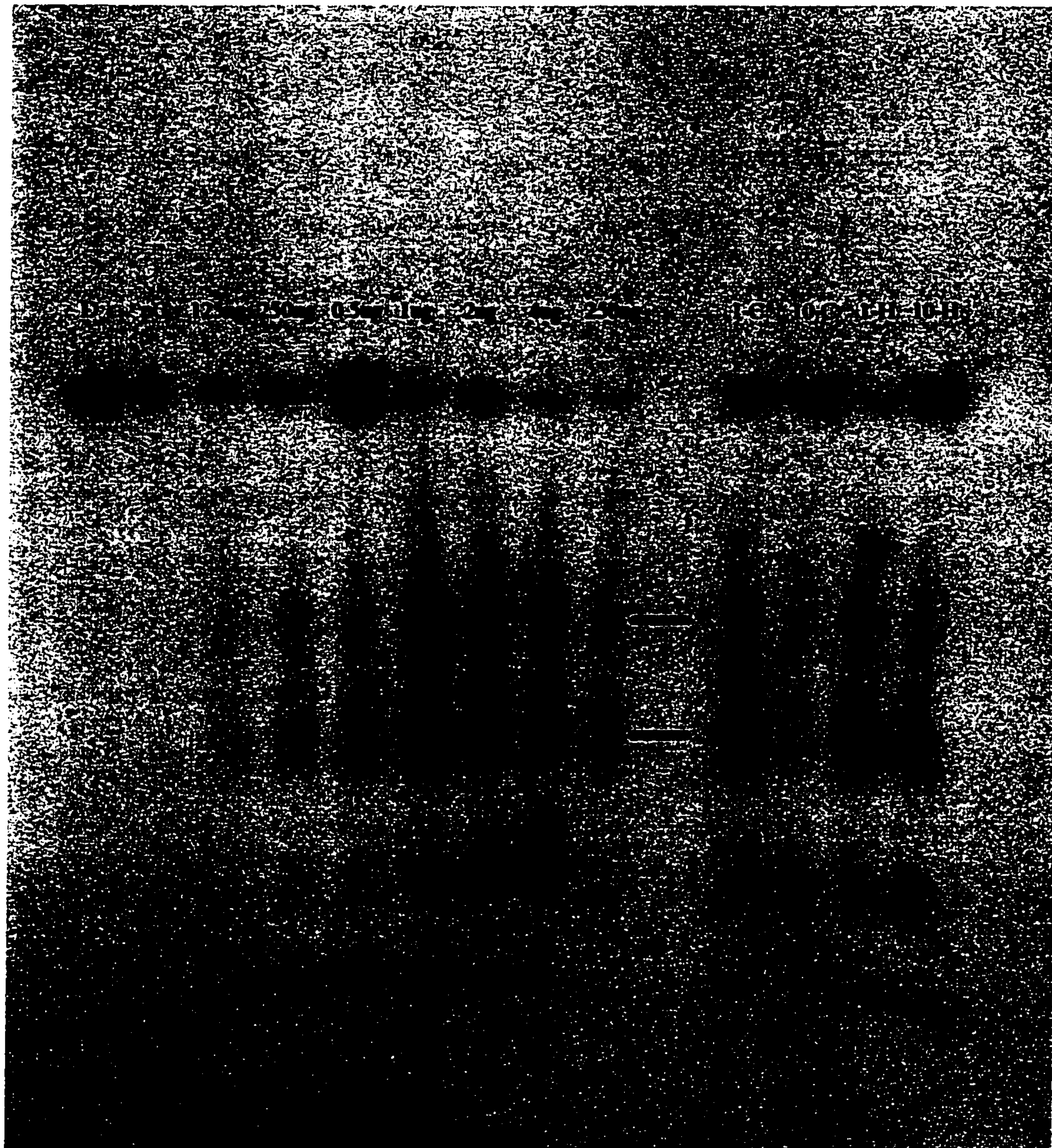
10 i l of each PCR mix was run on an agarose gel to confirm the presence of PCR product. After phenol/chloroform extraction, ethanol precipitation, and resuspension, DNA was quantified by agarose gel electrophoresis.

3.5 Electrophoretic mobility shift assays (EMSA).

3.5.1 Optimization of reaction conditions.

To optimize reaction conditions, trial EMSAs using the 20mer oligos containing either a CREA site or a HAP site were performed. In Figure 21, it can be seen that the lane containing just labeled CCGGGG DNA produces no shifted band, as compared to one shifted band seen in the next lane that contains both labeled DNA and crude bacterial lysate from cells not producing the GST:CREA fusion protein. Shifts performed with increasing amounts of lysate with GST:CREA, produce two bands not seen without the presence of the GST:CREA protein. The last four lanes demonstrate competition, or lack thereof, for the CREA binding site. In "1-C" and "10-C", the addition of ten-fold and one hundred-fold amounts of unlabeled CCGGGG as compared to the amount of labeled DNA, demonstrate clear competition between the labeled and unlabeled DNA for CREA binding. Increase in the amount of unlabeled specific competitor DNA clearly decreases the intensity of the two upper bands. Lanes "1-H" and "10-H", contain the same ratios of

