The role of the *acrB* and *creD* genes in carbon catabolite repression in *Aspergillus nidulans.*

by

Natasha Anne Boase, B.Sc. (Hons).

A thesis submitted in total fulfilment of the requirements for the degree of Doctor of Philosophy.

School of Molecular and Biomedical Science
Discipline of Genetics
The University of Adelaide
(Faculty of Sciences)

May 2004

Table of Contents

Abstract v
Declaration vii
Acknowledgements viii
List of Figures ix
List of Tables x
List of Abbreviations xi

CHAPTER 1: INTRODUCTION 1

1.1. *A. NIDULANS* AS A MODEL ORGANISM 1

1.2 CARBON CATABOLITE REPRESSION IN *S. CEREVISIAE* 3
- 1.2.1 Mig1p 3
- 1.2.2 Ssn6p-Tup1p corepressor complex 3
- 1.2.3 Snf1p and Glc7p-Reg1p 4
- 1.2.4 Mechanism of carbon catabolite repression 4

1.3 CARBON CATABOLITE REPRESSION IN *A. NIDULANS* 6
- 1.3.1 *creA*, phenotypic and molecular analyses 7
- 1.3.2 *creB* and *creC* 9
- 1.3.3 Signalling in carbon catabolite repression 11
- 1.3.4 *creD* and *acrB* 11
- 1.3.5 Model of carbon catabolite repression 12

1.4 AMYLASES 14
- 1.4.1 Acid-stable α-amylases 16
- 1.4.2 Acid-unstable α-amylases 17
- 1.4.3 Regulation of α-amylases 18

1.5 AIMS AND OBJECTIVES 20

CHAPTER 2: MATERIALS AND METHODS 21

2.1 MATERIALS 21
- 2.1.1 Enzymes 21
- 2.1.2 Radioactive isotopes and nucleic acids 21
- 2.1.3 Molecular weight markers 21
- 2.1.4 Bacterial strains 22
- 2.1.5 *A. nidulans* strains 22
- 2.1.6 Oligonucleotides 23
- 2.1.7 Vectors 25
- 2.1.8 *A. nidulans* DNA libraries 25
- 2.1.9 Kits and miscellaneous materials 25

2.2 SOLUTIONS, BUFFERS AND MEDIA 26
- 2.2.1 Solutions and buffers 26
- 2.2.2 Media 27

2.3 METHODS 27
- 2.3.1 Manipulation of *A. nidulans* 27

2.3.2 Nucleic acid isolation 27
2.3.3 Polymerase Chain Reaction (PCR) 28
2.3.4 DNA sequencing 28
2.3.5 Rapid amplification of 5' cDNA ends (5' RACE) 28
2.3.6 DNA gel electrophoresis and Southern transfer 28
2.3.7 RNA gel electrophoresis and Northern transfer 29
2.3.8 Ribonuclease Protection Assays 29
2.3.9 Radioactive labelling of probes 29
2.3.10 Bacterial-2-hybrid system 30
2.3.11 Amylase secretion plate tests 30
2.3.12 Sequence analyses 31
2.3.13 Phylogenetic analyses 31

2.4 NUCLEOTIDE SEQUENCE ACCESSION NUMBERS 32

CHAPTER 3: CLONING AND MOLECULAR ANALYSIS OF *CRED* 33

3.1 INTRODUCTION 33

3.2 CLONING *CRED* VIA COMPLEMENTATION OF THE *CRED34* PHENOTYPE ON ACRIFLAVINE USING COSMIDS AND COSMID DERIVATIVES 33
3.2.1 The *creD34* allele is recessive to the wild-type allele 33
3.2.2 Testing cosmids L21B08, P24E11, and relevant derivatives for complementation of the *creD34* mutant phenotype 34
3.3.3 Sequence analysis of pB3 38

3.3 COMPLETE COMPLEMENTATION OF THE *CRED34* PHENOTYPE 40

3.4 IDENTIFICATION OF *CRED* WITHIN BAC 22E20 41
3.4.1 Complementation analysis of five regions within 22Notrel1 43

3.5 MOLECULAR ANALYSIS OF *CRED* 45
3.5.1 Sequence analysis of *creD* and identification of the *creD34* mutation 45
3.5.2 Bioinformatic analysis of *creD* and identification of *apyA* 47
3.5.3 Multiple copies of *apyA* do not compensate for the absence of *creD* 49

3.6 PROTEINS THAT INTERACT WITH CRED AND APYA 49
3.6.1 Testing protein-protein interactions between CreD and CreA, ApyA and CreA 49
3.6.2 Identification of *hulA* and testing HulA for interactions with CreD and ApyA 53

3.7 DISCUSSION 55

CHAPTER 4: MOLECULAR CHARACTERIZATION AND ANALYSIS OF ACRB 61

4.1 INTRODUCTION 61

4.2 PHENOTYPIC ANALYSIS OF *ACRB2*, *ACRB14* AND *ACRB15* 61

4.3 THE *ACRB2* MUTATION SUPPRESSES ASPECTS OF THE *CREA*, *CREB* AND *CREC* MUTANT PHENOTYPES 65

4.4 MOLECULAR ANALYSIS OF *ACRB* 68

4.4.1 Sequence analysis of *acrB* 68
4.4.2 Molecular analysis of the *acrB* mutant alleles 72
4.5 Protein interactions involving AcrB 72
4.5.1 Screening for proteins that interact with AcrB 72
4.5.2 Testing AcrB for interactions with other known proteins 73
4.6 Discussion 73

CHAPTER 5: IDENTIFICATION OF AN AMYLASE CLUSTER 75

5.1 Introduction 75
5.2 Sequence analysis of the amlyase cluster 76
5.3 *amyB* is not part of the amylase cluster 82
5.4 Regulation of the amylases 82
5.4.1 Alpha-amylase secretion plate tests 83
5.5 Discussion 83

CHAPTER 6: CONCLUSIONS 89

APPENDIX A 95

REFERENCES 99

Abstract

This work describes the cloning and analysis of *creD*, and the characterization of the *acrB* gene, two components of a regulatory network controlling carbon source utilization in the filamentous fungus *Aspergillus nidulans* that involves ubiquitination and deubiquitination. Preliminary characterization of a novel amylase cluster in *A. nidulans* was also carried out as part of this project.

Carbon catabolite repression is a regulatory system which allows the utilisation of the most easily metabolised carbon source present (usually glucose), and results in repression of the expression of a large number of genes involved in the synthesis of various enzymes required for the utilisation of less favoured carbon sources. In *A. nidulans* three components of this regulatory network have previously been characterized; *creA,* which encodes a transcriptional repressor protein, *creB* which encodes a deubiquitinating enzyme, and *creC* which encodes a protein that contains five WD40-repeat motifs and a proline-rich region. CreB and CreC form a complex *in vivo* and it has been proposed that this complex acts to stabilize or modify the CreA repressor protein by removing ubiquitin moieties that either target CreA for destruction via the 26S proteasome or alter the function.

The *creD34* mutation had been identified as a suppressor of the effects of the *creC27* mutation, and it was also shown to suppress the effects of the *creB15* mutation, suggesting a role for *creD* in this regulatory network. The *creD* gene was cloned by complementation and physically analysed, and it encodes a protein that contains an arrestin_N and arrestin_C domain, and a PP*X*Y motif and two P*X*Y motifs known to be involved in protein-protein interactions. CreD is similar to the Rod1p and Rog1p proteins from *Saccharomyces cerevisiae,* and a search of the *A. nidulans* genome led to the identification of *apyA* in *A. nidulans*, another arrestin_N and arrestin_C domain containing protein with a single PP*X*Y motif. Rod1p and Rog3p interact with the HECT ubiquitin ligase Rsp5p in *S. cerevisiae*, and so the homologue of Rsp5p was identified in *A. nidulans* and designated *hulA*. Both CreD and ApyA interacted with HulA, as determined by the bacterial-2-hybrid system, with ApyA and HulA interacting more strongly than CreD and HulA.

Like *creD34*, the *acrB2* mutation results in an acriflavine-resistant phenotype, and so *acrB* was tested for suppression of the phenotypes due to *creA*, *creB* and *creC* mutations. Mutations in *acrB* result in altered utilization of various carbon sources in *A. nidulans*, and are heterogeneous in nature. The effects of the *acrB2* mutation are epistatic to those due to the *creB* and *creC* mutations, indicating a role for AcrB in this regulatory network. The *acrB* gene had been cloned via complementation and the physical analysis is reported here, with *acrB* encoding a

protein that contains three transmembrane domains and a coiled-coil region. The coiled-coil region of AcrB was subcloned into the bacterial–2-hybrid system for testing protein-protein interactions, but none were identified.

An amylase cluster was identified and preliminary characterization begun by R. Murphy. Here the sequence analysis of the cluster which contains three genes, *amyA*, an α-amylase, *agdA*, an α-glucosidase, and *amyR*, the transcriptional activator of amylolytic genes is presented. It was shown that a previously identified amylase in *A. nidulans*, *amyB*, did not form part of this amylase cluster.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Signed :

Natasha Boase
4th May, 2004

Data presented in this thesis have been published in the following papers:

Boase, N. A., Lockington, R. A., Adams, J. R. J., Rodbourn, L. and Kelly, J. M. (2003). Molecular characterization and analysis of the *acrB* gene of *Aspergillus nidulans*: A gene identified by genetic interaction as a component of the regulatory network that includes the CreB deubiquitination enzyme. *Genetics* **164,** 95-104.

Boase, N. A. and Kelly, J. M. (2004). A role for *creD*, a carbon catabolite repression gene from *Aspergillus nidulans*, in ubiquitination. *Molecular Microbiology* (In press).

Acknowledgements

Firstly, I would like to acknowledge my supervisor Dr. Joan Kelly, whose guidance and support is greatly appreciated. Joan, you are an inspiration both in and out of the lab, and I thank you for your candour and professionalism throughout my years here. You have set a great example of how to deal with people, and to do what needs to be done! Thankyou also for the prompt response when reviewing this thesis. I fear I have been spoilt under your supervision!

Thankyou to Robin, whose enthusiasm and amazing knowledge in the lab is irreplaceable. You seem to always have the answers to my questions, and provide a different perspective and renewed hope. Another thankyou for reading parts of this thesis.

Thanks to my fellow PhD student Jai. It has been great having you in the lab – and not in the least for the food (!) provided and late night chats. And I love the book!! Thanks to Damien F., good luck with your honours year and the many adventures I am sure lie ahead.

I'd also like to thank people past, who have made the year(s) pass quicker in the lab, including Julian, Damien A., Louise, Lidia and Jairus. And another thankyou to Rachael, who taught me so much in honours, I hope it has stood me in good stead...

A blanket thankyou to the genetics discipline/department. I have enjoyed my time here, and will miss many of you. Thanks to Michelle for the tutoring work, and for being so sociable. Thanks to current and past (Timmis lab) members; John and Damien for making prac demonstrating more interesting with a few footy discussions, and for the Friday night beers. Sas, for always being there for a chat/beer, and Coralie, for the accommodations in Lorne/Melbourne and the fun that seems to follow!

Wow, more people to thank. My fellow honours students, Janet and Mel, Marie, Jane and Bec. I'm glad we have all stayed in touch! Marie for the monthly schnitty at the Cath, Janet for just being you ! and Mel for keeping me sane and being such a good friend.

Thanks to the past and present members of the Alladins/Genies. A couple of premierships, and maybe even a wooden spoon this season ☹, but it has kept Tuesday nights interesting for the last few years.... And it seems our run of injuries may finally be over!

To my great mates Lyndall and Nat, thanks for your friendship over the last couple of decades! Celebratory drinks coming up!

Thankyou to Matt and Rach, and Les and Kay, for providing such great company each week as we watch our beloved footy team. Go the crows!!

The family unit must rate a mention. Mum and Dad, you have been so supportive for the last 30 (!) years, I love you both and am thankful for all that you have done. Nana and Papa, for always encouraging me to learn and spoiling me since I was born, Nanna and Aunty Rob for all the prayers and good wishes. Thanks to my big brother Andrew, and also to Amy and KB, two great chicks that I couldn't live without!

Thankyou to all of the Thewlis Family, who have helped me out in so many different ways for so many years now. A big thankyou to Carmel, whose courage is an inspiration! A special prayer for Paul and James, who we miss so very much. Thanks especially to James for making me feel like it was cool to be a scientist.

And lastly, Simon. Hmmm, so many things, so little room. Of course thankyou for putting up with me being a student so long, and just for putting up with me in general ☺. You are my best friend, and my biggest support. I am most looking forward to us doing all of those things that are on the "when you finish uni" list... or did you think it would never happen? Thankyou for being such a great partner.

List of Figures

Chapter 1

Figure 1.1 Lifecycle of *A. nidulans* 2
Figure 1.2 Model of carbon catabolite repression in *S. cerevisiae* 5
Figure 1.3 Model of carbon catabolite repression in *A. nidulans* 13
Figure 1.4 Enzymatic degradation of starch 15

Chapter 3

Figure 3.1 The *creD34* allele is recessive to the wild-type allele 35
Figure 3.2 Partial map of cosmids L21B08, P24E11 and derivatives tested for complementation of the *creD34* phenotype on acriflavine 36
Figure 3.3 Partial complementation of the *creD34* mutant phenotype 37
Figure 3.4 Map of the hypothetical genes contained within pB3, and ORF of the intergenic region pBd 39
Figure 3.5 Complementation of the *creD34* mutant phenotype 42
Figure 3.6 A schematic representation of the hypothetical genes within plasmid 22NotreI1 44
Figure 3.7 Location of the *creD34* mutation within the *creD* gene 46
Figure 3.8 Schematic representation of CreD and ApyA 48
Figure 3.9 Evolutionary relationships between CreD and other fungal arrestin domain containing proteins 50
Figure 3.10 Determining *apyA* copy number in tapyA transformant strains via Southern analysis 51
Figure 3.11 Multiple copies of *apyA* does not compensate for lack of *creD* 52
Figure 3.12 Evolutionary relationships between the HECT ubiquitin ligases in *S. cerevisiae* and *A. nidulans* 54
Figure 3.13 Investigation of protein-protein interactions between CreD and HulA, and ApyA and HulA as determined by the BacterioMatch II bacterial-2-hybrid system 56

Chapter 4

Figure 4.1 Phenotypic heterogeneity among *acrB* mutant alleles 62
Figure 4.2 Epistatic effects of the *acrB2* mutation on *creA*, *creB* and *creC* mutant phenotypes 66-67
Figure 4.3 Location of the *acrB* mutations within the *acrB* gene 69
Figure 4.4 Schematic of AcrB 71

Chapter 5

Figure 5.1 Schematic of the amylase cluster 77
Figure 5.2 Nucleotide sequence of the amylase cluster 78-81
Figure 5.3 Amylase activity secretion plate tests 84

Appendix A

Figure A.1 Nucleotide sequence of the pB3 plasmid 96-97

List of Tables

Chapter 2
Table 2.1 Strains of used *E. coli* in this study 22
Table 2.2 Strains of *A. nidulans* used in this study 22
Table 2.3 Oligonucleotide primers used in this study 23-24
Table 2.4 Plasmids used in this study 25

Chapter 4
Table 4.1 Phenotypic analysis of *acrB* mutant strains 63

Appendix A
Table A.1 Custom oligonucleotides designed for the pB3 plasmid 95

List of Abbreviations

aa	amino acid(s)
A	adenine
ANGIS	Australian National Genomic Information Service
[α-^{32}P]dATP	alpha-labelled deoxyadenosine triphosphate
B2H	bacterial - 2 - hybrid system
BLAST	basic local alignment search tool
bp, kb	base pairs, kilobase pairs
BSA	bovine serum albumin
C	cytosine
°C	degrees Celsius
cDNA	deoxyribonucleic acid complementary to ribonucleic acid
CCR	carbon catabolite repression
C-terminal	carboxyl-terminal
Da, kDA	Dalton(s), kiloDaltons
dATP	2'-deoxyadenosine-5'-triphosphate
dCTP	2'-deoxycytosine-5'-triphosphate
dGTP	2'-deoxyguanosine-5'-triphosphate
dH20	distilled water
DNA	deoxyribonucleic acid
dNTPs	2'-deoxynucleotide-5'-triphosphates
DTT	dithiothreitol
dTTP	2'-deoxythymidine-5'-triphosphate
EDTA	ethylenediaminetetraacetic acid
FSP	forward sequencing primer
G	guanine
g, mg, μg, ng	gram(s), milligram(s), microgram(s), nanogram(s)
IPTG	isopropyl–β-D-thiogalactoside
λ	bacteriophage lamda
L, ml, μl	litre(s), millilitre(s), microlitre(s)
M, mM	moles per litre, millimoles per litre
mA	milliAmperes
MEGA	Molecular Evolutionary Genetics Analysis
min, hr	minute(s), hour(s)
mRNA	messenger ribonucleic acid
MW	molecular weight
NCBI	National Center of Biotechnology Information
NMR	nitrogen metabolite repression
No.	number
N-terminal	amino-terminal
nt	nucleotide(s)
ORF	open reading frame
Orthologous	Two genes are orthologous if they diverged after a speciation event
Paralogous	Two genes are paralogous if they diverged after a duplication event
PCR	polymerase chain reaction
PEG	polyethylene glycol

5'-RACE	rapid amplification of 5' cDNA ends
®	registered
RNA	ribonucleic acid
rRNA	ribosomal RNA
RSP	reverse sequencing primer
SDS	sodium dodecyl sulphate
SSC	saline sodium citrate
™	Trademark
T	thymidine
TAE	Tris-acetate EDTA
Taq	*Thermus aquaticus*
TE	Tris-EDTA
Tris	Tris[hydroxymethyl] amino methane
U	Unit(s) of enzyme
Ub	ubiquitin
UTR	untranscribed region
UV	ultraviolet
X-gal	5'-bromo-4-chloro-3-indoyl-β-D-galactopyranoside
% (v/v)	percent volume per volume
% (w/v)	percent weight per volume

CHAPTER 1: INTRODUCTION

Many microorganisms are able to adapt to utilise a number of nutritional substrates to obtain carbon and nitrogen, and have developed mechanisms to regulate gene expression in response to their environment. In the filamentous fungus *Aspergillus nidulans* the preferred source of energy is the oxidation of glucose, and in the presence of a variety of potential carbon sources glucose is used preferentially to less readily metabolised carbon sources. Carbon metabolism is adjusted according to the immediate availability of glucose by the general process of carbon catabolite repression, and also via induction by pathway specific regulatory genes.

Carbon catabolite repression is a wide domain regulatory system that responds to the local nutritional environment to regulate carbon metabolism. In the presence of the preferred carbon source glucose, the structural genes that encode enzymes and permeases for utilising alternative sources of carbon such as ethanol, proline, acetamide and starch are repressed. Upon depletion of glucose these pathway specific enzymes are expressed to utilize the specific alternative carbon source present in the environment. Hence the organism conserves energy by only producing proteins as they are required. Carbon catabolite repression has been identified in bacteria and fungi, and has been most extensively studied in the yeast *Saccharomyces cerevisiae.*

1.1 *A. nidulans* as a model organism

A. nidulans is an ascomycete fungus that has proven to be a very successful experimental organism in which to study gene regulation and, in particular, carbon catabolite repression. *A. nidulans* has the ability to grow on simple defined solid media as compact colonies, with most strains producing mature colonies within 48 hours on defined medium, as well as the ability to grow in submerged culture. *A. nidulans* has both mitotic asexual and meiotic sexual life cycles, plus a parasexual cycle that can be manipulated under laboratory conditions to induce strains to cross (Figure 1.1). *A. nidulans* has a haploid genome under normal circumstances, allowing for the direct observation of recessive phenotypes. Diploids and heterokaryons can be forced from complementing auxotrophs (strains with different nutritional requirements from each other), and there is heterokaryon compatibility between all laboratory strains as they derived from a single source. Maintenance of *A. nidulans* can be by asexual reproduction via mycelial propagation, or through the germination of conidia. Large numbers of *A. nidulans* colonies can be screened easily in both a practical and economic sense, with strong selective techniques. The ease of genetic manipulation, the ability to perform mutagenesis, recombination,

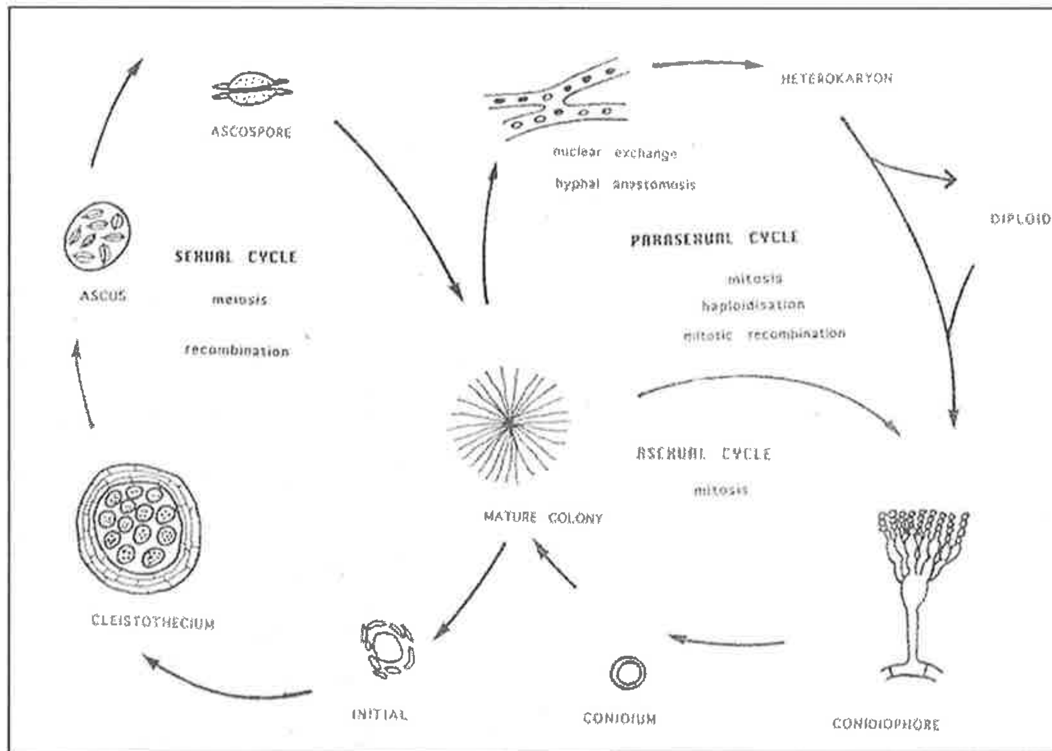

Figure 1.1 Lifecycle of *A. nidulans.*

Figure taken from Figure 4.1, Martinelli (1994)

complementation and dominance tests, and the efficiency of transformation (homologous and heterologous) make *A. nidulans* an ideal model system for the study of gene regulation. The sequence of the 30.1 Mb *A. nidulans* genome has recently been publicly released, complementing genetic mapping (Clutterbuck, 1997) and allowing bioinformatic analyses. *A. nidulans* has well documented biochemical pathways, under the regulation of pathway specific and global regulatory controls, permitting the investigation of regulatory networks.

1.2 Carbon catabolite repression in *S. cerevisiae*

Carbon catabolite repression has been extensively studied in the yeast *S. cerevisiae*, and the mechanisms involved have been well elucidated. The central components of carbon catabolite repression can be classified into positive regulators (i.e. the Snf1p protein kinase complex), and negative regulators of carbon source utilization (i.e. the Mig1p transcriptional repressor, the Ssn6p-Tup1p corepressor complex and the Reg1p-Glc7p protein phosphatase 1), as detailed below.

1.2.1 Mig1p

MIG1 encodes a transcriptional repressor containing two DNA binding Cys_2His_2 zinc-fingers in the amino terminal region, a nuclear export signal and a nuclear localization signal (Nehlin and Ronne, 1990). Mig1p binds DNA at the consensus sequence (5'-WWWWWWNSYGGGG-3'), found in carbon catabolite repressible promoters (Nehlin and Ronne, 1990). In the presence of high levels of glucose Mig1p rapidly moves into the nucleus where it is able to bind to the promoters of glucose-repressible genes. Mig1p recruits the Ssn6p-Tup1p corepressor complex to the promoters of many glucose-repressed genes to effect repression (Treitel and Carlson, 1995). Upon depletion of glucose, Mig1p is phosphorylated, dissociates with the corepressor complex and is exported from the nucleus to the cytoplasm via the nuclear exportin Msn5p (DeVit *et al.*, 1997; DeVit and Johnston, 1998). This glucose dependant sub-cellular localization of Mig1p responds quite rapidly to the presence/absence of glucose.

1.2.2 Ssn6p-Tup1p corepressor complex

Tup1p is a pleiotropic repressor containing six WD40 repeats (mediating protein-protein interactions), and Ssn6p is a pleiotropic repressor containing ten tetratricopeptide repeats (TPR). These two proteins associate in a high molecular weight corepressor complex, composed of one Ssn6p subunit and multiple Tup1p subunits (Varanasi *et al.*, 1996; Redd *et al.*, 1997). This corepressor complex utilizes different DNA binding proteins to repress transcription of genes regulated not only by glucose (via Mig1p, see above), but also by DNA damage, oxygen and cell

type. Distinct combinations of TPR motifs are required for Ssn6p to interact with these specific DNA binding proteins, and Ssn6p acts as a linker between the pathway specific DNA binding proteins and the Tup1p transcriptional repressor (Tzamarias & Struhl, 1995). The tethering of the corepressor complex to the promoter allows Tup1 to interact with histones H3 and H4, stabilizing adjacent nucleosomes and blocking the transcription machinery access to the promoter, thereby effecting repression (Edmendson *et al.*, 1996).

1.2.3 Snf1p and Glc7p-Reg1p

Snf1p is a serine/threonine protein kinase that is highly conserved among eukaryotes (Celenza and Carlson, 1986, 1989). Snf1p is part of a high molecular weight complex, including scaffolding proteins Sip1p, Sip2p and Gal83p, and complexes with Snf4p (Jiang and Carlson, 1997). In the presence of glucose, the Snf1p kinase is inactive, allowing Mig1p to affect repression. Snf1p contains a regulatory domain and a catalytic domain, and in the presence of glucose the regulatory domain of Snf1p autoinhibits the catalytic domain of Snf1p. In low or no glucose conditions, a conformational change results in the regulatory domain of Snf1p binding to Snf4p directly, freeing the catalytic domain of Snf1p to phosphorylate Mig1p (Jiang and Carlson, 1996). Ludin *et al.* (1998) proposed that the Snf1p kinase is phosphorylated to activate the catalytic activity of the Snf1p complex by an as yet unidentified kinase. The Glc7p-Reg1p protein phosphatase I responds to increasing levels of glucose by interacting with the active Snf1p complex, inducing a conformational change in the Snf1p complex back to the autoinhibited state, possibly by dephosphorylating the complex (Ludin *et al.*, 1998).

1.2.4 Mechanism of carbon catabolite repression

In the presence of high glucose (the repression signal), the Mig1p repressor protein is imported into the nucleus and binds at glucose-repressible promoters to effect repression by recruiting the Ssn6p-Tup1p corepressor complex (Figure 1.2). The Snf1p kinase that acts to phosphorylate Mig1p is inactive in high glucose conditions. Upon depletion of glucose (the derepression signal), the Snf1p kinase becomes active, due to the Snf4p activating subunit binding to the regulatory subunit of Snf1p, allowing the catalytic subunit of Snf1p to function. Snf1p phosphorylates Mig1p, resulting in the export of Mig1p from the nucleus to the cytoplasm, allowing expression of glucose repressible-genes.

The exact glucose (repressing) signal that affects the Snf1p complex catalytic activity is yet to be elucidated. Three kinases found in *S. cerevisiae*, hexokinases Hxk1p, Hxk2p and glucokinase Glk1p, are responsible for the phosphorylation of hexose sugars as the first

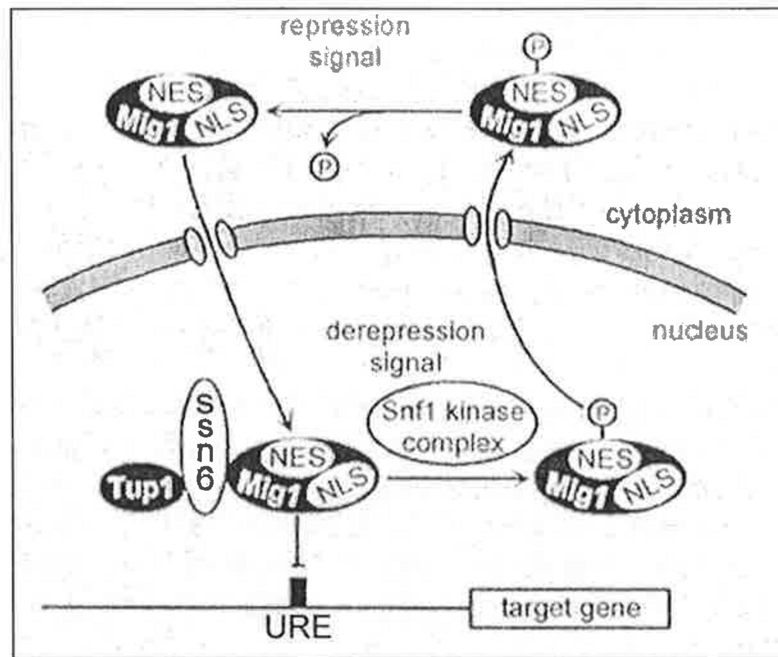

Figure 1.2 Model of carbon catabolite repression in *S. cerevisiae*.
NES Nuclear Export Sequence
NLS Nuclear Localization Sequence
URE Upstream Regulatory Element *i.e.*Mig1p binding site
Figure adapted from Figure 4, Schuller (2003).

irreversible step in glucose metabolism. These three kinases are expressed differentially (Herrero *et al.*, 1995) and are required for short-term glucose repression, but only Hxk2p is specifically required for long-term glucose repression (De Winde *et al.*, 1996). The catalytic activity of Hxk2p appears to be independent of its function in glucose signalling (Mayordomo and Sanz, 2001). Hxk2p is found in the cytoplasm and nucleus, and this nuclear localisation is necessary for glucose repression signalling (Herrero *et al.*, 1998). Hxk2p was shown to form part of a DNA binding complex at the *SUC2* (glucose-repressible) promoter to establish glucose repression (Herrero *et al.*, 1998). The phosphorylation status of Hxk2p also appears to be important for its function (Randez-Gil *et al.*, 1998). Sanz *et al.* (2000) identified a role for Hxk2p in regulating the activity of Snf1p via the phosphorylation of Reg1p, the regulatory subunit of protein phosphatase I (see Section 1.2.3), although the exact mechanism remains unknown. Hxk2p appears to have a central role in glucose sensing via the main glucose repression pathway, but the actual glucose signal that triggers repression is yet to be determined.

1.3 Carbon catabolite repression in *A. nidulans*

S. cerevisiae is a fermentative yeast that is highly adapted for growth in sugar rich environments, and has evolved to utilize glucose differently from most other organisms, preferring fermentation of glucose to ethanol rather than the oxidation of glucose, even under aerobic conditions. *A. nidulans* metabolizes glucose like most other eukaryotic organisms, undergoing aerobic oxidation via the mitochondrial pathways, and as such the regulatory system of carbon catabolite repression may be quite different to that found in *S. cerevisiae.*

In order to identify regulatory proteins that affect carbon catabolite repression in *A. nidulans* a genetic screen that exploits the interaction between carbon catabolite repression and nitrogen metabolite repression was devised (Arst & Cove, 1973). Nitrogen metabolite repression is a regulatory system whereby preferred nitrogen sources such as ammonium or glutamine are utilised preferentially to less favoured nitrogen sources, and is mediated by the positively acting regulatory gene *areA* (Arst & Cove, 1973). In the presence of glucose, and absence of ammonia or glutamine, the transcriptional activator protein AreA is required for the expression of genes required for the catabolism of alternative nitrogen sources. In the presence of glucose null alleles of *areA* are unable to grow on any nitrogen source other than ammonium as these strains cannot derepress genes that are subject to nitrogen metabolite repression. Some alternative carbon sources, such as proline and acetamide, can also serve as alternative nitrogen sources. The enzymes required for the metabolism of such compounds are therefore controlled by both carbon catabolite repression and nitrogen metabolite repression, and the relief of either repression leads

to expression. In the absence of glucose *areA* lack of function strains can grow on medium containing acetamide or proline as the only sources of carbon and nitrogen, but not in the presence of glucose. Arst and Cove (1973) exploited this interaction to identify mutations that suppress the effects of loss of function *areA* alleles such that the enzymes for acetamide or proline utilization are produced even in the presence of glucose (carbon catabolite repressing conditions). This genetic screen was used to identify regulatory proteins affecting carbon catabolite repression, encoded by the *creA*, *creB* and *creC* genes (Arst and Cove, 1973; Bailey and Arst, 1975, Hynes and Kelly, 1977).

1.3.1 *creA*, phenotypic and molecular analyses

The *creA* mutations constitute the majority of mutants identified by the genetic screen described above, and result in varying levels of derepression (i.e. inappropriate expression of genes that would normally be repressed) of a wide range of genes. For example, *creA*30 leads to high levels of derepression of acetamidase (AmdS) and alcohol dehydrogenase I (AdhI), but only very low levels of derepression of enzymes for proline utilization. On the other hand, *creA220* leads to high levels of derepression of enzymes for proline and acetamide use, but only very low levels of derepression of alcohol dehydrogenase I (Shroff *et al.,* 1997). The general trend towards derepression is observed in all mutant *creA* alleles, but individual pathways subject to carbon catabolite repression show a different spectrum of affects. The *creA* mutant alleles exhibit heterogeneous, pleiotropic phenotypes indicating a regulatory role for *creA*.

The *creA* gene in *A. nidulans* has been cloned and characterised (Dowzer and Kelly, 1989, 1991). The derived polypeptide of 416 aa contains several features characteristic of a transcriptional repressor, such as two zinc-fingers of the Cys_2His_2 class, an alanine rich region, and frequently appearing (STPXX) motifs. The consensus sequence bound by the CreA zinc finger domain is (5'-SYGGRG-3'), and the presence of an AT-rich sequence 5' to the consensus sequence can affect the binding of some of the sequences that fit this consensus (Kulmberg *et al.*, 1993; Cubero and Scazzocchio, 1994). Towards the C-terminal end of the protein a stretch of about 40 amino acids is completely conserved between a number of filamentous fungi and is similar to the Rgr1p repressor protein in *S. cerevisiae* (Drysdale *et al.,* 1993). However this Rgr1-like region is not functionally conserved with the yeast protein as the replacement of the Rgr1-like region in CreA by the corresponding Rgr1 region from *S. cerevisiae* failed to complement (Shroff, 1997).

Mutant *creA* alleles can be grouped into two broad classes, missense mutations in the DNA binding region, and frameshift or nonsense mutations that result in a truncated CreA

peptide. A null allele was constructed for *creA*, and this shared a similar mutant phenotype to a number of other identified mutations that resulted in truncated CreA protein (Shroff *et al.*, 1997). The most morphologically extreme mutation identified, which led to a very compact morphology, slow growth and a high degree of derepression, is *creA*306 that disrupts the recognition α-helix of the second zinc finger in the DNA binding domain. The *creA*306 allele is the only allele predicted to produce a full-length protein where DNA binding is completely abolished, and thus the extreme phenotype could be due to the CreA306 protein titrating other proteins that interact with CreA (Shroff *et al.*, 1997).

CreA mediates carbon catabolite repression for a number of alternative carbon utilising systems, and as such can be considered a global regulator. There appears to be a role for CreA in both repressing (presence of glucose) and nonrepressing (absence of glucose) conditions, as indicated by elevated levels of AdhI in *creA* mutants under nonrepressing (+/- induced) conditions (Mathieu and Felenbok, 1994; Shroff *et al.*, 1996). Northern and mutational analyses indicate that the *creA* gene itself is autoregulated, and it contains CreA binding sites in its own promoter (Shroff *et al.*, 1996; Strauss *et al.*, 1999). In addition to transcriptional regulation, Strauss *et al.* (1999) speculated that the CreA protein could be subject to regulation via degradation or protein inactivation under derepressing conditions as there was no efficient CreA binding to consensus oligomers in an electrophoretic mobility shift assay under these conditions.

The CreA protein has been identified in a large number of fungi, including *Aspergillus oryzae*, *Aspergillus niger*, *Sclerotinia sclerotiorum* and *Trichoderma reesei* (reviewed in Kelly, 2004). The zinc-finger region of CreA shows strong sequence similarity to the zinc-finger region of the *S. cerevisiae* Mig1p repressor (see Section 1.2.1). Cre1 from *S. sclerotiorum* is functionally related to CreA from *A. nidulans*, but cannot complement *MIG1* and *MIG2* deficiencies in *S. cerevisiae* (Vautard-Mey *et al.*, 1999). Vautard-Mey *et al.* (1999) raised polyclonal antibodies against a fusion protein, GST-Cre1, and demonstrated by subcellular fractionation that Cre1 is localised in the nuclei of glucose-grown hyphae, and in the cytoplasm when glucose is removed from the culture medium, similar to the glucose-dependant localisation of Mig1p in *S. cerevisiae*. Mutation of a putative AMPK phosphorylation site in CreA (Ser266) abolishes the repressor activity of the fusion GST-Cre1 protein, but its nuclear targeting was not affected (Vautard-Mey & Fevre, 2000). Phosphorylation has also been implicated in the DNA binding of Cre1 from *Hypocrea jecorina/Trichoderma reesei*, with the phosphorylation of Ser241 required to allow a DNA binding conformation of Cre1 (Cziferszky *et al.*, 2002). A phosphorylation-independent DNA binding mutation (S241A) resulted in the permanent carbon catabolite repression of cellobiohydrolase 1 expression (Cziferszky *et al.*, 2002). Thus phosphorylation of Cre1 results in

the ability to effect repression via DNA binding, whereas it relieves *S. cerevisiae* from Mig1 dependant repression by signalling export from the nucleus.

1.3.2 *creB* and *creC*

The genetic screen that identified the involvement of *creA* in carbon catabolite repression also identified the *creB* and *creC* genes as suppressors of *areA217* on medium contain glucose and acetamide. *creB* has previously been described as the *molB* mutation that conferred resistance to toxic concentrations of molybdate (Arst *et al.*, 1970; Arst, 1981). The *creB* and *creC* mutants share an almost identical phenotype, including resistance to molybdate and sensitivity to acriflavine in complete media. Mutations in *creB* and *creC* result in the derepression of a range of enzymes, such as acetamidase and alcohol dehydrogenase, that are usually subject to carbon catabolite repression, but the pathways affected are only a subset of those systems affected by mutations in *creA* (Hynes & Kelly, 1977). In addition, *creB* and *creC* mutations also lead to the failure to derepress enzymes for the metabolism of a range of carbon sources including quinate and proline, and lead to elevated levels of enzymes such as extracellular proteases in the presence or absence of glucose, indicating a role for CreB and CreC in both repressing and derepressing conditions (Hynes & Kelly, 1977). The shared pleiotropic mutant phenotype of *creB* and *creC* strains indicated that these genes may be regulatory, and the *creB creC* double mutant phenotype is not additive, indicating that these proteins may be part of a protein complex or act in the same pathway (Hynes & Kelly, 1977). There is no heterogeneity between different mutant alleles.

The *creB* gene was cloned utilising the linkage of *creB* to *acoB,* with a chromosome walk along chromosome II leading to the identification of a complementing cosmid. The *creB* gene encodes a 767 aa protein that is a functional deubiquitinating enzyme (Lockington & Kelly, 2001). CreB contains six deubiquitination (DUB) homolog domains, indicative of members of the ubiquitin processing protease (ubp) family, a coiled-coil region, most likely to be involved in substrate recognition, and four predicted PEST sequences, which are implicated as signals for proteolysis. There are highly conserved proteins to CreB in humans (UBH1), *Arabidopsis thaliana* (UBP3), *Pichia anomala* (UBP1), and hypothetical proteins from *Drosophila melanogaster* (AAF56066) and *Caenorhabditis elegans* (CAB54286), but interestingly no similar sequence in *S. cerevisiae* was identified.