Figure 21. Three hour exposure of EMSA performed with 20mer DNA fragments and crude bacterial lysate. Each lane contains 100 fmol of radioactively labeled CCGGGG 20mer probe. The constituents of each lane are listed below. "D" only contains labeled DNA. "pG" contains labeled DNA and 1 i g of crude lysate from cells transformed with the pGEX-2T vector. "125 ng", "250 ng", "0.5 i g", "1 i g", "2 i g", and "4 i g" each contain labeled DNA and the indicated amount of crude bacterial lysate containing GST:CREA. "1-C" contains labeled DNA, 1 i g crude lysate with GST:CREA, and 1 pmol unlabelled CCGGGG competitor DNA. "10-C" contains labeled DNA, 1 i g crude lysate with GST:CREA, and 10 pmol unlabelled CCGGGG competitor DNA. "1-H" contains labeled DNA, 1 i g crude lysate with GST:CREA, and 1 pmol unlabelled CCAAT competitor DNA. "10-H" contains labeled DNA, 1 i g crude lysate with GST:CREA, and 10 pmol unlabelled CCAAT competitor DNA. Black bars indicate bands shifted by GST:CREA.



unlabeled to labeled DNA, but the unlabeled DNA in these lanes is the non-specific competitor CCAAT oligonucleotide that has a binding site for the HAP-like protein. No competition for CREA binding is evident in these lanes.

Figure 22 shows EMSA results where shifts were performed with a crude bacterial lysate and with purified GST:CREA protein. As can be seen when comparing lane "pG" with all the lanes containing crude lysate, they all share one band in common. This band is missing when purified protein is used for the shifts. Two bands match between the lanes containing crude lysate and purified protein. These are the upper two bands that are visible in the crude lysate lanes, and the only two bands visible in the purified protein lanes. Lanes containing purified protein produce a lower band of higher intensity than the upper band. This is the reverse when lanes containing crude lysate are observed.

3.5.2 EMSA with mutated *glaA* promoter constructs and purified GST:CREA protein.

In Figure 23, radioactively labeled D10 (wild-type sequence) DNA was shifted using increasing amounts of purified GST:CREA protein. One band is steadily visible when the protein concentration ranges between 12.5 ng and 400 ng. Between 800 ng and 3.2 i g, a second intense band, and a fainter third band become visible. The four last lanes indicate that unlabeled D10 DNA competes for CREA:GST binding with labeled D10 DNA, but unlabeled E1 DNA displays almost no competition.

Figure 24 displays the type of shifts produced by each type of mutated *glaA* promoter construct. All mutants of CREA site 1 show one faint shifted band. For the CREA site 2

Figure 22. 2.5 hour exposure of EMSA performed with 20mer DNA fragments and both purified GST:CREA protein and crude bacterial lysate. Each lane contains 100 fmol of radioactively labeled 20mer CCGGGG probe. The constituents of each lane are listed below. "D" contains only labeled DNA. "pG" contains labeled DNA and 1 i g of crude lysate from cells transformed with the pGEX-2T vector. "0.25 i g P", "0.5 i g P", "1 i g P", and "2 i g P" all contain labeled DNA and the indicated amount of purified GST:CREA protein. "0.25 i g C", "0.5 i g C", "1 i g C", and "2 i g C" all contain labeled DNA and the indicated amount of crude bacterial lysate containing GST:CREA. "1-C" contains labeled DNA, 1 i g crude lysate with GST:CREA, and 1 pmol unlabelled CCGGGG competitor DNA. "10-C" contains labeled DNA, 1 i g crude lysate with GST:CREA, and 10 pmol unlabelled CCGGGG competitor DNA. "1-H" contains labeled DNA, 1 i g crude lysate with GST:CREA, and 1 pmol unlabelled CCAAT competitor DNA. "10-H" contains labeled DNA, 1 i g crude lysate with GST:CREA, and 10 pmol unlabelled CCAAT competitor DNA. Black bars indicate bands shifted by GST:CREA.

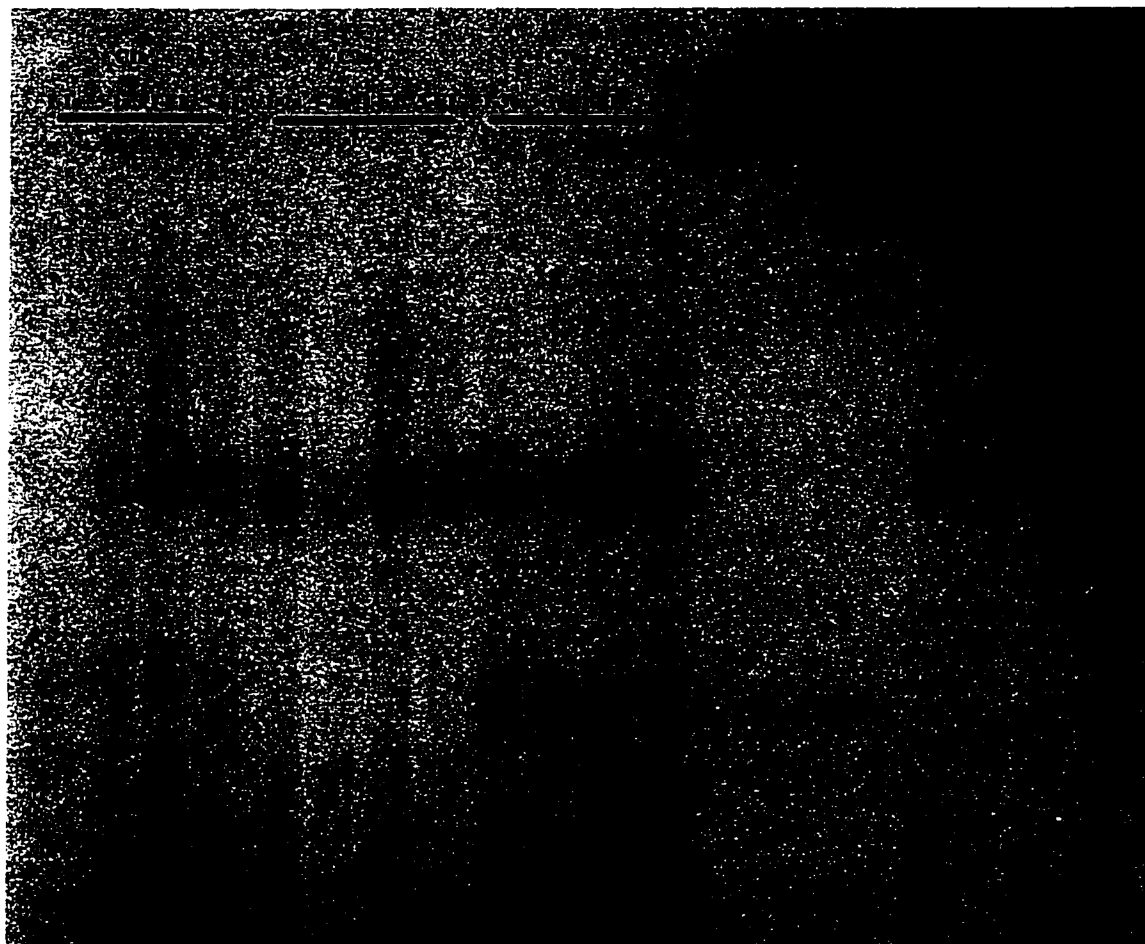
Figure 23. Five hour exposure of EMSA with D10 DNA (wild-type sequence *glaA* promoter region) and purified GST:CREA. Each lane contains 100 fmol of radioactively labeled D10 DNA. The constituents of each lane are listed below. "D" contains only labeled DNA. "pG" contains labeled DNA and 1 i g of crude lysate from cells transformed with the pGEX-2T vector. "12.5 ng", "25 ng", "50 ng", "0.1 i g", "0.2 i g", "0.4 i g", "0.8 i g", "1.6 i g", and "3.2 i g", contain labeled DNA and the indicated amount of purified GST:CREA protein. "D10-1" contains labeled DNA, 1 i g purified GST:CREA protein, and 1 pmol of unlabeled D10 competitor DNA. "D10-5" contains labeled DNA, 1 i g purified GST:CREA protein, and 5 pmol of unlabeled D10 competitor DNA. "E1-1" contains labeled DNA, 1 i g purified GST:CREA protein, and 1 pmol of unlabeled E1 competitor DNA (*glaA* promoter with no CREA binding sites). "E1-5" contains labeled DNA, 1 i g purified GST:CREA protein, and 5 pmol of unlabeled E1 competitor DNA. Black bars indicate bands shifted by GST:CREA. The "+" above a lane indicates that the sequence of the probe is incorrect. Please note that the dark shadow at the base of the gel is caused by free labeled probe, and that this gel has not been run as far as the others resulting in less resolution of the bands.

Figure 24. Three hour exposure of EMSA with purified GST:CREA protein and all of the versions of the 227 bp mutated *glaA* promoter region. Each lane contains 100 fmol of radioactively labeled DNA. The first lane "D10 *" contains only labeled D10 DNA. "D10 + pG" contains labeled D10 DNA and 3.2 ì g of crude bacterial lysate from cells transformed with the empty pGEX-2T plasmid. All the other lanes contain labeled DNA of the construct indicated at the top of the lane. and 3.2 ì g of purified GST:CREA protein. See Table 1 for a description of the various mutant probes. The "+" above a lane indicates that the sequence of the probe is incorrect. Black bars indicate bands shifted by GST:CREA.

mutants, only C2a shows an intense shifted band. The C2b and C2c constructs produce no shifted band. Both CREA site 3 mutants, $C3_1e$ and $C3_2g$, show an intense shifted band. For the set of double mutants of CREA site 2 and CREA site 3, C2/3-2a shows no shifted band, but C2/3-2b and C2/3-9 show a faint shifted band. The triple mutant E4 shows no shifted band, and the triple mutant E1 is almost the same, however, there is a barely perceptible smudge that could be interpreted as a faint band. The wild-type sequence constructs, D9 and D10, show two shifted bands: one lower intense band, and one higher faint band.