The *creC* gene was cloned by complementation utilising its proximity to the *glnA* gene (Todd *et al.,* 2000). *creC* encodes a 630 aa protein that contains five WD40 repeats, motifs involved in protein-protein interactions, a proline rich region, and a putative nuclear localisation

signal (Todd *et al.,* 2000). Highly conserved proteins are found in *Schizosaccharomyces pombe* (Yde3), mouse (DMR-N9), humans (DMR-N9), *A. thaliana* (T2N18.8) and *C. elegans* (C08B6.7). DMR-N9 in mouse and humans is associated with the myotonic dystrophy region (Mahadevan *et al.*, 1993; Shaw *et al.*, 1993). Again it is significant that there is no closely conserved CreC sequence in *S. cerevisiae*, although the WD40 region does show weak similarity with Tup1p (see Section 1.2.2). However the RcoA protein from *A. nidulans* shows considerably higher similarity to Tup1p than does CreC, and deletion of *rcoA* does not affect carbon catabolite repression (Hicks *et al.*, 2001). RcoA is involved in asexual development and sterigmatocystin production and the different function of the Tup1 homologue in *A. nidulans* compared to that in *S. cerevisiae* implies that a different mechanism of carbon catabolite repression exists in *A. nidulans* (Hicks *et al.*, 2001). The presence of CreB and CreC homologues in mouse and humans, but not in *S. cerevisiae*, supports that the regulatory mechanism CreB and CreC are involved in is conserved among multicellular eukaryotes but is not found in *S. cerevisiae.*

CreB and CreC have been shown to be present in a complex *in vivo*, and it has been proposed that the CreB/CreC deubiquitination complex removes ubiquitin moieties from CreA and other substrates, thereby modifying or stabilising these proteins (Lockington & Kelly, 2002). That CreB and CreC proteins are present in a protein complex *in vivo* under both repressing and derepressing conditions was shown by co-immunoprecipitation experiments (Lockington & Kelly, 2002). Overexpression of the CreB deubiquitination enzyme can compensate for the lack of the CreC, but not *vice versa*, indicating that CreB acts downstream of CreC (Lockington & Kelly, 2002). The role of CreC may be to stabilize or activate the CreB deubiquitination enzyme, perhaps by CreC masking the PEST region of CreB to stabilize or alter substrate recognition. The presence of proteins in a range of organisms strongly conserved with the CreB and CreC proteins identified in *A. nidulans* implies that these proteins may be a component of a regulatory network that is present in most eukaryotes.

Protein stability is a key regulatory mechanism in the control of metabolism, the cell cycle, cell growth and development, and disease. Selective degradation or stabilization of intracellular proteins by ubiquitin-dependent pathways is essential for correct regulation of many cellular processes. Protein ubiquitination is highly conserved amongst eukaryotes and involves the addition of the 76 amino acid peptide ubiquitin (Ub) to protein substrates. It entails a cascade of reactions involving three enzyme complexes: E1 (ubiquitin-activating enzymes), E2 (ubiquitin-conjugating enzymes) and E3 (ubiquitin-protein ligases) which act sequentially. There are a number of conjugating enzymes and protein ligases in the cell, and various combinations confer substrate specificity on the system. The best known role of ubiquitination is the marking of

proteins for destruction via the 26S proteasome, but ubiquitination can also promote endocytosis, membrane transport and transcriptional regulation (reviewed in Muratani and Tansey, 2003). This process can be opposed by the action of deubiquitinating enzymes that remove ubiquitin from specific substrates, thus stabilizing them. Perhaps the best-studied example of a regulatory deubiquitinating enzyme in a multicellular eukaryote is the *D. melanogaster* Fat facets deubiquitinating enzyme, which is required for correct eye development (Wu *et al.*, 1999). Genetic analysis has revealed that the critical substrate of Fat facets in the eye is Liquid facets (Cadavid *et al.*, 2000). Another example of a regulatory deubiquitinating enzyme is murine *DUB-1*, which regulates cellular growth pathways by controlling the level of expression or ubiquitination state of protein regulators at the G1/S phase transition (Zhu *et al.*, 1996), although the substrates still remain unknown.

1.3.3 Signalling in carbon catabolite repression

Recently Flipphi *et al.* (2003) investigated the role of hexose phosphorylating enzymes in the signalling of carbon catabolite repression in *A. nidulans*. Transcriptional repression by glucose was fully retained in strains lacking either hexokinase (*hxkA1*) or glucokinase (*glkA4*), but the double hexose kinase mutants (*hxkA1/glkA4*) showed transcriptional derepression in the presence of glucose and a complete absence of glucose phosphorylating activity (Flipphi *et al.*, 2003). So unlike the Hxk2p hexokinase in *S. cerevisiae* (see Section 1.2.4), neither hexose kinase in *A. nidulans* exhibits a unique regulatory function in carbon catabolite repression (Flipphi *et al.*, 2003).

1.3.4 *creD* and *acrB*

The *creD34* mutation was identified as a suppressor of the *creC27* mutant phenotype on glucose medium containing fluoroacetamide, and was also found to suppress the *creB15* mutant phenotype on this medium (Kelly & Hynes, 1977). The *creD34* mutation suppressed other aspects of the *creB* and *creC* mutant phenotypes, such as the derepression of the *facA* and *alcA* genes as analysed through the fluoroacetate and allyl alcohol sensitivity found in *creB15* and *creC27* strains (Kelly & Hynes, 1977). In a wild-type background the *creD34* mutation is more resistant than wild-type on glucose and fluoroacetate medium, suggesting that the *creD34* mutation leads to tighter repression of enzymes subject to carbon catabolite repression. The suppression of the *creB* and *creC* mutant phenotypes by *creD34* implies that *creD* is involved in an opposing process to the deubiquitination role of the CreB/CreC complex, such as the ubiquitination of the same target proteins. The *creD34* mutation results in resistance to the

presence of acriflavine in complete medium, and sensitivity to the presence of molybdate, the reverse of the phenotype conferred by the *creA*, *creB* and *creC* mutations.

The *acrB2* mutation was identified as a spontaneous resistant sector that allowed growth on complete medium containing acriflavine in a genetic screen with the joint aims of understanding acriflavine toxicity and obtaining extra tools for genetic mapping (Roper & Kafer, 1957). Not only did mutations in *acrB* result in acriflavine resistance, but also greater resistance than wild-type to the presence of malachite green and crystal violet. Arst (1981) pointed out the conserved genetic linkage of *creB* to *acrB*, and creC to *creD*, and suggested the possibility that the regions arose as a result of duplication. The molecular nature of *creB* and *creC* makes this origin unlikely (Todd *et al.*, 2000; Lockington and Kelly, 2001). However, the fact that *creD34* suppresses phenotypes due to *creB* and *creC*, and that *acrB* shares some phenotypes with *creD*, indicates that *acrB* and *creD* may encode proteins in a process opposing deubiquitination.

1.3.5 Model of carbon catabolite repression

A model for the interaction of CreA, CreB and CreC in carbon catabolite repression has been proposed (Lockington & Kelly, 2002) (Figure 1.3). In the presence of glucose (i.e. under repressing conditions), it is known that the CreA protein binds the consensus sequence in the promoters of genes that encode enzymes/permeases that allow the utilization of alternative carbon sources to effect carbon catabolite repression. CreA is recognised as a substrate of the CreB deubiquitinating enzyme via the coiled-coil region of CreB, and free CreB acts to remove ubiquitin chains from CreA, thereby stabilizing the CreA protein and allowing repression to occur. In the absence of glucose (i.e. under derepressing conditions), free CreB does not recognise CreA and the PEST sites of CreB are exposed, leading to the degradation of CreB via a PEST-mediated pathway. CreA remains ubiquitinated and is either degraded via the 26S proteasome, or is modified and no longer binds to the promoter. In both the presence and absence of glucose, there is some CreB complexed with CreC, which has a role in the turnover of enzymes/permeases required for proline and quinate utilization as an example.

Carbon catabolite repression in *A. nidulans* involves a regulatory de/ubiquitination network, which is not present in *S. cerevisiae*, but is present in higher eukaryotes, highlighting the importance of using *A. nidulans* to further dissect this conserved regulatory ubiquitination network.

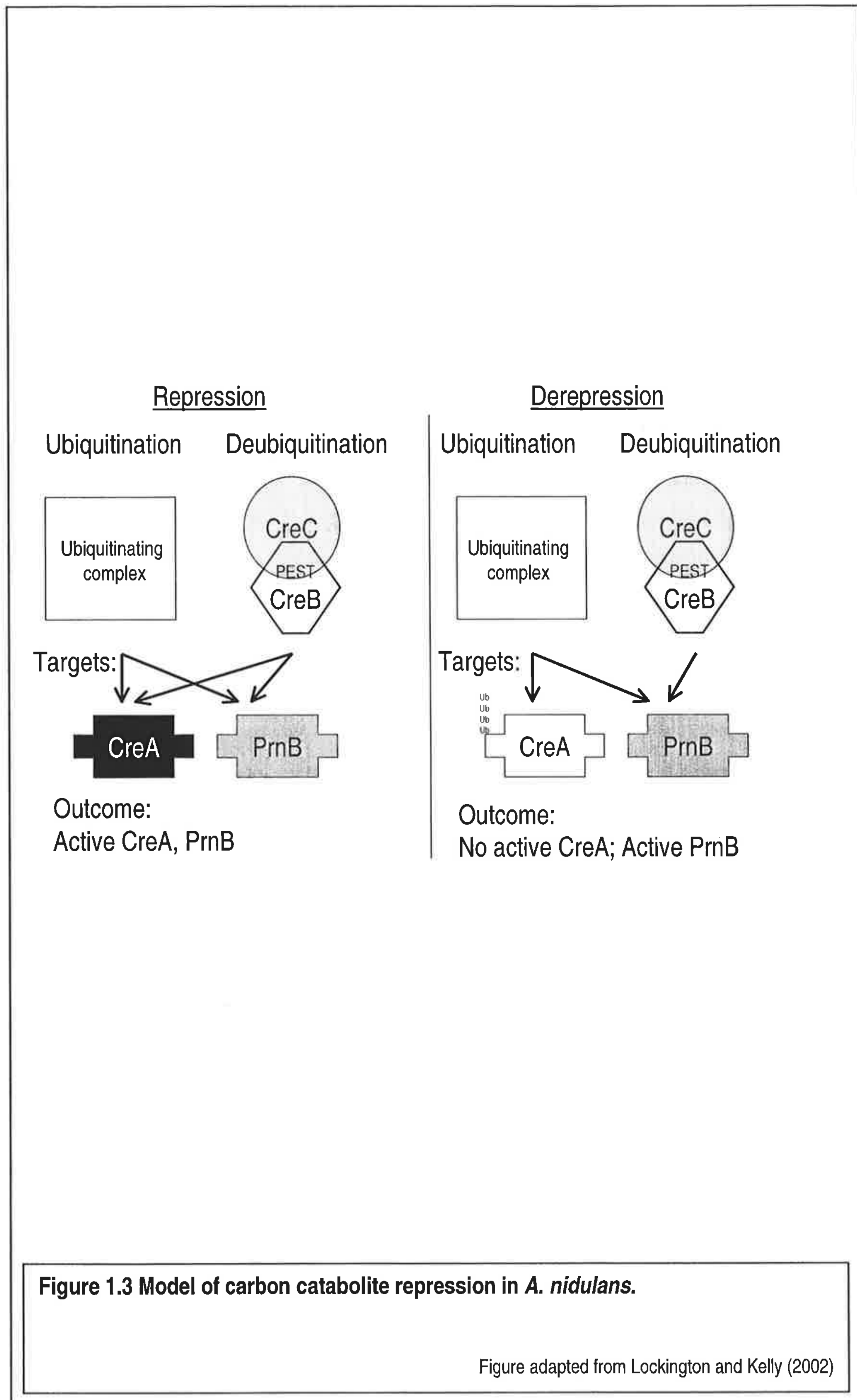

Figure 1.3 Model of carbon catabolite repression in *A. nidulans*.

Figure adapted from Lockington and Kelly (2002)

1.4 Amylases

Starch is a major storage polysaccharide in plants consisting of α-1,4-glucan chains with α–1,6 branching. Filamentous fungi produce α-amylases, glucoamylases and α-glucosidases that act in synergy to hydrolyse starch to glucose. The major group of starch hydrolysing enzymes are α–amylases (EC 3.2.1.1, 1,4-α-glucan glucanohydrolase), which are often secreted. α-amylases are the first group of starch-degrading enzymes to act on starch by catalysing the random internal endoamylolytic cleavage of α–1,4-glycosidic bonds, resulting in the production of glucose, maltose and other branched oligosaccharides from their substrates. Starch is subsequently degraded by glucoamylase and α-glucosidase. Glucoamylase (EC 3.2.1.3, glucan 1,4-α-glucosidase), and α-glucosidase (EC 3.2.1.20, also known as maltase) are both exoamylases that cleave the α–1,4-glycosidic bonds from non-reducing ends, and some glucoamylases are also capable of hydrolysing the branching α–1,6-glycosidic bonds when the next bond is an α–1,4 glycosidic bond (Figure 1.4). Starch serves as an alternative carbon source, so carbon catabolite repression would be expected to play an important role in the breakdown of starch, and therefore in the regulation of α–amylase activity and the other amylolytic enzymes that work cooperatively to reduce starch to glucose so that it can be utilised as a carbon source.

Alpha-amylases act not only on starch, but also on glycogen and related polysaccharides and oligosaccharides, to catalyse the random cleavage of internal α-1,4-glycosidic bonds. This results in the production of glucose, maltose, dextrins and other branched oligosaccharides from their substrates, which can then be further reduced. There are two major types of α–amylase activities, acid-unstable (neutral) and acid-stable α–amylases (Minoda and Yamada, 1963). Many of the physiochemical properties of these two types are similar, but the acid-stable α–amylase protein retains enzymatic activity after treatment with acid, and acid-unstable amylases do not. Structurally, the separate domains of the different amylase types are the same, but there is a rotation of the COOH-terminal domain in acid-unstable α–amylases relative to the acid-stable enzyme (Boel *et al.*, 1990). Acid-unstable α–amylases have been detected and characterised in a number of *Aspergillus* species, whereas characterisation of the acid-stable form is limited.

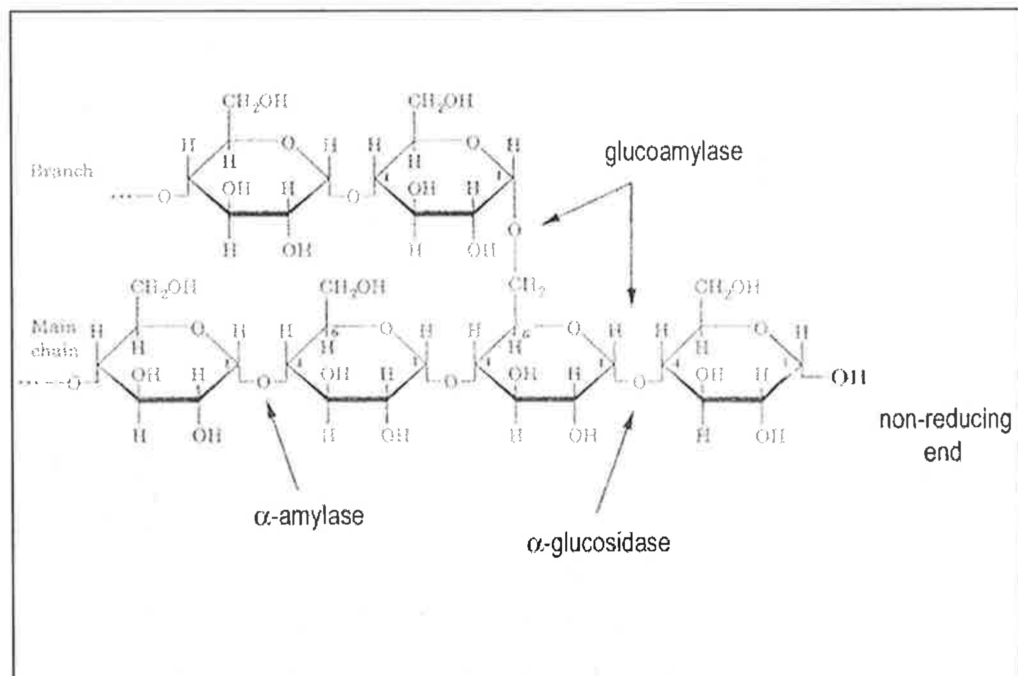

Figure 1.4 Enzymatic degradation of starch. The position of action of the starch-degrading enzymes α-amylase, glucoamylase and α-glucosidase. α-amylase cleaves internal α-1,4-glycosidic bonds; α-glucosidase cleaves external α-1,4-glycosidic bonds; glucoamylase cleaves external α-1,4-glycosidic bonds, and α-1,6-glycosidic bonds when adjacent to an α-1,4-glycosidic bond.

Figure adapted from Stryer (1998)

1.4.1 Acid-stable α-amylases

Acid-stable α-amylases were first detected after the purified crystalline form of α-amylase from *A. niger* was treated with acid (pH 2.5) at 37°C for 30 minutes and retained some amylolytic activity (Minoda and Yamada, 1963). The catalytic action of acid-stable and acid-unstable α-amylases is quite similar, but the dextrinizing unit per mg enzyme protein of the acid-unstable α-amylase was about six times greater than that of the acid-stable α-amylase. Thermal stability and pH stability are also quite different, with the acid-stable α-amylase being more stable at high temperatures and low pHs, retaining about 90% catalytic activity even at pH 2.2, while the acid-unstable enzyme is inactivated completely at pH 3.0 (Minoda *et al.*, 1968).

The acid-stable form of α-amylase from *A. kawachii* was characterised, and had an optimal pH range much lower than the acid-unstable α-amylase (pH ranges of 2.0-6.5 and 4.5-9.5 respectively), but the same optimal pH (pH 5.0). The acid-stable α-amylase also has a much higher molecular weight (approximately 85 kDa compared to 50 kDa for the acid-unstable) (Sudo *et al.,* 1993). The production of acid-stable α-amylase by *A. kawachii* seems to be solid-state culture-specific and is affected by the moisture content in solid medium, whereby increasing the moisture content of the solid medium decreases the production of acid-stable α-amylase (Nagamine *et al.*, 2003). Productivity and specific activity of acid-stable α-amylase is low, and does not appear to be induced by starch or malto-oligosaccharides, but commences when the concentration of intracellular content storage glycogen begins to decrease and an inducer such as dextrin is present. The decrease in content storage glycogen itself may not directly act on the synthesis of acid-stable α-amylases, but the carbon starvation conditions that are required for the decrease in glycogen may be important (Sudo *et al.*, 1993). Aside from the difference in stability at low pHs, acid-stable α-amylase activity can be distinguished from the acid-unstable α-amylase activity by cleaving the third glycosidic bond of malto-oligosaccharides instead of the second glycosidic bond cleaved by the acid-unstable α-amylases (Suganuma, 1997).

Woeldike (1989, SWISSPROT Accession number P56271) isolated the gene encoding acid-stable α-amylase from *A. niger*, and the deduced amino acid sequence from the cloned DNA showed that the acid-stable α-amylase of *A. niger* was distinct from the acid-unstable α-amylase. Amino acid identity between the two proteins was approximately 80% (Boel *et al.,* 1990), and there were extensive structural similarities between the separate domains of the acid-unstable and acid-stable amylase. However, a rotation of the COOH-terminal domain in acid-unstable α–amylase was observed relative to the acid-stable enzyme. An acid-stable α-amylase from *A. kawachii*, AsaA, shares 97% aa identity with the *A. niger* acid-stable α-amylase between

aa 1-479, and beyond this contains a C-terminal region rich in serine and threonine which may function as a raw starch binding site (Kaneko *et al.*, 1996).

1.4.2 Acid-unstable α-amylases

Cloning of the α–amylase (often referred to as *Taka–amylase A*) genes began in several strains of *A. oryzae* (Gines *et al.*, 1989; Tada *et al.*, 1989; Tsukagoshi *et al.*, 1989; Wirsel *et al.*, 1989), and were subsequently cloned in a similar fashion from *A. niger var. awamori* (Korman *et al.,* 1990) and other fungi to identify multiple copies of α–amylase genes. The amino acid sequence of an α–amylase gene from *A. oryzae* (Toda *et al.*, 1982) was used to synthesize DNA probes to hybridise to genomic copies of the α–amylase genes via Southern analysis. The three cloned α-amylase genes from *A. oryzae* all encode a 499 aa polypeptide, whereas the two α-amylase genes from *A. niger* encode a 499 and 498 aa protein respectively. The cloned sequences of the α–amylase genes contain nine exons and eight introns of almost identical sequence in identical positions. Some of the predicted polypeptide products show a few amino acid differences from the published amino acid sequence (Toda *et al.*, 1982) that are not expected to alter the functional structure of the protein [ranging from 8 aa differences (Gines *et al.*, 1989) to 12 aa differences (Tsukagoshi *et al.*, 1989)]. Therefore α–amylase seems to be encoded by a small multi-gene family, consisting of three genes in *A.oryzae*, *amy1*, *amy2* (*taaG2*) and *amy3*, and two genes in *A. niger*, *amyA* and *amyB*.

The 5' non-coding regions of *A. oryzae amy1*, *amy2* and *amy3* genes, and *A. niger amyA* and *amyB* genes are essentially the same, including a typical TATA box at -32 bp and a CCAAT element at -125 bp. Also, a putative GT rich motif was observed at positions –40 bp through to –27 bp. The 3' non-coding region of the three *A. oryzae* genomic DNAs have identical sequence to 69 bp downstream from the translation termination codon, after which they differ significantly from each other (Tsukagoshi *et al.*, 1989). Surprisingly, the α–amylase genes in *A. niger* show 98% DNA sequence identity to those of *A. oryzae* in their coding regions, as well as almost identical intervening sequences and 5' untranslated regions.

Like *A. oryzae* α–amylases, *A. niger amyA* encodes a mature peptide of 478 aa, while *amyB* encodes a 477 aa protein after the 21 aa signal peptide has been cleaved. All identified amylase genes contain this 21 aa hydrophobic secretory signal, and eight identical introns. *amyA* and *amyB* are identical in DNA sequence except for the last 9 bp of coding sequence and their 3' untranslated regions. The first 69 bp of the *amyB* 3' non-coding region is identical to that of *A. oryzae* 3' regions, and the point at which they all diverge is the same. The nucleotide sequence of the α–amylase gene of *A. shirousamii* is also almost identical to that of *A. oryzae,* including the 5'

non-coding region, with only 2 bp substitutions present in the coding regions. This surprising degree of sequence identity suggests that there has been much evolutionary pressure to maintain the α–amylase gene sequence and organisation, or it may be an example of a recent horizontal transfer of genes.

1.4.3 Regulation of α–amylases

The Taka-amylase A gene (*taaG2)* in *A. oryzae* has a high-level expression, and is induced by both starch and maltose, and repressed by an excess of glucose (Errat *et al.*, 1984). Regulation of the amylolytic genes has been best studied using *taaG2* in *A. oryzae* and when introduced into *A. nidulans,* to identify three regulators of amylolytic gene expression; the wide-domain regulators CreA and AnCP/AnCF, and the pathway specific activator AmyR, as detailed below.

Analysis of the 5' promoter region of α–amylase *taaG2* in *A. oryzae* revealed CreA consensus binding sites (5'-SYGGRG-3'), and potential CreA binding sites are found in promoters of other starch induced genes, such as α-glucosidase. Kato *et al.* (1996) demonstrated that CreA binds to the α-amylase promoter of *taaG2 in vitro*, indicating that the CreA protein is involved in the repression of the α–amylase genes in response to glucose.

Nuclear proteins that recognise the CCAAT element and the adjacent upstream sequence in the promoter region of the *A. oryzae* α–amylase gene were identified by Nagata *et al.* (1993) using *A. nidulans* as an intermediate host. An A. nidulans CCAAT element binding protein was identified (AnCP) that not only binds to the CCAAT element of α–amylase (*taaG2*) in *A. nidulans*, but also binds to promoters of other genes that are unrelated to amylase which contain this CCAAT element (Kato *et al.*, 1997). When this CCAAT sequence was mutated *in vivo,* AnCP binding was completely abolished, and an approximate four fold decrease of α–amylase expression resulted, suggesting a positive function for the CCAAT element. Even though the amylase activity was lowered when this CCAAT element was mutated, amylase activity was still significantly induced by starch and repressed by glucose, indicating that AnCP function is not necessary for the starch induction of α–amylase expression.

A number of CCAAT binding proteins have been independently identified in *A. nidulans* that modulate gene expression, including AnCP binding to the promoter of *taaG2*, AnCF (*A. nidulans* CCAAT factor) binding to the *amdS* promoter to affect acetamidase expression (van Heeswijck and Hynes, 1991), and PENR1, binding to *acvA*, *aatA*, and *ipnA* affecting the expression of the penicillin biosynthesis (Then Berg *et al.*, 1996; Litzka *et al.*, 1996). These proteins complexes are homologous to the *S. cerevisiae* Hap complex involved in transcriptional

activation, and are minimally composed of three subunits, HapB, HapC, and HapE (Papagiannopoulos *et al.*, 1996; Kato *et al.*, 1998; Steidl *et al.*, 1999). Recently, a gene encoding a novel transcriptional activator, *hapX*, was isolated from *A. nidulans* and shown to interact with the Hap complex (Tanaka *et al.*, 2002). *A. oryzae* also contains a protein complex, AoCP, comprised of HapB, HapC and HapE that binds specifically to a CCAAT element in the promoter of *taaG2* (Tanaka *et al.*, 2000, 2001). CCAAT elements are found in the promoters of many genes, including other amylolytic genes, such as the α-glucosidase genes of *A. nidulans* and *A. oryzae*, indicating that they may also be modulated by the AnCP/AnCF (Hap) complex.

amyR encodes a transcriptional activator of the starch-induced genes that was independently cloned from *A. oryzae* by two separate groups (Petersen *et al.*, 1999; Gomi *et al.*, 2000). The *amyR* disruptant strain showed restricted growth on starch medium, and produced low levels of amylolytic enzymes such as α-amylase and glucoamylase compared with a wild-type strain, indicating that AmyR is the positive regulator involved in starch/maltose induced expression of these genes (Gomi *et al.*, 2000). AmyR contains a $Zn(II)_2Cys_6$ DNA binding binuclear cluster in the N-terminal region (aa 28 – 54), similar to the Gal4p transcription factor from *S. cerevisiae*, and a leucine zipper-like heptad motif at aa 351 which can form coiled-coil structures, possibly acting as a homo-dimerization domain (Petersen *et al.*, 1999; Gomi *et al.*, 2000). Petersen *et al.* (1999) identified two different AmyR binding sites, both containing the triplet CGG that is known to interact with a Gal4p-like $Zn(II)_2Cys_6$ binuclear cluster (Reece and Ptashne, 1993). One site, (5'-CGG-N_8-CGG-3') is a conventional Gal4p-type binding site, whilst the second site, (5'-CGGAAATTTAAA-3') differs from most other Gal4p-type binding sites as it only has one CGG triplet. AmyR binding sites were identified in a number of starch-induced promoters from both *A. niger* and *A. oryzae* (Petersen *et al.*, 1999), and previously identified regions required for starch induction including Region IIIa (Minetoki *et al.*, 1996) and the starch responsive element (SRE) (Tani *et al.*, 2000) contain these AmyR binding sites.

Transcriptional analysis revealed that the *amyR* gene was expressed at the same level in the presence of glucose as in the presence of maltose, compared to the amylolytic genes that were only highly transcribed in the presence of maltose (Gomi *et al.*, 2000). Consistent with this Petersen *et al.* (1999) observed that AmyR had an effect on the α-amylase promoter under both inducing and non-inducing conditions, and hypothesised that the presence of an inducer, such as maltose, transforms AmyR into a better activator.

Gomi *et al.* (2000) identified an amylase cluster, containing *amyR*, *agdA* (encoding α-glucosidase) and *amyA* (encoding α-amylase) clustered within a 12 kb fragment in *A. oryzae*.

The *amyR* and *agdA* genes share a 1.5 kb upstream region, and are transcribed in opposite directions.

1.5 Aims and Objectives

In *A. nidulans* three genes involved in carbon catabolite repression have previously been characterized; *creA,* which encodes a transcriptional repressor protein, *creB* which encodes a deubiquitinating enzyme, and *creC* which encodes a protein that contains five WD40-repeat motifs and a proline-rich region. CreB and CreC form a complex *in vivo* and it has been proposed that this complex acts to stabilize the CreA repressor protein by removing ubiquitin moieties that target CreA for destruction via the 26S proteasome or alter its function. The *creD34* mutation was identified as a suppressor of the effects of the *creC27* mutation, and it was also shown to suppress the effects of the *creB15* mutation, suggesting a role for CreD in this regulatory network. The *acrB2* mutation shared some phenotypes with *creD34*, but others had not been tested.

The main objective of this project was to further elucidate the de/ubiquitination network regulating carbon catabolite repression. The first aim was to clone and characterize *creD*, and determine its role in carbon catabolite repression. The gene was cloned by complementation and sequenced. Bioinformatic analysis led to a proposed function, and a bacterial-2-hybrid screen was performed to identify proteins that interacted with CreD. The second aim was the characterization of the *acrB* gene. *acrB* was tested for the ability to suppress the effects of the *creB* and *creC* mutations, indicating a role in carbon catabolite repression. Sequence analysis and phenotypic analyses of three mutant *acrB* alleles were performed to define important regions of the gene.

An amylase cluster was identified and preliminary characterization begun by R. Murphy. Here the sequence analysis of the cluster was completed and bioinformatic analysis performed. It was shown that a previously identified amylase in *A. nidulans*, *amyB*, did not form part of this amylase cluster.

This work describes the cloning and analysis of *creD*, and the characterization of the *acrB* gene, two components of a regulatory network controlling carbon source utilization in the filamentous fungus *A. nidulans* that involves ubiquitination and deubiquitination.

CHAPTER 2: MATERIALS AND METHODS

2.1 Materials

General laboratory chemicals and growth media were of analytical research grade and were purchased from a range of commercial manufacturers.

2.1.1 Enzymes

Restriction enzymes (and corresponding 10 X buffers) were purchased from Boehringer-Mannheim, New England Biolabs (NEB) and Promega. The remaining enzymes were obtained from the following manufacturers:

Klenow fragment	Geneworks, Adelaide, Australia	#KF-1
T4 DNA ligase	Boehringer-Mannheim, Germany	#481220
Taq DNA Polymerase	Geneworks, Adelaide, Australia	#BTQ-1
PfuTurbo® DNA Polymerase	Stratagene, La Jolla, CA	#600250
Lysing enzymes	Sigma Chemical Co., St. Louis, MO	#L-1412

All enzymes were used following the manufacturers instructions, in the appropriate reaction buffers.

2.1.2 Radioactive isotopes and nucleic acids

[α-^{32}P] dATP (3000 Ci/mmol)	Radiochemical Centre, Amersham, UK	#AA0004
[α-^{32}P] dATP (3000 Ci/mmol)	PerkinElmer Life Sciences Inc., Boston, MA	#ADA-32
[α-^{32}P] dUTP (3000 Ci/mmol)	PerkinElmer Life Sciences Inc., Boston, MA	#ARU-32L
Mixed random decamers	Geneworks, Adelaide, Australia	
Ultrapure dNTP Set	Amersham Pharmacia Biotech Inc., Piscataway, NJ	#27-2035-02

2.1.3 Molecular weight markers

100 bp DNA ladder	New England Biolabs, Beverly, MA	#323-1
1 kb DNA ladder	New England Biolabs, Beverly, MA	#323-2
100 bp DNA ladder	Promega, Madison, WI	#G2101
λ DNA restricted with *HindIII*	Finnzymes, Finland	#F301S

2.1.4 Bacterial strains

The strains of *Escherichia coli* used in this study are presented are in Table 2.1.

Table 2.1 Strains of *E. coli* used in this study.

Strain	Genotype	Reference
NM522	*supE* Δ(*lac-proAB*) Δ(*mcrB-hsdSM*)5 F'[*proAB lacI^q lacZ* Δ*M15*]	Gough and Murray (1983)
DH5α	F$^-$ *mcrA* Δ*lacU169* (ϕ*80d lacZ* Δ*M15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Hanahan (1983)
B2H I reporter	Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 recA1 gyrA96 relA1 lac* [F' *lacI^q bla lacZ* Kan^r]	Stratagene #200180
B2H II reporter	Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 hisB supE44 thi-1 recA1 gyrA96 relA1 lac* [F' *lacI^q HIS3 aadA* Kan^r]	Stratagene #200192

2.1.5 *A. nidulans* strains

The genotypes of *A. nidulans* strains used in this study are presented in Table 2.2.

Table 2.2 *A. nidulans* strains used in this study.

Strain	Genotype	Origin
wild-type	*biAl;riboB2*	Derived from Pateman *et al.* (1967)
creA204	*biA1 creA204; niiA4*	Derived from Hynes and Kelly (1977)
creA204b	*yA1 creA204, riboB2*	Derived from Hynes and Kelly (1977)
creB1937	*yA1 pabA1;creB1937;riboB2*	Lockington and Kelly (2001)
creC27	*biA1;creC27*	Hynes and Kelly (1977)
creD34	*creD34; riboB2*	Derived from Kelly and Hynes (1977)
acrB2Y	*yA2;acrB2;choA1*	FGSC869
acrB2G	*biA1;acrB2;riboB2*	Derived from *acrB2Y*
acrB14	*pabaA1;acrB14;pantoB100*	Supplied by HN Arst
acrB15	*pabaA1;acrB15*	Supplied by HN Arst
creA204 creD34	*creA204; creD34; niiA4*	Derived from Kelly and Hynes (1977)
creB15 creD34	*creB15 creD34;riboB2 niiA4*	Derived from Kelly and Hynes (1977)
creC27 creD34	*creC27 creD34;riboB2 niiA4*	Derived from Kelly and Hynes (1977)
creA204 acrB2	*creA204; acrB2; niiA4*	This work
creB1937 acrB2	*yA$^-$ pabaA1;creB1937 acrB2*	This work
creC27 acrB2	*yA2;creC27 acrB2;choA1*	This work
creD34 acrB2	*creD34 acrB2;niiA4*	This work
creC27creD34/ creC27 creD$^+$		Diploid from *creC27* and *creC27creD34* strains (this work)

Table 2.2 Legend. Gene symbols are described in Clutterbuck (1993, 1997). FGSC = Fungal Genetics Stock Center. All strains are *veA1* mutants.

2.1.6 Oligonucleotides

Custom oligonucleotides used in this study (Table 2.3) were all purchased from Geneworks (Adelaide, Australia).

Table 2.3 Oligonucleotide primers used in this study

Primer name	Primer sequence (5' to 3')	Use	Location
APYF	TTTCTGCTCGTCTTGGTC	22Notrel1 PCR of APY region	Ch 3
APYR	AGGTGTGGCTGTGTTGTC	22Notrel1 PCR of APY region	Ch 3
PHF	GACAACACAGCCACACCT	22Notrel1 PCR of PH region	Ch 3
PHR	AAGCCTCAACCTCTCCAC	22Notrel1 PCR of PH region	Ch 3
MidF	GAAACGCCAGACACAGAC	22Notrel1 PCR of Mid region	Ch 3
MidR	TGGAGAAGTCGTGATTGG	22Notrel1 PCR of Mid region	Ch 3
STF	TACTCCGACAGCAACCAT	22Notrel1 PCR of ST region	Ch 3
STR	AGTTCGTTGGCAGTCTCA	22Notrel1 PCR of ST region	Ch 3
NovelF	GCCTGAATTACCAGTTGC	22Notrel1 PCR of novel region	Ch 3
NovelR	CGGAGTAGATTGCTGTGG	22Notrel1 PCR of novel region	Ch 3
DF1	GTGGAGGTGAGCAAGAAG	*creD* PCR/sequence	Ch 3
DF2	CAAGCTGATCCCTCTCCT	*creD* PCR/sequence	Ch 3
DF3	GGTATCGGCATCAGAACA	*creD* PCR/sequence	Ch 3
DF4	ACGTGCTCCTGAAACAGA	*creD* PCR/sequence	Ch 3
DF6	GCGGAGGCGGCAGTGCAA	*creD* PCR/sequence	Ch 3
DR1	CTATGACTTCGGGAATGG	*creD* PCR/sequence	Ch 3
DR2	GGTTAGCGCATCCATACA	*creD* PCR/sequence	Ch 3
DR3	ATACTTGCGGCCTATTTC	*creD* PCR/sequence	Ch 3
DR4	ATGAGCACCAACAGAAAC	*creD* PCR/sequence	Ch 3
DR5	GTACGCGAGAACTGAAAC	*creD* PCR/sequence	Ch 3
BThulAFRI	ACGAATTCACGACCTAATCTCAGT	*hulA* B2H PCR/sequence	Ch 3
BThulARBam	AAGGATCCTGACCGGAAGTAGATGAG	*hulA* B2H PCR/sequence	Ch 3
TRGcreDFBam	ATGGATCCCGTGGAGGTGAGCAAGAAG	*creD* B2H PCR/sequence	Ch 3
TRGcreDRXho	TATCTCGAGCTATGACTTCGGGAATGG	*creD* B2H PCR/sequence	Ch 3
TRGapyAFRI	ATGAATTCCTGGGAAGATCAAGATG	*apyA* B2H PCR/sequence	Ch 3
TRGapyARXho	AATCTCGAGTCAAGCCACTAATCCTG	*apyA* B2H PCR/sequence	Ch 3
hulASEQF	GAGCGGAGAGCGCACCAG	sequence pBThulA	Ch 3
hulASEQR	GCTGTTGCTGGATGGTGG	sequence pBThulA	Ch 3
apyACF	CAAACCGCAATCTGAAGC	PCR for cloning *apyA*	Ch 3
apyACR	TGCCTAAACTGCCTCAAG	PCR for cloning *apyA*	Ch 3
KBB	GCGAAGGGAAATGCAGAC	*acrB* PCR/ sequence	Ch 4
KB	ATGGGCTAAATCGAGAGC	*acrB* PCR/ sequence	Ch 4
AK	GCAATGGATGGTTTCTGC	*acrB* PCR/ sequence	Ch 4
SF	CTCGCCTCGCCAATAGAT	*acrB* PCR/ sequence	Ch 4
KF	CGCGTGGATACAGAAGCG	*acrB* PCR/ sequence	Ch 4
AS	CAATCTCTGCCTTGGCGG	*acrB* PCR/ sequence	Ch 4
SB	AAGAGGCTACCGCCGCTC	*acrB* PCR/ sequence	Ch 4
SBB	GCGAGCGGTGGCTTGGCA	*acrB* PCR/ sequence	Ch 4
KBF	GACACGCCACCATTCGTC	*acrB* PCR/ sequence	Ch 4
SBF	CTTGGGCACATTGGCGTC	*acrB* PCR/ sequence	Ch 4
acrB1	AGTCTTGGCTGCTATCCG	*acrB* PCR/ sequence	Ch 4
acrB2	TCCTTATGCGCGGCAGTG	*acrB* PCR/ sequence	Ch 4
acrB3	GGGCAACGGTGGTCATCA	*acrB* PCR/ sequence	Ch 4
acrB4	TCTTCCTAACCGTACCCG	*acrB* sequence	Ch 4

Primer Name	Primer sequence (5' to 3')	Use	Location
acrB5	AACCAGTCACTACACGGA	*acrB* PCR/ sequence	Ch 4
acrB6	GATCCAGCAGTACCGTTGTATGTT	nested 5' RACE	Ch 4
acrBT5aiRI	CGAATTCACCTCCGCAACCCTCA	*acrB* B2H PCR/sequence	Ch 4
acrBT3BAM2	GTGGATCCTAGTTGGCTGGCTGGC	*acrB* B2H PCR/sequence	Ch 4
acrTRG5aiBAM	CTGGATCCTCTCCGCCACCCTCAT	*acrB* B2H PCR/sequence	Ch 4
acrTRG3RI	TGGAATTCAGTTGGCTGGCTGGC	*acrB* B2H PCR/sequence	Ch 4
AMYO1	CAGCCAGCACTGTCGCC	amylase cluster sequence	Ch 5
AMYO2	TACGGCATTCCGATGTT	amylase cluster sequence	Ch 5
AMYO3	AAGATCATGGCCCGTCT	amylase cluster sequence	Ch 5
AMYO4	TCTTTCCACTCGCCTTCG	amylase cluster sequence	Ch 5
AMYO5	CACAGAACGCCCTAACCC	amylase cluster sequence	Ch 5
AMYO6	CGGCGGACTCGGCAACTG	amylase cluster sequence	Ch 5
AMYO7	CCGTCTGTCCTGAGTACA	amylase cluster sequence	Ch 5
AMYO8	TGACCTCTGCTACAGTCG	amylase cluster sequence	Ch 5
AMYO9	AGATCTTGGCCCGTCTGG	amylase cluster sequence	Ch 5
AMYO10	TCCTGGCGTCGTGGTTGG	amylase cluster sequence	Ch 5
AMYO11	GGCAGAATGAGACCTACC	amylase cluster sequence, *amyA* PCR for riboprobe	Ch 5
AMYO12	CATGGGATAGTTCAGCAC	amylase cluster sequence	Ch 5
AMYO13	GGTAGGTCTCATTCTGCC	amylase cluster sequence	Ch 5
AMYO14	CGCCCACCTGCGTGAAGC	amylase cluster sequence	Ch 5
AMYO15	CTAGAATCCTCGGTAGCG	amylase cluster sequence	Ch 5
AMYO16	GGGAAGAGGTACGAGTGG	amylase cluster sequence	Ch 5
AMYO17	ACCTCCGCTTACGCTTCC	amylase cluster sequence	Ch 5
AMYO18	CGAGCAGTTCTTCACCGG	amylase cluster sequence *amyR* PCR for riboprobe	Ch 5
AMYO19	GCCTCACCAGCCGCTCCG	amylase cluster sequence *amyR* PCR for riboprobe	Ch 5
AMYO20	CTTGGTACATCTATCGCC	amylase cluster sequence	Ch 5
AMYO21	CATCTACGGACAGCACGC	amylase cluster sequence	Ch 5
AMYO22	TGCCGAGTACTGCTGACG	amylase cluster sequence	Ch 5
AMYO23	GAAGAGTCTTAACGTCCC	amylase cluster sequence	Ch 5
AMY024	GGTACGACTACAGGCCGC	*amyA* PCR for riboprobe	Ch 5
AMY5	ATCCAGTACTAACGCCGC	*amyB* PCR for riboprobe	Ch 5
AMY8	CCGTATGCGGTCTGCACC	*amyB* PCR for riboprobe	Ch 5
pTRGforward	CAGCCTGAAGTGAAAGAA	sequence pTRG constructs	Chs 3,4
pTRGreverse	ATTCGTCGCCCGCCATAA	sequence pTRG constructs	Chs 3,4
FSP	GTAAAACGACGGCCAGT		
RSP	AACAGCTATGACCATG		
T7	AATACGACTCACTATAG		
T3	ATTAACCCTCACTAAAG		
SP6	GATTTAGGTGACACTATAG		

Table 2.3 Legend FSP (Forward Sequence Primer), RSP (Reverse Sequence Primer), T7, T7 and SP6 are all commercially designed primers, and were used for sequencing inserts in standard cloning vectors such as pBluescript and pGEMTeasy. pTRGforward and pTRGreverse are commercially designed primers for sequencing inserts of pTRG constructs in the BacterioMatch bacterial-2-hybrid system.