Competition binding studies between labeled DNA and unlabeled DNA with GST:CREA protein were performed on one construct from each family of mutant promoter regions (one from C1, one from C2, one from C3, one from C2/3, and one from D), and can be seen in Figure 25. With the mutant C1a, definite competition is observed between labeled and unlabeled C1a DNA, but not between labeled C1a and unlabeled E1 DNA. Constructs C2a and $C3_1e$ follow this pattern. Interestingly, double mutant C2/3-2a shows no shifted band and therefore no competition between labeled C2/3-2a DNA and unlabeled C2/3-2a DNA or unlabeled E1 DNA. Finally, labeled D9 DNA shows competition with unlabeled D9 DNA and no competition with unlabeled E1 DNA. In all cases where competition is visible, an increase in the amount of competitor DNA is related to a decrease in band intensity.

Figure 25. Three hour exposure of EMSA performed with partially purified protein and various mutant versions of the 227 bp mutated *glaA* promoter region. Every lane contains 1.6 i g of purified GST:CREA protein. Each group of four lanes spanned by a black bar contains 100 fmol radioactively labeled DNA per lane, of the construct indicated above. The constituents of each lane are indicated below. In group C1a: "1-1a" contains labeled C1a DNA, protein, and 1 pmol unlabeled C1a DNA; "5-1a" contains labeled C1a DNA, protein, and 5 pmol unlabeled C1a DNA; "1-E1" contains labeled C1a DNA, protein, and 1 pmol unlabeled E1 DNA; "5-E1" contains labeled C1a DNA, protein, and 5 pmol unlabeled E1 DNA. Every group of four is modeled as above. The first two lanes contain either 1 pmol or 5 pmol unlabeled DNA of the same construct as the labeled DNA, and the last two lanes contain either 1 pmol or 5 pmol of unlabeled E1 construct DNA. The "+" above a lane indicates that the sequence of the probe is incorrect. Black bars indicate bands shifted by GST:CREA.



4. Discussion

4.1 The *A. niger* genomic library.

The construction and characterization of this library, though not directly linked to the later experiments of this thesis, is still an important contribution to the *A. niger* research group. Using this library, other researchers will be able to pull out non-coding sequences of DNA, such as the consensus binding sequence of CREA, and discover novel promoters regulated by this protein, as well as using the library to clone genes from *A. niger*. Approximating the *A. niger* genome size as 30 000 kb, the genomic library contains twenty *A. niger* genome equivalents.

4.2 Expression of the GST:CREA fusion protein of *A. niger*.

GST fusion proteins of the DNA binding region of CREA have been constructed for *A. nidulans*, *A. oryzae*, and *H. grisea*. GST-fusions encorporating the full length CREA protein have not been reported in current literature, and neither has any form of construct involving the *creA* gene of *A. niger* (Kulmburg *et al.*, 1993; Kato et al., 1996; Takashima *et al.*, 1998). My construct encoding the full length CREA protein of *A. niger* is novel both for the species and the protein in question. Kulmburg *et al.* reported the presence of bands of lower molecular weight when purifying their truncated CREA fusion to GST (Kulmburg *et al.*, 1993). They assumed that these bands were due to degradation of the fusion protein, and since no pictures of the SDS-PAGE gels were shown, it is impossible to determine the purity of the GST:CREA fusion product used for EMSA (Kulmburg *et*

al., 1993). All other investigators performing EMSA with truncated CREA fusions with GST, obtained their plasmid construct from Kulmburg's group (Cubero and Scazzocchio, 1994; Espeso and Peñalva, 1994; Panozzo *et al.*, 1998; Strauss *et al.*, 1999; Mathieu *et al.*, 2000; Orejas *et al.*, 2001). I too observed the presence of numerous bands of lower molecular weight in the purified product, and believe that there is a complementary hypothesis as to the reason for occurrence of these bands.

In the pGEX-2T plasmid, the GST-tag exists as an amino-terminal fusion. When the bacterial lysate is purified over the Glutathione Sepharose 4B column, all proteins containing a GST-tag are bound to the column. This includes all partial translations of the fusion protein that possess a functional GST region able to bind to the column. The full length fusion protein is expected to run at approximately 74 kDa. Western blotting using a polyclonal anti-GST antibody produced many bands of lower molecular weight, indicating that all these bands contain GST epitopes recognizable to the antibody. This evidence argues against degradation and for partial translation, since proteolytic degradation would destroy secondary and tertiary structure required for glutathione binding.