2.1.7 Vectors

Vectors used in this project are presented in Table 2.4

Table 2.4 Plasmids used in this study.

Plasmid	Description	Manufacturer
pPL3	*riboB*+ in pUC19, used as selectable marker for cotransforming *A. nidulans*	Oakley *et al.* (1987)
pBluescript II SK+	General cloning vector	Stratagene
pGEM-Teasy	Cloning of PCR products	Promega
pBT	Bait vector of BacterioMatch B2H system	Stratagene
pTRG	Target vector of BacterioMatch B2H system	Stratagene

2.1.8 *A. nidulans* DNA libraries

The *A. nidulans* gDNA Bacterial Artificial Chromosome (BAC) library used was made by Dr. Ralph Dean, Director, Center for Integrated Fungal Research, North Carolina State University (for reference see Zhu *et al.,* 1997) and kindly provided by Dr. A. Andrianopoulos, Department of Genetics, University of Melbourne. The *A. nidulans* chromosome II specific gDNA cosmid library is described in Brody *et al.* (1991).

The *A. nidulans* cDNA library (kindly provided by S. Osmani) was constructed from pooled RNA isolated from a range of developmental stages of mycelia grown in glucose conditions, and was used for determination of intron positions and as the basis for the construction of the bacterial-2-hybrid (B2H) cDNA library (R. Lockington, *unpubl. data*).

2.1.9 Kits and miscellaneous materials

BacterioMatch™ Two-hybrid System Vector Kit	Stratagene, La Jolla, CA	#240065
DNeasy® Plant Mini Kit	Qiagen, Valencia, CA	#69104
First Choice™ RLM-RACE Kit	Ambion Inc., Austin, TX	#1700
Hybond-N+ nylon membrane	Amersham, UK	#RPN303B
Nucleospin® RNA Plant Kit	Machery-Nagel, Germany	#740949
PSIΨClone BAC DNA kit	Princeton Separations, Adelphia, NJ	#PP120
QIAquick PCR Purification Kit	Qiagen, Valencia, CA	#28104
Riboprobe® Combination System – T3/T7 Kit	Promega, Madison,WI	#P1450

RPA III Ribonuclease Protection Assay Kit	Ambion Inc., Austin, TX	#1415
Wizard® *Plus* SV Minipreps DNA purification systems Kit	Promega, Madison, WI,	#A1460
X-ray film	Fuji RX Medical X-ray Film, Fuji Photo Film Co., Ltd., Tokyo, Japan	#03E220

2.2 Solutions, buffers, and media

2.2.1 Solutions and buffers

All solutions and buffers were prepared using millipore water and where appropriate, were autoclaved. Solutions not able to be autoclaved were sterilised by filtrations through a 0.2 μm filter. Solutions and all other buffers routinely used in this study were as follows:

Blotto	10% skim milk powder, 2 mg/ml sodium azide
100 X Denhardts solution	2% (w/v) Ficoll 400, 2% (w/v) polyvinylpyrrolidone, 2% (w/v) BSA
10 X Loading dye	0.25% bromophenol blue, 0.25% xylene cyanol FF, 25% Ficoll 400
10 X MOPS buffer	200 mM MOPS, 50 mM anhydrous NaAc, 10 mM EDTA, adjusted to pH 7.0
10 x Oligolabelling buffer	0.5 M Tris-HCl pH 6.9, 0.1 M $MgSO_4$, 1 mM DTT, and 0.6 mM each of dCTP, dGTP, and dTTP
Phenol/chloroform	50% (v/v) phenol, 48% (v/v) chloroform, 2% (v/v) isoamyl alcohol, buffered with an equal volume of Tris-HCl (pH 8.0), 0.2% (v/v) β-mercaptoethanol
Prehybridisation (formamide) buffer	40% formamide (v/v), 4 x SSPE, 1% SDS, 5% blotto, 100 μg sonicated salmon sperm DNA
Prehybridisation (phosphate) buffer	10 mM sodium orthophosphate, 5 X SSC, 2% blotto, 0.4% SDS, 100 μg sonicated salmon sperm DNA
RNA loading buffer (MOPS)	50% (v/v) formamide, 37% (v/v) formaldehyde, 1 x MOPS buffer, 1 x loading dye, 10 mg/ml ethidium bromide
RNA loading buffer (phosphate)	50% (v/v) formamide, 12% (v/v) formaldehyde, 10 mM sodium orthophosphate, 1 x loading dye, 10 mg/ml ethidium bromide
1 x SSC	0.15 M NaCl, 0.015 M $Na_3C_6H_5O_7.2H_2O$, pH 7.2
1 x SSPE	0.18 M NaCl, 10 mM NaH_2PO_4, 1 mM EDTA, pH 7.4
1 x TAE	40 mM Tris base, 20 mM NaAc, 2 mM EDTA, pH 7.8 with glacial acetic acid
1 X TE	10 mM Tris-HCl (pH 8.0), 1 mM EDTA

2.2.2 Media

Aspergillus complete and minimal media were as described by Cove (1966). Carbon sources were added at a final concentration of 1% (w/v) unless stated otherwise. Nitrogen sources were added to a final concentration of 10mM unless otherwise stated.

Standard bacterial media (L-broth, Luria Bertani Agar etc.) were prepared as described in Sambrook *et al.* (1989). Where required, ampicillin or chloramphenicol were added to a final concentration of 50 μg/ml or 170 μg/ml respectively, unless otherwise stated.

2.3 Methods

Standard molecular techniques were performed as outlined in Sambrook *et al.* (1989).

2.3.1 Manipulation of *A. nidulans*

Growth testing and meiotic analyses were performed using the methods described by Cove (1966). Genetic manipulations were carried out using techniques described by Clutterbuck (1974). Transformation experiments were carried out using the method of Tilburn *et al.* (1983), typically adding 100 mg Lysing Enzymes (Sigma Chemical Co.) for protoplasting. Transformants from co-transformation experiments were selected using the *riboB*$^{+}$ selectable marker plasmid pPL3 on media lacking riboflavin. The genotypes of all double mutant strains constructed during this project were verified by molecular analyses.

2.3.2 Nucleic acid isolation

A. nidulans genomic DNA was isolated using the DNeasy® Plant Mini Kit according to the manufacturers instructions. Prior to the purchase of this kit, gDNA was isolated from wet mycelium by the method of Lee and Taylor (1990).

BAC DNA was isolated from *E. coli* using the PSIΨClone BAC DNA kit according to the manufacturers instructions.

Plasmids were isolated from *E. coli* using the Wizard® *Plus* SV Minipreps DNA purification systems Kit according to the manufacturers instructions.

Cosmids and fosmids were isolated from *E. coli* using the Wizard® *Plus* SV Minipreps DNA purification systems Kit according to the manufacturers instructions, except for the final elution step which was performed with 50 μl of nuclease-free water pre-warmed to 65°C and incubated at room temperature for 5 minutes before the final centrifugation.

Total RNA was isolated from *A. nidulans* mycelia grown in specified conditions, using the Nucleospin® RNA Plant Kit according to the manufacturers instructions.

2.3.3 Polymerase Chain Reaction (PCR)

Optimal conditions for amplification of DNA fragments via the Polymerase Chain Reaction (PCR) were determined for each set of template and primers. Generally, these reactions were as outlined in Innis and Gelfand (1990), with 100 ng of double-stranded DNA and 100 ng of each specific oligonucleotide primer in a final reaction volume of 50 μl. DNA was typically amplified in a programmable PTC-200 DNA Engine (MJ Research, Inc.) as follows: an initial denaturation at 95°C for two minutes, followed by 25–30 cycles using these conditions: denaturation at 95°C for one minute, a 30 second annealing step and extension at 72°C for one minute. The annealing temperature was modified to suit specific template/primers sets and ranged between 48°C and 60°C. The extension time also varied depending on the size of the fragment to be amplified, allowing a 1 kb/min synthesis rate.

2.3.4 DNA sequencing

PCR products were purified using the QIAquick PCR Purification Kit prior to sequencing, and plasmid DNA required no further purification after isolation as describe above. Sequencing reactions were performed using a Prism Big Dye Terminator Kit (Applied Biosystems-Perkin Elmer) as outlined in the manufacturers instructions. Dye terminator reactions were electrophoresed on an ABI Prism Model 377 automated sequencing machine at the IMVS sequencing facility (Institute of Medical and Veterinary Science, Frome Rd, Adelaide SA 5005).

2.3.5 Rapid amplification of 5' cDNA ends (5'-RACE)

5'-RACE was performed on total RNA from D-glucose-grown mycelium using the First Choice™ RLM-RACE Kit (Ambion), following the manufacturers instructions.

2.3.6 DNA gel electrophoresis and Southern transfer

DNA for Southern blot analysis was separated by gel electrophoresis through 1% agarose in 1 x TAE buffer, depurinated by treatment with 0.25 M HCl for 10 minutes and transferred to Hybond N+ membrane (Amersham) by alkaline transfer in 0.4 M NaOH for four hours or 20 x SSC overnight, as recommended by the manufacturer. Following overnight transfer in 20 X SSC, filters were treated with 0.4 M NaOH for 20 minutes. DNA filters were prehybridised for at least 90 minutes at 42°C in prehybridisation (formamide) buffer. Denatured labelled DNA probe was added directly to the prehybridisation mix and incubated at 42°C overnight. Filters were washed with 2 x SSC, 0.1% SDS for at least 15 minutes at 42°C, and then 0.5 x SSC, 0.1 % SDS for 20 minutes at 42°C and then exposed to X-ray film or to a phospho-imager plate.

Sometimes an additional 10 minute high stringency wash with 0.1 x SSC, 0.1% SDS was required. Membranes were stripped of probe by placing the membrane in a 400 ml solution of boiling hot 0.5% SDS for 10 minutes, then rinsed in 2 X SSC and stored in plastic wrap at 4°C.

2.3.7 RNA gel electrophoresis and Northern transfer

Two methods of RNA gel electrophoresis were used during the course of this project. Using the "MOPS" method, 1 μg - 5 μg total RNA was electrophoresed through 1.2% agarose - 0.6 M formaldehyde gels in 1 x MOPS running buffer. RNA samples were added to RNA loading buffer (MOPS) and heated to 68°C for 15 minutes prior to loading. Alternatively, using the "phosphate" method 1 μg - 5 μg total RNA was electrophoresed in 1.5% agarose - 0.6M formaldehyde - 10 mM sodium orthophosphate gels in 10 mM sodium orthophosphate running buffer. RNA samples were added to RNA loading buffer (phosphate) and heated to 68°C for 15 minutes prior to loading.

RNA was transferred to Hybond N+ (Amersham) by alkaline transfer in 0.04 M NaOH for 2-3 hours. Denatured radioactively labelled probe (DNA or RNA) was added directly to the prehybridisation mix and incubated overnight. Filters were washed with 2 x SSC, 0.1% SDS for at least 15 minutes, and then 0.5 x SSC, 0.1 % SDS for 20 minutes and then autoradiographed. Sometimes an additional 10 minute wash with 0.1 x SSC, 0.1% SDS was required. Two hybridisation conditions were used on the Northern filters. Hybridisations and washes were all performed at 42°C when prehybridisation (formamide) buffer was used, and at 65°C when prehybridisation (phosphate) buffer was used.

2.3.8 Ribonuclease Protection Assays

Ribonuclease Protection Assays were typically performed on 5 μg total RNA using the RPA III Ribonuclease Protection Assay Kit, according to the manufacturers instructions.

2.3.9 Radioactive labelling of probes

For DNA probes, 50-200 ng DNA was radioactively labelled with [α-^{32}P] dATP by primer extension of random decamer oligonucleotides using the method of Hodgson and Fisk (1987). Radiolabelling by PCR was also performed on PCR templates as per the standard PCR protocol, except that the dNTPs are replaced by a mixture of dCTP, dGTP and dTTP, and [α-^{32}P] dATP, and typically the cycle number was reduced to 8 cycles.

For RNA probes, 200–1000 ng RNA was radioactively labelled with [α-^{32}P] dUTP using the Riboprobe® Combination System –T3/T7 Kit according to the manufacturers instructions.

2.3.10 Bacterial-2-hybrid system

Protein-protein interactions were investigated utilising Stratagene's BacterioMatch® I Two-Hybrid System Vector Kit (B2H). B2H I reporter cells supplied with the kit were made competent by the $CaCl_2$ method, and transformed with relevant constructs and incubated overnight at 30°C (Sambrook *et al.,* 1989). Interactions were determined to be positive as measured by growth on LB-agar plates supplemented with 750 μg/ml carbenicillin, 15 μg/ml tetracycline, 34 μg/ml chloramphenicol, and 50 μg/ml kanamycin, and validated by a blue colour reaction on X-gal indicator plates grown overnight at 37°C (LB-agar plates supplemented with 40μg/ml X-gal [5-bromo-4-chloro-3-indolyl-β-D-galactoside] and 20μg/ml IPTG [isopropyl-β-D-thiogalactopyranoside] as a competitive inhibitor).

The BacterioMatch® II Validation reporter cells became available in late 2003 and are compatible with constructs made for use with the B2H I reporter cells. The B2H II reporter cells were made competent by the $CaCl_2$ method, and transformed with relevant constructs and incubated overnight at 37°C (Sambrook *et al.,* 1989). Interactions were determined to be positive as measured by growth on "Selective Screening Medium" consisting of minimal media plus 5 mM 3-amino-1,2,4-triazole (3-AT), 25 μg/ml chloramphenicol and 12.5 μg/ml tetracyline, and validated by growth on "Dual Selective Screening Medium" consisting of minimal medium plus 5 mM 3-AT and 12.5 μg/ml streptomycin, 25 μg/ml chloramphenicol and 12.5 μg/ml tetracyline, with all media prepared as outlined by the manufacturer. The pBTLGF2 and pTRGGAL11^{P} constructs provided with the B2H kit served as a positive control.

2.3.11 Amylase secretion plate tests

Culture media (100 ml) were inoculated with conidia (from either wild-type or *creA204b* mutant strains) and grown in conditions as described below:

16 hrs at 37°C in medium containing either 1% D-glucose (v/v), 0.1% D-fructose (v/v), 1% starch (v/v), 1% maltose (v/v), 1% D-glucose (v/v) and 1% starch (v/v), or 1% D-glucose and (v/v) 1% maltose (v/v)

16 hrs at 37°C in medium containing 0.1% D-fructose (v/v), then either addition of 1% starch (v/v) or 1% maltose (v/v) and reincubation for 4 hrs at 37°C

16 hrs at 37°C in medium containing 1% glucose (v/v), then addition of 1% starch (v/v) and reincubation for 4 hrs at 37°C

Prior to harvesting the mycelia, 20 μl of the liquid culture medium (containing any secreted α-amylase) was plated onto 0.5% starch plates, and incubated at 37°C for 1 hr, 4 hrs or 16 hrs. Then the starch plates were flooded with iodine, which stains starch blue, and the amount

of α-amylase activity was deduced from the amount of clearing seen on the starch plate, where the starch has been broken down by α-amylase.

The harvested mycelia from these overnight cultures were used to isolate RNA using methods described above.

2.3.12 Sequence analyses

DNA sequences obtained from automated sequencing reactions were visualised using the Chromas V2.23 (Technelysium) programme. Sequence analyses were performed using programmes accessed via the Australian National Genomic Information Service (ANGIS, http://www1.angis.org.au/pbin/WebANGIS/wrapper.pl) facility or NEBcutter Version 2 (http://tools.neb.com/NEBcutter2/index.php). Database searches were performed via the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov). The *A. nidulans* genome sequence was accessed via the Whitehead Institute's Center for Genome Research website (http://www-genome.wi.mit.edu/annotation/fungi/aspergillus/index.html).

2.3.13 Phylogenetic analyses

Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 2.1 (Kumar *et al.*, 2001). Sequences used for analyses were identified as proteins containing arrestin domains via a BlastX (Altschul *et al.*, 1990) search using *creD*, initially against the NCBI database (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi), and then against individual organism genome sequence databases with results as follows: *A. fumigatis* AfApy1 (AFGDP80TR.69) obtained from The Institute for Genomic Research (TIGR) *A. fumigatis* genome database (http://www.tigr.org/tdb/e2k1/afu1/), *N. crassa* NcrApy1 (EAA28290.1) obtained from the *N. crassa* genome database (http://www-genome.wi.mit.edu/annotations/fungi/neurospora/), *S. pombe* SpomApy1 (SPCC584.15C) and SpomApy2 (SPAC31A2.12) obtained from the *S. pombe* genome database (http://www.genedb.org/fenedb/pombe/index.jsp), and *S. cerevisiae* ScRod1p (YORO18W) and ScRog3p (YFR022W) obtained from the *S. cerevisiae* genome database (http://yeastgenome.org/). The alignments of amino acid sequences were generated using the ClustalW programme (Thompson *et al.*, 1994) accessed via ANGIS. Four regions of high sequence conservation within the arrestin domains were used for phylogenetic analysis, corresponding to aa 41 to 68, 107 to 162, 181 to 226 and 282 to 331 of CreD. Phylogenetic trees were constructed using the maximum parsimony method (Fitch, 1971) with the close-neighbor interchange search (Nei and Kumar, 2000) or using the neighbor-joining method (Saitou and Nei, 1987), and bootstrap analysis was performed with 1000 replicates.

2.4 Nucleotide sequence accession numbers.

The sequence of the *creD* gene has been deposited in GenBank under Accession no. AY458430. The sequence of the *apyA* cDNA has been deposited in GenBank under Accession no. AY460113. The sequence of the *acrB* gene has been deposited in GenBank under Accession no. AF485329. The sequence of the amylase cluster, including *amyR*, *agdA* and *amyA*, has been deposited in GenBank under Accession no. AF208225.

CHAPTER 3: CLONING AND MOLECULAR ANALYSIS OF *creD*

3.1 Introduction

In *A. nidulans creB* encodes a deubiquitinating enzyme that forms a complex that includes the WD40-motif containing protein encoded by *creC*, and mutations in these genes lead to altered carbon source utilization. The *creD34* mutation suppresses the phenotypic effects of mutations in *creC* and *creB*, and thus *creD* is likely to be involved in the ubiquitination aspect of the regulatory network that involves the CreB/CreC deubiquitination complex. To further dissect this network the *creD* gene was cloned and characterized.

Strains containing the *creD34* mutation are resistant to the presence of acriflavine in complete media in comparison to wild-type strains, which are sensitive. The genetic distance between *creD* and *creC* has been previously mapped at between 3-6 cM (Kelly & Hynes, 1977), and *creC* has been located within the chromosome II specific cosmid library (Brody *et al.*, 1991) to cosmid P24E11 (Todd *et al.*, 2000). Therefore, the initial approach taken to clone *creD* was to walk in the cosmid library, and test adjacent cosmids for complementation of the *creD34* phenotype on acriflavine.

According to the published chromosome II cosmid library map, cosmid L21B08 overlaps the *creC* containing cosmid P24E11. In a preliminary study L21B08 apparently complemented the *creD34* mutant phenotype on acriflavine (R. Todd, *pers. comm.*), and this closeness to *creC* is in keeping with the expected distance of about 30 kb between *creC* and *creD* according to previous comparisons of physical and genetic maps. Thus the cosmid L21B08 provided the starting point for the identification of *creD*.

3.2 Cloning *creD* via complementation of the *creD34* phenotype on acriflavine using cosmids and cosmid derivatives

3.2.1 The *creD34* allele is recessive to the wild-type allele

In order for the strategy of cloning *creD* by complementation to be successful it was necessary to confirm that the *creD34* allele was recessive to the *creD*$^+$ allele for the phenotype observed on medium containing acriflavine. The *creC27* mutant strain is sensitive to the presence of acriflavine in complete medium, whilst the *creC27creD34* double mutant strain is resistant to the presence of acriflavine in complete medium (Boase *et al.*, 2003). The *creD34* phenotype has previously been shown to be recessive to the wild-type allele on a range of media (Kelly and

Hynes, 1977), but the dominance properties on acriflavine containing medium was not tested. In order to confirm that the *creD34* mutation was recessive to the wild-type allele for its phenotype on acriflavine a *creC27creD34/creC27creD*$^+$ diploid was constructed and the phenotype was tested (Figure 3.1). The diploid was as sensitive as *creC27* on acriflavine containing medium. Thus the affect of the *creD34* mutation is recessive to the wild-type allele on medium containing acriflavine, allowing a complementation approach to clone *creD*.

3.2.2 Testing cosmids L21B08, P24E11, and relevant derivatives for complementation of the *creD34* mutant phenotype

Cosmid L21B08 was retested and partial complementation of the *creD34* phenotype on acriflavine was confirmed. To narrow down the complementing region, L21B08 was digested with the restriction enzyme *SalI,* and religated to excise an approximate 10 kb *SalI* fragment. This subclone, L2-Sal, containing approximately 22 kb of genomic DNA insert (Figure 3.2), was used to co-transform a *creD34; riboB2* strain together with the *riboB*$^+$ containing plasmid pPL3. L2-Sal complemented the *creD34* mutant phenotype on medium containing acriflavine, but the sensitivity to acriflavine was not restored completely to the level of wild-type (Figure 3.3).

Preliminary Southern analysis indicated that the cosmids L21B08 and P24E11 overlapped each other (R. Todd, *pers. comm.*), and the order of known genes *glnA* and *creC* in P24E11 allowed orientation of the cosmid (Figure 3.2), such that a previously constructed plasmid, p8A (R Sladic, *unpubl. data*), containing a 12 kb genomic DNA fragment (Figure 3.2) was then used to test for complementation of the *creD34* phenotype. p8A partially complemented the *creD34* phenotype on acriflavine, and so previously constructed derivatives of this plasmid, pX1 (10 kb gDNA insert) and pB3 (5 kb gDNA insert), were tested and also shown to partially complement the *creD34* mutant phenotype on acriflavine containing medium (Figures 3.2 and 3.3). The pB3 complementing region was refined further by digesting pB3 with *EcoRI* and the five resulting *EcoRI* fragments (designated a,b,c,d and e) of pB3 were subcloned into the pBluescript II SK+ vector individually and in various combinations (Figure 3.2). Additional subclones of the pB3 plasmid were constructed by subcloning a *BglII* - *ClaI* fragment (pBb'ecd) and an *EcoRI* – *EcoRV* fragment (pBabdc') (Figure 3.2). All plasmids lacking the 'pBd' *EcoRI* fragment of pB3 did not show complementation of the *creD34* mutant phenotype on acriflavine, whilst all plasmids containing the 'pBd' *EcoRI* fragment of pB3 showed complementation (Figures 3.2 and 3.3).

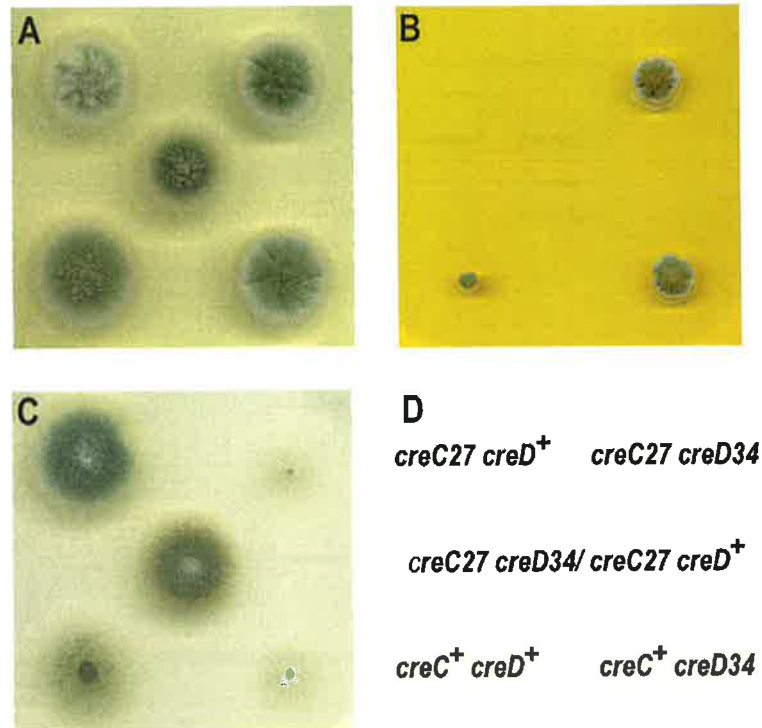

Figure 3.1 The *creD34* allele is recessive to the wild-type allele. All strains were grown at 37°C on either (**A**) 1% complete medium for 2 days (**B**) 1% complete medium plus 0.001% acriflavine for 2 days or (**C**) 1% glucose plus 10 mM pyrrolidinone medium for 3 days. The relevant genotypes are indicated by the key (**D**). For full genotypes see Materials and Methods, Table 2.2.

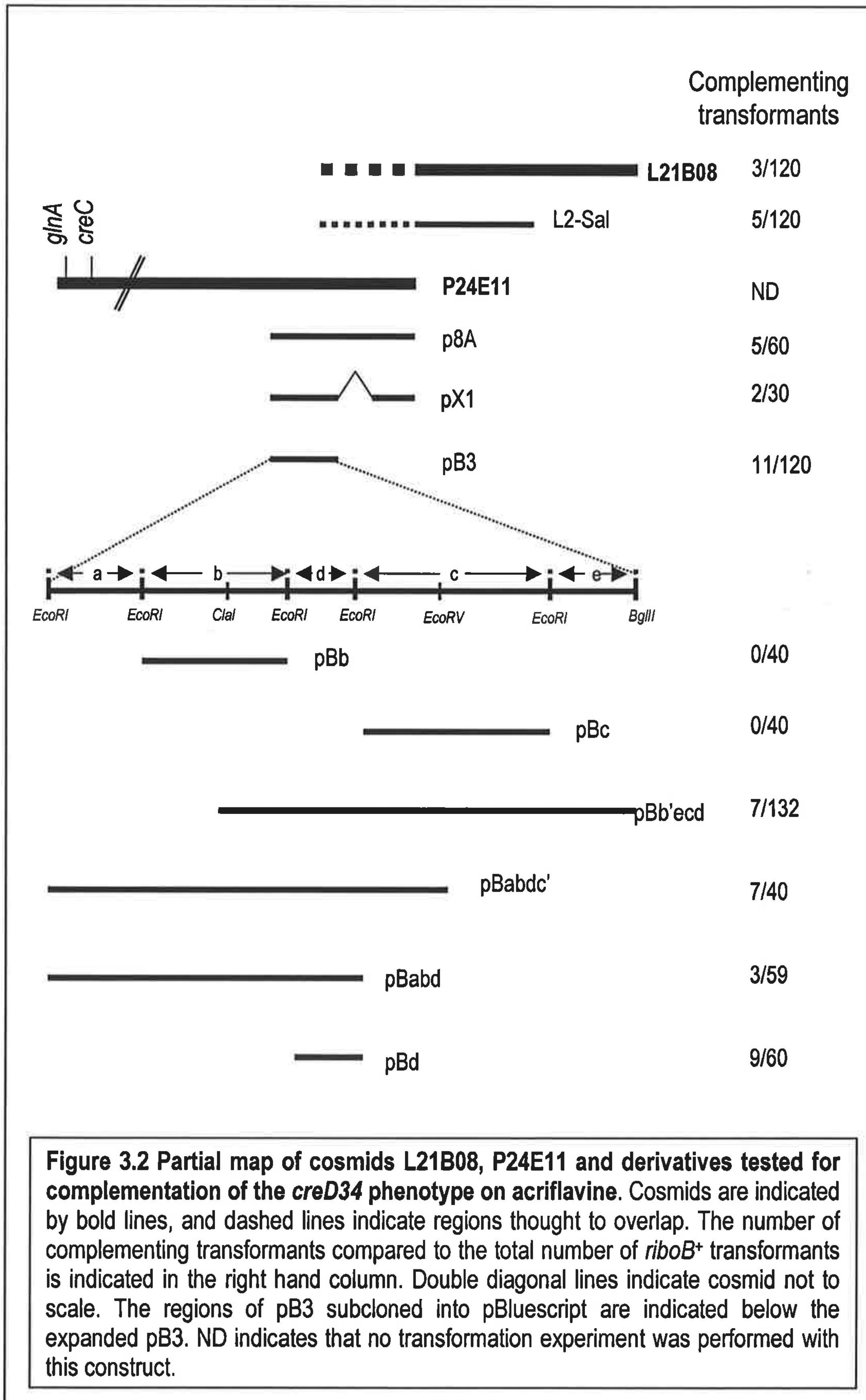

Figure 3.2 Partial map of cosmids L21B08, P24E11 and derivatives tested for complementation of the *creD34* phenotype on acriflavine. Cosmids are indicated by bold lines, and dashed lines indicate regions thought to overlap. The number of complementing transformants compared to the total number of *riboB*$^{+}$ transformants is indicated in the right hand column. Double diagonal lines indicate cosmid not to scale. The regions of pB3 subcloned into pBluescript are indicated below the expanded pB3. ND indicates that no transformation experiment was performed with this construct.

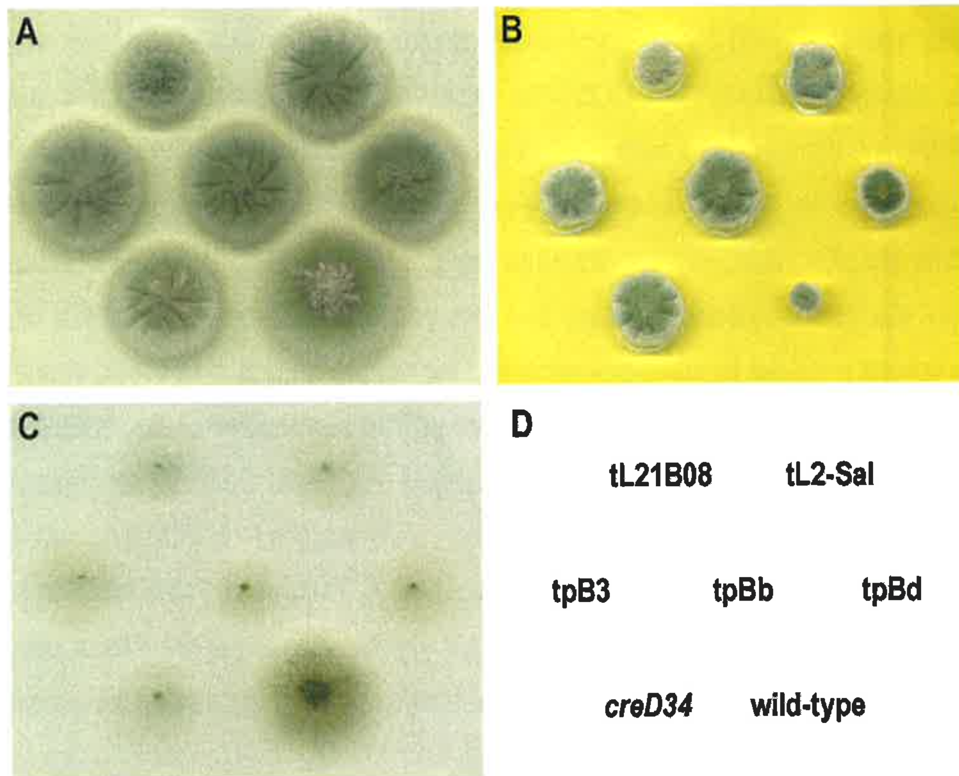

Figure 3.3 Partial complementation of the *creD34* mutant phenotype. All strains were grown at 37°C on either (**A**) 1% complete medium for 2 days (**B**) 1% complete medium plus 0.001% acriflavine for 2 days or (**C**) 1% glucose plus 10 mM pyrrolidinone medium for 3 days. The strains are indicated by the key (**D**). For full genotypes see Materials and Methods, Table 2.2. Transformants of the *creD34* strain are designated tX, where X represents the transforming DNA. Each transformant shown is representative of a number of similar transformants generated in the same transformation experiment.

3.2.3 Sequence analysis of pB3

The five *EcoRI* fragments of pB3 were sequenced using standard forward and reverse sequence primers and specific oligonucleotides were designed to generate the remaining sequencing after assembly of these five regions. This analysis revealed the presence of two hypothetical genes within the 5 kb plasmid pB3 (Figure 3.4, Appendix A), largely contained within either the 2349 bp 'pBb' or 1651 bp 'pBc' subcloned *EcoRI* fragments.

The pBb region contains a hypothetical gene that encodes a protein homologous to the hypothetical protein Gon1p from *S. cerevisiae*, as determined by BlastX analysis, and shares 49.9% aa identity with Gon1p. The *GON1*-like gene encodes a putative protein that contains an ATP/GTP binding domain and a calcium-binding domain, as determined by protein motif analysis in GCG (accessed via ANGIS), and only a short 33 bp 5' untranslated region for this protein is found within the pBb clone, with the promoter region mainly residing in the pBd region. BlastX analysis of the pBc sequence revealed that this subclone contains the majority of a gene encoding a protein similar to the hypothetical protein Ysc84p from *S. cerevisiae*, sharing 52.3% amino acid identity, and containing an SH3 domain. The start codon of the *YSC84*-like gene resides 13 bp beyond the pBc region, in the pBd fragment. Prior to the results of the complementation analysis, the pBb and pBc regions were amplified via PCR from wild-type and *creD34* mutant derived gDNA and sequenced (Appendix A, Table A.1). There were no sequence differences found between the wild-type and *creD34* alleles within either the pBb or pBc regions, which is consistent with the subsequent lack of complementation of the *creD34* mutant phenotype seen for these subclones.

The pBd intergenic region between the *GON1*-like and *YSC84*-like genes partially complemented the *creD34* mutant phenotype on acriflavine medium, suggesting that the *creD34* mutant phenotype may be the result of a promoter mutation for either or both of these genes. Thus the 'd' *EcoRI* fragment of pB3 was sequenced from the *creD34* mutant strain in order to identify the molecular nature of the *creD34* mutation. The pBd region was amplified via PCR from gDNA of both a wild-type and *creD34* mutant strain and sequenced (Appendix A, Table A.1). No mutation was present within pBd, hence neither the *GON1*-like or *YSC84*-like gene encode CreD. Thus it is likely that pBd contains a sequence that acts as a suppressor of the phenotypic effects of the *creD34* mutation on acriflavine, rather than *creD* itself.

The pBd region is 701 bp in length and contains several small open reading frames (Figure 3.4). In multiple copies this region can suppress the *creD34* mutant phenotype on medium containing acriflavine, perhaps due to the titration of regulatory proteins. The four open reading frames of pBd larger than 20 amino acids in length were analysed by BlastX against the

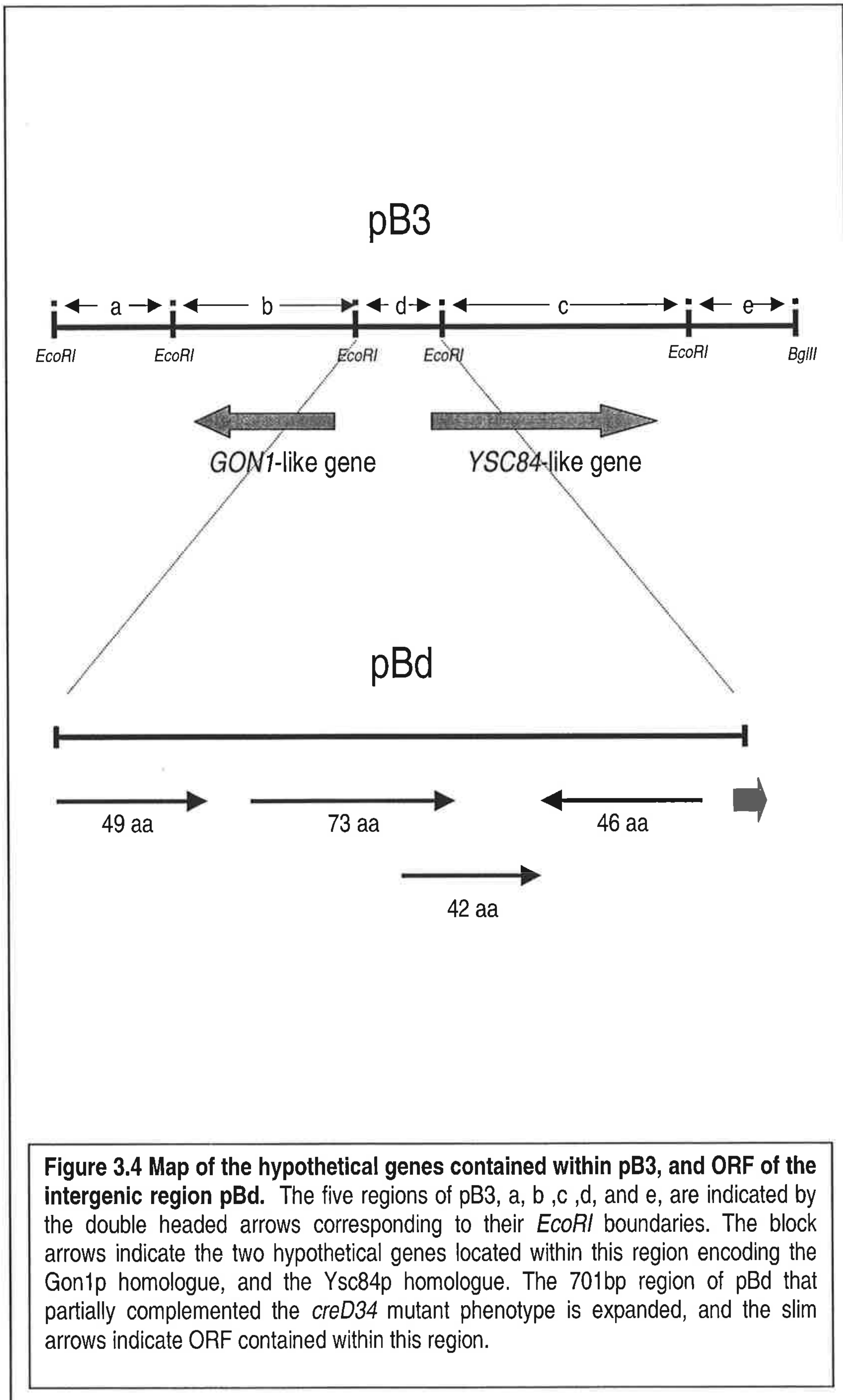

Figure 3.4 Map of the hypothetical genes contained within pB3, and ORF of the intergenic region pBd. The five regions of pB3, a, b ,c ,d, and e, are indicated by the double headed arrows corresponding to their *EcoRI* boundaries. The block arrows indicate the two hypothetical genes located within this region encoding the Gon1p homologue, and the Ysc84p homologue. The 701bp region of pBd that partially complemented the *creD34* mutant phenotype is expanded, and the slim arrows indicate ORF contained within this region.

NCBI database, and no highly similar sequences were identified, and thus they are unlikely to encode a protein.

Since regions of DNA (such as pBd and L2-Sal) could suppress the *creD34* mutant phenotype of resistance to acriflavine in complete medium, transformants that appeared resistant to the presence of acriflavine were tested on a wider range of media. A *creC27creD34* double mutant strain was transformed with the pBd suppressor plasmid, and the phenotype was tested both on medium containing acriflavine and medium containing allyl alcohol, as the *creC27* strain is sensitive to the presence of allyl alcohol and the *creC27creD34* double mutant strain is resistant to the presence of allyl alcohol. As expected, the pBd plasmid again suppressed the phenotype of *creD34* on acriflavine in a *creC27* background, but did not complement the *creD34* mutant phenotype on medium containing allyl alcohol, indicating that this suppression of the *creD34* mutant phenotype is specific for acriflavine. Arst *et al.* (1982) showed that the *creD34* mutation led to weaker growth than wild-type on medium containing 10 mM pyrrolidinone as a sole nitrogen source. The *creD34* phenotype on pyrrolidinone as a sole nitrogen source was shown to be recessive in the *creC27creD34/creC27creD*$^{+}$ diploid (Figure 3.1C), allowing this phenotype to be used in conjunction with the phenotype on acriflavine to test for complementation. None of the transformants that complemented the *creD34* mutation on acriflavine containing medium showed complementation for growth on medium containing 10 mM pyrrolidinone as a sole nitrogen source (Figure 3.3C), again indicating that a suppressor of the *creD34* phenotype on media containing acriflavine alone had been successfully cloned, but was of no further interest.

3.3 Complete complementation of the *creD34* phenotype

As the transformants that appeared to complement the *creD34* phenotype on medium containing acriflavine failed to complement on medium containing pyrrolidinone as a sole nitrogen source, the P24E11 and L21B08 cosmid were tested no further. Other cosmids that were also adjacent to P24E11 according to the published cosmid map were tested for complementation (23B07, 22B11, 22A10, 23E03 and 21G07) but none complemented the *creD34* phenotype on medium containing pyrrolidinone as a sole nitrogen source. Since these cosmids should span the 3-6 map units between *creC* and *creD*, the ordering of the cosmid library in this region was checked. All cosmids used thus far in transformation experiments, including P24E11 and L21B08, were digested with *SacII,* transferred to a nylon membrane and probed with oligolabelled pB3. The only cosmid containing pB3 was P24E11, the cosmid from which pB3 was derived. Therefore the cosmid library map was not in order in this particular region. To identify cosmids that were

adjacent to P24E11, the p8A plasmid was oligolabelled and used to probe the chromosome II specific cosmid library. The only positive hybridisations were to the cosmids already known to contain *creC* and p8A, P24E11 and P23D04 (Todd *et al.*, 2000). This indicated the presence of a gap in the cosmid library, which may contain *creD.* The partial complementation on acriflavine medium of L21B08 and L2-Sal transformants is therefore the result of a different region of the genome than that seen in pBd transformants, as these regions do not overlap. Hence it would appear that there are a number of regions in the genome that can partially suppress the acriflavine resistance phenotype due to *creD34* mutation.

Since the walking approach in the chromosome II cosmid library was shown to be impossible due to the presence of gaps, the pBc region was radiolabelled by PCR and used to probe an *A. nidulans* BAC library (Zhu *et al.*, 1997) to identify a large region likely to contain *creC* and *creD.* Forty-eight BAC clones hybridised to the pBc probe, including BAC 22E20. BAC 22E20 was shown to complement the *creD34* phenotype on complete medium containing acriflavine and on glucose medium with pyrrolidinone as a nitrogen source (Figure 3.5). To identify cosmids that were contained within the region spanned by the complementing BAC, BAC 22E20 was oligolabelled and used to probe the chromosome II specific cosmid library. Again, only the cosmids known to contain pBc (and *creC*) hybridised, indicating that there are large gaps in the cosmid library. As BAC 22E20 complements the *creD34* phenotype, and yet does not correspond to any cosmids other than the *creC* containing cosmids shown not to complement, the chromosome II specific cosmid library does not contain *creD.* The cosmid library was therefore not a useful tool in identifying *creD,* and the BAC 22E20 became the new starting point for the identification of *creD* via complementation.

3.4 Identification of *creD* within BAC 22E20

The BAC 22E20 contains approximately 70 kb of *A. nidulans* genomic DNA. In order to refine the complementing region BAC 22E20 was digested with *NotI* and religated to excise an approximate 40 kb *NotI* fragment, resulting in plasmid 22Notrel1 which contains 30.79 kb of genomic DNA. Plasmid 22Notrel1 was used to transform a *creD34; riboB2* strain and shown to complement the *creD34* mutant phenotype on medium containing acriflavine and also on D-glucose medium with pyrrolidinone as a nitrogen source to wild-type levels (Figure 3.5). Preliminary restriction analysis of 22Notrel1 was underway when the sequence of the *A. nidulans* genome became publicly available via the Whitehead Institute's Center for Genome Research website (http://www-genome.wi.mit.edu/annotation/fungi/aspergillus/index.html). This resource was utilized to analyse the *A. nidulans* genomic sequence contained within 22Notrel1. The

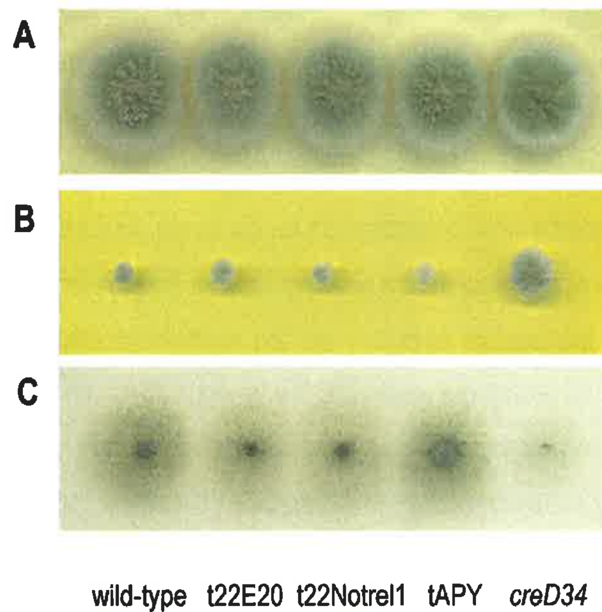

Figure 3.5 Complementation of the *creD34* mutant phenotype. All strains were grown at 37°C on either (**A**) 1% complete medium for 2 days (**B**) 1% complete medium plus 0.001% acriflavine for 2 days or (**C**) 1% glucose plus 10 mM pyrrolidinone medium for 3 days. Each strain is identified in the key below the plates, with transformants of the *creD34* strain designated tX, where X represents the transforming DNA. Each transformant shown is representative of a number of similar transformants generated in the same transformation experiment.

commercially available sequencing primer SP6 was used to sequence in from one end of 22Notrel1 to locate the region within the *A. nidulans* genome (Figure 3.6) as the SP6 promoter is found in the BACwich vector used to construct the BAC library (Zhu *et al.*, 1997). The 30,790 bp *HindIII – NotI* genomic DNA fragment of the 22Notrel1 plasmid was searched for open reading frames longer than 20 aa using the NEBcutter ORF programme (see Section 2.3.12), and reading frames longer than 100 aa were analysed against the NCBI database using BlastX. Known and hypothetical genes identified were roughly mapped within the 30 kb insert (Figure 3.6). The *GON1*-like and *YSC84*-like genes were contained within 22Notrel1 (see 3.2.3), which is not unexpected as BAC 22E20 was identified by hybridisation to sequence within the *YSC84*-like gene (pBc). As the pB3 plasmid contains these genes and was shown not to complement the *creD34* phenotype (Figure 3.3) only putative genes downstream of this region were of interest. The *CBHP2*-like and *ZDS2*-like genes had been previously shown to be within p8A (R. Sladic, *unpubl. data*), which was shown here not to complement the *creD34* phenotype (see 3.2.3, Figure 3.3). The remaining five regions containing large open reading frames within 22Notrel1 were chosen for PCR amplification to test for complementation of the *creD34* mutant phenotype as they were sufficiently far away from *creC* according to expected distances from genetic mapping comparisons to physical maps.

3.4.1 Complementation analysis of five regions within 22Notrel1

Within the 22Notrel1 plasmid the five regions containing large open reading frames not previously tested for complementation were designated APY, PH, mid, ST and novel (Figure 3.6). The APY region consists of four adjacent ORFs (51 aa, 186 aa, 143 aa and 292 aa in length) that show similarity to the arrestin domain and PY motif containing proteins, Rod1p and Rog3p from *S. cerevisiae*, and both the 51 aa and 186 aa ORFs contain an arrestin domain as determined by a conserved domain search in NCBI (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Primers APYF and APYR were designed to flank this hypothetical gene with at least an additional 600 bp upstream and downstream of the proposed gene start and stop codons, to produce a 3328 bp PCR product. The PH region consists of a single 402 aa ORF that contains a pleckstrin homology (PH) domain as determined by a conserved domain search using NCBI, and is most highly similar to the *N. crassa* hypothetical protein EAA28289.1. Primers PHF and PHR were designed to flank this hypothetical gene with an additional 1 kb of sequence upstream of the proposed start codon, and 660 bp downstream of the stop codon, and result in a 2896 bp PCR product. The ST region consists of two ORFs (93 aa and 358 aa) that are both highly similar to

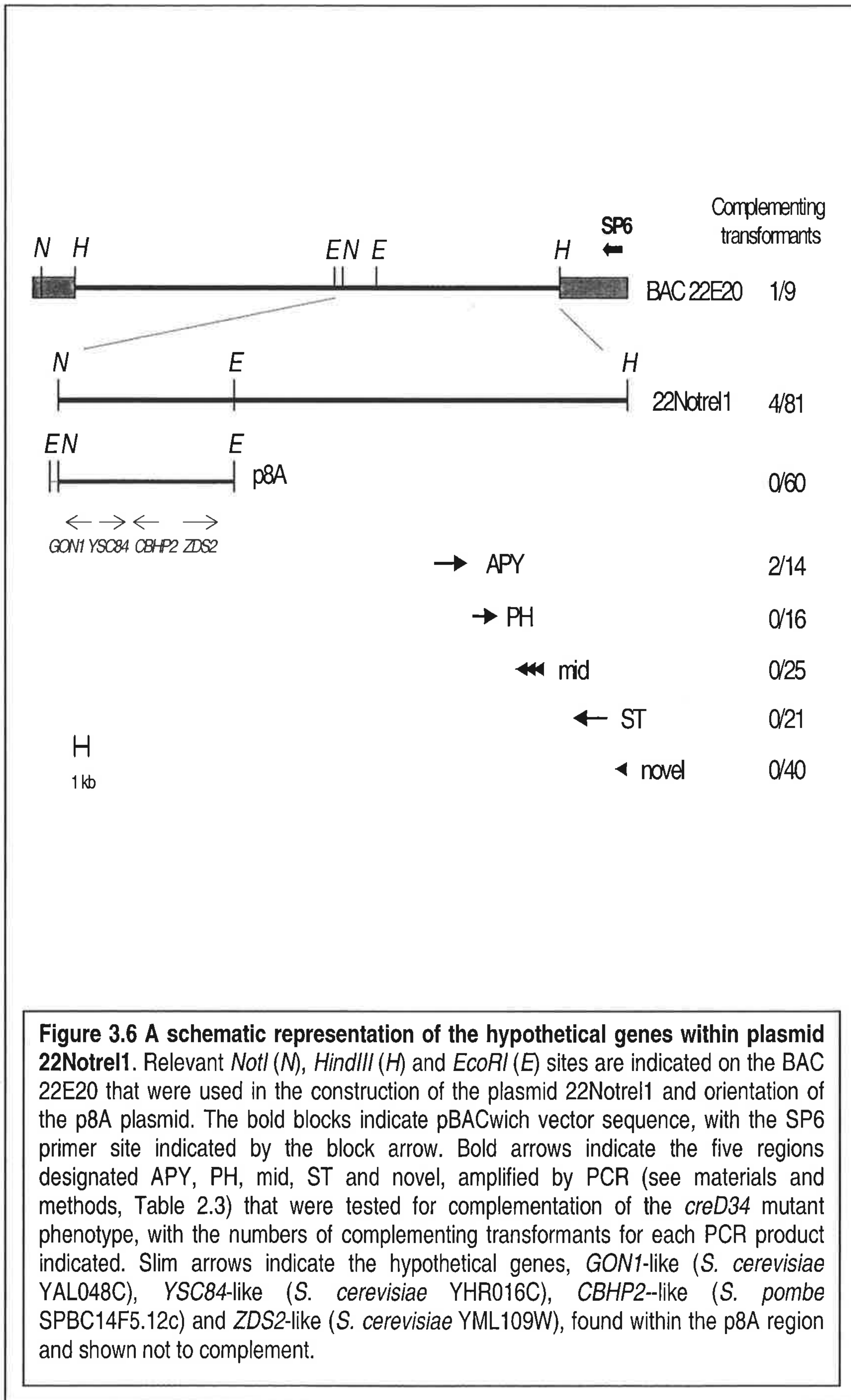

Figure 3.6 A schematic representation of the hypothetical genes within plasmid 22Notrel1. Relevant *NotI* (*N*), *HindIII* (*H*) and *EcoRI* (*E*) sites are indicated on the BAC 22E20 that were used in the construction of the plasmid 22Notrel1 and orientation of the p8A plasmid. The bold blocks indicate pBACwich vector sequence, with the SP6 primer site indicated by the block arrow. Bold arrows indicate the five regions designated APY, PH, mid, ST and novel, amplified by PCR (see materials and methods, Table 2.3) that were tested for complementation of the *creD34* mutant phenotype, with the numbers of complementing transformants for each PCR product indicated. Slim arrows indicate the hypothetical genes, *GON1*-like (*S. cerevisiae* YAL048C), *YSC84*-like (*S. cerevisiae* YHR016C), *CBHP2*--like (*S. pombe* SPBC14F5.12c) and *ZDS2*-like (*S. cerevisiae* YML109W), found within the p8A region and shown not to complement.

the putative membrane transporter CAC36925.1 from *S. pombe*, with the 358 aa ORF containing a sugar transporter (ST) domain as identified by a conserved domain search using NCBI. Primers STF and STR were designed to flank this hypothetical gene with an additional 1.5 kb upstream of the start of the first ORF, and 600 bp downstream of the putative stop codon, and produce a 3504 bp PCR product. The mid region consists of three adjacent large ORFs (159 aa, 171 aa and 282 aa in length) that have no conserved protein domains or highly similar sequences in the NCBI database. Primers midF and midR were designed to flank these three ORFs with an additional 1.5 kb upstream of the start of the first ORF, and 500 bp downstream of the end of the third ORF, and result in a 3857 bp PCR product. The novel region consists of a single 300 aa ORF, which has no highly similar sequence in the NCBI database as determined from BlastX analysis, and no conserved protein motifs. Primers novelF and novelR were designed to flank this ORF plus an additional 500 bp downstream of the stop codon, and result in a 1418 bp PCR product. Using the *A. nidulans* genome database, ORFs upstream of this 300 aa novel ORF were analysed to identify a possible homologue, but immediately upstream ORFs were also not similar to any sequence in the NCBI database.

These five regions, APY, PH, ST, mid and novel, were amplified by PCR and the resulting products were used to co-transform a *creD34; riboB2* strain with the *riboB*$^{+}$ containing plasmid pPL3 (Figure 3.6). The APY PCR product complemented the *creD34* mutant phenotype on both complete medium containing acriflavine and on glucose medium containing pyrrolidinone as a nitrogen source to wild-type levels, whilst none of the other regions resulted in complementation (Figure 3.5). Therefore, the APY PCR region that encodes a protein with similarity to the Rod1p and Rog3p proteins from *S. cerevisiae* complements the phenotype due to the *creD34* mutation.

3.5 Molecular analysis of *creD*

3.5.1 Sequence analysis of *creD* and identification of the *creD34* mutation

Primers were designed using sequence data from the Whitehead *A. nidulans* genome database to span the length of the 3328 bp complementing region encoding the arrestin domain and PY motif containing protein. Genomic DNA from this region was amplified via PCR using these custom oligonucleotides, and the resulting PCR products sequenced in both directions. CreD is encoded within a 2028 bp sequence which was found to contain four introns, as determined by sequence analysis of PCR products amplified from cDNA, and the intron sequences conformed to the consensus 5' and 3' junction (GT-AG) and lariat (RCTRAC)

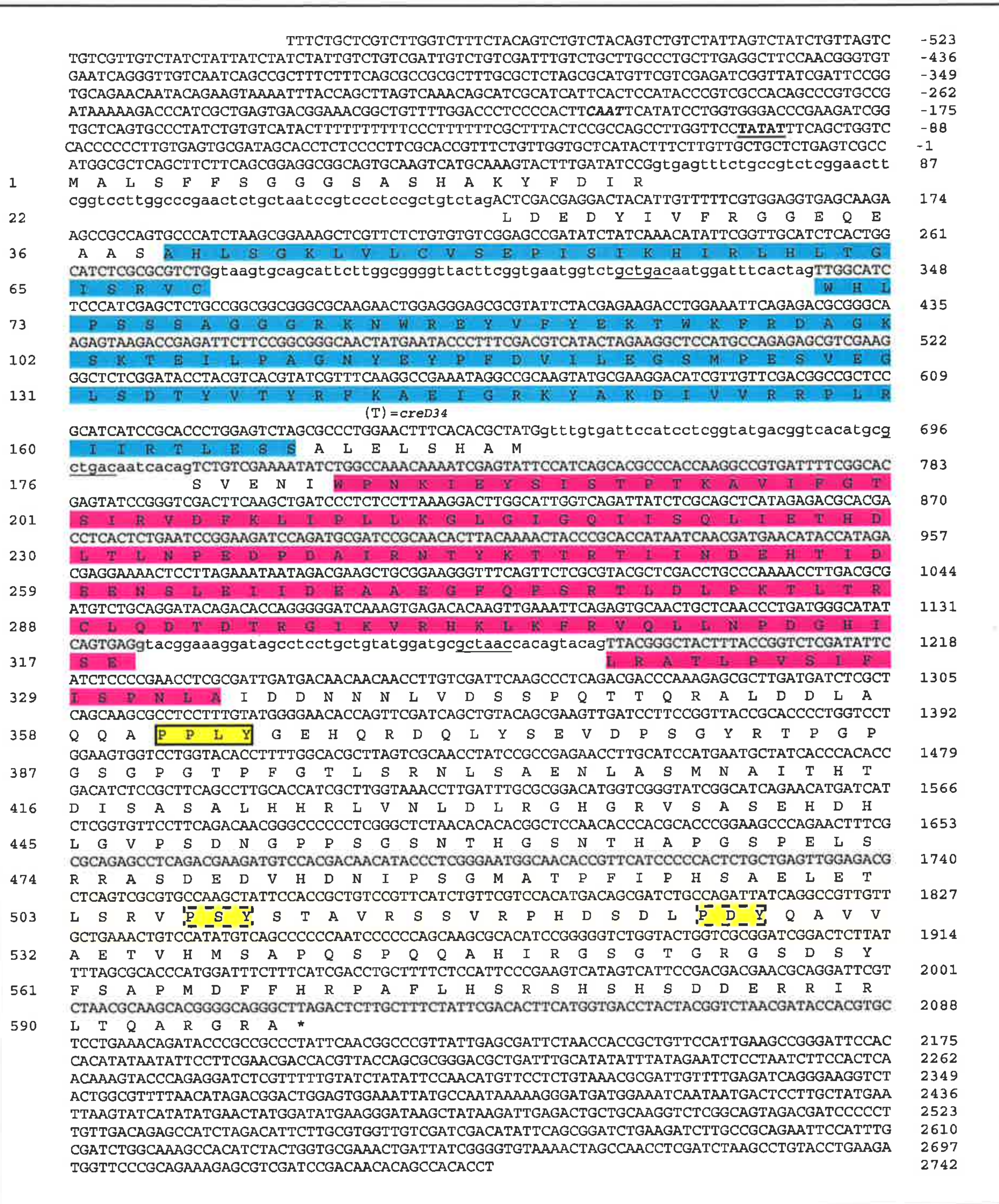

Figure 3.7 Location of the *creD34* mutation within the *creD* gene. The complete sequence of the *creD* genomic region is shown. The four introns are depicted in lowercase letters and lariat sequences are underlined. A putative TATA box is underlined and in boldface type, and CAAT motif bolded and italicised. The amino acid sequence is shown in single letter code beneath the DNA sequence, with numbered amino acid residues indicated on the left-hand side, and numbered nucleotide positions indicated on the right-hand side, with the first nucleotide of the start codon as +1. The position and nature of the *creD34* mutation is shown as a single base change at nt + 642 which results in a stop codon. An asterisk indicates the stop codon of *creD*. The amino acids of the arrestin_N domain are shaded blue, and amino acids of the arrestin_C domain are shaded pink. The PPXY motif is highlighted yellow and boxed in a bold line, and the minimal core PXY motifs are highlighted yellow and boxed in a broken line. The *creD* sequence has been submitted to the GenBank database under Accession number AY458430.

sequences for filamentous fungal introns (Gurr *et al.*, 1987) (Figure 3.7). The predicted translation start point of CreD shows similarity to the Kozak consensus sequence (Kozak, 1987), and a putative TATA box is located 103 bp upstream of the ATG, and a CAAT motif 206 bp upstream of the ATG. The *creD* gene encodes a 597 aa protein that contains an arrestin_N domain (aa 39 - 168) (PF00339), an arrestin_C domain (aa 181 – 334) (PF027252) (SMART protein motif analysis programme: Schultz *et al.*, 1998), a single PP*X*Y motif (aa 361 - 364) and two P*X*Y motifs (aa 507 – 509 and 525 – 527).

Particularly in view of the partial suppressors initially identified, the *creD* gene was sequenced from PCR products produced from genomic DNA obtained from a wild-type strain and a strain with the *creD34* mutation to identify the molecular nature of the *creD34* mutation and confirm that this gene was indeed *creD*. Primers were designed using sequence data from the Whitehead *A. nidulans* genome database (see Table 2.3). The *creD*$^{+}$ and *creD34* alleles were sequenced in their entirety and *creD34* was found to contain a G-to-T transition at nt 642 that causes the codon for glutamic acid at aa 170 to become an ochre stop codon, thereby truncating the remaining 427 amino acids. This mutant gene product lacks the PY motifs and the arrestin_C domain (Figures 3.7 and 3.8).

3.5.2 Bioinformatic analysis of *creD* and identification of *apyA*

Analysis using BlastX revealed CreD was most similar to the hypothetical protein SPAC31A2.12 from *Schizosaccharomyces pombe* (1.8e^{-60}), with the most similar analysed proteins to CreD being Rod1p and Rog3p from *S. cerevisiae*, which both contain an arrestin_N and arrestin_C domain, two PP*X*Y motifs and a P*X*Y motif. BestFit alignments of CreD and Rod1p, and CreD and Rog3p, indicate they share 32.8% aa identity and 55.0% aa similarity, and 34.9% aa identity and 56.9% aa similarity respectively, across the length of the proteins (Figure 3.9A). As there were two such highly conserved sequences to CreD in *S. cerevisiae*, *creD* was used to search the *A. nidulans* genome database and a second gene, *apyA*, encoding a hypothetical protein containing an arrestin_N and arrestin_C domain and PP*X*Y motif was identified on chromosome VI (locus AN3265.2) (Figure 3.8). *apyA* was amplified from cDNA using PCR and sequenced, and the sequence concurred with the *A. nidulans* Whitehead genome database prediction of an 1107 bp ORF interrupted by four introns. The *apyA* cDNA sequence data was entered into the GenBank database under the Accession number AY460113. The hypothetical ApyA is 368 aa in length, and shares 44.2% aa identity, and 64.9% aa similarity with CreD. BestFit alignments of ApyA and Rod1p, and ApyA and Rog3p, indicate they share 29.9% aa identity and 54.0% aa similarity, and 31.3% aa identity and 55.5% aa similarity respectively,

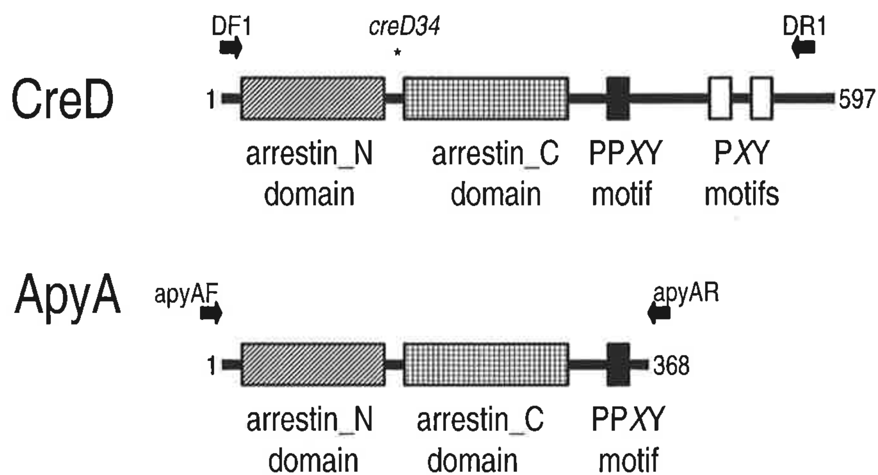

Figure 3.8 Schematic representation of CreD and ApyA. The diagonally striped box indicates the arrestin_N domain, the gridded box represents the arrestin_C domain, the filled box represents the PP*X*Y motif and the open box represents the P*X*Y motifs. Block arrows indicate the primers used to amplify regions of CreD and ApyA ligated into the pTRG vector (see materials and methods). The CreD34 protein is truncated at the position marked with an asterisk.

across the length of the proteins (Figure 3.9A). To determine the phylogenetic and evolutionary relationship between these genes, maximum parsimony analysis was undertaken which included arrestin domain containing proteins from *Aspergillus fumigatis* (Apy1), *Neurospora crassa* (Apy1) and *S. pombe* (Apy1 and Apy2), identified by Blast searches in the respective genome databases. Interestingly, the two *S. cerevisiae* proteins grouped strongly together (with 96% support of bootstrap replicates), indicating *ROD1* and *ROG3* are likely to be paralogous genes resulting from a gene duplication event within the *S. cerevisiae* lineage (Figure 3.9B). Similarly, *creD* and *apyA* appear to have arisen due to a gene duplication within the *A. nidulans* lineage (Figure 3.9B), a duplication event that did not occur in *N. crassa* as only a single protein containing an arrestin domain was identified in the complete genome sequence of *N. crassa*. These results account for the almost identical similarities of ApyA and CreD to Rod1p and Rog3p.

3.5.3 Multiple copies of *apyA* do not compensate for the absence of *creD*

Since the *creD34* mutant phenotype is apparent in an *apyA*$^+$ background, the two proteins cannot simply have identical functions unless their concentration is limiting. To investigate whether multiple copies of the *apyA* gene can compensate for the lack of a functional *creD* gene, a genomic clone of the *apyA* gene was generated by cloning the PCR product generated using primers apyACF and apyACR (see Table 2.3), and DNA of fosmid 8052C12 (known to contain *apyA*) as the template, into pBluescript. This clone, pBapyA, was introduced into a *creD34; riboB2* strain by co-transformation with pPL3 containing the *riboB*$^+$ gene. The transformants were analysed by hybridisation of Southern transfers, and strains with a low number (2) and a high number (5+) of copies of *apyA* were identified (Figure 3.10). These strains were tested for sensitivity to the presence of acriflavine in complete medium and growth on pyrrolidinone as a nitrogen source, and they showed the same phenotype as *creD34* (Figure 3.11). Thus, multiple copies of *apyA* cannot compensate for the absence of functional *creD*.

3.6 Proteins that interact with CreD and ApyA

3.6.1 Testing protein-protein interactions between CreD and CreA, ApyA and CreA

The Stratagene BacterioMatch I bacterial-2-hybrid system (B2H I) is an easy and rapid alternative to yeast two-hybrid systems for the detection of protein-protein interactions *in vivo*. All screening and validation steps are performed in bacteria. A protein of interest (the bait) is fused to the full-length bacteriophage λcI repressor protein (λcI) within the bait vector (pBT), and binds to the λ operator. The target protein is fused to the N-terminal domain of the α-subunit of RNA

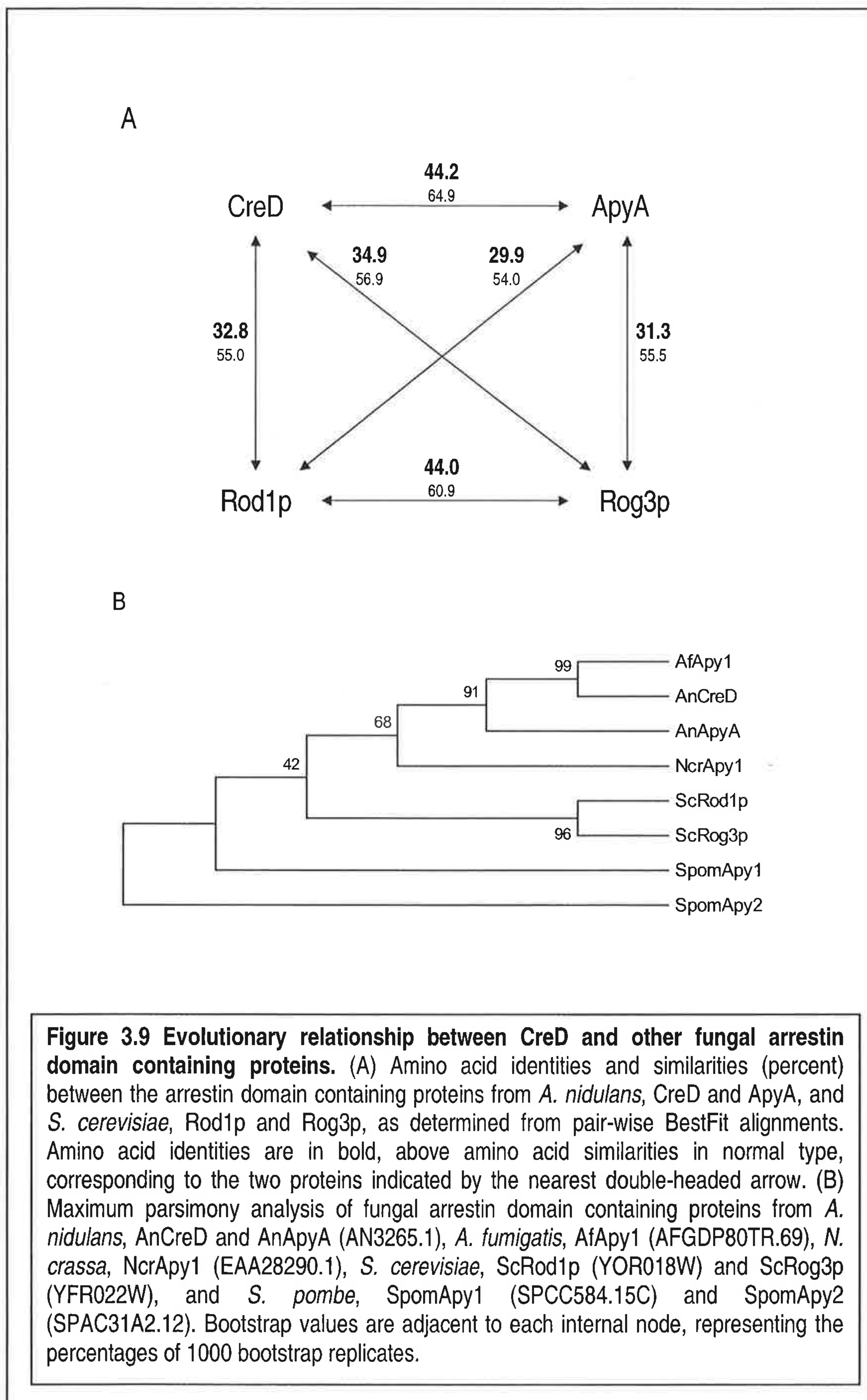

Figure 3.9 Evolutionary relationship between CreD and other fungal arrestin domain containing proteins. (A) Amino acid identities and similarities (percent) between the arrestin domain containing proteins from *A. nidulans*, CreD and ApyA, and *S. cerevisiae*, Rod1p and Rog3p, as determined from pair-wise BestFit alignments. Amino acid identities are in bold, above amino acid similarities in normal type, corresponding to the two proteins indicated by the nearest double-headed arrow. (B) Maximum parsimony analysis of fungal arrestin domain containing proteins from *A. nidulans*, AnCreD and AnApyA (AN3265.1), *A. fumigatis*, AfApy1 (AFGDP80TR.69), *N. crassa*, NcrApy1 (EAA28290.1), *S. cerevisiae*, ScRod1p (YOR018W) and ScRog3p (YFR022W), and *S. pombe*, SpomApy1 (SPCC584.15C) and SpomApy2 (SPAC31A2.12). Bootstrap values are adjacent to each internal node, representing the percentages of 1000 bootstrap replicates.

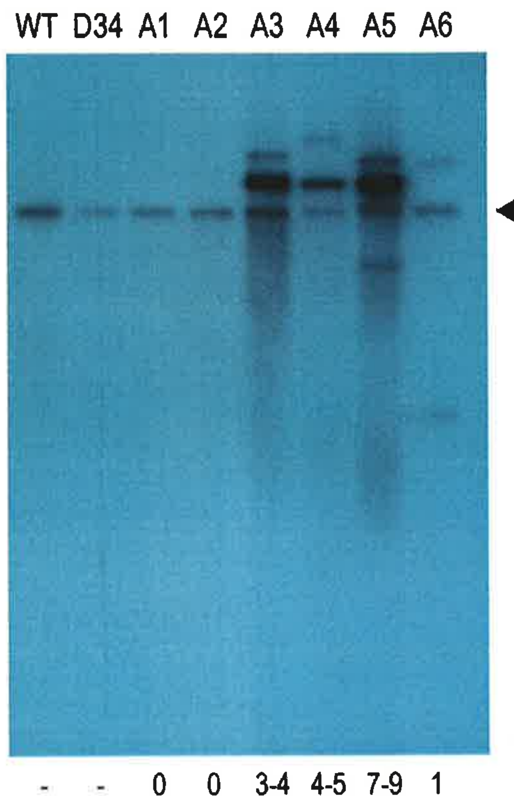

Figure 3.10 Determining *apyA* copy number in tapyA transformant strains via Southern analysis. Autoradiograph exposed overnight to a Southern filter probed with *apyA* that contained genomic DNA from WT, a wild-type strain, D34, the *creD34* parent strain, and A1-A6, six individual cotransformants of *riboB*$^+$ and *apyA*. The arrow indicates the endogenous copy of *apyA*. Additional copy numbers of *apyA* were estimated by the number of additional bands and strength of hybridisation. Phosphorimaging quantitative analysis of the additional bands relative to the level of hybridisation to the endogenous *apyA* band in each individual transformant track was also performed to help determine additional copy numbers of *apyA* in each transformant, and these estimates are shown below the autoradiograph.

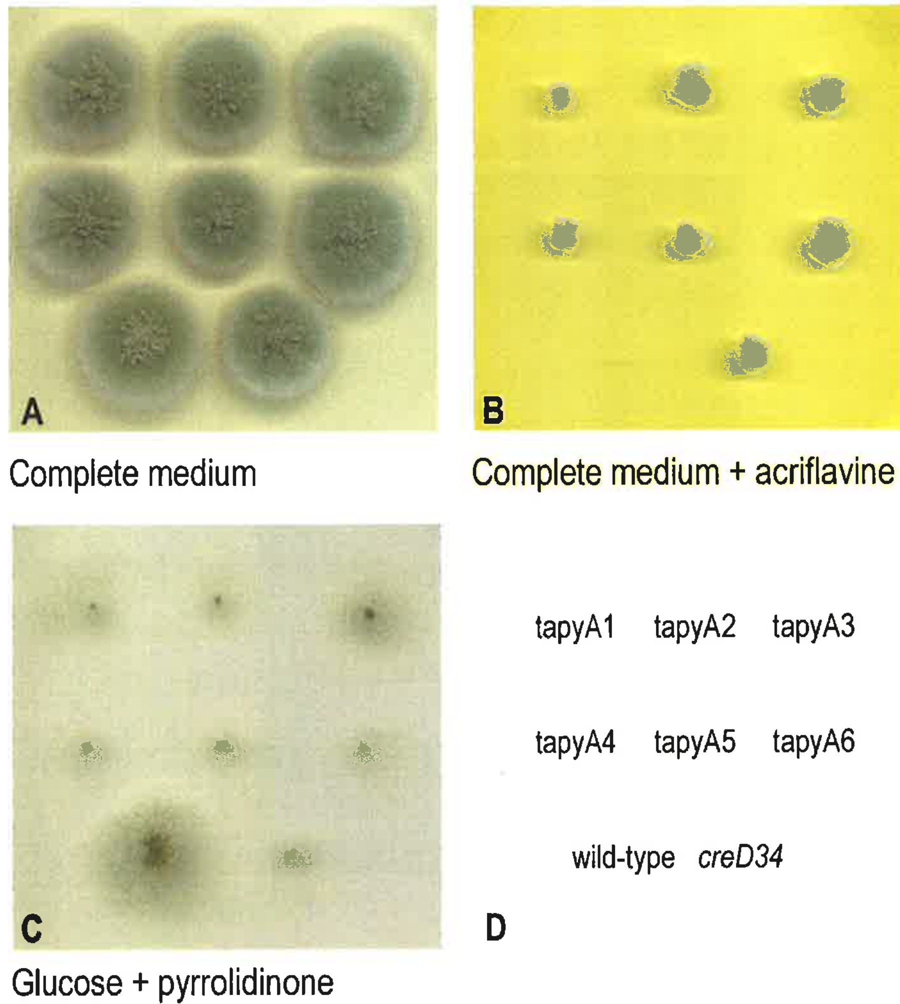

Figure 3.11 Multiple copies of *apyA* does not compensate for lack of *creD*. All strains were grown at 37°C on either (**A**) 1% complete medium for 2 days (**B**) 1% complete medium plus 0.001% acriflavine for 2 days or (**C**) 1% glucose plus 10 mM pyrrolidinone medium for 3 days. The strains are indicated by the key (**D**). Six individual *riboB*$^+$ *apyA* cotransformants are designated tapyA1-6.

polymerase within the target vector (pTRG). When the bait and target interact, they recruit and stabilize the binding of RNA polymerase close to the promoter and activate the transcription of the *amp*^{r} reporter gene. A second reporter gene, *lacZ*, is expressed from the same promoter and provides an additional screen to validate the bait and target interaction

To investigate whether CreD or ApyA interact with CreA, the bacterial-2-hybrid system (B2H I) was used to detect protein-protein interactions. The entire coding region of *apyA* cDNA was ligated into the target vector, to produce the construct pTRGapyA, and the majority of the coding region of *creD* cDNA, incorporating amino acids 31 to 580 that encompass the arrestin domains and PY motifs, was ligated into the target vector to produce the pTRGcreD construct. The previously made pBTcreA construct incorporates aa 139 to 416 of CreA into the pBT vector, which includes all of the CreA protein beyond the DNA binding region (R. Lockington, *unpubl. data*). All vectors were tested for auto-activation of the reporter system, and deemed suitable for use in the B2H I system as no auto-activation was observed. The pBTLGF2 and pTRGGAL11^p constructs provided with the B2H I kit served as a positive control. Neither CreD and CreA, nor ApyA and CreA directly interacted in this system.