The phenomenon of partial translation is intimately associated with the coding sequence of *creA*. Refer to Figure 5 for the complete nucleotide sequence of the gene, as well as the translated amino acid sequence below. As indicated in bold type, the DNA sequence of *creA* will be transcribed into an mRNA containing 28 codons, whose corresponding tRNAs are specific to eukaryotic cells, and are not found in *E. coli* unless artificially

inserted. These codons will cause stalling of the ribosome if the host bacterial cell does not carry plasmids encoding for these rare tRNAs. I tried to remedy this problem lowering the incubation temperature during induction of the fusion-protein in the hope that the resultant decrease in metabolic rate would allow for translation of these rare codons. Through many attempts at this process, I found no increase in the yield of full length protein produced. As a next step, I decided to try the Rosetta© strain of cells, that contain plasmids encoding the rare tRNAs. Due to time constrains, only one trial with these cells was performed, and consequently I was not able to tailor the induction conditions to maximize expression, however, these cells provide a promising means to reduce the formation of partially translated products.

Because of the population of GST:CREA proteins of varying lengths, it is difficult to discern which of these proteins retain DNA-binding ability and which do not. Proteins over 40 kDa contain the full zinc finger DNA-binding domain (26 kDa for GST plus 14 kDa for the zinc finger domain), and therefore should possess the ability to bind the consensus sequence. Referring to the eluate lanes in Figure 20, there is one very intense band located above 30 kDa, as well as three dark bands observed to run between 45 kDa and 66 kDa. The band above 30 kDa corresponds to a translated protein that was truncated around amino acid 70 of CREA. Amino acids 73, 75 and 79 are all prolines coded for by a eukaryotic-specific codon. The bands between 45 kDa and 66 kDa correspond to translated proteins truncated between amino acids 180 and 300 of CREA. There are three groups of eukaryotic-specific codons that correspond to proteins of 45kDa to 66kDa: amino acids 83 and 87 of CREA; amino acids 252 and 257 of CREA;

amino acids 289, 290, 293, 300, 309, 311, 315, 326 and 327 of CREA. The large number of closely spaced eukaryotic-specific codons between amino acids 289 and 327 of CREA, may help to explain the low yield of full length fusion proteins. This plethora of proteins of varied length renders it very difficult to calculate the concentration of active protein in the sample.

4.3 EMSA with GST:CREA and mutated *glaA* promoter regions.

4.3.1 EMSA with 20mer oligos

Shifts performed with crude lysate containing GST:CREA produced three distinct shifted bands. The lowest band corresponded with the band in the pG lane, indicating non-specific interaction. The two upper bands can be considered to be as a result of GST:CREA binding, and were consistently visible across an added protein range of 125 ng to 4 ì g. Both bands showed equal susceptibility to competition with unlabeled CCGGGG DNA. The existence of two bands can be interpreted in two ways: the GST:CREA protein is binding as a both a monomer and dimer, or proteins of two vastly different sizes are binding the 20mer oligonucleotide. Although the current purification process removes the non-specific binding activity from the eluted sample, the presence of partial translation products attached to the GST-moiety does not allow conclusions regarding possible dimerization to be drawn from the presence of two shifted bands.

When purified GST:CREA was run beside crude lysate, it could be definitively stated that the upper two bands were of consequence, though it should be noted that the purified protein produced two bands of approximately equal intensity, whereas the crude lysate produced an upper band of greater intensity than the lower band. This correlates with the purified protein eluate containing a larger proportion of proteins in the 40 kDa to 55 kDa range.

4.3.2 EMSA with mutated *glaA* promoter regions

Since the work for this thesis was completed, most of the constructs I worked with have been sequenced. The constructs with the correct mutated CREA binding site sequence(s) are: C1a, C1c, C1f, C2b, C2c, $C3_1$e, and $C3_2$g, C2/3-2a, and E4 (unpublished data). Construct C2a contained no mutation in the CRE2 site, and was therefore identical to wild-type construct (unpublished data). Sequencing demonstrated that E1 was actually a double mutant not a triple mutant (unpublished data). Sequencing of C2/3-2b is in progress, but preliminary results indicate that the construct does not contain the correct mutated CREA binding sites (unpublished data).

Shifts performed with D10 (wild-type) DNA and purified GST:CREA produced three bands. The upper and lower bands were of greater intensity than the middle band. It is tempting to conclude that three bands correspond to three binding sites, however the following fact must be considered: the optimization reactions performed with the 20mer oligonucleotides, which contain only one binding site, produced two bands. Competition studies performed with unlabeled D10 DNA show that only the upper two bands are

subject to competition. From this, we can speculate as to the mechanisms of GST:CREA binding to the consensus sequences. It is possible that only one molecule of protein binds to each site, and that there are two lower affinity sites and one higher affinity site. This would explain the appearance of three bands, but only two bands appear to be susceptible to competition. Another option is that the bottom band is an artifact, and only the upper two bands are significant. (For the sake of clarity, these two bands will be referred to as "upper" and "lower", and the bottom band will be ignored hereafter.) In this case, there is the possibility that the two bands are caused by different sizes of protein binding to the DNA. It is also possible that only two sites bind GST:CREA. One molecule of the protein can bind to either of these sites with low affinity, but if both sites are bound, the two protein molecules stabilize each other and bind with greater affinity to the DNA. This would explain the faint lower band, and the more intense upper band. It should be noted that this phenomenon could just be a result of the two GST moieties dimerizing, and not indicative of the *in vivo* situation.