3.6.2 Identification of *hulA* and testing HulA for interactions with CreD and ApyA.

The Rod1p and Rog3p proteins have been shown to interact via their PY motifs with the WW domains of the ubiquitin ligase Rsp5p in *S. cerevisiae* (Andoh *et al.*, 2002). In order to identify the *A. nidulans* homologue of the Rsp5p protein, the *A. nidulans* genome was searched using the Rsp5p sequence in a tBlastN search, and six HECT ubiquitin ligases were identified and designated as follows; *hulA* (AN1339.2), *hulB* (AN0444.2), *hulC* (AN1746.2), *hulD* (AN1874.2), *hulE* (AN1966.2), and *hulF* (AN3999.2). Rsp5p was used to also search the *S. cerevisiae* genome via a tBlastN search, and five HECT ubiquitin ligases were identified as follows; Rsp5p, Tom1p, Hul4p, Hul5p and Ufd4p. Phylogenetic analysis was undertaken on all 11 proteins to identify orthologous genes from the two species, and the Rsp5p homologue in *A. nidulans* was identified as the hypothetical protein HulA (AN1339.2), strongly supported with 100% of bootstrap replicates of the Neighbour-Joining dendrogram (Figure 3.12).

Like Rsp5p, the *A. nidulans* HulA protein contains a HECT ubiquitin ligase domain, a C2 domain and three WW domains. To examine whether the interactions between the PY motif containing proteins and the WW domains of HulA also occurred in *A. nidulans*, constructs were made to test the direct protein-protein interactions between CreD and HulA, and ApyA and HulA via the BacterioMatch I bacterial-2-hybrid system. The pBThulA plasmid was constructed by the PCR amplification of the *A. nidulans hulA* gene from cDNA using the primers BThulAFRI and

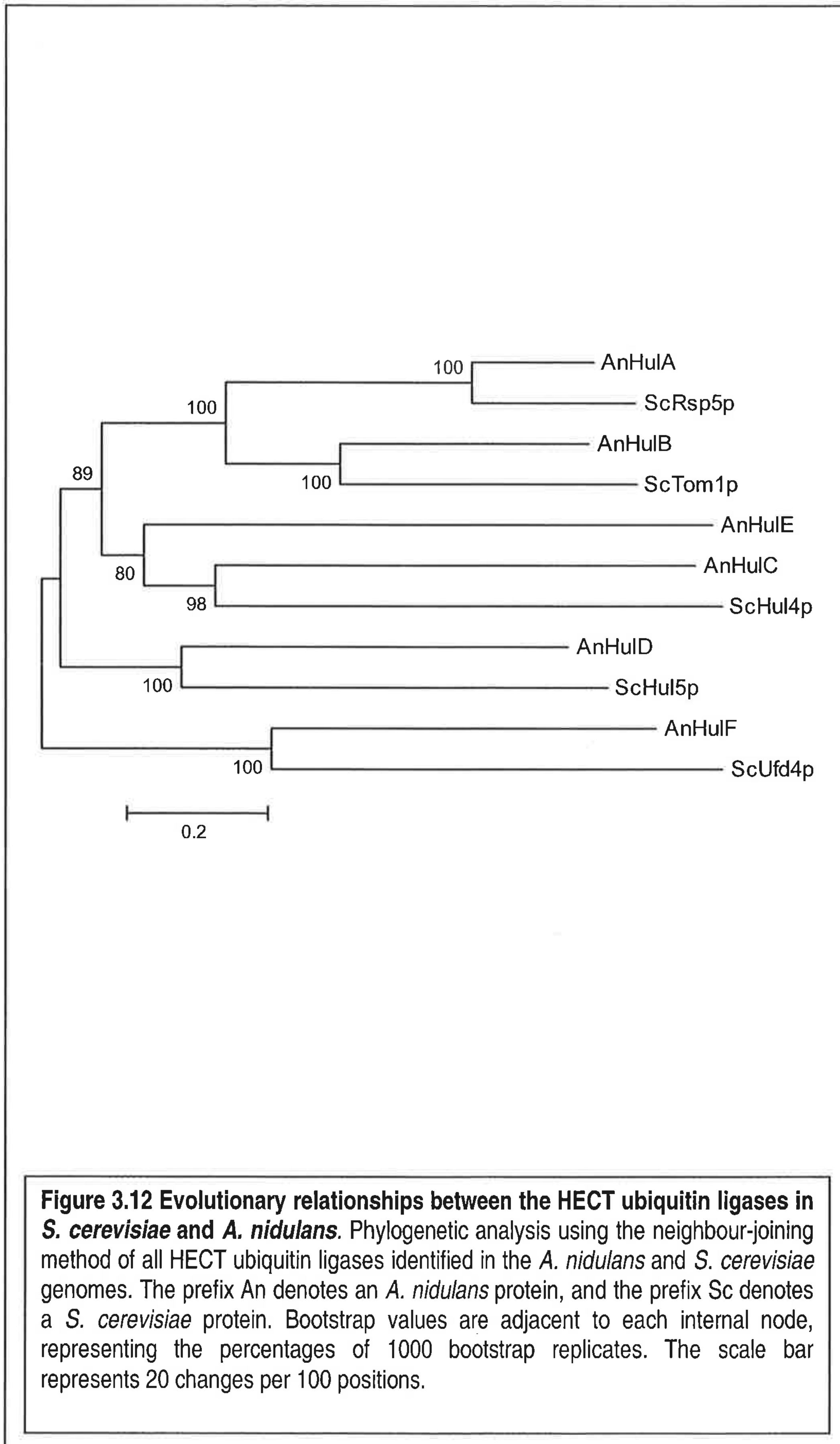

Figure 3.12 Evolutionary relationships between the HECT ubiquitin ligases in *S. cerevisiae* and *A. nidulans*. Phylogenetic analysis using the neighbour-joining method of all HECT ubiquitin ligases identified in the *A. nidulans* and *S. cerevisiae* genomes. The prefix An denotes an *A. nidulans* protein, and the prefix Sc denotes a *S. cerevisiae* protein. Bootstrap values are adjacent to each internal node, representing the percentages of 1000 bootstrap replicates. The scale bar represents 20 changes per 100 positions.

BThulARBam, incorporating amino acids 218 to 445 that encompass the three WW domains of HulA into the pBT vector. The ApyA protein was shown to interact with the WW domains of HulA, as indicated by growth on carbenicillin containing medium and validated by a blue colour reaction on X-gal indication plates (data not shown). However, the CreD protein did not interact with HulA in the B2H I system.

The BacterioMatch II bacterial-2-hybrid system was released in late 2003, and is compatible with plasmids from the B2H I system, but utilizes a different reporter system that has a reduced background and is thus more sensitive for detecting protein-protein interactions. The new *HIS3-aadA* reporter cassette allows the detection of protein-protein interactions by transcriptional activation of the *HIS3* reporter gene, which allows growth on minimal medium lacking histidine. The competitive inhibitor of the His3 enzyme, 3-amino-1,2,4-triazaole (3-AT), was used to reduce the background. Potential positive interactions can be verified by the secondary reporter *aadA* gene, which confers streptomycin resistance. To differentiate between the positive and negative controls for the B2H I system, the carbenicillin concentration used to test for protein-protein interactions was three times more than that recommended by Stratagene. To confirm the positive interaction seen between ApyA and HulA, and to test the possibility that a weaker interaction between CreD and HulA that may have fallen beneath background detection levels in the B2H I system exists, the new B2H II reporter cells were transformed to test for interactions between these proteins. The ApyA protein was shown to interact strongly with the WW domains of HulA, as indicated by strong growth on medium lacking histidine (*HIS3* activation) and validated by resistance to streptomycin (*aadA* activation) (Figure 3.13). Interaction of the CreD protein with HulA was clearly detected, but it was weaker as indicated by weaker growth on medium lacking histidine (+/- streptomycin) (Figure 3.13).

3.7 Discussion

In the initial stages of this work, complementation of the *creD34* mutant phenotype was only assessed on complete medium containing acriflavine. On this test, p8A as well as other regions of the genome not in the region of *creD* (*eg*. cosmid L21B08) apparently complemented. For p8A, this region was narrowed down to a 701 bp region that contained the regulatory regions of the divergently transcribed *GON1*–like and *YSC84*-like genes, but sequence analysis showed no sequence differences between the wild-type and *creD34* strains. Thus certain regions of the genome, when in high copy number, can suppress the acriflavine resistance phenotype conferred by the *creD34* mutation. However, when these transformants were tested on glucose medium containing pyrrolidinone as a sole nitrogen source, no complementation was observed, and they

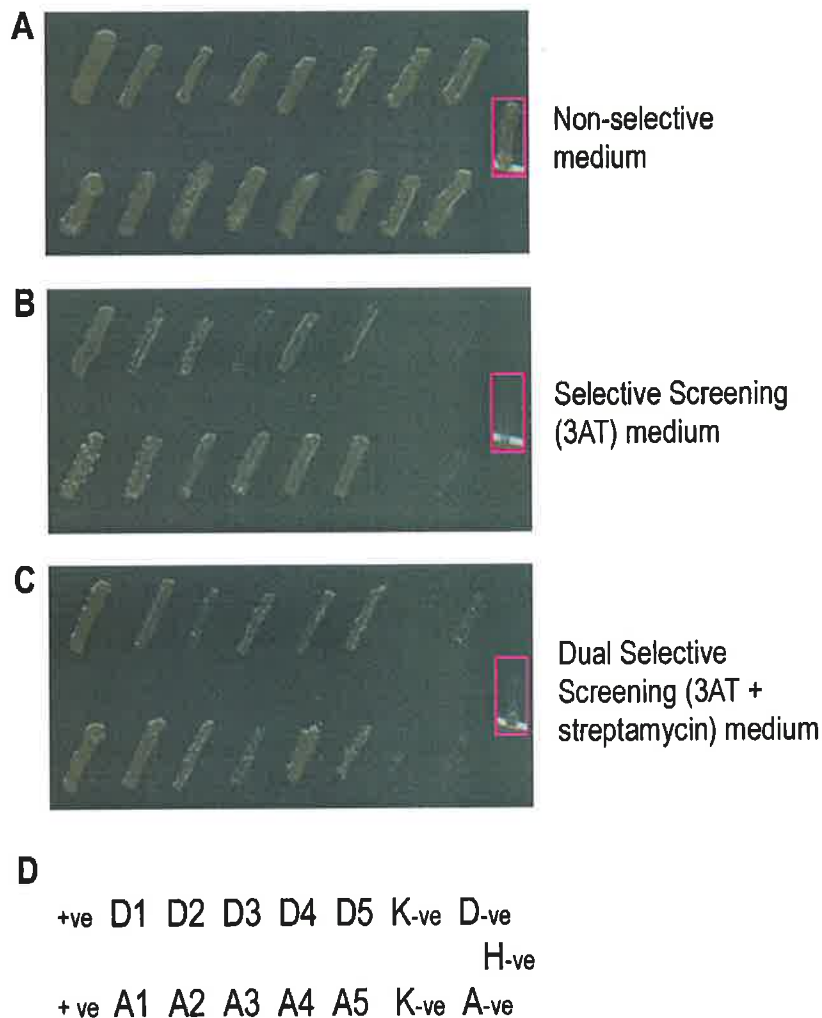

Figure 3.13 Investigation of protein-protein interactions between CreD and HulA, and ApyA and HulA as determined by the BacterioMatch II bacterial – 2 – hybrid system. (A) Non-selective screening (minimal) medium supplemented with 25 μg/ml chloramphenicol and 12.5 μg/ml tetracycline. (B) Selective Screening Medium containing minimal medium plus 5 mM 3-amino-1,2,4-triazole (3-AT), 25 μg/ml chloramphenicol and 12.5 μg/ml tetracycline (C) Dual Selective Screening Medium containing minimal medium plus 5 mM 3-AT and 12.5 μg/ml streptomycin, 25 μg/ml chloramphenicol and 12.5 μg/ml tetracycline. (D) Key to strains: +ve indicates the positive control strain, cotransformed with pBTGLF2 and pTRGGa11p. D1-5 are five independent strains cotransformed with pBThulA and pTRGcreD. K–ve indicates a negative control strain, cotransformed with pBT and pTRGGal11p. D-ve indicates a negative control strain cotransformed with pBT (no insert) and pTRGcreD. A1-5 are five independent strains cotransformed with pBThulA and pTRGapyA. A-ve indicates a negative control strain cotransformed with pBT (no insert) and pTRGapyA. H-ve indicates a negative control strain cotransformed with pBThulA and pTRG (no insert). All plates incubated at 37°C overnight.

grew as poorly as the *creD34* mutant strain (Figure 3.3). These early experiments were utilising the ordered chromosome II specific cosmid library of *A. nidulans*, however it was shown here that this was not suitable for the identification of *creD* as the cosmid library is unordered in this region and contains large gaps.

The *creD* gene encodes a protein that contains an arrestin_N and arrestin_C domain, a PP*X*Y motif and two P*X*Y motifs, and is highly similar to the Rod1p and Rog3p proteins from *S. cerevisiae*. The presence of two such similar sequences to *creD* in the *S. cerevisiae* genome led to the search for another arrestin domain containing protein in *A. nidulans*, which identified a gene we have designated *apyA*, encoding an arrestin_N and arrestin_C domain and PP*X*Y motif containing protein. In *S. cerevisiae* Rod1p and Rog3p have been shown to interact with the ubiquitin ligase Rsp5p, and so to determine whether this interaction occurred in *A. nidulans* the homologue of Rsp5p was identified, and the gene encoding this HECT ubiquitin ligase was designated *hulA*. CreD and ApyA were tested for interactions with HulA via the bacterial-2-hybrid system, and strong interaction was found between ApyA and HulA, with weaker but clear interaction between CreD and HulA.

Rod1p and Rog3p from *S. cerevisiae* are the two most highly similar characterized proteins to CreD and ApyA, and both of these proteins contain an arrestin_N and arrestin_C domains, two PP*X*Y motifs and a P*X*Y motif. PY motifs [PP*X*Y] are small proline rich sequences that are commonly bound by WW domains (Chen and Sudol, 1995), and P*X*Y has been shown to be a minimal basic core that can bind WW domains with reduced affinity (Chen *et al.*, 1997). *ROD1* was identified in a high copy plasmid library screen selecting for increased resistance to *o* – dinitrobenzene (*o*-DNB). The Rod1p protein also plays a physiological role in the resistance to calcium and zinc (Wu *et al.*, 1996). Loss of Rod1p causes cells to become hypersensitive to *o*-DNB, calcium, zinc and diamide (Wu *et al.*, 1996). *ROG3* was identified as a suppressor mutation of the temperature sensitivity found in a *mck1 mds1* double mutant strain, where *mck1* and *mds1* encode homologues for glycogen synthase kinase 3 (Andoh *et al.*, 2002). Rod1p and Rog3p share 44% sequence identity with each other, and both of these proteins have been shown to interact via their PY motifs with the WW domains of the E3 ubiquitin ligase Rsp5p (Andoh *et al.*, 2002). The Rsp5p ubiquitin-protein ligase contains a HECT (homologous to the E6-Associated Protein Carboxyl Terminus) domain that catalyses ubiquitin transfer, and three WW domains. WW domains are protein modules approximately 40 amino acids in length which contain two highly conserved tryptophans about 20 – 23 aa apart (Bork and Sudol, 1994). These small domains fold in a three stranded, anti-parallel β-sheet, which creates a hydrophobic pocket for binding proline rich ligands motifs. WW domains of the Rsp5p type bind PP*X*Y motifs (Chen and

Sudol, 1995). Rsp5p is not only involved with protein degradation, but other cellular functions such as transcriptional regulation and transcriptional-coupled repair (Chang *et al.*, 2000), and is a trans-activator for the steroid-hormone receptor family of transcription factors (Imhof and McDonnell, 1996). The human homologue of Rsp5p, Nedd4, has been shown to regulate the turnover of the epithelial plasma membrane sodium channel (ENaC). Mutations of the PY motif in the ENaC disrupt the interaction with the WW domains of Nedd4 and results in Liddle's syndrome, a severe hereditary form of hypertension in humans (Schild *et al.*, 1996).

The CreD protein is predicted to be involved in the ubiquitination aspect of the regulatory network involving CreA, CreB, CreC and AcrB from the analysis of the epistatic interactions of the various mutation combinations (Kelly and Hynes, 1977; Boase *et al.*, 2003). The proposed model of carbon catabolite repression regulated via CreA modification or stability involves the CreB deubiquitination enzyme acting in a protein complex with (at least) the CreC protein to modify or stabilize CreA under carbon repressing conditions (Lockington and Kelly, 2002). Strauss *et al.* (1999) have also implicated protein modification or stability of CreA in the mechanism of carbon catabolite repression. The CreB/CreC complex is predicted to have substrates additional to CreA, as *creB* and *creC* mutations have a broader range of mutant phenotypes under both repressing and derepressing conditions than *creA* (Lockington and Kelly, 2001). Both *acrB* and *creD* mutations have been shown to suppress aspects of the *creB* and *creC* mutant phenotypes (Kelly and Hynes, 1977; Boase *et al.*, 2003) implicating *creD* and *acrB* as having a role in ubiquitination.

The interaction in *S. cerevisiae* between Rod1p and Rog3p with the E3 ubiquitin-protein ligase Rsp5p, and the presence of two homologous proteins, CreD and ApyA, in *A. nidulans* would suggest that there might also be an interaction between CreD and HulA, and ApyA and HulA in *A. nidulans.* Bacterial-2-hybrid (B2H II) analysis of CreD and ApyA with the WW domains of HulA from *A. nidulans* indicated that ApyA and the WW domains of HulA interact strongly, and that CreD and HulA interact, but less strongly, in this system. The ApyA protein of *A. nidulans* shares 44% amino acid identity with CreD, and both contain an arrestin_N and arrestin_C domain and a single PP*X*Y motif. CreD contains the low-affinity minimal core motifs, PSY and PDY, plus the PP*X*Y motif, PPLY, whereas ApyA contains a single PY motif, PPAY. These different motifs may have different affinities for any given binding partners, or interact with different partners. The introduction of extra copies of the *apyA* gene did not suppress the phenotype due to the *creD34* mutation, and thus in the absence of CreD the elevated levels of *apyA* expression cannot compensate. Thus the ApyA protein of *A. nidulans* is a newly identified protein proposed to be involved in the ubiquitination aspect of this regulatory network.

There are a number of HECT ubiquitin ligases in the *A. nidulans* genome, and since CreD did not interact as strongly as ApyA with HulA, it is possible that CreD also interacts with another HECT ubiquitin ligase. A BlastX search of the *A. nidulans* genome revealed the presence of six such HECT ubiquitin ligases, and a BlastX search of the *S. cerevisiae* genome identified five HECT ubiquitin ligases. Phylogenetic analyses on all 11 proteins revealed that each of the five *S. cerevisiae* HECT ubiquitin ligases had a clear orthologue in *A. nidulans*. The additional HECT ubiquitin ligase in *A. nidulans* HulE (AN1874.1) may be a partner for CreD. *creD* and *apyA* appear to be the result of a gene duplication event, so it is reasonable to think that they may have evolved different specificities for the same substrates, as well as different substrates.

Both CreD and ApyA contain an arrestin_N and arrestin_C domain. Arrestins specifically bind to phosphorylated activated forms of their associated G protein-coupled receptors preventing further G protein activation, often redirecting signalling to other pathways (reviewed in Palczewski, 1994). This suggests that CreD is involved in signalling by recognising appropriately phosphorylated substrates via its arrestin domains. The arrestin domain could target CreD and ApyA to phosphorylated substrates, and then recruit a HECT type ubiquitin ligase to ubiquitinate the substrate, providing a link between ubiquitination and phosphorylation in protein regulation and stability. There is currently no mutant available for *apyA*. The similarity in sequence to *creD* would suggest a role in the carbon catabolite repression pathway, and the interaction with HulA implies ApyA could also be involved in the ubiquitination aspect of the network that involves the CreB/CreC complex.

CHAPTER 4: MOLECULAR CHARACTERIZATION AND ANALYSIS OF *acrB*

4.1 Introduction

The *acrB2* mutation was isolated in the filamentous fungus *A. nidulans* as a spontaneous resistant sector on complete medium containing acriflavine in a genetic screen with the joint aims of understanding acriflavine toxicity and obtaining extra tools for genetic mapping (Roper and Kafer, 1957). As outlined in the previous chapter, the *creD34* mutation also confers resistance to acriflavine and was identified as a suppressor of the effects of the *creC27* mutation. This aspect of phenotypic similarity between *acrB2* and *creD34* led us to analyze whether the *acrB2* mutation is also a suppressor of the phenotypes due to *creB* and *creC* mutations, which would indicate a role for AcrB in a regulatory ubiquitination pathway that is required for correct regulation of carbon metabolism.

4.2 Phenotypic analysis of *acrB2*, *acrB14* and *acrB15*

As selected, the three *acrB* mutant strains analysed, *acrB2, acrB14* and *acrB15*, were resistant to the presence of acriflavine in complete medium, with similar levels of resistance between *acrB* mutant strains and the *creD34* strain (Figure 4.1B). All three alleles conferred a slightly altered colony morphology on complete medium. All three *acrB* mutant strains were more sensitive to the presence of 11 mM molybdate in synthetic complete medium than the wild-type strain, but the alleles were heterogeneous, with the *acrB2* strain being the most sensitive to the presence of molybdate, and the *acrB15* strain showing lesser sensitivity (Figure 4.1D). Thus the effects of the *acrB* mutant alleles when the toxic compounds acriflavine and molybdate are added to the medium are the same as those for *creD34*, and the opposite to the effects of *creB* and *creC* mutant alleles which confer increased sensitivity to acriflavine and resistance to molybdate added to complete medium (Arst, 1981; Hynes and Kelly, 1977; Kelly and Hynes, 1977).

When tested for their ability to utilize various carbon sources, *acrB* mutant strains were found to have pleiotropic phenotypes with respect to carbon source utilization. They show decreased ability to utilize a number of different sugars as sole carbon sources in comparison to both the wild-type and the *creD34* mutant strain, including fructose, cellobiose, trehalose and starch. In general, the *acrB14* strain showed a slightly more extreme phenotype than the *acrB2* and *acrB15* strains on alternative carbon sources (Figure 4.1E-F; Table 4.1). The reduced growth

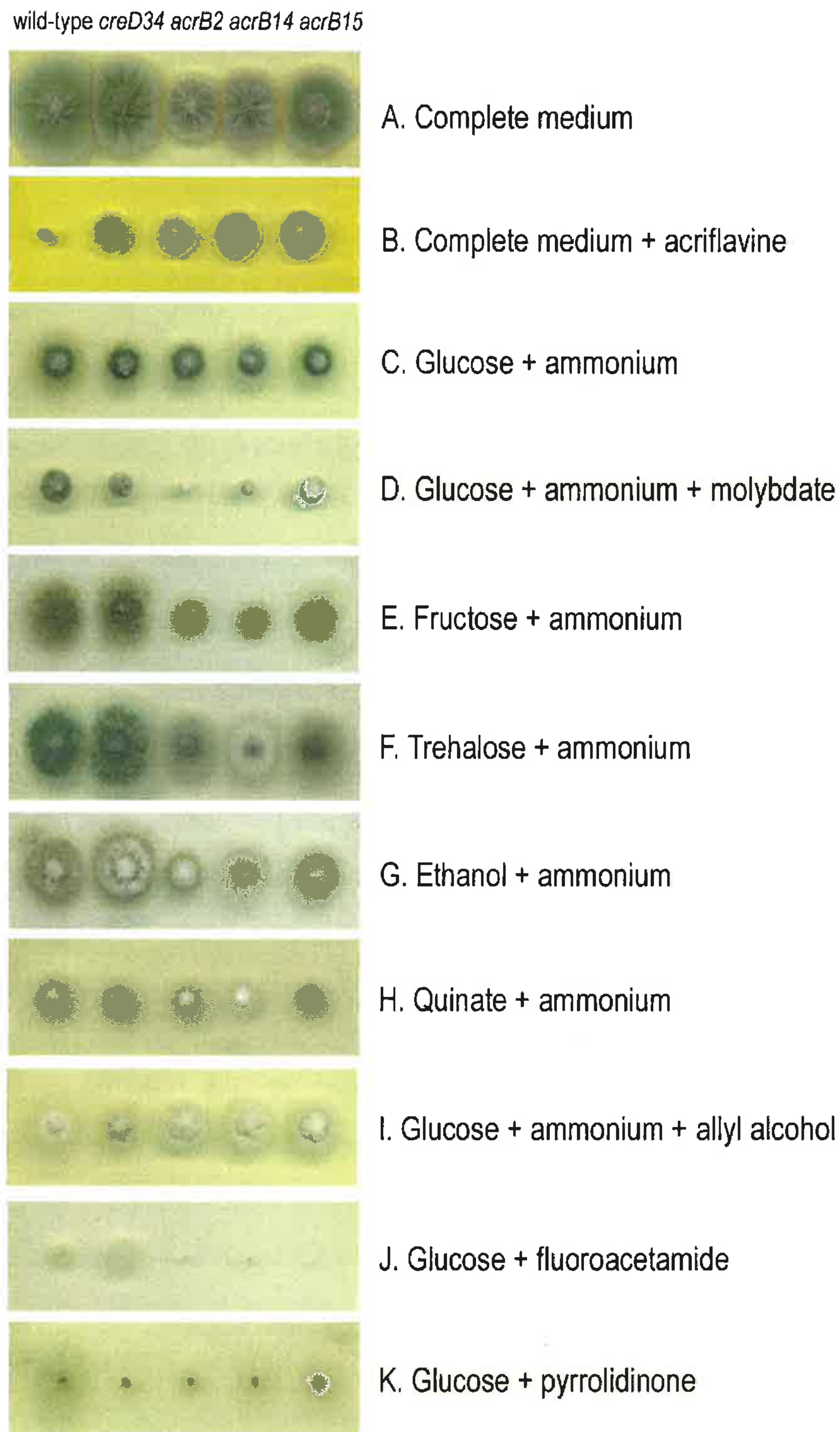

Figure 4.1 Phenotypic heterogeneity among *acrB* mutant alleles. Strains containing *acrB* mutant alleles, *acrB2* (strain *acrB2G*), *acrB14* and *acrB15*, were grown at 37°C for 2-3 days, along with a strain containing *creD34* and a wild-type strain on the following media: (A) 1% Complete medium; (B) Complete medium plus 0.001% acriflavine; (C) 1% D-glucose plus 10 mM ammonium L-tartrate; (D) 1% D-fructose plus 10 mM ammonium L-tartrate; (E) 1% D-glucose plus 10 mM ammonium L-tartrate plus 11 mM molybdate; (F) 1% D-trehalose plus 10 mM ammonium L-tartrate; (G) 0.5% ethanol plus 10mM ammonium L-tartrate; (H) 50 mM quinate plus 10 mM ammonium L-tartrate; (I) 1% D-glucose plus 10 mM ammonium L-tartrate plus 10mM allyl alcohol; (J) 1% D-glucose plus 10 mg/ml fluoroacetamide; (K) 1% D-glucose plus 10 mM 2-pyrrolidinone.

Table 4.1 Phenotypic analysis of *acrB* mutant strains.

Growth condition	Wild-type	*creD34*	*acrB2G*	*acrB14*	*acrB15*
Complete	10	10	9	9	9
Glucose	10	10	10	10	10
Cellobiose	10	10	5	5	5
Fructose	10	10	4	3	4
Galactose	8	8	4	3	4
Lactose	8	8	4	4	4
Maltose	10	10	6	5	6
Raffinose	10	10	6	6	6
Starch	10	10	8	8	8
Sucrose	10	10	7	7	7
Trehalose	10	10	6	5	6
Xylose	10	10	7	7	7
Acetate	9	10	6	5	8
Ethanol	10	10	5	4	5
Glucuronate	8	8	7	7	7
Glycerol	10	10	9	9	9
Succinate	3	3	3	3	2
Quinate	10	10	9	9	9
Acriflavine	1	8	9	10	10
Molybdate	10	5	2	3	4
Fluoroacetamide	3	5	0	0	0
2.5 mM allyl alcohol	10	10	10	10	10
10 mM allyl alcohol	9	9	10	10	10
β-alanine	2	1	2	1	2
GABA	8	8	2	2	2
Pyrrolidinone	8	8	2	3	1
Glutamate	8	8	4	4	4
Proline	8	8	4	4	5
Threonine	2	2	2	2	2
β-alanine + glucose	10	5	6	6	6
GABA + glucose	10	8	5	5	5
Pyrrolidinone + glucose	10	1	3	5	7
Glutamate + glucose	10	10	9	9	9
Proline + glucose	10	10	10	10	10
Threonine + glucose	10	4	10	7	10
Glucose + acetamide	10	9	10	9	10
Acetamide	5	5	4	4	4
Acetamide + $[NH4]_2T$	5	5	4	4	4
Glucose + acrylamide	1	1	1	1	1

Table 4.1 Legend. For full genotypes, see Materials and Methods, Table 2.2. Plates were incubated at 37°C for two or three days. Carbon sources were added to carbon free medium containing 10 mM ammonium L-tartrate at 1.2% (acetate), 1% (D-cellobiose, D-fructose, D-galactose, D-glucose, glycerol, lactose, D-maltose, D-raffinose, starch, succinate, sucrose, D-trehalose, and D-xylose), 0.5% (D-glucuronate and ethanol) or 50 mM ((-)-quinate). Molybdate was added to synthetic complete medium at 11 mM. Fluoroacetamide was added to 1% D-glucose medium at 10 mg/ml. Allyl alcohol was added to synthetic complete medium at 2.5 mM or 10 mM. Acetamide, β-alanine, GABA (γ-amino-n-butyric acid), L-glutamate, L-proline, 2-pyrrolidinone, and L-threonine were added at 10 mM to 1% glucose medium, and at 50 mM when added to medium without another carbon source. Ammonium L-tartrate and acrylamide were added at 10 mM. Growth is scored on a numerical scale with 10 the strongest growth, and 1 the weakest. Scores are not directly comparable between plates.

on various sugars may indicate a failure of uptake, or a failure to express genes encoding enzymes required for their utilization. The ability to utilize other alternative carbon sources such as quinate and glycerol was only slightly affected by the *acrB* mutations, with the exception of ethanol where the *acrB* mutants grew considerably less well than wild-type (Figure 4.1G-H; Table 4.1).

Carbon catabolite repression of *alcA* and *amdS* was assessed using plate assays (Figure 4.1I-J). Test medium containing 2.5 mM or 10 mM allyl alcohol in the presence of glucose was used to investigate the expression of the alcohol dehydrogenase encoded by *alcA*. The *acrB* mutant strains appeared to be slightly more resistant than the wild-type and *creD34* strains, indicating even tighter regulation of carbon catabolite repression of *alcA* in these mutant strains compared to wild-type. The *acrB* mutant strains were hypersensitive on medium containing glucose plus 10 mg/ml fluoroacetamide, indicating an increased level of the acetamidase expressed from the *amdS* gene compared to the wild-type strain, which is the opposite of the phenotype conferred by the *creD34* mutation. The lack of growth of the *acrB* mutant strains on medium containing glucose plus 10 mM acrylamide as a nitrogen source (Table 4.1) indicates that this increased level of *amdS* expression is not due to constitutive expression, as acrylamide is a substrate of the acetamidase but not an inducer of *amdS* (Hynes and Pateman, 1970). Further, wild-type growth of the *acrB* mutant strains on acetamide as either a carbon, nitrogen or carbon and nitrogen source indicates that the *amdS* gene is fully inducible (Table 4.1). Thus there is a defect in the glucose-mediated repression of *amdS* but no other effect on expression.

Growth on media containing the ω-amino acids 10 mM β-alanine, 10 mM GABA, or 10 mM pyrrolidinone as nitrogen sources was significantly decreased in the *acrB* mutant strains compared to the wild-type strain, as was the case for a *creD34* mutant strain (Table 4.1). Similarly, growth of the mutant *acrB* strains on media containing 50 mM GABA and 50 mM pyrrolidinone acting as both a carbon and nitrogen source was also decreased compared to wild-type and the *creD34* mutant strain, but an effect was not observed on medium containing 50 mM β-alanine. Interestingly, the *acrB* mutant strains varied in their growth on medium containing 10 mM pyrrolidinone as a nitrogen source, with the *acrB15* strain having stronger growth, compared to the intermediate growth of the *acrB14* strain, and the weaker growth of the *acrB2* strain (Figure 4.1K; Table 4.1). This heterogeneity between alleles indicates that not all three alleles result in complete loss of function.

Growth on media containing other amino acids as nitrogen sources, such as 10 mM proline and 10 mM glutamate, was unaffected by the *acrB* mutations. However, decreased utilization of 50 mM proline and 50 mM glutamate when they were used as both carbon and

nitrogen sources was observed for the *acrB* mutant strains in comparison to the wild-type strain (Table 4.1).

4.3 The *acrB2* mutation suppresses aspects of the *creA*, *creB* and *creC* mutant phenotypes

The *creD34* mutation was identified as a suppressor of the hypersensitivity to fluoroacetamide conferred by the *creC27* mutation. The effects of the *creD34* mutation are pleiotropic in that it also leads to suppression of the effects of *creC27* on other enzymes subject to carbon catabolite repression, such as alcohol dehydrogenase I, but it does not suppress the effects of *creC27* that are apparent in derepressing conditions, such as the poor growth on D-quinate medium (Kelly and Hynes, 1977; Kelly, 1980). The *creD34* mutation not only suppresses some of the phenotypes conferred by the *creC27* mutation, it also suppresses some of the phenotypic effects of the *creB15* mutation, and, weakly, of the *creA204* mutation (Kelly and Hynes, 1977; Kelly, 1980). These similarities between the phenotypes of *creD* and *acrB* mutant strains led us to construct double mutant strains to test whether the effects of mutation in *acrB* are also epistatic to some of the phenotypes conferred by *creB* and *creC* mutations.

The presence of the *acrB2* mutation reversed the sensitivity conferred by the *creA204*, *creB1937* and *creC27* mutations to the presence of acriflavine in complete medium (Figure 4.2B). The *creD34 acrB2* double mutant strain was even more resistant to the presence of acriflavine than either of the individual mutant strains, indicating an additive effect. The resistance of the *creB1937* and *creC27* strains to the presence of molybdate was reversed in the relevant *acrB2* double mutant strains, and to a greater degree than in the *creD34* double mutant strains (Figure 4.2D).

The *acrB2* mutation results in reduced growth compared to wild-type on sole carbon sources such as fructose, maltose, starch, cellobiose, ethanol and acetate (Figure 4.2E-J). This reduced growth is also seen in the *acrB2creB1937* and *acrB2creC27* double mutants strains compared to the single *creB1937* and *creC27* mutant phenotypes respectively. The *creD34* mutation had no effect on growth of mutant strains on different sugars, and neither *creD34* or *acrB2* mutations influenced the growth of *creA204* mutants on media containing alternate sole carbon sources (Figure 4.2E-J).

The inability of the *creB1937* and *creC27* mutant strains to utilize D-quinate as an alternative carbon source was partially suppressed by the *acrB2* mutation in the double mutant strains, an effect not observed in the equivalent *creD34* double mutant strains (Figure 4.2K).

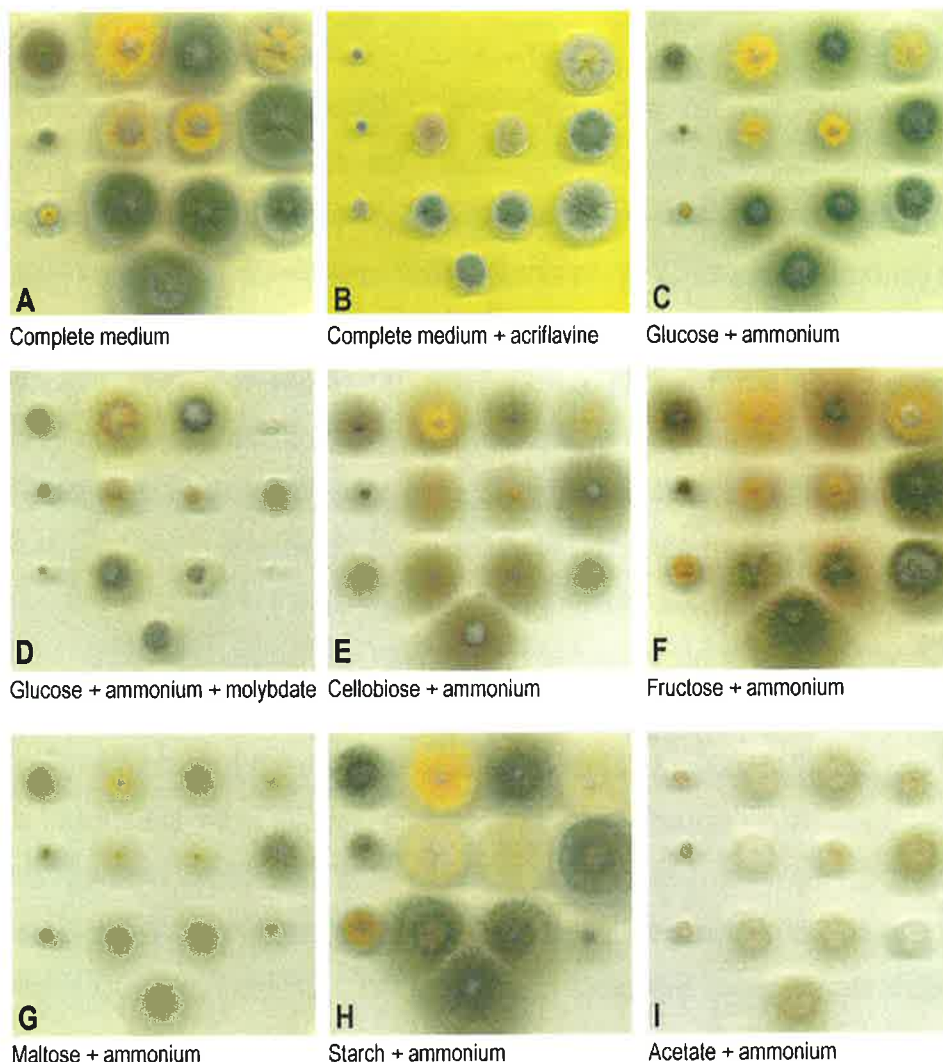

Figure 4.2 Epistatic effects of the *acrB*2 mutation on *creA*, *creB* and *creC* mutant phenotypes. Strains indicated in (R) were grown at 37°C for 2-3 days on the following media: (A) Complete medium; (B) Complete medium plus 0.001% acriflavine; (C) 1% D-glucose plus 10 mM ammonium L-tartrate; (D) 1% D-glucose plus 10 mM ammonium L-tartrate plus 33 mM molybdate; (E) 1% D-cellobiose plus 10 mM ammonium L-tartrate; (F) 1% D-fructose plus 10 mM ammonium L-tartrate; (G) 1% D-maltose plus 10 mM ammonium L-tartrate; (H) 1% D-starch plus 10 mM ammonium L-tartrate; (I) 1.2% acetate plus 10 mM ammonium L-tartrate.

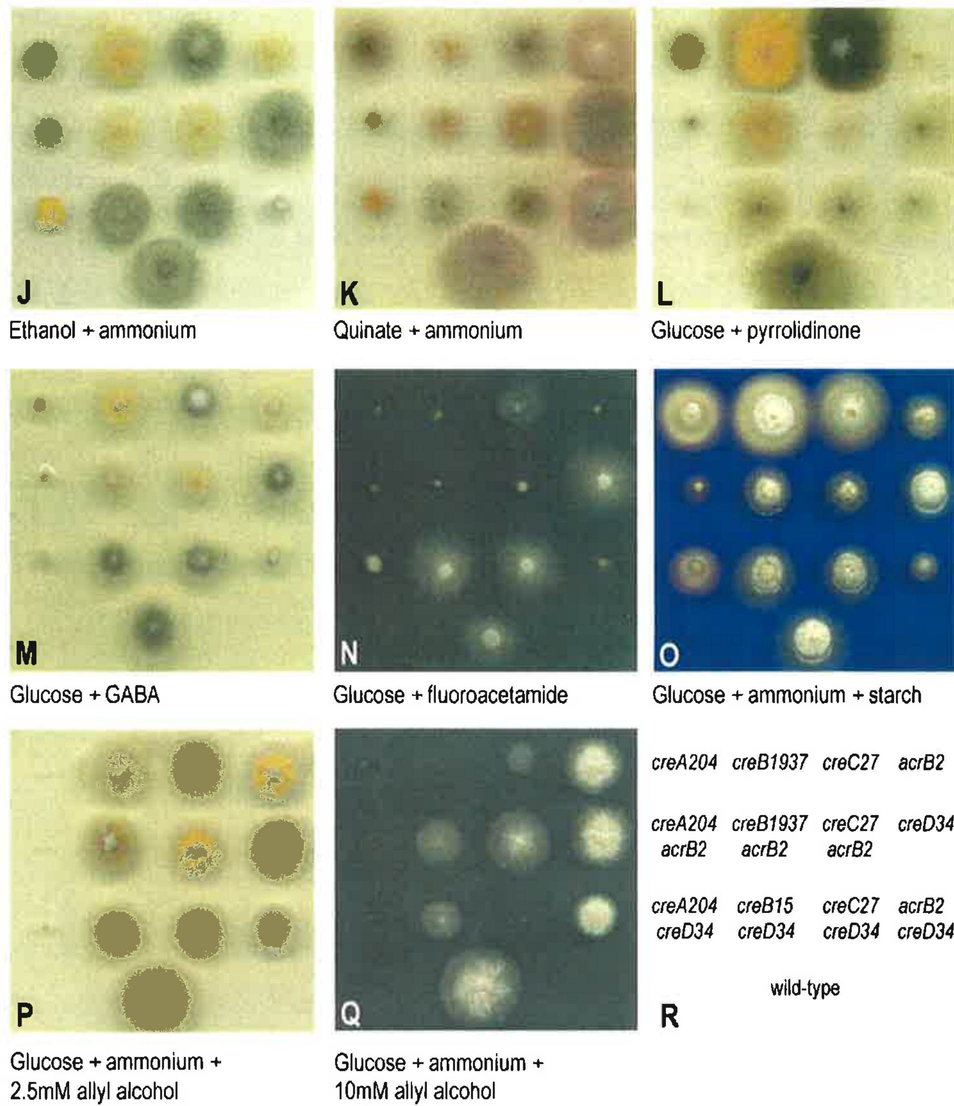

Figure 4.2 Epistatic effects of the *acrB2* mutation on *creA*, *creB* and *creC* mutant phenotypes (continued)... (J) 0.5% ethanol plus 10 mM ammonium L-tartrate; (K) 50 mM (-)-quinate plus 10 mM ammonium L-tartrate; (L) 1% D-glucose plus 10 mM 2-pyrrolidinone; (M) 1% D-glucose plus 10 mM GABA; (N) 1% D-glucose plus 10 mg/ml fluoroacetamide; (O) 1% D-glucose plus 1% D-starch plus 10 mM ammonium L-tartrate stained with iodine; (P) 1% D-glucose plus 10 mM ammonium L-tartrate plus 2.5 mM allyl alcohol; (Q) 1% D-glucose plus 10 mM ammonium L-tartrate plus 10 mM allyl alcohol; (R) Key to strain placement.

creB and *creC* mutant strains grow even more strongly than wild-type on medium containing 10 mM pyrrolidinone as the nitrogen source, but in double mutant strains with *acrB2* this phenotype is reversed. This is also seen to a lesser degree on glucose medium containing 10 mM GABA as the nitrogen source (Figure 4.2L-M).

Carbon catabolite repression of *alcA*, *amdS* and the (secreted) amylase genes were assessed using plate assays (Figure 4.2N-Q). The *acrB* mutant strains were as hypersensitive on medium containing glucose plus 10 mg/ml fluoroacetamide as the *creA204*, *creB1937* and *creC27* strains, as were the corresponding double mutants. Note that the *creD34* mutation does suppress the phenotype of the *creB15* and *creC27* mutant strains, highlighting again the differences between the *acrB2* and *creD34* mutant phenotypes.

Secreted amylase can be detected on a starch plate stained by iodine, with clearing of the blue colour indicating the starch has been broken down. The *creA204*, *creB15* and *creC27* strains all inappropriately secrete amylase in the presence of glucose, which can be detected by the size of the cleared halos, and the *acrB2* mutation suppressed this phenotype (Figure 4.2O).

The *acrB2* mutation reversed the toxic effects of allyl alcohol added to D-glucose medium in *creC27* and *creB1937* backgrounds, the same phenotypic suppression previously seen for the *creD34* mutation except slightly stronger. The *acrB2* mutation did not suppress the *creA204* mutant phenotype of sensitivity to allyl alcohol in the presence of glucose.

4.4 Molecular analysis of *acrB*

The close genetic linkage of *acrB* and *creB* (Clutterbuck, 1997) was taken advantage of to clone the *acrB* gene. The BAC clone which contains *creB*, BAC 4B1 (Lockington and Kelly, 2001), was tested for complementation of the *acrB2* mutation in a transformation assay testing for the reversal of the resistance to acriflavine in complete medium conferred by the *acrB2* mutation. Both the BAC 4B1 and the subclone p4AcrB, containing a 6 kb *ClaI* fragment within the BAC 4B1 region, were found to complement the *acrB2* mutation (J. Kelly and R. Lockington, *unpubl. data*). Thus *acrB2* was located in the middle of the 6 kb insert of p4AcrB, and this plasmid provided the starting point for the molecular analyses of the *acrB* gene presented here.

4.4.1 Sequence analysis of *acrB*

Since this work was undertaken before the *A. nidulans* genome sequence was publicly available, the 6 kb insert of p4AcrB was sequenced using custom oligonucleotides (see Table 2.3), and AcrB is encoded within a 3111 bp sequence of the p4AcrB plasmid (Genbank Accession No. AF485329, Fig. 4.3). The *acrB* gene contained a single intron whose position was

```
        CTTGACAGTAGCTGCCTGTAATGTCAGGCCATAAAGTTTCTGCCTTAGGCAATTTTGAGGATGTTTCCAATCTTTAGGTCGTTTC  -349
        GTCTTTTTCGAACCAGGTCGTCTCTTCCTCTCCTTCTTCCCGCCTCTACACTCCCCCACGCCACGTCTTATCCCCTTCACTCTCCCC  -262
                                                          #
        CTCTATCTTATTATTCGAATTCTTCCCCTCTTTGCTCTGTGTATTCTCATATCTACTTGTTTCTTACGTTTCTCCCTTTTTGTCATT  -175
        CATTCATTTTAACTGCAGAGACCCTGGCCTCGAATCCAGTCTTTTCCAAACTCCCTTCCGCCCTTTCCCGCGAACTTTTTTTGGAAC  -88
        TGCCATCGACTTCATTTCCACCTGGTCCAAGTCCTTGGCTGCTATCCGCCAAGCCGACTCGGACCTGGGTGTTGGGTTCCCTCGTCC  -1
        ATGCCACGTTCATCGGCTACGGCTAGGAAGAGTCATAGCAATCGACAAGACCATGGCGGCGGCGGTTCTGGGAAGAAGCCAAGCAAA  87
1       M  P  R  S  S  A  T  A  R  K  S  H  S  N  R  Q  D  H  G  G  G  G  S  G  K  K  P  S  K
        CAAAAGTCAAGTGGGCATTTAAACGCAACATACAACGGTACTGCTGGATCCGAAACTGGACCGTCTTCGCAGGTCGACTGGCCATCG  174
30      Q  K  S  S  G  H  L  N  A  T  Y  N  G  T  A  G  S  E  T  G  P  S  S  Q  V  D  W  P  S
        CATCGTTCTGGAGACCAATCGATCGCAGCAGCCGCAGCCAAATCTAACGGACCGGTCGATAGCTTGAAAGCGGACACCAACGGACGT  261
59      H  R  S  G  D  Q  S  I  A  A  A  A  A  K  S  N  G  P  V  D  S  L  K  A  D  T  N  G  R
        GGTTATCCGGGCGGATATGCGAAGGGAAATGCAGACATGTCTTACGGGCAGACGAATGGTGGCGTGTCGCCGAATGGTGGACTCGCC  348
88      G  Y  P  G  G  Y  A  K  G  N  A  D  M  S  Y  G  Q  T  N  G  G  V  S  P  N  G  G  L  A
        GGGCCGGCTTCACGTCGTACGGATAAGTCGGTCACTGGGACCAAGAGGACAACTTCGAATGCGTCGGTGAATCCGTTCCAGCTGGCA  435
117     G  P  A  S  R  R  T  D  K  S  V  T  G  T  K  R  T  T  S  N  A  S  V  N  P  F  Q  L  A
        TCCACCATTCTCCGATCGTGTCCAATGTACGACACTATCGCCATCTTGATCTTTCTGCTTCAGCTCCCGCCTATGGTTCTCACTTTG  522
146     S  T  I  L  R  S  C  P  M  Y  D  T  I  A  I  L  I  F  L  L  Q  L  P  P  M  V  L  T  L
        GTTCAATTCTTGTTTGCGTCCTTGACATTTATGCCTCCCAGCGGCACCGCTTCTGGATCCTTCACCTCCAACTTCGATATTTTCCAG  609
175     V  Q  F  L  F  A  S  L  T  F  M  P  P  S  G  T  A  S  G  S  F  T  S  N  F  D  I  F  Q
                        (T+T)=acrB14
        GGACCCGCCGGAACCCCCTCGCTCGGTACCATGATTGCAATGGATGGTTTCTGCCTGCTTGTATGGGGCCTCTTTATGTGGACGTGG  696
204     G  P  A  G  T  P  S  L  G  T  M  I  A  M  D  G  F  C  L  L  V  W  G  L  F  M  W  T  W
        GCCCAGAATTTTGCTCTCGATTTAGCCCATGTCCAGGTTGCCATCACTCTGGGCGGTGGAGGTGCAGGGAAAAATGGTGGTGTCAAT  783
233     A  Q  N  F  A  L  D  L  A  H  V  Q  V  A  I  T  L  G  G  G  G  A  G  K  N  G  G  V  N
        GCGCTCTGCGTCGGTATTGTTCTGATTCTGCATCTCATACGCAGCAAAGGAATACAGGATTTTGTCGTCGGCCATCTTGTTTCAGCA  870
262     A  L  C  V  G  I  V  L  I  L  H  L  I  R  S  K  G  I  Q  D  F  V  V  G  H  L  V  S  A
        AAAATCATTAGCCCCGATTTACTGTCGCATTATTCTTACCTCATGCCCGCCGAATTCAAGCGCACCGAATCGCAATCATCCCCGAGT  957
291     K  I  I  S  P  D  L  L  S  H  Y  S  Y  L  M  P  A  E  F  K  R  T  E  S  Q  S  S  P  S
        TGGATCCGGAGCCTGCTTGCTGTACATATTCTGGCCCAGGCGGGTACTGCGATGGCGAGGCGATCGATGACTAAAAATAGGACCCCG  1044
320     W  I  R  S  L  L  A  V  H  I  L  A  Q  A  G  T  A  M  A  R  R  S  M  T  K  N  R  T  P
        GCCCCATCACGATCAGGCAAACGCGTGGATACAGAAGCGTCTGCCGGCTCACAAACCCAGATCGACTCGGCGTTCGAATCCGCGGCC  1131
349     A  P  S  R  S  G  K  R  V  D  T  E  A  S  A  G  S  Q  T  Q  I  D  S  A  F  E  S  A  A
        AGCGTTTCTTCCTATCTAGGCCCCGACGGGCAGATTATCACTGCCGCGCATAAGGACGGCAGGGATCGTTTGATATCGGCAAAAAAA  1218
378     S  V  S  S  Y  L  G  P  D  G  Q  I  I  T  A  A  H  K  D  G  R  D  R  L  I  S  A  K  K
        CGACGAAGGCAGGCAAATCAAGTCAGGAGCCGGCAACCTTTTTGGGCTGCACTGGCAAGCACAAAAGTCACGGTCATGAGGGAGTAT  1305
407     R  R  R  Q  A  N  Q  V  R  S  R  Q  P  F  W  A  A  L  A  S  T  K  V  T  V  M  R  E  Y
        GAACATTCTAGGGCCTTGTCAAAAACTGCTAGAGGACTTGCTACGACGGAGGACGATCTTCAAGGCGTTTCTTTGGACGATGGACTT  1392
436     E  H  S  R  A  L  S  K  T  A  R  G  L  A  T  T  E  D  D  L  Q  G  V  S  L  D  D  G  L
        GTTTGGATTACGTATGTGGATAGCTCGACGATTAAGTTTGCAGCTGGGGATTTTGCGTCTTCGGACGACCATTCCGCGTCAGGTGTC  1479
465     V  W  I  T  Y  V  D  S  S  T  I  K  F  A  A  G  D  F  A  S  S  D  D  H  S  A  S  G  V
        TGCGAAGCAGGCCGTGTGAGCAGCGAGGATGCGGAGCCGTTTTACGTCTGCGTCAATGGTGCGCCATGGGCAACGGTGGTCATCACT  1566
494     C  E  A  G  R  V  S  S  E  D  A  E  P  F  Y  V  C  V  N  G  A  P  W  A  T  V  V  I  T
        AAAGAGCATGATCCTTCAAAAGCGTCTAATACAATCTATTGGCGAGGCGAGATATCAGGTCTTGCACCCAATTGCGCGTATACTTGC  1653
523     K  E  H  D  P  S  K  A  S  N  T  I  Y  W  R  G  E  I  S  G  L  A  P  N  C  A  Y  T  C
                                                                                  (_)=acrB2
        TCTTTTGTTAAATGCGATACGGATGAGGAAATCTGCGCCATGAGTGTCAAGACCCCTGCGGCCAATGATGCAGAACAAGgtaagata  1740
552     S  F  V  K  C  D  T  D  E  E  I  C  A  M  S  V  K  T  P  A  A  N  D  A  E  Q  A
                                                                              (__)=acrB15
        ttcgttctatggtttccgattcgtctcgatgtggtctaatttgtgccttccatagCCAATTCGGTGCCGGCCCCTCCGCAACCCTCA  1827
578                                                                N  S  V  P  A  P  P  Q  P  S
        TATCGACCATCCTCCCCAACAACCACGCTGAAGAACTCGATCATCAATGCTGAGGCGAAACTGAACGAAAAGCGTGCTCGACTCCGA  1914
589     Y  R  P  S  S  P  T  T  T  L  K  N  S  I  I  N  A  E  A  K  L  N  E  K  R  A  R  L  R
        AAGGCCAAAAATGACCACAAGCTTGCTATTTCTAAGATCAAGAAGGAGCTGGACAATTACACCAATCGTCTTCAGAGCGGCACGGAT  2001
618     K  A  K  N  D  H  K  L  A  I  S  K  I  K  K  E  L  D  N  Y  T  N  R  L  Q  S  G  T  D
        GAAAACAGGCAGAAGCAACGCTCTCTTCAATTGGAAAGGAACATTCGACAAACTGAAGAGGCTACCGCCGCTCTGGACAACCAGATC  2088
647     E  N  R  Q  K  Q  R  S  L  Q  L  E  R  N  I  R  Q  T  E  E  A  T  A  A  L  D  N  Q  I
        GATAACTTGGGTAATGTTCCTGACGATGAGTATCAGGAGTGGGTTGAACAGAAGGCAAAGTACGAACGTGAATTGGAGCTCCTCAAA  2175
676     D  N  L  G  N  V  P  D  D  E  Y  Q  E  W  V  E  Q  K  A  K  Y  E  R  E  L  E  L  L  K
        TCCGCCAAGGCAGAGATTGCTGCCACGCGTACCGCCAATGCTCGCGAGTTATCTTCATTGGAATCCGAGTTGAACTCTACCACGCAA  2262
705     S  A  K  A  E  I  A  A  T  R  T  A  N  A  R  E  L  S  S  L  E  S  E  L  N  S  T  T  Q
        CGGCGCGAACGTCTGCAGGGTCGCCGAACCAGAGTGAATGAGCAGTACGAACGGATCATCTCGGCCAACGCACAGGGTCTCAATGAG  2349
734     R  R  E  R  L  Q  G  R  R  T  R  V  N  E  Q  Y  E  R  I  I  S  A  N  A  Q  G  L  N  E
        CGAGAGCGCCGCGCTGCAGAGCAGTTTGCCCGGGAACAAGATCAGTCGAAGTTGGAGCAAAGTTTCAACGAACAATTCGCGAGCATC  2436
763     R  E  R  R  A  A  E  Q  F  A  R  E  Q  D  Q  S  K  L  E  Q  S  F  N  E  Q  F  A  S  I
        AGTCAATCAGTGCAGGATTATCAGCTGCGCACCAGTCAATTGTGGCAACAGTGTACCGCCGTCGAACAAGCCCTCCAGCAGCAGTTG  2523
792     S  Q  S  V  Q  D  Y  Q  L  R  T  S  Q  L  W  Q  Q  C  T  A  V  E  Q  A  L  Q  Q  Q  L
        CTCATGGAGCCCGCTCCGCTAACACCCGAAGGCGAGCTGCCCGGTACTAGTACGTTTGCCGACGCGCCCAGCGTGCCCTTGGGCACA  2610
821     L  M  E  P  A  P  L  T  P  E  G  E  L  P  G  T  S  T  F  A  D  A  P  S  V  P  L  G  T
        TTGGCGTCAAATATGCCAAGCCACCGCTCGCTACTAGGACAGAGCTTTCCGCCGCTCAAGTCTAGTCCTCTGCAGCACTATGCTTCG  2697
850     L  A  S  N  M  P  S  H  R  S  L  L  G  Q  S  F  P  P  L  K  S  S  P  L  Q  H  Y  A  S
        CCAATTGGAACTGCTCCGTCTCATCCGACTAGTCCAATCGCCGCTCCATCCTACCAGCCTTTCTCCAGCTCGCCATTTGGTAACGCG  2784
879     P  I  G  T  A  P  S  H  P  T  S  P  I  A  A  P  S  Y  Q  P  F  S  S  S  P  F  G  N  A
        GCATCCTTCCTTGACCCGGACTTTGTCTACCGCGACCGTTCGTTCTCCAACCGCTCCGCACGCAGCAGCCTCTATGGCTCTGAGTTC  2871
908     A  S  F  L  D  P  D  F  V  Y  R  D  R  S  F  S  N  R  S  A  R  S  S  L  Y  G  S  E  F
        CCGGACGCGATAACGGCCCGCCGTGTCCCCTTTGGCGTTGATCCTTTCGAGCTCGGTAACGAGAAACGACGCGGTTCGGGATCTGAC  2958
937     P  D  A  I  T  A  R  R  V  P  F  G  V  D  P  F  E  L  G  N  E  K  R  R  G  S  G  S  D
        AGCACCCCGCTCAACGGTCCATCTGGCCTACGTCCTATCTCCAGTCCTTTCCAGCGAGCCGCCAGTCGCGCAAGTGGAACCGGCAGT  3045
966     S  T  P  L  N  G  P  S  G  L  R  P  I  S  S  P  F  Q  R  A  A  S  R  A  S  G  T  G  S
        GGCGGAAGCGGAGGAAGTGGTAGTGGTAGTGGCAGTCCTAGCTCTGCCCGTGGAAAGGGAAACTAAACGGAACTGCTTGCCAGCCAG  3132
995     G  G  S  G  G  S  G  S  G  S  G  S  P  S  S  A  R  G  K  G  N  *
        CCAACTCAGCCGTCTGCAGTGACGACAAAAGCATTGGAGATTGATCCTTACTCGCCTGCCCATATGTCCATATTTGGTCCTCGGACA  3219
        TGGCTCGGACTGTCCGTGTAGTGACTGGTTATATAGTTTGCCGTCGCCAACGGTCTATAGTATAGATCCTTGTGCTCGCCGCGTAGA  3306
        GCCACATCTTATATTTTGGCGAAGAAAAAGAG                                                         3338
```

Figure 4.3 Location of the *acrB* mutations within the *acrB* gene. The complete sequence of the *acrB* genomic region is shown (GenBank database Accession No. AF485329). The single intron of 63 bp is depicted in lower case letters and the 5' and 3' splice junction and larial sequences are underlined. A putative TATA box and a putative CCAAT box are underlined and in bold letters. The major start point of transcription is indicated above the DNA by the # symbol. Bold italics indicate single theoretical CreA and AreA binding sites in the 5' untranslated regions. The amino acid sequence is shown in single letter code beneath the DNA sequence, with numbered amino acid residues indicated on the left hand side, and numbered nucleotide positions indicated on the right hand side, with the first nucleotide of the start codon as +1. The position and nature of the DNA sequence changes in the *acrB2*, *acrB14* and *acrB15* alleles are shown, with (_) indicating a deletion of the base directly below, and (+) indicating an addition of a base at the position below

identified by sequence analysis of PCR amplified cDNAs, and the intronic sequence conforms to the consensus 5' and 3' junction and lariat sequences for fungal introns (Gurr *et al.*, 1987). To determine the transcription start point (tsp) of the gene, 5'RACE (rapid amplification of cDNA ends) of RNA purified from D-glucose grown mycelium was carried out using nested primers. Direct sequencing of the nested PCR product and of several subclones of this product mapped a single major tsp –221 bp from the translation start (Figure 4.3). A TATA-like sequence is located 33 bp upstream of the tsp, and a CCAAT box 155bp upstream of the tsp. The *acrB* gene is not present in the *A. nidulans* EST database (containing EST information for cDNA clones from a mixed vegetative and 24hr asexual development culture of *A. nidulans* strains FGSCA26 constructed in lambda Zap by Dr. R. Aramayo at Texas A and M University (USA) http://www.genome.ou.edu/asperg.html) indicating a low level of expression under these conditions.

The *acrB* gene encodes a 1015 amino acid polypeptide. AcrB is predicted to contain three trans-membrane domains near the N terminus, at amino acid residues 158-180 (I), 214-236 (II) and 257-274 (III), and a coiled-coil region, at amino acid residues 596-753 (SMART protein motif analysis program, Schultz *et al.*, 1998) (Figure 4.4). The trans-membrane domain prediction program, TMHMM2 (Krogh *et al.*, 2001), predicts the most probable location and orientation of trans-membrane helices from sequence data, and this program indicated that the N-terminus of the AcrB protein, and the region between the second and third trans-membrane domains are most likely cytoplasmic, with the C-terminal region of the protein predicted to be external.

There is a hypothetical protein in contig 1168 of the TIGR *Aspergillus fumigatus* Genome Database that shows 76% identity with the *A. nidulans* sequence, but these are very closely related ascomycete species. Database searches (SwissProt/SpTrEMBL/PDB) using the BlastP program (Altschul *et al.*, 1990) revealed that the AcrB protein was not highly similar to any protein in these databases. AcrB shows most sequence similarity with hypothetical proteins, one from *N. crassa*, Q9C2R3 (31% identity over 1015 amino acids) and one from *S. cerevisiae*, YG2K (24% identity over 861 amino acids). The hypothetical *N. crassa* protein contains three putative trans-membrane domains and a coiled-coil domain, and although it shows low similarity along the length of the protein with AcrB, specific regions such as the trans-membrane domains share higher sequence identity (55% identity over 24 amino acids). The *S. cerevisiae* hypothetical protein has a similar protein structure to AcrB, with three trans-membrane domains and a coiled-coil region, but the sequence similarity is uniformly low throughout the protein.

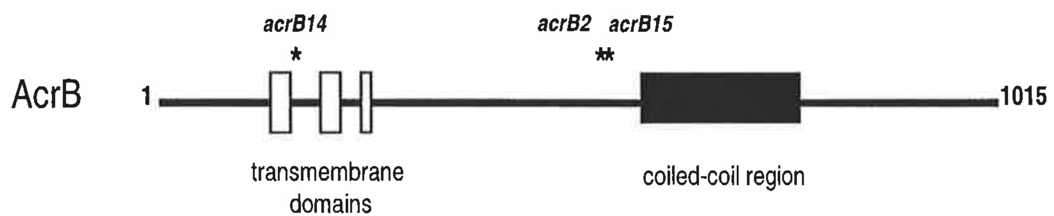

Figure 4.4 Schematic of AcrB. The open boxes indicate the three transmembrane domains, and the filled box indicates the coiled-coil region of AcrB. Asterisks mark the location of the relevant *acrB* mutations.

4.4.2 Molecular analysis of the *acrB* mutant alleles

Three mutant alleles of *acrB* were sequenced in order to identify functional regions of the AcrB protein and to show that *acrB* rather than a suppressor of *acrB2* had been isolated. We used a PCR approach using primers spanning the *acrB* gene followed by direct sequencing of the PCR products (see Table 2.3). Direct sequencing of *acrB* PCR products from wild-type DNA was undertaken for comparison. The *acrB2* mutation is a single base pair deletion at nt +1734 at the 5' splice site of the only intron (Figures 4.3, 4.4). This would be predicted to disrupt the splicing of the intron resulting in a frameshift after amino acid 577, and the additional residues, GRYSFYGFRFVSMWSNLCLP, before terminating. The *acrB15* mutant allele contains a two base pair deletion at nt +1815, which results in a frameshift after amino acid 584, with an additional amino acid sequence of ATLISTILPNNHAEELDHQC before truncating. Both *acrB2* and *acrB15* gene products contain all three trans-membrane domains but lack the coiled-coil region. The *acrB14* allele contains a C to T transition plus a one base pair insertion (T) at nt +627 that results in a frameshift after amino acid 209 and truncates 84 amino acids later. The additional amino acids are FARYHDCNGWFLPACMGPLYVDVGPEFCSRFSPCPGCHHSGRWRCREKWW CQCALRRYCSDSASHTQQRNTGFCRRPSCFSKNH. The *acrB14* mutant gene product lacks two of the three trans-membrane domains and the coiled-coil region and, since it is also recessive to the wild-type allele, probably represents a null allele.

4.5 Protein interactions involving AcrB

4.5.1 Screening for proteins that interact with AcrB

The Stratagene BacterioMatch I Bacterial-2-hybrid system (B2H) was used to identify proteins that interact with AcrB. The coding region of *acrB* cDNA that incorporated the coiled-coil region of AcrB (aa 582 – 1015) was ligated into the bait vector to produce the pBTccAcrB construct. This vector was shown not to auto-activate and then used to screen an *A. nidulans* cDNA library in the search for protein-protein interactions. There were fifty-five carbenicillin resistant colonies that indicated an interaction may be occurring, but upon validation all proved to be false positives. The *A. nidulans* cDNA library used in this screen was created by digesting a previously made yeast-2-hybrid cDNA library with *EcoRI* and *XhoI*, and subcloning these fragments into the target vector of the B2H system (R. Lockington, *unpubl. data*). Sequencing of some of the false positives indicated that the pTRG-Library plasmids often had multiple partial inserts that were out of frame. The availability of this cDNA library prompted its initial use, but future work would require the construction of a B2H specific *A. nidulans* cDNA library.

4.5.2 Testing AcrB for interactions with other known proteins

To investigate whether AcrB homo-dimerises, the coiled-coil region of *acrB* that was ligated into the bait vector (see Section 4.5.1) was also incorporated into the target vector to produce the construct pTRGccAcrB. No interaction was detected. The entire length of the *creB* gene had previously been ligated into the target vector to produce the construct pTRG-BFLAG (R. Lockington, *unpubl. data*). The pBTccAcrB vector was used to investigate a direct interaction with CreB (pTRG-BFLAG), CreD (pTRGcreD, see Section 3.6.1), and ApyA (pTRGapyA, see Section 3.6.1). The pTRGccAcrB vector was used to investigate a direct interaction with CreA (pBTcreA, see Section 3.6.1) and HulA (pBThulA, see Section 3.6.1). The AcrB protein did not interact with any of these known proteins in this system.

4.6 Discussion

Phenotypic analysis of the effects of three *acrB* alleles has shown that mutations in *acrB* have a broader range of phenotypes than the resistance to toxic compounds such as acriflavine, crystal violet, and malachite green previously described (Roper and Kafer, 1957) as they grow poorly on a range of sole carbon sources including sugars, starch and ethanol, and they also grow poorly on some ω-amino acids such as β-alanine, GABA and pyrrolidinone as either sole carbon and nitrogen sources, or as nitrogen sources in the presence of D-glucose as a carbon source. Genetic analysis has shown that the *acrB2* mutation is a suppressor of the phenotypes conferred by mutations in *creB* and *creC*. Strains containing the *creB1937* and *creC27* mutations show derepressed expression of the *alcA* gene, in that there is inappropriate expression in the presence of D-glucose, and the presence of the *acrB2* mutation completely reverses the effects of these mutations as judged by plate testing. The reduced growth of the *acrB2* containing strain on medium containing ethanol as a sole carbon source may indicate that this is due to a failure to induce *alcA* in either repressing or in derepressing conditions due to the loss of *acrB*, and that this failure in induction overrides the derepression caused by *creA*, *creB* and *creC* mutations. The *creB1937* and *creC27* mutations lead to poor growth on D-quinate and D-glucuronate as sole carbon sources, and the *acrB2* mutation partially repairs this defect. These phenotypes are consistent with a regulatory defect associated with the membrane and signalling, but not with a simple interpretation that only permeases and transporters are affected, as in the case of ethanol, a fat soluble compound, no permease is required or exists.

The amino acid sequence of AcrB indicates that the protein contains membrane spanning domains, indicating the probability of its residing in the membrane. There is also a coiled-coil region that would allow the prediction that the protein forms either a homo-dimer or hetero-dimer. However, using the BacterioMatch I bacterial-2-hybrid system to directly test

whether the coiled-coil region is responsible for AcrB forming a homodimer suggested that the coiled-coil region alone is not enough to form a homo-dimer. AcrB did not interact with CreA, CreB, CreC, CreD, ApyA or HulA in the BacterioMatch I bacterial-2-hybrid system, but it would be worthwhile validating these results in the newly released BacterioMatch II bacterial-2-hybrid system that is more sensitive and has a reduced background.

The striking aspect of the sequence is the lack of sequence similarity with proteins in databases from other eukaryotes, including those for which the entire genome has been sequenced, and it is clearly not a highly conserved homologue of any gene in a characterized signalling pathway. If functional homologues exist they are very diverged in sequence from AcrB and thus unrecognizable. The CreB and CreC proteins are involved in a deubiquitination network that removes ubiquitin from target proteins, and although these target proteins are not yet directly identified, the mutant phenotypes indicate that they are proteins involved in carbon metabolism and its regulation. Unlike AcrB, CreB and CreC are well conserved amongst most eukaryotes other than yeast. The phenotypic evidence suggests that AcrB is involved in an opposite process to deubiquitination, such as an ubiquitin ligase pathway, since a failure to add ubiquitin to substrates could suppress the phenotypic effects of mutations that affect the removal of ubiquitin moieties. Thus the ubiquitination/deubiquitination network in *A. nidulans* that involves CreB, CreC and AcrB provides an ideal genetic and molecular genetic system in which to unravel regulation of protein stability.

CHAPTER 5: IDENTIFICATION OF AN AMYLASE CLUSTER

5.1 Introduction

Alpha–amylases are enzymes involved in the breakdown of starch and maltose, and a number of α–amylase genes have been identified in Aspergilli species other than *A. nidulans*. Shroff *et al.* (1997) observed an over-expression of secreted α-amylase activity in the presence of glucose in *creA* mutant strains in comparison to the wild-type strain, as indicated by a clear halo on medium containing starch and glucose flooded with iodine. This suggested the presence of secretable α-amylase activity in *A. nidulans* that is carbon catabolite repressible. In *creA* mutant strains this α-amylase activity is derepressed, and inappropriately expressed at high levels.

In an effort to identify α-amylase genes in *A. nidulans*, the *A. oryzae* Amy2 protein sequence was used to identify sequences in the *A. nidulans* EST database that when translated shared a high amount of sequence identity to Amy2 (R. Murphy, *unpubl. data*). This led to the subsequent cloning and characterization of the α-amylase *amyB* (GenBank Accession No. AF208224). The EST database used to identify *amyB* was compiled by sequencing random clones from a cDNA library made from mycelia grown in the presence of 1% glucose (carbon repressing) and sodium nitrate as the nitrogen source (R. Aramayo, *pers. comm.*). That a number of *amyB* ESTs are present in this database implies that *amyB* is unlikely to be under the control of carbon catabolite repression. AmyB lacked the typical fungal signal peptide found in the first 21 aa of most other identified α-amylases, suggesting that AmyB is not secreted. Therefore the derepressed, secretable α-amylase activity found in the *creA* mutant strains was likely to be the result of another α-amylase in *A. nidulans*, and a second approach to identify α-amylase genes was adopted.

The α-amylase gene from *A. oryzae*, *amy2*, was used as a probe in a low stringency Southern of an *A. nidulans* λ gDNA library. One positive λ clone was identified, and designated λOAMY1, which contained a 23 kb genomic *SacI* fragment inserted into the *XhoI* site of λGEM11 (R. Murphy, *unpubl. data*).

This work preceded the public release of the *A. nidulans* genome database.

5.2 Sequence analysis of the amylase cluster

The λOAMY1 clone was used to create a number of subclones of this region, including OAMY1RI5.0SK+1/5, OAMY1Sac3.5SK+1, OAMY1BamSK+1, OAMY1Bam0.7SK+1, and OAMY1SacSK+1/3, which were then used to sequence the 10 kb region encoding the amylase cluster using standard forward and reverse sequence primers, and a series of custom oligonucleotides (Figure 5.1) (see Materials and Methods, Table 2.3). The amylase cluster consists of *amyR*, a putative transcriptional activator of the amylase genes, *agdA*, an α–glucosidase, and *amyA*, an α-amylase gene (Figures 5.1 and 5.2).

amyR encodes a 662 aa putative transcriptional activator of the amylase (and maltose) genes, and contains a $Zn(II)_2Cys_6$ binuclear cluster DNA binding domain (pfam00172; aa 13 –53) and a fungal specific transcription factor domain (pfam04082; aa 252 – 351), a region identified by homology with a number of other fungal transcription factors including the transcriptional activator XlnR, and the regulatory protein Gal4p from *S. cerevisiae.* The fungal $Zn(II)_2Cys_6$ binuclear cluster domain is a Cys-rich motif found in a number of fungal transcriptional regulatory proteins that is involved in zinc-dependant binding of DNA, whereby two Zn atoms are bound by six Cys residues (Pan and Coleman, 1990). The 5' upstream region of *amyR* contains 5 putative CreA binding sites, and thus may be under the control of carbon catabolite repression. AmyR shares 69.8% aa identity with AmyR from *A. oryzae,* and 72.7% aa identity with AmyR from *A. niger.*

amyA encodes a 498 aa α-amylase enzyme with an α–amylase catalytic domain (pfam00128; aa 28-382), and is classified as a member of family 13 of the glycosyl hydrolases, a group of enzymes that hydrolyse the glycosidic bonds between carbohydrates. AmyA contains a fungal cleavable signal peptide at aa 1 – 16 (von Heijne, 1986). The 5' upstream region of *amyA* contains 5 putative CreA binding sites, and a CCAAT element that may be bound by the AnCP/AnCF complex (see Section 1.4.3). No AmyR binding sites were found in the 5' non-coding region of *amyA*, in contrast to the promoters of the amylase genes found in *A. oryzae* and *A. niger.* No ESTs corresponding to *amyA* were identified in the *A. nidulans* EST database constructed from sequences obtained from RNA isolated from glucose-grown mycelia suggesting that these CreA binding sites may be functional in carbon catabolite repression. AmyA shares 69.1% aa identity with Amy1 from *A. oryzae*, 68.9% aa identity with AmyA from *A. niger*, and only 61.2% aa identity with AmyB from *A. nidulans.*

AgdA is an α–glucosidase enzyme, 993 aa in length, encoded by *agdA* containing 4 exons and 3 introns in conserved positions in comparison to α-glucosidases in other filamentous