Looking at the results gleaned from running all the mutant constructs in parallel, it is possible to make a few more inferences as to the interaction of GST:CREA and the CREA consensus binding sequences. Only data from constructs whose sequences were confirmed to be correct are considered. When either CREA site 1 or 2 is mutated, only a faint band is visible. Mutating CREA site 3 has no impact on GST:CREA binding, and leads to the conclusion that sites 1 and 2 are important. A double mutant of CREA sites 2 and 3 produces no shifted band. For the triple mutants no shifted bands were visible indicating that CREA was not able to bind, which was the expected result. It could be

concluded that a double mutant behaves in the same manner as a triple mutant, however, competition studies performed with E1 -- a double mutant of CRE sites 1 and 2 -- show that when a fifty-fold excess of unlabelled E1 DNA is added, slight competition is observed. This behaviour must be studied more before any conclusions can be drawn.

Competition studies with unlabeled self-DNA and unlabeled E1 (double mutant) DNA bolster the above conclusions. CRE site 1 mutants produce faint shifted bands. CRE site 3 mutants produce intense bands that are similar to those produced by wild-type DNA. Mutants of CRE sites 2 and 3 show once again no shifted bands. A very slight amount of competition is visible with E1 DNA, which is attributed to it being a double, not a triple mutant.

I suggest that CRE sites 1 and 2, which are oriented tail to tail and separated by 9 bp, as well as CRE site 3, are all individually able to bind CREA with low affinity. If a molecule of protein is bound to two of the three sites, the interactions between the proteins stabilize the DNA-protein complex. The complex formed between CREA proteins bound to CRE sites 1 and 2 is most energetically favourable, although CREA proteins bound to both CRE sites 1 and 3 or CRE sites 2 and 3 may also stabilize each other. I attribute the faint upper bands seen in Figures 24 and 25 in the lanes containing wild type DNA to the binding of three molecules of CREA to the three CRE sites in the promoter. The existence of a faint lower band and an intense upper band in Figure 23 is discounted due to this result not being duplicated. The binding of three molecules occurs only when high concentrations of CREA are present. Various researchers investigating

the mechanisms of CREA-promoter interaction have observed this phenomenon of paired CREA sites (Cubero and Scazzocchio, 1994; Espeso and Peñalva, 1994; Panozzo *et al.*, 1998; Strauss *et al.*, 1999; Mathieu *et al.*, 2000). As always, this suggestion is tempered by the knowledge that phenomenon may be due to GST-GST dimerization.

4.4 Future avenues of study.

4.4.1 GST:CREA fusion protein

In further studies using this fusion protein, removal of all the partially translated proteins from the purified product should be effected via chromatography or other means. Another option would be to insert the *creA* coding sequence into a plasmid that codes for a carboxy-terminal affinity tag. In this way, only proteins that contain the completely translated construct will possess the affinity tag for purification, removing the problem of partially translated products. I believe that the most effective method will incorporate both of these strategies. First, the CREA protein should be expressed as a carboxy-terminal fusion construct. Second, this construct should be transformed into the Rosetta© strain, and induction conditions tailored to produce maximum yield. Once we are in possession of full length CREA proteins, we can perform studies to conclude whether CREA exists in a monomeric or dimeric state.

4.4.2 CREA binding assays with mutated *glaA* promoter sequences

Once the sequences of all the constructs are confirmed, and once the full set of constructs is complete, more definitive conclusions can be made as to the binding of CREA to the various mutants. Along with the *in vivo* studies of these constructs that are currently underway, a better understanding of the interactions between CREA and the *glaA* promoter region of *A. niger* will be developed.

Literature Cited

Aleksenko, A. and Clutterbuck, A.J. 1997. Autonomous plasmid replication in *Aspergillus nidulans*: AMA1 and MATE elements. Fungal Genet. Biol. 21: 373-387

Archer, D.B. and Peberdy, J.F. 1997. The molecular biology of secreted enzyme production by fungi. Crit. Rev. Biotechnol. 17(4): 273-306.