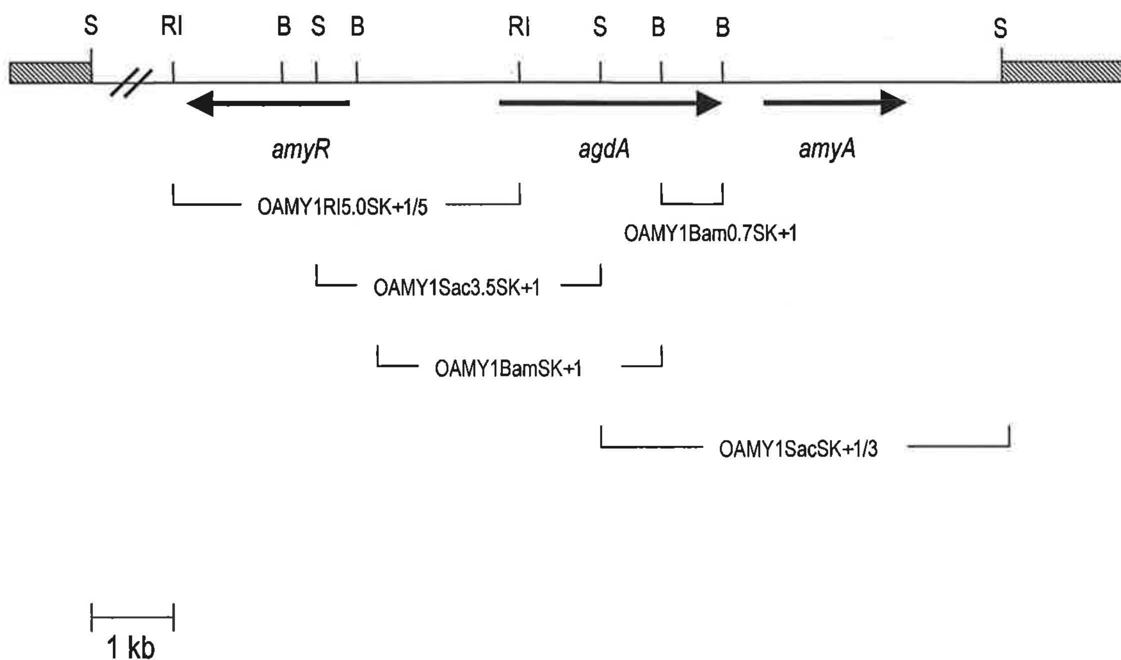

Figure 5.1 Schematic of the amylase cluster. The shaded boxes represent the λGEM11 vector, with the lines in between representing the 23 kb gDNA insert; the double slash indicates this region is not to scale. Relevant restriction sites are indicated above the gDNA insert, with RI = *EcoRI*, B = *BamHI*, S = *SacI*. The arrows indicate the positions of the open reading frames for *amyR*, *agdA*, and *amyA*. The bars underneath the gDNA insert represent the different fragments that were subcloned and used as templates for sequencing the region. The figure is roughly to scale as indicated by the scale bar.

```
1     GAATTCAAGAAAATTAGTTGTGCAGTCCTTGTGAATTATGCGCACATGTATTCACCGAAGTACAAAGCAGAGCTATAG
79    AAGTATAAAGTAAATTATATTGTTCGTTCCAATGAGTCTATAGGTGGTATGTCATATCGTGTCATGTCAGCTGAGTCT
157   CGTACTTCCTGGCAATAGGTATGAGGGTGGAGAGATAGTGATAATAAGGATGATAGATACATGGAGGGAGTAAAGACG
235   TTGGGAAGATACGCAGATGTGCAAAACACAGTCCAATATCTCCGACAGGTTGATAGATATATAAAAACCTAGAATCCT
313   CGGTAGCGATATATATACCCTTTCTTTGTATGGGCGAATTAGGGAGGAGAGCAGAAGGGTGAGGAGGAGGAGGAG
391   GAGGAGGAGGAGGAGGTAGTGGTGGTGGTGGAGCCAGTCCAGAAGTCAAGTCAATAACCTCCCAGCTTGAAACGAGAG
                                                            *  T  L  L  R  G  A  Q  F  S  L     653
469   TTCTGCAGCCATACTAGTGGACGAATGCATTCGCGAGACCGAAAGGGGTACTGACGATGATCCAGACAATGAAGGCGG
        E  A  A  M  S  T  S  S  H  M  R  S  V  S  L  P  V  S  S  S  G  S  L  S  P  P     627
547   CTGATCACTAGTCTGGTTCTGGTCGTTCGTATCACCAGAAGTTATTTCCCATGCACACGAATTTGTTGTCCTGCTGCC
        Q  D  S  T  Q  N  Q  D  N  T  D  G  S  T  I  E  W  A  C  S  N  T  T  R  S  G     601
625   TGTGCTAGAATCCATAGTATTGTCATCAGCCGCGGGCCATGGTCCAGAGATATCTAAAGTTTCAGAAGCAATTGAAGA
        T  S  S  D  M  T  N  D  D  A  A  P  W  P  F  S  I  D  L  T  E  S  A  I  S  S     575
703   AGAGCTAGCGGAATTGAGAAGCGGAGGATTAATCTGGTGGAAGTCACTGATAGAAAGCGGCGTCGGGGAATCAAAGCC
      ..S  S  A  S  N  L  L  P  P  N  I  Q  H  F  D  S  I  S  L  P  T  P  S  D  F  G     549
781   GAGGGTGCCTTTGCAGCGCTCAAGGAGAGATGGGAAGAGGTACGACTGGGAGCCACGGATGCGGGAGAGGGTGGAAAG
        L  T  G  K  C  R  E  L  L  S  P  F  L  Y  S  Q  S  G  R  I  R  S  L  T  S  L     523
859   GATACCCCAGAGGAGCTCGCGGGGGTCGACGGTTGACTCGGCGAGGCTGTTTGCAGCTTTTGTGGAGAGAGAACGTGA
        I  G  W  L  L  E  R  P  D  V  T  S  E  A  L  S  N  A  A  K  T  S  L  S  R  S     497
937   GACGTCGGCGATAGATGTACCAAGATCGAAGAGCTTTTGTTCctgcgtaatcgtcagccttaatcgaagtaaattgct
       V  D  A  I  S  T  G  L  D  F  L  K  Q  E                                          483
1015  aaagtggtgaagactacCATCCCAATCCCATGCGCGTCGACAGCCCCCTGCGATGCCTCCGCAATGACAGACATAACA
                    M  G  I  G  H  A  D  V  A  G  Q  S  A  E  A  I  V  S  M  V     463
1093  GCCTTGCCCACCATTACAGGAAGATGGAACGGGAGGACCGTCTCATTCCGCGAGCCTGGCTGCGATGCACGCGTCATG
       A  K  G  V  M  V  P  L  H  F  P  L  V  T  E  N  R  S  G  P  Q  S  A  R  T  M     437
1171  GAGAGTTTCCACATCATGGCCTGAAGCCACTGCTGTGTGATGAGGATATCAACCTTTTGGATTTCCGAGACACCCTCG
       S  L  K  W  M  M  A  Q  L  W  Q  Q  T  I  L  I  D  V  K  Q  I  E  S  V  G  E     411
1249  AGCGAGATCGGTTTGCTGAGAGTTGCCTGGATGGCTGATGTAGGCGGCATCTCGGAGGAGCCGTCTCGGTTATTTCCA
       L  S  I  P  K  S  L  T  A  Q  I  A  S  T  P  P  M  E  S  S  G  D  R  N  N  G     385
1327  CCTCCGCTTACGCTTCCCCCTGCTGAGACCCAGTCGTAAAGGTTGACGCTGAGGCTCTCGAAGACGCCGATGAGGTTG
       G  G  S  V  S  G  G  A  S  V  W  D  Y  L  N  V  S  L  S  E  F  V  G  I  L  N     359
1405  ATGAAGCCGTAGGCGAGAATCGGGTCGTCGGAGCACAGGACTTGCGGTTTGTGGATGGAGTTGCGGAGCATGACAGGT
       I  F  G  Y  A  L  I  P  D  D  S  C  L  V  Q  P  K  H  I  S  N  R  L  M  V  P     333
1483  TTCGCTTGTTGGAGGGCATAGCCTctgcagcacttgtcagtttatggtctattcagctctacggaggtaaatacCGTT
       K  A  Q  Q  L  A  Y  G                                                  R     324
1561  CTGTGATGAAGAGCAGCCAGAAGACCCGTCGCTTCTGTTCTGCTTCTTCGGTGTTCAACTCGGCGTATGTTGATTCGC
      E  T  I  F  L  L  W  F  V  R  R  K  Q  E  A  E  E  T  N  L  E  A  Y  T  S  E     298
1639  GGTGCAAGCCCAATGCGAAAACCATTGATGTCGCCTGGCAAAGATAGAACCATGCGTGGTCCTGGCGGTCTAGATTTC
      R  H  L  G  L  A  F  V  M  S  T  A  Q  C  L  Y  F  W  A  H  D  Q  R  D  L  N     272
1717  CGTATGAAGCAAATAGGAAGAATGAGGTTAGAAGACTCTCAACATTCATATCCTCAATTGGGTCGCAATCTTTCCTTG
      G  Y  S  A  F  L  F  F  S  T  L  L  S  E  V  N  M  D  E  I  P  D  C  D  K  R     246
1795  CGCGCACGGCTTCGGCGAGCAGTTCTTCACCGGACATTGATGAGTTCTCACCGGCCTGAAAATGCGAAGGATCCGCGA
      A  R  V  A  E  A  L  L  E  E  G  S  M  S  S  N  E  G  A  Q  F  H  S  P  D  A     220
1873  CGGGCGTTGCTCCGTCGAGTTTCAGCTGGATATGTGTGGCGGCACATAGGGAGGCCAGAAAGGCGTACCGCTGGGGAG
      V  P  T  A  G  D  L  K  L  Q  I  H  T  A  A  C  L  S  A  L  F  A  Y  R  Q  P     194
1951  TTAGACGCTCTGGATGGTGGCAGTCCTGCTGGAGCTCTTCTCTTCGGACCACGGGCATGATAGGAAACATATATTTCA
      T  L  R  E  P  H  H  C  D  Q  Q  L  E  E  R  R  V  V  P  M  I  P  F  M  Y  K     168
2029  AGTAGACATTGACGTGCGCCAGAAGCACAGGTGCGGTCAGAATCCGCGGCGAACGAAACAGCACCGCAGTCGGAGAGT
      L  Y  V  N  V  H  A  L  L  V  P  A  T  L  I  R  P  S  R  F  L  V  A  T  P  S     142
2107  CCGATAGGGAGTTCAGAGAATCCGGCGAAGAGACGAGTTCTGGTGGAGGCAGCCGGCTGAATGAGTCAGGGTGTTCAG
      D  S  L  S  N  L  S  D  P  S  S  V  L  E  P  P  P  L  R  S  F  S  D  P  H  E     116
2185  GCGGCAGATATTGCGGCTCTTCCATGAAGGGCGGAGACGTGGGCGAGCCGCCGACCGGGTAGCTCGTAGGCTCTGCAT
      P  P  L  Y  Q  P  E  E  M  F  P  P  S  T  P  S  G  G  V  P  Y  S  T  P  E  A      90
2263  ACCATTCCCTCTCTGCAGAGAATCGTGGAGGATCATACTGGAGAATCTGTCGGTCTCTGGCCACTAGTGGATGGATAG
      Y  W  E  R  E  A  S  F  R  P  P  D  Y  Q  L  I  Q  R  D  R  A  V  L  P  H  I      64
2341  GCGCTAGAGGGTATAGAGTCCGAAATTTCGGCCCTTTGCGACGCAGCACATCGCTGTAGGAGCAGGAAAGGAGCAGGC
      P  A  L  P  Y  L  T  R  F  K  P  G  K  R  R  L  V  D  S  Y  S  C  S  L  L  L      38
2419  GCTGGCACTTGTCGCACGGGAGCTCTCGTGAGCACTTGATTTTCCTCCGGCGACAGTTGTCACAGGCCTGCTTGAAAG
      R  Q  C  K  D  C  P  L  E  R  S  C  K  I  K  R  R  R  C  N  D  C  A  Q  K  F      12
2497  GTTTGTGTTTCGGAGCGGCTGGTGAGGCCATGGTGCACGATGGCCGGGCCAAAGGCGAACGATATGAGCAGTGCCGGT
      P  K  H  K  P  A  A  P  S  A  M                                                     1
2575  AGATGAGAAGCTAGTCTTGAAGAGGCATGCGGGTCATGTCGGATTGCTCGTCAACAGTATAAAATCGCTATAGTCCTT
2653  GCAGACGTACGCATGGGTGAAAATCTGGGGAAAGCGAGCGGGGGAGGGAGAGGATAAAGAGAGACGGGCAGGAGTAAT
2731  CGGATAGCAGTGACACTAACTCCACTATCCGCTTGCCCTCCGTATAGTACTCCGTAGTGGCCTGCAGCCGTAGTCAGG
2809  CTCAGCCCCGCGTTACACCATTGCTTCTCGGCTTCGTCTCTTAGCAACCCTTAACAGTTCATTTAAGGCTTCTTCGCG
2887  TTAGCTTTGCCGGTGGACTGCACTCCAGCCTGCCACCTTTTGCCCAGAGAAAAGAGCGAGGTGACGGCTACAGCCAGG
2965  CCCGAGAGCCGGCTGTCAGAAACCCCTGAATCCTGGATCCGGGATTGTTGCCCAGTATGACGTCTTGGCAGCATGCGC
3043  GATTTTCTAGAAACCTCCCCTGCATGAAGCCCATTTGCCTGTGTGATCCTTGGTCTAGCGGTCGTCCCCTGCGCTGGC
3121  CATTCGTCAGTGTCGAGTAAGACATGAACGCCTCTTTAGAAATCGGGCTTCTGTGGCTCACTTAAGGGCCCAAGTCTG
3199  TCGGGAAGCCGAGGAACAGCAGCATGACCCAAGGAACCGATGCAGTGCCAACCACAGATGCCGAATCCTCGTACATCG
3277  TCTTGGTGGTGTTGATGTGCAGCGACCGTTCTCCACCGTGGGTTCTGGGAGGGGGCCCGATTCCAGTCTCAATCTCGA
3355  GGTCTGTTTGAGTTTTAGGATAACGAAGCGCTGGATGTCTGGATTTCATGATTTCATGTGATACTCATGGTGCCGAGA
```

```
3433 GCATAGCTAGCAGTCCAGACTCCGGATTATCCAGGAAATCGATCAAGTCGACTGTTCTCAGTCTGGTCGGGGACTGGG
3511 GTTAGCGGGAAGGCTGGAGACTGGAGTTGCTCACCGGTGGTAATTTTAGGCATGGCACGCCCGTAACATTACGCACAC
3589 AATACGGTATTCTGGGGATGAACCCTGCTCTGGGTAGATACTCCGTACAGAATAGCTTGGCTGAGAGTTAACACTCCA
3667 CAATCTCCTCTCTTCATCCTCTTCCATCACCCGGTGCTTCGACCTTCGTATCCTGCATCCTTGATATCAACTATCCTC
3745 GACTATAACAGCTTCGTTATAGGCTCGGTCTTTTGTCGGAGATATGACCAATCAGCATTTTGCAGTGCCTCCACCCCT
3823 CGAACGGATCTGAGTCCCAAGCTTTGGCCGGGTTGATCGCGGGGGACCGGTCGTCCCCGCATCACCCACTGGCTTGCT
3901 GCCTGCTGGCTTCTGGCTTCAACATGGGCTCTCTCTTCTGTTCATCCTCATCATACGCCCCTTGCATGCATCGTTGCG
3979 TCTAGGATTCTAAAGTCAAATGGATGATCGCCGCGGGACGTGCCACAGCATGGTTTGACTCTTGCTTGTCTTGTGAAT
4057 CCGTGATTTTCCCGTCGAGTTCGAAACCTCGAAGATTAAATTGAGTCAAGGCCGTGGTTGTTGGATGATCTGATCCAT
4135 CGTACGATTGGAAAAGCCTCCTATCCCCACTATGGTCCGCTTCCTGCACCTAGCTGGGACCCTTCCGGTCCTAGCCTC
                                                 M  V  R  F  L  H  L  A  G  T  L  P  V  L  A  S     16
4213 TGGCGCAGTCCAGGACGCCCTCCGGCCAGTTGCCGAGTCCGCCGCGACCGTAACCGCAACCGCGACCGTAGCCGGGCA
      G  A  V  Q  D  A  L  R  P  V  A  E  S  A  A  T  V  T  A  T  A  T  V  A  G  Q     42
4291 ACAGGCTCAATTTACACTCTCGGATTATGTGGATGTCGGGGCAGACTTGATTGCCAATGTGGACGATCCCGAAGCCGT
      Q  A  Q  F  T  L  S  D  Y  V  D  V  G  A  D  L  I  A  N  V  D  D  P  E  A  V     58
4369 CAATGCCCAGTCCGTCTGTCCTGGGTACAAGGCTTCCGATATACAACAAACGGACCTGGGCTTTACCGCCAGCCTACG
      N  A  Q  S  V  C  P  G  Y  K  A  S  D  I  Q  Q  T  D  L  G  F  T  A  S  L  R     84
4447 GCTGGCCGGCGAGCCCTGCAATGTATACGGGACGGACGTCGAGTCTTTGACTTTGGAGATGCAGTATCAGGATACGGA
      L  A  G  E  P  C  N  V  Y  G  T  D  V  E  S  L  T  L  E  M  Q  Y  Q  D  T  D    110
4525 CCGCTTGAACATCCAAATCACACCGACCTACGTGGACGCATCCAACGCGTCCTGGTATATTCTGCCAGAGGAATTCGT
      R  L  N  I  Q  I  T  P  T  Y  V  D  A  S  N  A  S  W  Y  I  L  P  E  E  F  V    136
4603 CCCCCGACCGAAACCGGCCGCAGGTGCGTCCGAGTCGCACAGCGACTTCGCCGTGACATGGTCAAACGAGCCAACTTT
      P  R  P  K  P  A  A  G  A  S  E  S  H  S  D  F  A  V  T  W  S  N  E  P  T  F    162
4681 CAACTTTCAGGTCACCCGGAAGTCGACAGGCGAGGTGCTCTTCGATACGGCCGGTTCCGTATTGGTTTTCGAAAACCA
      N  F  Q  V  T  R  K  S  T  G  E  V  L  F  D  T  A  G  S  V  L  V  F  E  N  Q    188
4759 ATTCATCGAGTTTGTGACATCTTTGCCGGAGGAATATAACCTGTATGGTCTAGGAGAACGCATCAACCAGCTCCGTCT
      F  I  E  F  V  T  S  L  P  E  E  Y  N  L  Y  G  L  G  E  R  I  N  Q  L  R  L    214
4837 GTTGCGAAACGCCACACTGACCTCCTATGCGGCTGACATTGGCAATCCGATTGATGCgtacgttcggttactctcatc
      L  R  N  A  T  L  T  S  Y  A  A  D  I  G  N  P  I  D  A                         233
4915 cttgccccatgctaatggtatgcagCAACATCTACGGACAGCACGCATTTTACGTAGACACAAGATACTTCTCAGTTG
                               N  I  Y  G  Q  H  A  F  Y  V  D  T  R  Y  F  S  V  D    251
4993 ACGAGGCTGGAAAACACACATACGTGAAAAGTAGTGAGGCTGACCCTTCGGCAACATATACCTCCTACTCGCATGGGG
       E  A  G  K  H  T  Y  V  K  S  S  E  A  D  P  S  A  T  Y  T  S  Y  S  H  G  V   277
5071 TTTTCCTCAGAAACTCTCACGGCCATGAAGTTGTGTTGAATCCGCAGGGCCTGACTTGGCGGACGATTGGAGGAAGCA
       F  L  R  N  S  H  G  H  E  V  V  L  N  P  Q  G  L  T  W  R  T  I  G  G  S  I   303
5149 TCGACCTCACCCTCTACTCGGGGCCGACTGTAGCAGAGGTCACAAAGCAATACCAGCGCAGCACCGTTGGTCTCCCCG
       D  L  T  L  Y  S  G  P  T  V  A  E  V  T  K  Q  Y  Q  R  S  T  V  G  L  P  A   329
5227 CTATGCAAAAGTACGACACGCTTGGTTTCCATCAATGCCGGTGGGGTTACAACAACTGGTCGGTCTTTGCAGATGTTC
       M  Q  K  Y  D  T  L  G  F  H  Q  C  R  W  G  Y  N  N  W  S  V  F  A  D  V  L   355
5305 TTGCGAATTTCGAGAAATTTGAGATTCCGTTGGAATATCTCTGgtatgcacgaatacgcacctatttgaatcctagct
       A  N  F  E  K  F  E  I  P  L  E  Y  L  W                                        369
5383 tacatctttagGGCCGACATTGACTATATGCATGGTTATAGAAATTTTGAAAATGACGAGTATAGATTTCCATATAAC
                   A  D  I  D  Y  M  H  G  Y  R  N  F  E  N  D  E  Y  R  F  P  Y  N   391
5461 GAGACCAAAGTTTTCTTGGACAAGCTTCACGCAGGTGGGCGCCATTTTGTCCCAATTGTGGACGCGGCGCTGTATATC
      E  T  K  V  F  L  D  K  L  H  A  G  G  R  H  F  V  P  I  V  D  A  A  L  Y  I    417
5539 CCTAATCCGCAAAACGCTTCAGATTCgtaagtttttcgtttttattgccgagtactgctgacgatcttagTTATGAAA
      P  N  P  Q  N  A  S  D  S                                              Y  E  T    429
5617 CGTACACTCGAGGCGCAGCTCGAGACGTCTTCCTGAAGAATCCAGATGGCAGCCTCTACATCGGCGCTGTATGGCCTG
      Y  T  R  G  A  A  R  D  V  F  L  K  N  P  D  G  S  L  Y  I  G  A  V  W  P  G    455
5695 GTTATACAGTCTTTCCCGACTGGCATCATCCCGATGCGGCTGATTTCTGGGCCAACGAGCTCGTGACTTGGTATGAGA
      Y  T  V  F  P  D  W  H  H  P  D  A  A  D  F  W  A  N  E  L  V  T  W  Y  E  K    481
5773 AAGTCAAATTCGATGGAGTGTGGTACGACATGAGTGAGGTATCGTCTTTCTGTGTGGGAAGCTGTGGGTCTCGAAACC
      V  K  F  D  G  V  W  Y  D  M  S  E  V  S  S  F  C  V  G  S  C  G  S  R  N  R    507
5851 GGACACTCAACCCCGTTCACCCGCCTTTCAGATTGCCAGGTGAACCGGGCAACGTTGATTACGAGTATCCTGAAGGAT
      T  L  N  P  V  H  P  P  F  R  L  P  G  E  P  G  N  V  D  Y  E  Y  P  E  G  F    533
5929 TTGAGCTGTCCAACGCGACAGAGGCTGCTTCGGCATCTGCAGCTTCTTCGAGCCAGGCCGCAACCACCGCCACAGAGA
      E  L  S  N  A  T  E  A  A  S  A  S  A  A  S  S  S  Q  A  A  T  T  A  T  E  T    559
6007 CTACGACATCCACCAGCTCTTACTTGCGTACTTCGCCCACTCCTGGTGTCCGTAATGTCAACTATCCACCTTATGTGA
      T  T  S  T  S  S  Y  L  R  T  S  P  T  P  G  V  R  N  V  N  Y  P  P  Y  V  I    585
6085 TCAACCATGTCCAGACGGGCCATGATCTTGCTGTGCATGCAGTCTCGCCAAATGCTACTCACGTTGACGGTTATCACG
      N  H  V  Q  T  G  H  D  L  A  V  H  A  V  S  P  N  A  T  H  V  D  G  Y  H  E    611
6163 AGTACGACGTGCACAGCCTCTACGGACATATGGGTATCCAGGCCACCTACCGAGGTTTGACTCAGATTGCGCCCAGGA
      Y  D  V  H  S  L  Y  G  H  M  G  I  Q  A  T  Y  R  G  L  T  Q  I  A  P  R  K    637
6241 AGCGTCCATTCATCATTGGCCGCTCGACGTTTGCTGGCTCTGGAAAATGGGCCGGCCATTGGGGCGGTGACAACTACT
      R  P  F  I  I  G  R  S  T  F  A  G  S  G  K  W  A  G  H  W  G  G  D  N  Y  S    663
6319 CCCGGTGGTCATCCATGTACTTTTCAATCTCGCAAGCGTTGCAGTTTTCTCTCTACGGCATTCCGATGTTCGGAGTCG
      R  W  S  S  M  Y  F  S  I  S  Q  A  L  Q  F  S  L  Y  G  I  P  M  F  G  V  D    689
6397 ATACGTGCGGGTTTAGTGGAAATACCGCGGAGGAACTCTGCAACCGCTGGATGCAGTTGTCGGCTTTCTTCCCGTTCT
      T  C  G  F  S  G  N  T  A  E  E  L  C  N  R  W  M  Q  L  S  A  F  F  P  F  Y    715
6475 ATCGTAACCATAATGTTCTTGGGACAATCCCTCAGGAGCCTTATCAATGGGCATCGGTCATTGATGCCACCAAGAAGG
      R  N  H  N  V  L  G  T  I  P  Q  E  P  Y  Q  W  A  S  V  I  D  A  T  K  K  A    741
6553 CGATGAGGATCCGATATGCTTTGCTGCCTTATTTCTACACTTTGATGCATGACGCGCATACTACGGGCTCTACTGTGC
      M  R  I  R  Y  A  L  L  P  Y  F  Y  T  L  M  H  D  A  H  T  T  G  S  T  V  L    767
```

```
6631 TGCGAGCGCTGGCCTGGGAGTTTCCGGACGATCCATCCCTGGCTGCGATTGACAATCAGTTCCTGGTCGGCCCCTCCA
     R A L A W E F P D D P S L A A I D N Q F L V G P S I                           793
6709 TCCTGGTTACTCCAGTCCTCGAGCCTCAAGTAAGTACCGTTAAAGGCGTCTTTCCTGGAGTAGGGCAGGGAGAAGTCT
     L V T P V L E P Q V S T V K G V F P G V G Q G E V W                           819
6787 GGTACGATTGGTACACACAAACTGCAGTTGACGCCCAGCCGGGAGTCAACACGACGATCGACGCGCCCCTCGGCCATA
     Y D W Y T Q T A V D A Q P G V N T T I D A P L G H I                           845
6865 TACCGGTCTATGTCCGGGGCGGAAGTATCCTGCCAATGCAAGAGCCTGCTTTGACGACCCGCGACGCCCGGAAGACGC
     P V Y V R G G S I L P M Q E P A L T T R D A R K T P                           871

6943 CGTGGGCACTGTTAGTCGCACTCGGTAAAGACGGAACAGCTTCAGGCCATCTATACCTCGATGATGGGGAGAGTATCC
     W A L L V A L G K D G T A S G H L Y L D D G E S I H                           897
7021 ATCCAAAGGTGTCGCTTAACGTCAAATTTAGGGCCACACAGACCGCCCTAACCGTATCATCCGAAGGCGAGTGGAAAG
     P K V S L N V K F R A T Q T A L T V S S E G E W K E                           923
7099 AAGCCAACCCATTGGCGAATGTAACCATTTTAGGTGTCCTTGAAAACCCGGTCTCGGTCACATCTAACGGACAGCAAG
     A N P L A N V T I L G V L E N P V S V T S N G Q Q V                           949
7177 TGCCTGCAGAATATGACGCGCAGTCGCGGATCCTGGTTATTACCGGGCTGAACCAATTTACAAACAATGGTGCTTGGG
     P A E Y D A Q S R I L V I T G L N Q F T N N G A W G                           975
7255 GGCAAGATTGGACCCTTCGGTGGTGAAGCGGTTTGGTATAGGTGGATAAGGGCCGGAGTCACCGGTGTGATTTAGTTT
     Q D W T L R W *                                                               983
7333 GAAGTACCGGTAGTGAAATGCGACTCCAGCAGGTACATTAGGGCGTTCAGACCAATAATACTTGGGATGTTCCCCGCA
7411 CTCTGCCGGCTTGCAAGGTTCTCCACTGCACCCCTGGTAATCTCCTGGCGTCGTGGTTGGACTGGCGTGGCCATTCTT
7489 TGCCAACCATTGCAGAAACCTCCGACAAAGCCGGATACTTGGGGGATGCTCAGCTGGTCGTAGGATACATGTGATAAA
7567 TATGGGCACCCCACCGTAGTGTTAGGTGTTGTAGAGTGTAGACCATCAACAAGATCCTTCTGCAACGATGCGGTCGCC
                                                                         M R S P   4
7645 ACTCTTTCTCTCATTTGCGGCGACAGTGCTGGCTGCGACGCCTGCCGAATGGCGCTCGCAGTCAATCTACTTTTTACT
     L F L S F A A T V L A A T P A E W R S Q S I Y F L L                           30
7723 CACCGACCGGTTTGCCCGCACCGACAACTCGACAACCGCCGAATGTGATACTAGTGCAgtggtgagcagcctggacct
     T D R F A R T D N S T T A E C D T S A                                         49
7801 agattttgattctaattcagagactgaccctgcgcagAAGTACTGCGGCGGGACATGGCAGGGAATCATTAACCAGCT
                                           K Y C G G T W Q G I I N Q L             63
7879 GGACTACATCCAGGGGATGGGCTTCACAGCCATCTGGATCACCCCAGTGACTGCCAATCTCGAGGATGGGCAGCATGG
     D Y I Q G M G F T A I W I T P V T A N L E D G Q H G                           89
7957 GGAGGCATACCATGGGTACTGGCAGCAGGATATgtaggtctccttcatatccagtggcataaaactaacgataagtgc
     E A Y H G Y W Q Q D I                                                         100
8035 agATATGCGTTGAACCCGCACTTTGGCACTCAAGACGACCTCCGAGCACTGTCTGACGCGCTGCACGACCGGGGAATG
       Y A L N P H F G T Q D D L R A L S D A L H D R G M                           125
8113 TACCTTATGGTCGACGTGGTTGCCAATCATTTTgtgagtttgaactgcaggtgtgcttgcttttatagatgccaatac
     Y L M V D V V A N H F                                                         136
8191 tcatggtccatgtagGGCTACGACGCCCCGGCCGCGTCGGTCGACTACAGCGCCTTCAACCCCTTTAACTCGGCAGAC
                    G Y D A P A A S V D Y S A F N P F N S A D                      157
8269 TACTTCCACACTCCCTGCGATATTACGGACTACGACAACCAGACCCAGGTCGAGGATTGCTGGCTGTACACCGACGCT
     Y F H T P C D I T D Y D N Q T Q V E D C W L Y T D A                           183
8347 GTCAGTCTTCCAGATGTCGATACCACCAACGAGGAGGTCAAGGAGATTTGGTACGACTGGGTGGGTGACCTTGTGTCT
     V S L P D V D T T N E E V K E I W Y D W V G D L V S                           209
8425 GACTACTCTAgtaaagcccctcccctatcctcccagtactgtatacatatctatgaaggtgttctcacatccaaaaca
     D Y S I                                                                       213
8503 aagTCGACGGCCTTCGCATCGACACCGCTCGACACGTACAGAAGGACTTCTGGCGCGACTACAACGATGCCGCGGGCG
          D G L R I D T A R H V Q K D F W R D Y N D A A G V                        238
8581 TGTACTGCGTCGGCGAGGTCTTCCAGGGCGATCCCGATTACACATGCGGGTACCAGGAGGTTATGGACGGGGTGCTGA
       Y C V G E V F Q G D P D Y T C G Y Q E V M D G V L N                         264
8659 ACTATCCCATgtactttcccactatttatcccgattacctgcatgcccgctaactgtaccagCTACTACCCCCTGTTG
       Y P I                                                     Y Y P L L         272
8737 CGCGCTTTCAGCTCCACATCTGGCAGTCTCAGTGATCTAGCCAACATGATCGAAACGGTAAAGTACACCTGCTCAGAC
     R A F S S T S G S L S D L A N M I E T V K Y T C S D                           298
8815 GCTACCTTGCTGGGCAACTTCATCGAGAACCACGATAACCCACGCTTTGCCTCgtacgcttcctcaacccaatttccc
     A T L L G N F I E N H D N P R F A S                                           316
8893 cggttcttctaggcaatgccactgaccagacaatacagGTACACCGACGACATCTCCCTCGCCAAGAACGTCGCCGCC
                                            Y T D D I S L A K N V A A              329
8971 TTCGTGATCCTCTCCGACGGGATCCCCATCATCTACGCCGGTCAAGAACAGCACTACTCCGGCGCAGGAGACCCGGCA
     F V I L S D G I P I I Y A G Q E Q H Y S G A G D P A                           355
9049 AACCGTGAGGCGACCTGGCTCTCCGGATACGACTCGACGAGCGAGCTGTACCAGTTCATTTCGAAGACGAACCAGATC
     N R E A T W L S G Y D S T S E L Y Q F I S K T N Q I                           381
9127 CGGAATCATGCGATCTGGCAGAATGAGACCTACCTGTCTTACAAAgtatctctctcctttcccctacataccgcggcc
     R N H A I W Q N E T Y L S Y K                                                 396
9205 ccttgagtaatggggtttatataatagagctgacggcgaatgaatgaacagAACTATGCTATCTACAACGAAAACAAC
                                                         N Y A I Y N E N N         405
9283 GTCCTTGCCATGCGCAAAGGATTCGACGGGTCGCAGATCATCACAATCCTTACGAACGCTGGCGCTGACGCTGGTTCA
     V L A M R K G F D G S Q I I T I L T N A G A D A G S                           431
9361 TCGACTGTCTCGGTTCCGAACACCGGGTTCACAGCTGGTGCGGCAGTCACTGAGATCTATACCTGCGAGGACATTACA
     S T V S V P N T G F T A G A A V T E I Y T C E D I T                           457
9439 GTCTCGGGCAGCGGTGAAGTGTCAGTGCCTATGGAGAGCGGCTTGCCGAGGGTTCTGTATCCGAAGGCGAAGCTGGAA
     V S G S G E V S V P M E S G L P R V L Y P K A K L E                           483
```

```
9517  GGGAGCGGGATTTGCGGCCTGTAGTCGTACCCGATGATGAAACATAATGGACAAGTTGGTGAAAGGTTGGGCCAGAGG
      G  S  G  I  C  G  L  *                                                          491
9595  TTGGACGGTCGTGAGGGTTATATTGGGGACGGTGAATATATATAGGAACATCGTTCCTTATAGTGCGGATGGTGTGGT
9673  AAGTATAGACTTTGGATGAAATTCGTTTTTGGATAATTGTGTCAGTATATGATTTGGTGTGGTCTTTTCGTTTGATTT
9751  TAGCATA
```

Figure 5.2 Nucleotide sequence of the amylase cluster. The amino acid sequences of AmyR, AgdA and AmyA are shown in single letter code beneath the DNA sequence, with numbered amino acid residues indicated on the right-hand side, and numbered nucleotide positions indicated on the left-hand side. The peptide encoded by *amyR* is highlighted in yellow, the peptide encoded by *agdA* is highlighted in green, and the peptide encoded by *amyA* is highlighted in purple. Introns are depicted in lower case letters. Consensus **CreA binding sites** are indicated by **bold-face type**, CCAAT elements are highlighted in blue, and the AmyR binding sites in the *agdA* promoter are and shaded grey. The $Zn(II)_2Cys_6$ binuclear DNA binding domain of AmyR is boxed with a continuous line, and the fungal specific transcription factor of AmyR is boxed with a dotted line.

Sequences were deposited in Genbank with Accession numbers as follows:

Amylase cluster nucleotide sequence	AF208225
AmyA protein sequence	AAF17103.1
AgdA protein sequence	AAF17102.1
AmyR protein sequence	AAF17101.1

fungi. AgdA contains a glycosyl hydrolase family 31 domain (pfam01055; aa 578-953), another member of the glycosyl hydrolases that hydrolyse glycosidic bonds. The 5' region of *agdA* contains 11 putative CreA binding sites, 2 putative AmyR binding sites of the (5'-CGG-N_8-CGG-3') type, and a CCAAT element that may be the target of the AnCP/AnCF. AgdA shares 73.6% aa identity with AgdA from *A. oryzae*, and 72.4% aa identity with AglA from *A. niger.*

5.3 *amyB* is not part of the amylase cluster

To determine whether the previously characterized *amyB* gene resides within the amylase cluster region, a BAC library was screened to identify whether a member of the amylase cluster and *amyB* colocalized to a single BAC clone (which typically contain between 50-150 kb gDNA). An *A. nidulans* BAC library was screened using *amyA* and then *amyB* as a probe. *amyA* hybridised to 33 BAC clones, and *amyB* hybridised to 24 BAC clones. No overlapping hybridisation was seen. Therefore, *amyB* cannot be within approximately 150 kb of *amyA* and does constitute part of the amylase cluster containing *amyA*.

Verification of these results came with the public release of the *A. nidulans* genome in March 2003. *amyB* resides on contig 1.55, linkage group VI, and the amylase cluster containing *amyA* on contig 1.32, linkage group VII.

5.4 Regulation of the amylases

To determine the transcriptional regulation of the amylases identified in *A. nidulans*, Northern analyses were undertaken. RNA was prepared from wild-type and *creA204* mutant strains under various growth conditions, including overnight (16 hr) incubation in media containing as a carbon source; 0.1% fructose, 1% glucose, 1% starch, 1% maltose, and also 0.1% fructose for 16 hrs followed by 1% starch for 4 hrs, 0.1% fructose for 16 hrs followed by 1% maltose for 4 hrs, and 1% glucose for 16 hrs followed by 1% starch for 4 hrs. RNA gel electrophoresis and Northern hybridisations were performed using two different methods (see Section 2.3.7), but under no experimental condition was any hybridisation detected when *amyA*, *amyB,* or *amyR* radiolabelled DNA or RNA probes were used (see Table 2.3), indicating low levels of transcript under the conditions tested.

In order to increase the sensitivity of the assays, RNase protection assays (RPA) were performed on the same RNA samples used in the Northern analysis, using *amyA*, *amyB* and *amyR* as probes. Again, no transcripts were detected under these conditions.

5.4.1 Alpha-amylase secretion plate tests

The absence of detectable transcripts in Northern and RPA experiments suggested that conditions considered to induce α-amylase activity, i.e. the presence of starch/maltose, were not inducing transcription of the amylolytic genes to detectable levels. Simple secretion plate tests were performed to identify conditions that resulted in secreted α-amylase activity. By spotting 20 μl of an overnight *A. nidulans* culture onto a 0.5% starch plate, incubating the plate at 37°C (for 1 hr, 4 hrs, and 16 hrs), and then flooding the plate with iodine, the amount of α-amylase activity can be determined by the size of the cleared halo (see Section 2.3.11).

For the wild-type strain, the only condition in which secreted amylase activity was detected was in mycelium grown in medium containing starch as the sole carbon source (Figure 5.3). In a strain containing the *creA204* mutation, secreted amylase activity was detected in all conditions containing either starch or maltose. These results are consistent with the presence of secretable α-amylase activity under the control of CreA.

5.5 Discussion

The amylase cluster in *A. nidulans* consists of the transcriptional regulator AmyR, an α-glucosidase AgdA, and an α-amylase AmyA. This is the same arrangement of genes as observed in *A. oryzae* (Gomi *et al.*, 2000). The sequence of all three components of the amylase cluster was submitted to GenBank on 24/11/1999 (Accession No. AF208225). Tani *et al.* (2001a) independently reported on the cloning and characterization of the *amyR* gene from *A. nidulans*, and had submitted the sequence of the *amyR* gene only to GenBank on 3/6/99 (Accession No. AB024615). Tani *et al.* (2001a) designed oligonucleotides to the 5'- and 3'- ends of the *A. oryzae amyR* gene, and then PCR amplified *amyR* using *A. oryzae* gDNA as a template. The zinc binuclear cluster domain region of the *A.oryzae amyR* PCR product was used in Southern analyses to probe an *A. nidulans* gDNA library, and a 12 kb *EcoRI* fragment containing *amyR* and the 5'-part of the *agdA* gene was identified and sequenced. The two AmyR sequences entered into GenBank for *A. nidulans* are identical.

Amylase secretion plate tests indicate that there is secretable α-amylase activity in the wild-type strain when grown for 16 hrs in medium containing 1% starch, but RNA isolated from wild-type strains grown in this condition lacked detectable levels of *amyA* and *amyR* transcripts. The 4 hr induction period by starch/maltose did not result in observable amylase activity by secretion plate tests. Preliminary Northern and RPA investigations into the regulation of the

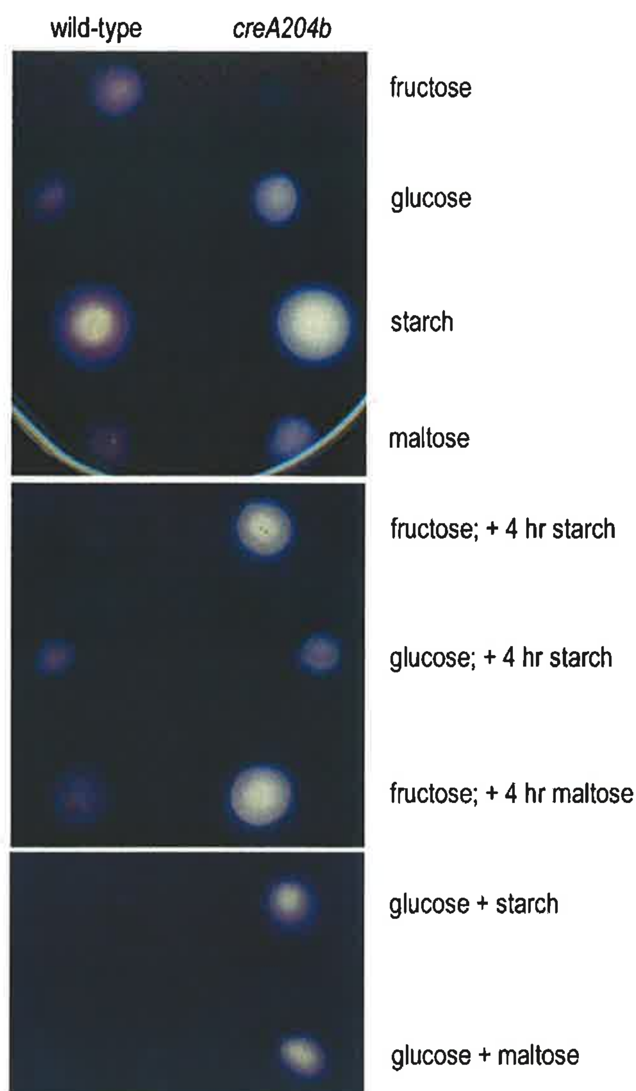

Figure 5.3 Amylase activity secretion plate tests. 20μl of overnight culture was spotted onto a 0.5% starch plate and stained with iodine after a 16 hr incubation at 37°C. Clearing indicates starch breakdown, and therefore amylase activity. Growth conditions were, in order from top to bottom, 0.1% fructose, 1% glucose, 1% starch, 1% maltose, 0.1% fructose for 16 hrs followed by 1% starch for 4 hrs, 1% glucose for 16 hrs followed by 1% starch for 4 hrs, 0.1% fructose for 16 hrs followed by 1% maltose for 4 hrs, 1% glucose and 1% starch, 1% glucose and 1% maltose. All cultures were grown for 16 hours (unless stated otherwise) at 37°C, with ammonium tartrate as a nitrogen source.

amylase cluster had revealed that the level of *amyR* and *amyA* transcripts were below the level of detection for all of the conditions tested in both wild-type and *creA204* strains.