Arst, H.N., Jr. and Cove, D.J. 1973. Nitrogen metabolite repression in *Aspergillus nidulans*. Mol. Gen. Genet. 126: 111-141

Bautista, L.F., Aleksenko, A., Hentzer, M., Santerre-Henriksen, A. and Nielsen, J. 2000. Antisense silencing of the *creA* gene in *Aspergillus nidulans*. Appl. Environ. Microbiol. 66(10): 4579-4581

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

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 Genet. Biol. 27(2-3): 243-252

Carlile, M.J. and Watkinson, S.C. 1994. Genetic variation and evolution in, The Fungi. Academic Press. London. UK

Cassart, J.-P., Östling, J., Ronne, H. and Vandenhaute, J. 1997. Comparative analysis in three fungi reveals structurally and functionally conserved regions in the Mig1 repressor. Mol. Gen.Genet. 255: 9-18

Celenza, J.M. and Carlson, M. 1984. Cloning and genetic mapping of SNF1, a gene required for expression of glucose-repressible genes in *S. cerevisiae*. Mol. Cell. Biol. 4: 49-53

Celenza, J.M. and Carlson, M. 1986. A yeast gene that is essential for release from glucose repression encodes a protein kinase. Science. 233: 1175-1180

Celenza, J.M. and Carlson, M. 1989. Mutational analysis of the *Saccharomyces cerevisiae* SNF1 protein kinase and evidence for functional interaction with SNF4 protein. Mol. Cell. Biol. 9: 5034-5044

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

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

Dzikowska, A., Swianiewicz, M. Talarczyk, A. Wisniewska, M., Goras, M., Scazzocchio, C. and Weglenski, P. 1999. Cloning, characterization and regulation of the ornithine transaminase (*otaA*) gene of *Aspergillus nidulans*. Curr. Genet. 35(2): 118-126

El-Baradi, T. and Pieler, T. 1991. Zinc finger proteins: what we know and what we would like to know. Mech. Dev. 35:155-169

Espeso, E.A. and Peñalva, M.A. 1994. In vitro binding of the two-finger repressor CreA to several consensus and non-consensus site at the *ipnA* upstream region is context dependent. FEBS Lett. 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

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

Gancedo, J.M. 1992. Carbon catabolite repression in yeast. Eur. J. Biochem. 206: 297-313

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

Hawker, L.E. 1971. The physiology of reproduction in fungi. Hafner Publishing Company. New York, NY

Hohmann, S. and Mager, W.H. 1997. Yeast stress responses. R.G. Landes Company. Austin, Texas

Hynes, M.J. and Kelly, J.M. 1977. Pleiotropic mutants of *Aspergillus nidulans* altered in carbon metabolism. Mol. Gen. Genet. 150: 193-204

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

Kato, M., Sekine, K. and Tsukagoshi, N. 1996. Sequence-specific binding sites in the Taka-amylase A G2 promoter for the CreA repressor mediating carbon catabolite repression. Biosci. Biotechnol. Biochem. 60(11): 1776-9

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

Kato, M., Hayashi, K., Kobayashi, T. and Tsukagoshi, N. 2000. A simple and rapid method for the preparation of a cell-free extract with CCAAT-binding activity from filamentous fungi. Biosci. Biotechnol. Biochem. 64(2): 455-457

Kimura, K., Wanatabe, T., Sunagawa, T. and Usami, S. 1999. Citric acid production from xylan and xylan hydrolysate by semi-solid culture of *Aspergillus niger*. Biosci. Biotechnol. Biochem. 63(1): 226-228

Kulmburg, P., Sequeval, D., Lenouvel, F., Mathieu, M. and Felenbok, B. 1992. Identification of the promoter region involved in autoregulation of the transcriptional activator ALCR in *Aspergillus nidulans*. Mol. Cell. Biol. 12(5): 1932-1939

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*. Mol. Microbiol. 7(6): 847-857

Litzka, O., Papagiannopolous, P., Davis, M.A., Hynes, M.J. and Brakhage, A.A. 1998. The penicillin regulator PENR1 of *Aspergillus nidulans* is a HAP-like transcriptional complex. Eur. J. Biochem. 251(3): 758-767

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

Lyven. http://www.lyven.com/products.htm

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): 123-131

Mercado, J.J., Vincent, O. and Gancedo, J.M. 1991. Regions of the promoter of the yeast *fbp1* gene implicated in catabolite repression may bind the product of the regulatory gene *Mig1*. FEBS Lett. 291: 97-100

Mulder, E.G. 1938. Influence of copper on growth of microorganisms. Ann. Ferment. 4: 513-533

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

Nardelli, J., Gibson, T.J., Vesque, C. and Charnay P. 1991. Base sequence discrimination by zinc-finger DNA-binding domains. Nature. 349:175-178

NCBI Taxonomy Browser.
http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/index.cgi

Nehlin, J.O. and Ronne H. 1990. Yeast MIG1 repressor is related to the mammalian early growth response and Wilms' tumour finger proteins. EMBO J. 9: 2891-2898

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. 10: 3373-3377

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

Orejas, M., MacCabe, A.P., Perez-Gonzalez, J.A., Kumar, S. and Ramon, D. 2001. The wide-domain carbon catabolite repressor *CreA* indirectly controls expression of the *Aspergillus nidulans xlnB* gene, encoding the acidic endo-beta-(1,4)-xylanase X(24). J. Bacteriol. 183(5): 1517-1523

Östling, J. and Ronne, H. 1998. Negative control of the Mig1p repressor by Snf1p-dependent phosphorylation in the absence of glucose. Eur. J. Biochem. 252: 162-168

Östling, J., Cassart, J.-P., Vandenhaute, J. and Ronne, H. 1998. Four hydrophobic amino acid residues in the C-terminal effector domain of the yeast Mig1p repression are important for its in vivo activity. Mol. Gen. Genet. 260:269-279

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. and 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

Pàrraga, G., Horvath, S.J., Eisen, A., Taylor, W.E., Hood, L., Young, E.T. and Klevit, R.E. 1988. Zinc-dependent structure of a single-finger domain of yeast ADR1. Science. 241(4872): 1489-92

Pavletich, N.P. and Pabo, C.O. 1991. Zinc finger-DNA recognition: Crystal structure of a Zif268-DNA complex at 2.1Å. Science. 252(5007): 809-817

Pontecorvo, G. 1953. The genetics of *Aspergillus nidulans*. Adv. Genet. 5: 141-238

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

Ruijter, G.J.G., Vanhanen, S.A., Gielkens, M.M.C., van de Vondervoort, P.J.I. and Visser, J. 1997. Isolation of *Aspergillus niger creA* mutants and effects of the mutations on expression of arabinases and L-arabinose catabolic enzymes. Microbiol. 143: 2991-2998

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

Sakai, A., Shimizu, Y., Kondou, S., Chibazakura T. and Hishinuma, F. 1990. Structure and molecular analysis of *RGR1*, a gene required for glucose repression in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 10: 4130-4138

Santerre Henriksen, A.L., Even, S., Müller, 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. Miccrobiology. 145(3): 729-734

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: 950-959

Shroff, R.A., O'Connor, S.M., Hynes, M.J., Lockington, R.A. and Kelly, J.M. 1997. Null Alleles of *creA*, the regulator of carbon catabolite repression in *Aspergillus nidulans*. Fungal Genet. Biol. 22: 28-38

Sophianopoulou, V., Suárez, T., Diallinas, G. and Scazzocchio, C. 1993. Operator derepressed mutations in the proline utilization gene cluster of *Aspergillus nidulans*. Mol. Gen. Genet. 236: 209-213

Stahl, U. and Tudzynski, P. eds. 1992. Molecular biology of filamentous fungi. VCH. Weinheim, Federal Republic of Germany

Steidl, S., Papagiannopoulos, O.L., Andrianopoulos, A., Davis, M.A., Brakhage, A.A. and Hynes, M.J. 1999. *AnCF*, the CCAAT binding complex of *Aspergillus nidulans*, contains products of the *hapB*, *hapC*, and *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.J., 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

Takashima, S., Nakamura, A., Hidaka, M., Masaki, H. and Uozumi, T. 1998. Isolation of the *creA* gene from the cellulolytic fungus *Humicola grisea* and analysis of CreA binding sites upstream from the cellulase genes. Biosci. Biotechnol. Biochem. 62(12): 2364-70

Thom, C. and Raper, K. B. 1945. A manual of the *Aspergilli*. Williams & Wilkins. Baltimore, MD

Tomee J.F., and van der Werf, T.S. 2001. Pulmonary aspergillosis. Neth. J. Med. 59(5): 244-58

Treitel, M.A. and Carlson, M. 1995. Repression by SSN6-TUP1 is directed by MIG1, a repressor activator protein. Proc. Natl. Acad. Sci. USA. 92:3132-3136

Tudzynski, B., Liu, S. and Kelly, J.M. 2000. Carbon catabolite repression in plant pathogenic fungi: isolation and characterization of the *Gibberella fujikuroi* and *Botrytis cinerea creA* genes. FEMS Microbiol.Lett. 184(1): 9-15

U. S. Food and Drug Administration Center for Food Safety & Applied Nutrition -- **Office of Food Additive Safety**. http://www.cfsan.fda.gov/~rdb/opa-gras.html

Vautard, G., Cotton, P. and Fèvre, M. 1999. The glucose repressor CRE1 from *Sclerotinia sclerotiorum* is functionally related to CREA from *Aspergillus nidulans* but not to the Mig proteins from *Saccharomyces cerevisiae*. FEBS Lett. 453: 54-58

Verdoes, J.C. 1994. Molecular genetic studies of the over production of glucoamylase in *Aspergillus niger*. Vrije Universiteit Academisch Proefschrift. Amsterdam, Nederlands

Verdoes, J.C., Punt, P.J., Stouthamer, A.H. and van den Hondel, C.A. 1994. The effect of multiple copies of the up stream region on expression of the *Aspergillus niger* glucoamylase-encoding gene. 145(2): 179-187

Wang, J., Sirenko, O. and Needleman, R. 1997. Genomic footprinting of Mig1p in the MAL62 promoter. J. Biol. Chem. 272(7): 4613-4622

Webster, J. 1970. Introduction to fungi. Cambridge University Press. London, UK

Worral, J.J. ed. 1999. Structure and dynamics of fungal populations. Kluwer Academic Publishers. Dordrecht, the Netherlands