Tani *et al.* (2001a) disrupted the *amyR* gene, and these disruptant strains grew poorly on media containing starch, maltose and iso-maltose, while growth on glucose, fructose, cellobiose and a number of other carbon sources were not affected, suggesting that only amylolytic enzyme expression was affected. The effect of the *amyR* disruption was also examined at the transcriptional level, using *taaG2* from *A. oryzae* as a reporter gene since *A. nidulans* produces such low amounts of amylase, leading to difficulties in detection of amylase activity and mRNA (Chikamatsu *et al.*, 1999; Tani *et al.*, 2001a). The *amyR* disruptant strains showed no detectable amylase activity, and the *taaG2* gene was not transcribed in the absence of AmyR in comparison to a wild-type (*amyR*$^{+}$) strain carrying the *taaG2* gene that was transcribed at detectable levels by Northern analysis when grown for 36 hrs in medium containing 2% starch plus 2% glycerol. Tani *et al.* (2001a) determined that *amyR* expression was regulated by carbon sources, with the most abundant *amyR* transcription observed when grown on starch, followed by maltose, then glucose and glycerol as determined by Northern analysis. This is in contrast to *amyR* transcription levels in *A. orzyae* which were equal in maltose- and glucose-grown conditions (Gomi *et al.*, 2000). As in *A. oryzae*, *amyR* expression in *A. nidulans* is negatively controlled by CreA (Tani *et al.*, 2001a).

The inability to detect *amyA* transcripts in Northern and RPA experiments carried out in this project is consistent with the difficulties many other groups have faced when studying amylases in *A. nidulans*, which is why the *taaG2* gene from *A. oryzae* has often been used as an amylase reporter gene, as noted above. The growth conditions which resulted in detectable *amyR* transcript by Tani *et al.* (2001a) included a 25 hr incubation at 37°C in media containing 2% carbon source (*eg.* starch), compared to the 16 hr incubation in 1% starch carried out in this project. Recently it was shown that isomaltose, converted from maltose by α–glucosidases, is the most effective inducer for amylase synthesis in *A. nidulans* (Kato *et al.*, 2002b).

The lack of an AmyR binding site in the promoter region of *amyA* is consistent with the low levels of amylase expression seen in *A. nidulans*. As previously mentioned, an amylase cluster also exists in *A. oryzae* (GenBank Accession No. AB021876) with a similar arrangement of genes to the amylase cluster identified here. However, the region between the stop codon of AgdA, and the start codon of AmyA in *A. oryzae* is significantly longer than in *A. nidulans* (2447 bp compared to 388 bp respectively), and there is an AmyR binding site in the promoter region (nt -261) for *amyA* in *A. oryzae* that is missing in *A. nidulans*, indicating that the *A. oryzae* amylase can be induced by the pathway specific regulator AmyR whereas the *A. nidulans amyA* cannot.

The *amyR* disruptant strains showed a similar mutant phenotype to the *malA1* mutant strain of *A. nidulans* (Roberts, 1963), with poor growth on starch and maltose, and Tani *et al.* (2001a) showed that the genetically identified *amyR* and *malA* are the same gene.

The *S. cerevisiae* transcriptional activator for genes involved in maltose utilization, Mal63p, was aligned with AmyR and five homologous regions were identified [Zn (aa 13-54), MH1 (aa 152-214), MH2 (aa 234-375), MH3 (aa 419-496) and MH4 (aa 516-542)] (Tani *et al.*, 2001a). The MH2 region of AmyR is required for the transcriptional activation of the amylase genes, and a truncated AmyR protein that lacked the MH3 and MH4 regions functioned as a constitutive activator, implying this region may be involved in negatively regulating transcription (Tsukagoshi *et al.*, 2001).

Two AmyR binding sites in *A. oryzae* were identified as (5'-CGG-N_8-CGG-3') and (5'-CGGAAATTTAA-3') (Petersen *et al.*, 1999), and presumably these sites will be recognised by AmyR in *A. nidulans*. Tani *et al.* (2001b) identified one putative AmyR binding site in the *agdA* promoter nt -396 to –383 (where the translation start site is +1), and showed that this AmyR binding site is capable of induction, and that AmyR binds to this region. Most $Zn(II)_2Cys_6$ binuclear cluster containing proteins bind as homodimers to the CGG repeats of the binding site, and mutational analyses showed that both CGG triplets in the *agdA* promoter AmyR binding site are required for high level induction by AmyR (Tani *et al.*, 2001b). Tani *et al.* (2001b) failed to recognise a second AmyR binding site in the *agdA* promoter, located at nt –303 to –316, but their results are not compromised by this as the experiments performed did not include this second AmyR binding site.

With the public release of the *A. nidulans* genome in 2003, it made broad searches of the genome with specific domains possible. The glycosyl hydrolase 31 domain of AgdA is found within the genome an additional five times, including AgdB (Kato *et al.*, 2002a). AmyA contains an α–amylase domain, and this domain is shared with AmyB and seven hypothetical proteins. It is interesting to note that two of these hypothetical proteins, AN3309.2 and AN3308.2, physically lie adjacent to each other in the genome. None of the proteins containing an α–amylase domain also contain a starch-binding domain, except for AmyB. Three hypothetical proteins were identified in addition to AmyB that contain the starch-binding domain. AmyR contains both a $Zn(II)_2Cys_6$ domain and a fungal specific transcription factor domain, both common domains within the genome (123 and 110 examples respectively, with a proportion of these containing both, such as FacB).

The amylase cluster identified here contains the transcriptional activator AmyR that has been more thoroughly worked on by Tani *et al.* (2001a, 2001b). Preliminary investigations into amylase regulation had begun when the regulation of/by *amyR* was described by Tani *et al.*, (2001a), and as such no further work on the amylase cluster was performed. The presence of a number of additional amylolytic proteins to those already characterized, based on genome sequence searches, leaves room for the characterization of these amylolytic genes and whether they are functional in *A. nidulans*.

CHAPTER 6: CONCLUSIONS

In the unicellular eukaryote *S. cerevisiae*, carbon catabolite repression involves the DNA binding protein Mig1p (Nehlin and Ronne, 1990), which recruits the Ssn6p-Tup1p general corepressor complex to effect repression (Tzamarias and Struhl, 1995). Mig1p nuclear localization is regulated by the protein kinase Snf1p, which phosphorylates Mig1p in the absence of glucose to trigger nuclear export of Mig1p and relieves repression (Devit *et al.*, 1997). Mig1p shares a high degree of sequence similarity in its DNA binding domain with CreA, the master DNA-binding repressor required for carbon catabolite repression in *A. nidulans*. However, the mechanism of carbon catabolite repression in the multicellular eukaryote *A. nidulans* is quite different, as CreA does not appear to recruit a general corepressor complex, and the Tup1p homologue in *A. nidulans*, RcoA, is only weakly involved in carbon catabolite repression (Hicks *et al.*, 2001), and no highly conserved sequence to Ssn6p is found in the *A. nidulans* genome.

Lockington and Kelly (2002) proposed a model for carbon catabolite repression in *A. nidulans* that includes the transcriptional repressor CreA and a deubiquitinating enzyme complex comprising at least CreB and CreC that are involved in a regulatory de/ubiquitination network. In this model, the CreB/CreC complex acts directly on CreA to affect the stability or activity of CreA, and thus also effects carbon catabolite repression. In the presence of glucose, the coiled-coil region of the CreB deubiquitinating enzyme recognises ubiquitinated CreA as a substrate, and cleaves the ubiquitin chains from CreA. This results in CreA no longer being targeted for degradation via the 26S proteasome or having its function altered, leaving CreA active to bind to the promoter of glucose-repressible genes and effect repression. In the absence of glucose, the levels of CreB are reduced via PEST-mediated degradation of CreB, resulting in CreA remaining ubiquitinated, leading to the degradation of CreA or an alteration of its activity, such that carbon catabolite repression is relieved. Consistent with this, *creA* mRNA levels do not directly correlate with CreA-mediated repression, implying that protein modification and/or stability of CreA could be an important component of the carbon catabolite repression mechanism (Arst *et al.*, 1990; Shroff *et al.*, 1996; Strauss *et al.*, 1999). In addition, Strauss *et al.* (1999) noted that CreA contained a sequence similar to a consensus sequence that triggers ubiquitination in yeast, implying that CreA could be modified by ubiquitination in the absence of glucose, resulting in inactivity/degradation of CreA.

The CreB/CreC complex also plays a role in the turnover of some enzymes and permeases (such as those in the proline and quinate utilization pathways) in both the presence and absence of carbon catabolite repression, as *creB* or *creC* mutations lead to the failure to

express enzymes for the utilization of L-proline (as an example) in both the presence and absence of glucose (Lockington and Kelly, 2001). Regulation of the proline utilization cluster would suggest that this action is likely to be directly on the permeases or enzymes themselves, rather than via the pathway specific regulatory proteins involved in induction, as the proline permease PrnB is directly regulated by CreA (Cubero *et al.,* 2000), whereas the pathway specific transcriptional activator PrnA has constitutive nuclear localisation and is not dependant on proline induction (Pokorska *et al.*, 2000). The quinate permease QutD is another likely target for the CreB/CreC complex, and analysing the stability of these permeases, and whether they are ubiquitinated *in vivo* could help define the role for CreB/ CreC in their regulation.

The CreB/CreC regulatory deubiquitination complex may not act via CreA to affect carbon catabolite repression, but could directly act to stabilize permeases, transporters or glucose sensing proteins by the removal of ubiquitin, thereby altering the cellular concentration or localisation of signalling molecules. Examples of permease and transporter ubiquitination are found in yeast, with the general amino acid permease, Gap1p, targeted by the ubiquitin ligase Npi1p for ubiquitination and subsequent degradation in response to the addition of ammonium (Springael and Andre, 1998; Rotin *et al.*, 2000). Although direct or indirect effects of CreB/CreC on CreA are discussed above as alternatives, both models could apply as they do not conflict with each other. In order to further characterize the roles of CreB and CreC in the cell, we analysed phenotypic suppressors of some or all of the phenotypes of *creB* and *creC* mutant alleles, in order to identify proteins involved in the ubiquitination of substrates that are deubiquitinated by the CreB/CreC complex. Genetic analysis here has shown that the *acrB2* mutation is a suppressor of the phenotypes conferred by mutations in *creB* and *creC*, including the suppression of the derepressed expression of *alcA* (encoding AdhI) in *creB* and *creC* mutant strains, as judged by growth on medium containing allyl alcohol. These phenotypes are consistent with a regulatory defect associated with the membrane and signalling, but not with a simple interpretation that only permeases and transporters are affected, as for ethanol, a fat soluble compound, no permease is required or exists.

The *creD34* mutation arose as a spontaneous phenotypic suppressor of *creC27* that led to tighter repression of enzymes subject to carbon catabolite repression in the presence of glucose, and was also shown to suppress mutations in *creB* and, partially, in *creA* (Kelly and Hynes, 1977). These epistatic interactions of *creD* with *creA*, *creB*, and *creC* implicated *creD* as being involved in this regulatory network in an opposing process to deubiquitination, such as ubiquitination. *creD* was cloned and characterized here, and encodes a protein that contains an arrestin_N and arrestin_C domain, a PP*X*Y motif and two P*X*Y motifs, and is highly similar to the

Rod1p and Rog3p proteins from *S. cerevisiae*. The presence of two such similar sequences to *creD* in the *S. cerevisiae* genome led to the search for another arrestin domain containing protein in *A. nidulans*, which identified a gene we have designated *apyA*, encoding an arrestin_N and arrestin_C domain and PPXY motif containing protein. In *S. cerevisiae* Rod1p and Rog3p have been shown to interact with the ubiquitin ligase Rsp5p, and so to determine whether this interaction occurred in *A. nidulans* the homologue of Rsp5p was identified, and the gene encoding this HECT ubiquitin ligase was designated *hulA*. CreD and ApyA were shown to interact with HulA via the bacterial-2-hybrid system, with a strong interaction between ApyA and HulA, and a weaker but clear interaction between CreD and HulA. If CreB/CreC are involved in the deubiquitination of CreA, among other substrates, then CreD/ApyA/HulA are clearly involved in the ubiquitination aspect of this network. The epistatic effects of *creD* with *creB* and *creC* are consistent with this model, whereby the lack of ubiquitination partially compensates for an inability to deubiquitinate given targets. The mutant phenotype of *creD*, of even tighter repression of enzymes subject to carbon catabolite repression, could be explained according to the proposed model of Lockington and Kelly (2002), whereby the decreased ubiquitination of CreA in a *creD* mutant background results in more efficient deubiquitination of CreA by CreB in the presence of glucose compared to wild-type, resulting in an increase of stabilized CreA to effect repression. To further determine the role of the newly identified ApyA and HulA proteins in this system, these genes should be disrupted and their phenotypes analysed to see if carbon catabolite repression is perturbed. As ApyA interacts strongly with the ubiquitin ligase HulA, it would appear that ApyA is involved in ubiquitination and like *creD*, would be capable of suppressing the phenotypic effects of the *creB* and *creC* mutations. CreD may interact more strongly with another of the six HECT ubiquitin ligases found in *A. nidulans* than with HulA, and interactions with these other ubiquitin ligases should be tested, starting with HulE, which is the only HECT ubiquitin ligase not to have a clear orthologue in *S. cerevisiae*.

Arst (1981) observed that the phenotype of acriflavine resistance and molybdate sensitivity of a strain containing the *creD34* mutation was similar to that of strains containing mutations in *acrB*. Mutations in *acrB* had not previously been tested for their effects on carbon catabolite repression, and we found that like *creD34*, *acrB2* can suppress the derepressed phenotype of mutations in *creC*, *creB* and partially in *creA*, implicating AcrB in the ubiquitination aspect of carbon catabolite repression. AcrB was cloned and contains three membrane spanning domains and a coiled-coil region, and thus may function in signal transduction (Boase *et al.*, 2003). AcrB was not highly similar to any known protein, and we failed to detect any protein-protein interactions using the BacterioMatch I bacterial-2-hybrid system to screen an *A. nidulans*

cDNA library, nor in direct testing of AcrB with CreA, CreB, CreC, CreD, ApyA and HulA. Whilst these results indicate the lack of a strong interaction of AcrB with these proteins, these experiments need to be repeated with the more sensitive BacterioMatch II reporter cells to identify any proteins that interact with AcrB, possibly through the coiled-coil region. AcrB mutants grew poorly on a range of sole carbon sources and on some ω-amino acids, indicating an inability to derepress the enzymes required for the utilization of these sole carbon sources in the absence of glucose, possibly due to a lack of ubiquitination. To identity the role of AcrB, overexpression of AcrB could be investigated by transforming a construct containing *acrB* under the control of the high-level constitutive promoter *gpdA* into *A. nidulans.* Similarly, the phenotypic effects on carbon catabolite repression caused by increased concentrations of CreD in the cell could also be determined in this manner. However, overexpression of mRNA may not translate to altered protein levels, and as such Westerns would be required to establish protein levels. Presumably increasing the amount of ubiquitination under carbon repressing conditions would result in a decreased amount of stabilized CreA to effect repression.

De/ubiquitination networks are emerging as an essential control mechanism in the regulation of development. A well-studied case of regulation via the de/ubiquitination of substrates is that of eye development in Drosophila. E3 ubiquitin ligases have a number of roles during Drosophila eye development, including control of cell proliferation, specification, differentiation and death. For the correct development of photoreceptors to occur in the developing eye of *D. melanogaster*, the Fat Facets deubiquitination enzyme is required to deubiquitinate epsin (encoded by the *liquid facets* gene), thereby stabilising epsin and promoting endocytosis, which prevents the misspecification of photoreceptor cells in the developing eye (Chen and Fischer, 2000; Cadavid *et al.*, 2000, Chen *et al.*, 2002). In the absence of Fat facets, epsin remains ubiquitinated and is degraded, promoting cells to inappropriately adopt a photoreceptor fate. A mouse homologue of Fat facets, Fam, has been identified and it has been suggested to play a role in mouse development via its deubiquitination of epsin (Chen *et al.*, 2002; Oldham *et al.*, 2002). This is one of the few cases where the deubiquitination enzyme, the ubiquitinated substrate, and the role they play are known, although the E3 ligase that targets epsin for ubiquitination is yet to be determined.

There have been very few cases of mutations with defined phenotypes affecting both the ubiquitination and deubiquitination pathways of the same regulatory ubiquitination system. The evidence reported here implicates CreD, AcrB, ApyA and HulA in the ubiquitination aspect of the regulatory network including the CreB/CreC deubiquitinating complex that is involved in carbon catabolite repression. CreD, ApyA and HulA, have homologues in yeast, namely Rod1p/Rog3p

and Rsp5p, whilst the deubiquitination aspect of this network, the CreB and CreC complex, do not have any highly similar sequences in yeast, but do in other filamentous fungi and higher eukaryotes such as humans, again highlighting the differences between carbon catabolite repression in *A. nidulans* compared to *S. cerevisiae.* The linking of deubiquitination and the regulation of carbon metabolic pathways in *A. nidulans* has allowed us to study the key regulatory role of de/ubiquitination in this process. Ubiquitin and the proteasome are being increasingly implicated in gene control, including the control of transcriptional activators by regulating their cellular location, their association with other proteins, and their stability and thus abundance. The clear mutant phenotypes involving *creD, acrB, creB and creC* represents a chance to identify roles and substrates of a regulatory ubiquitination/deubiquitination network that may well extend beyond carbon catabolite repression in *A. nidulans.*

Chapter 6: Conclusions

APPENDIX A

The pB3 plasmid partially complemented the *creD34* mutant phenotype, and in an attempt to identify *creD* and the nature of the *creD34* mutation, the pB3 region was sequenced from both a wild-type strain and a *creD34* mutant strain (see Section 3.2.3). Custom oligonucleotides were designed to PCR amplify this region in sections, and to sequence these sections (see Table A.1). The amplified sequences of individual regions were aligned to give the complete sequence of pB3 (Figure A.1). The pB3 region did not contain the *creD34* mutation.

Table A.1 Custom oligonucleotides designed for the pB3 plasmid

PRIMER NAME	PRIMER SEQUENCE (5' TO 3')
D1	CCCAGTCCACCAATCCTCC
D2	ATTGGCGTCCCTACGTTCG
D3	CGCTGAAGCCGCTGGTGC
D4	CGTAATGTCAACGAGGCC
D5	TGGAACTTCAGGTAACGC
D6	GAATGTGCGGACGCGGCG
D7	TGTTATGATCTGGCGACG
D8	CAGATGGCTACAGGTTGG
D9	CCTTTGCATCGAATAGGG
D10	GCCTCCGTTTGCATCAGC
D11	CTGATCGAGTAGCTTCGC
D12	GGTTGAGCTTCCAGGCGG
D13	CATTACAGCATGGACGCC
D14	GGAACTTGGTGTGGCTGG
D15a	CCCAGCCTTCCTCGCCTC
D15b	GGCATCGTTGAGTCCACC
D16	TCATGGTCCGCGCCGTCC
D17	TCCATGGTATCCTGTCCG
D18	TGATGTGGTCTGGGAGGG
D19	ATCAGGACGGTGACTTGG
D20	GAGGTGCTGATAGGACGG
D21	TAGCGTAAGGGTTCCGGC
D22	AGAACCCGGCGAGCGTGC
D23	AATCTGGAGGCGACGAGG
D24	CAGAGCATATCGTTACGC
D25	TGAGTCGTTCGGTTCGGC
D26	GGGAGGTAATAAGGCTGG

```
1     GAATTCGCGGCCGCGGTTTGTTCGAGTTCGCCGACTGTTTGAGAGACGGCACGCAAGGAAATAAGAGTG
70    AGCTAACCAGAAATGGATTCACAGACGACACGTGCAATGCCGGAGACATTCAGATTGTAGATGAAGTAG
139   TGAGATGGCGTTAAAGACAAAAGGCGTTCAAGTTCACAGGAAGGGGAGAGGAGTCTATCGACTTCCAAG
208   CCAACGGGCAGCGGCAGTGTGGAATTCCGGGGAACTTGGTGTGGCTGGTTGGTGGAACTTCTGCGTGGG
277   GCAAGCTACCGATGGCGCCGACATAGTACATACATGCACTGACATGCCAGGAAACTATACATAGAGGCT
346   CGTGCTTCACATTTGCCATTTTAGACTTAATATTCCATTCTGGAGTGACAAAGTGGACGCCTGTGTCAG
415   GTTCTTAACCTTAATGGCTGTCTATTCGACAAAGTGCCTGTAGCTCAACACATTGCGAAGATTTCATAT
484   ACATTGCTCGGTATTCGCAATAAATCTATCCAAAGCAACGCAATAATTGATTTCAGGGGTATCGAGTTC
553   CATCATTACAGCATGGACGCCCTCAACCAGGCGTCAGACTCCACTACCACTGACTCGTCGCCAGATCAT
                                                   *  V  G  S  G  S  V  R  R  W  I  M   623
622   AACAGCAGCAGCTCCCGCACAAACCACAGCTCCGAGAGCAATTCCCCACGACATCCATTTGCCCTCGAC
        V  A  A  A  G  A  C  V  V  A  G  L  A  I  G  W  S  M  W  K  G  E  V   600
691   GTCTTCCTCGCTCCGAGGGAATGCTGTACTCGGTTCCATTGCGGCTTCTGCGATGTGCACAAACACCTC
        D  E  E  S  R  P  F  A  T  S  P  E  M  A  A  E  A  I  H  V  F  V  E   577
760   TTGAATGGAACTCCAGGTAACGCTCACATGAAGCGGCGGACCAGGCATATTCAATAGAGCTGTGTATTC
        Q  I  S  S  W  T  V  S  V  H  L  P  P  G  P  M  N  L  L  A  T  Y  E   554
829   ATGAGGCTGATGTTCAGCGCGCTGGGTTGTACGGTCTAGGTCAGCCTTGAGTGCGATGTACACACTCGG
        H  P  Q  H  E  A  R  Q  T  T  R  D  L  D  A  K  L  A  I  Y  V  S  P   531
898   AAGCTCCTCCAGGTGTGGATACTTCGCCCTGAGAGCGGGTATATATGCGAATGAATCGGGGTCTGACGA
        L  E  E  L  H  P  Y  K  A  R  L  A  P  I  Y  A  F  S  D  P  D  S  S   508
967   GTCATATGTGTATACGATCACATCACACTGATCGAGTAGCTTCGCTTGATTTTCAAGAATCGCAGGTTC
        D  Y  T  Y  V  I  V  D  C  Q  D  L  L  K  A  Q  N  E  L  I  A  P  E   485
1036  TAACTCGCCGAGCTCATCCATGATCAGGTAACATTGTTTTCCGCCTGGAAGCTCAACCGTGTTGACTGC
        L  E  G  L  E  D  M  I  L  Y  C  Q  K  G  G  P  L  E  V  T  N  V  A   462
1105  GGTACGCGGTTGAATTGTGGGGTGATAAGTGGTACTAAATCCGCGGGACAGAAAAGCATCGAGAAGGGC
        T  R  P  Q  I  T  P  H  Y  T  T  S  F  G  R  S  L  F  A  D  L  L  A   439
1174  AGATTTGCCAGACCCCGGAGCTCCGACTATGTGGCCAAGGACGACATTGCGACCGACGCGTCCTGGACG
        S  K  G  S  G  P  A  G  V  I  H  G  L  V  V  N  R  G  V  R  G  P  R   416
1243  TTTTCTCCGCTTACGGGGTCGTGTTACCTTAAGAGCCGCGGTCGTGGAGGGGTTGCTCCGATCTGAAGA
        K  R  R  K  R  P  R  T  V  K  L  A  A  T  T  S  P  N  S  R  D  S  S   393
1312  CTCGAATCCCAAGTACGCTAGATATTCCAGAGTCGTCTTGGGAGAAGTAAAGGTTGTCATACTCCACTG
        E  F  G  L  Y  A  L  Y  E  L  T  T  K  P  S  T  F  T  T  M  S  W  Q   370
1381  TGCTAGCCAGCCCTGAAGAGTGACATGGCCAGCCTCGTTTCTGACAGTGCATGAAGGAAATGAGCCGTC
        A  L  W  G  Q  L  T  V  H  G  A  E  N  R  V  T  C  S  P  F  S  G  D   347
1450  AGCCCAAGAGGCGGGCAATCCAGGGGTGGGTGCGAACAATGATGCTAACTCGGCATCGTTGAGTCCACC
        A  W  S  A  P  L  G  P  T  P  A  F  L  S  A  L  E  A  D  N  L  G  G   324
1519  ATCATTGTCTTTGTCTGAGAGAAGGAAAAGGTTGACAAAGAACCGATAGCCCTCAGGAGACAATTCCGC
        D  N  D  K  D  S  L  L  F  L  N  V  F  F  R  Y  G  E  P  S  L  E  A   301
1588  TGATGCAAACGGAGGCACCTCAAACTTTGGGTGGAGATAGCTCTCTTGCAACGAAAGGCTATCGGTGTA
        S  A  F  P  P  V  E  F  K  P  H  L  Y  S  E  Q  L  S  L  S  D  T  Y   278
1657  CTGAAATGCGCGAAGTATGATCCAAACTGTTTCATGGCGTCCTTTCTCGGCATACATCTTGTTCAGATG
        Q  F  A  R  L  I  I  W  V  T  E  H  R  G  K  E  A  Y  M  K  N  L  H   255
1726  AATAAATCCCCGGCAATCTATCCCGGACGGCGTCACAGAGTCAGGATGTGTTTTCTGAATTGTCTCTTT
        I  F  G  R  C  D  I  G  S  P  T  V  S  D  P  H  T  K  Q  I  T  E  K   232
1795  TATATGGACCAAGTCTTCTTCGCTCAGAGGTTTCTCAAAACACCTCATCTGAAAATCCTTTATCTCTTT
        I  H  V  L  D  E  E  S  L  P  K  E  F  C  R  M  Q  F  D  K  I  E  K   209
1864  ATCTGAGAGATAACCGTCCCGGTCTTTATCACTGAGGTAGAAGATGCGTTGTAACGCGGCGACAGCCGC
        D  S  L  Y  G  D  R  D  K  D  S  L  Y  F  I  R  Q  L  A  A  V  A  A   186
1933  CGGTTTTAAGGCTGATTCCTTTGCATCGAATAGGGGTGCGATAGGGTGTGTGACCGCTTTCTGGCAAAG
        P  K  L  A  S  E  K  A  D  F  L  P  A  I  P  H  T  V  A  K  Q  C  L   163
2002  GAAAAAGGCCTCGTTGACATTACGATGTTCTCGGGCACTGGTCCGAATGCATGAATCGATTTCCTTGAA
        F  F  A  E  N  V  N  R  H  E  R  A  S  T  R  I  C  S  D  I  E  K  F   140
2071  CTCTGACATCAACGGCAACATCTCGTCCTCGATGACTTGCGTCTCGGTATGATCGGCAGCAAGATCGGA
        E  S  M  L  P  L  M  E  D  E  I  V  Q  T  E  T  H  D  A  A  L  D  S   117
2140  CTTGTTGGCACACAGCACGACAGGCACATTCACGCCGAGAGAGCGGAAATAGGGCAGCCAGAAAAGCGC
        K  N  A  C  L  V  V  P  V  N  V  G  L  S  R  F  Y  P  L  W  F  L  A    94
2209  AACCCTCTCATAGCTATAGTGATCAGAATAGACAAGCAATATCACGTTCGATTTCCGAATTTCTCTAGC
        V  R  E  Y  S  Y  H  D  S  Y  V  L  L  I  V  N  S  K  R  I  E  R  A    71
2278  CAAGTTGCTCCGTTCTTGCGGCACAGCAGAGGTATCGACCACCGTTGTCGTTGTAACATTCTCGGGGGT
        L  N  S  R  E  Q  P  V  A  S  T  D  V  V  T  T  T  T  V  N  E  P  T    48
2347  TCCAATGGTAGGAGGGATAGTGATCTGTGGCAGAATGGGCTGGATCTTGTTCGTCACAAACACGCCTTT
        G  I  T  P  P  I  T  I  Q  P  L  I  P  Q  I  K  N  T  V  F  V  G  K    25
2416  TACGAGGGAGGTAATAAGGCTGGACTTTCCAGTTCCCTCATCCCCGCAAACACAAATTCGCActagatg
        V  L  S  T  I  L  S  S  K  G  T  G  E  D  G  C  V  C  I  R             6
2485  gtgtcagtatctttgttctttcgtggacgcctcaacaaaagagagctccaacCTGTAGCCATCTGAACA
                                                        V  T  A  M             1
2554  AATAAAGAAGCACTTGAAGAAGCGAATTCTAAGGCCCTCTGAGATCATTGCTGAACTTGGAGTACACCC
2623  TACACCGCCATGAGGCCAGGAGGCTGGCGCGTCAGTTCAGGTTAGGAACCAGAGCAGCTAGCAAGTCCG
2692  ACGGAAGTTTGGGCGATCGCCGCGGTCTTGTGGCGGTGCGATGCGTTTAAGCCTCAGGCCTCCAAATGG
2761  ACAAGCTTGACACACATTCGGGACCAACATGAGCCTTGGTCTCAGTGGGTGACCAGCATACTGCAACAG
2830  AGCATATCGTTACGCTATTTCAATATGGGACTAGATTAATCTGTATTCTGTACCACATGCCGCGCCCGC
2899  CCAAAGATGATGTAACGGAGTTAGCCAAGTGCCGAACCGAACGACTCAGCCACAATCTGGGCCATCTAC
```

```
2968 CCGATTTTACGAGAAGCTTATGGACACCACTTCGTAGTTGCTATCCTTAAGCATAGCTGCTTGAAGACT
3037 GTATCGCCATCTGTCTATAGTTCCCCCTGTGAAGCGCAGTACTGAACGGGGAAATCAGCGGCTCAGCCT
3106 AGCATCCATTCTTAGCTCCACTTCTAGGCCCTCCCATTTGGCAAGGCCTAACGTCGAACATAATATTAC
3175 AGTTTCGACAACGTTCGAAACTCTCCCCAGCCTTCCTCGCCTCTTTACAACAACCATTCCCTCAAGAAC
3244 CCGGCGAGCGTGCGCAAGATGCCTCTTGGAATTCACAACCCTCTCCCTTCGTCGCTATCGAgtaagtcc
                       M  P  L  G  I  H  N  P  L  P  S  S  L  S  S          15
3313 tcgtcgcctccagatttccccgcagttacccacaggcgacaacttgggcgtgagctgactgctaatcct
3382 caatttctctcggtagGCGAATGCAAAAAGGCTGGCAAAATCCTAGCATCTTTTGTTGACCCGCGGCAA
                       E  C  K  K  A  G  K  I  L  A  S  F  V  D  P  R  Q    32
3451 GCATTCGGGCCTGACAAGGTCATCCCCCCTGAGGTATTAGCTGGTGCTAAGgtacgcaaaggatactcg
     A  F  G  P  D  K  V  I  P  P  E  V  L  A  G  A  K                      49
3520 gtgaagctgatagctgaattatggctgattgcttggcttctagGGTCTTGCCATTTTGACCGTCCTGAA
                                                G  L  A  I  L  T  V  L  K    58
3589 AGCAGGGTTTTTGGGTTCCGGTCGATTTGGTTCGGGTATCGTCGTTGCTCGACTAGCTGATGGATCATG
      A  G  F  L  G  S  G  R  F  G  S  G  I  V  V  A  R  L  A  D  G  S  W    81
3658 GTCCGCGCCGTCCGCGATCGCAACCGCAGGTGCTGGATTTGGAGGTCAGATTGGATTTGAGTTGACCGA
      S  A  P  S  A  I  A  T  A  G  A  G  F  G  G  Q  I  G  F  E  L  T  D   104
3727 TTTCGTCTTCATCTTGAACGATGCTGCCGCCGTCCGCACATTCTCTCAAGCCGGAACCCTTACGCTAGG
      F  V  F  I  L  N  D  A  A  A  V  R  T  F  S  Q  A  G  T  L  T  L  G   127
3796 AGGTAATGTTTCGATTGCCGCTGGGCCGGTTGGTCGAAACGCTGAAGCCGCTGGTGCTGCTAGTACGAA
      G  N  V  S  I  A  A  G  P  V  G  R  N  A  E  A  A  G  A  A  S  T  K   150
3865 GGGTGTGGCCGCTGTATTCTCATATTCCAAAACCAAGGGACTTTTCGCCGGTGTCAGCTTAGAAGGAAG
      G  V  A  A  V  F  S  Y  S  K  T  K  G  L  F  A  G  V  S  L  E  G  S   173
3934 CATGCTGGTCGAACGTAGGGACGCCAATGAGAGACTCTACAATAGCCGAGTATCTGCTCGCCAGCTCCT
      M  L  V  E  R  R  D  A  N  E  R  L  Y  N  S  R  V  S  A  R  Q  L  L   196
4003 CAGTGGTACTATCCCGCCTCCACCCGCCGCCGAACCTCTAATGCGTGTGCTGAACTCCCGCGCTTTCTA
      S  G  T  I  P  P  P  P  A  A  E  P  L  M  R  V  L  N  S  R  A  F  Y   219
4072 CGGCGTGCGCACAAACGGCGACTCTATGTACAACGATATCCCCGTCTACGATGATCGTCATGATGATGT
      G  V  R  T  N  G  D  S  M  Y  N  D  I  P  V  Y  D  D  R  H  D  D  V   242
4141 GGTCTGGGAGGGTCGCAGGGGAGAGGCATATGGCGAGGGGATCAGACGTGATCGGACAGGATACCATGG
      V  W  E  G  R  R  G  E  A  Y  G  E  G  I  R  R  D  R  T  G  Y  H  G   265
4210 ACCTTCAGATGACTACGAATACCACGACCGGCCGCGCCGTGCAACGACGTGGGCTGACGATATATATGA
      P  S  D  D  Y  E  Y  H  D  R  P  R  R  A  T  T  W  A  D  D  I  Y  D   288
4279 TCGTCCTGCTGGAGGCTTGAGTCGTTCCTCCACCGCCCGATTTAGCAGCCGCAACGACACCTTTGATAC
      R  P  A  G  G  L  S  R  S  S  T  A  R  F  S  S  R  N  D  T  F  D  T   311
4348 CTACAATAGACAGCGGAGCAACACCTATGATGATGATTATGTGTACTCTGACCGCAAACCCAGTCGTCC
      Y  N  R  Q  R  S  N  T  Y  D  D  D  Y  V  Y  S  D  R  K  P  S  R  P   334
4417 TACAGCACCCAAGCCCgtttcgacagcgcactggcaggccgtccgttacgcgagGACCAGGCCATTGCG
      T  A  P  K  P                                        D  Q  A  I  A    344
4486 CTGTACACATTCGACGCGGATCAGGACGGTGACTTGGGTTTCAAGAAGGGGGAGATTATCACAATCATC
     L  Y  T  F  D  A  D  Q  D  G  D  L  G  F  K  K  G  E  I  I  T  I  I    367
4555 AAACGCACCGAAAAGAAGGAGGATTGGTGGACTGGGCGGATCGGTGACCGCGTCGGCATATTCCCTGCG
     K  R  T  E  K  K  E  D  W  W  T  G  R  I  G  D  R  V  G  I  F  P  A    390
4624 TAAGATACCTTGCTCTGAAATGGCTGGACTGCTTGCTAACTCAGTGCCTTTAGGAATTATGTTGACGCA
     *                                                                      391
4693 GCTTAAGACTTTCGCCCATGACCATTTATTTATATGCAACGAAATTCATGACCTTGCCCGCTTATACCT
4762 CTCCTTTGGCAATTTTCATTTATACCGATCGAGTGTGTTTCAGTCTTTTTTTGGAATCCTTGGGCATTT
4831 TATCCCATCTCGTCCGAGCTTCTCTTTTCAGCTCTTGCGTAAGCATTGGAGTTTGAAGGTGGTTTTACG
4900 AGTGCGTACAGGCGAAAGAGAATTCGGCGCTTCTCAGAATATCAAGTTCAAATTCCCCTGCCTTCCCAC
4969 TTATTATCCGTCCTATCAGCACCTCTAGCAGCTCTTGTGATTATTATTACTGTCCAATATACAGGTTTT
5038 CTTTGATACAGGCTTGTCTCGGAGAAATTAACATTAAAGAGGTAGATAAGCGTCATCTGCAGGTTATTC
5107 TCGCAGTCGTTGAGGCTGGGACCCGTCACCTGCAAGGAGAGGAGATCT
```

Figure A.1 Nucleotide sequence of the pB3 plasmid. The amino acid sequences of the Gon1p-like protein, highlighted in grey, and the Ysc84p-like protein are shown in single letter code beneath the DNA sequence, with numbered amino acid residues indicated on the right-hand side, and numbered nucleotide positions indicated on the left-hand side. *EcoRI* sites are underlined and bolded. Putative introns are in lower case letters.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Andoh, T., Hirata, Y. and Kikuchi, A. (2002). PY motifs of Rod1 are required for binding to Rsp5 and for drug resistance. *FEBS letters* **525**, 131-134.

Arst, H. N. (1981). Aspects of the control of gene expression in fungi. *Symposia of the Society for General Microbiology* **31**, 131-160.

Arst, H. N. and Cove, D. J. (1973). Nitrogen metabolite repression in *Aspergillus nidulans*. *Molecular and General Genetics* **126**, 111-141.

Arst, H. N., Brownlee, A. G. and Cousen, S. A. (1982). Nitrogen metabolite repression in *Aspergillus nidulans*: A farewell to *tamA*? *Current Genetics* **6**, 245-257.

Arst, H. N., MacDonald, D. W. and Cove, D. J. (1970). Molybdate metabolism in *Aspergillus nidulans*. I. Mutations affecting nitrate reductase and-or xanthine dehydrogenase. *Molecular and General Genetics* **108**, 129-145.

Arst, H. N., Tollervey, D., Dowzer, C. E. A. and Kelly, J. M. (1990). An inversion truncating the *creA* gene of *Aspergillus nidulans* results in carbon catabolite derepression. *Molecular Microbiology* **4**, 851-854.

Bailey, C. and Arst, H. N. (1975). Carbon catabolite repression in *Aspergillus nidulans*. *European Journal of Biochemistry* **51**, 575-577.

Boase, N. A., Lockington, R. A., Adams, J. R. J., Rodbourn, L. and Kelly, J. M. (2003). Molecular characterization and analysis of the *acrB* gene of *Aspergillus nidulans*: A gene identified by genetic interaction as a component of the regulatory network that includes the CreB deubiquitination enzyme. *Genetics* **164**, 95-104.

References

Boel, E., Brady, L., Brzozowski, A. M., Derewenda, Z., Dodson, G. G., Jensen, V. J., Petersen, S. B., Swift, H., Thim, L. and Woldike, H. F. (1990). Calcium binding in α-amylases: An X-ray diffraction study at 2.1-A resolution of two enzymes from *Aspergillus. Biochemistry* **29**, 6244-6249.

Bork, P. and Sudol, M. (1994). The WW domain: a signalling site in dystrophin? *Trends in Biochemical Sciences* **19**, 531-533.

Brody, H., Griffith, J., Cuticchia, A. J., Arnold, J. and Timberlake, W. E. (1991). Chromosome-specific recombinant DNA libraries from the fungus *Aspergillus nidulans. Nucleic Acids Research* **19**, 3105-3109.

Cadavid, A. L. M., Ginzel, A. and Fischer, J. A. (2000). The function of the Drosophila Fat facets deubiquitinating enzyme in limiting photoreceptor cell number is intimately associated with endocytosis. *Development* **127**, 1727-1736.

Celenza, J. L. 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. L. and Carlson, M. (1989). Mutational analysis of the *Saccharomyces cerevisiae* Snf1 protein kinase and evidence for interaction with the Snf4 protein. *Molecular and Cellular Biology* **9**, 5034-5044.

Chang, A., Cheang, X. E. and Sudol, M. (2000). Rsp5 WW domains interact directly with the carboxyl-terminal domain of RNA Polymerase II. *The Journal of Biological Chemistry* **275**, 20562-20571.

Chen, H. and Sudol, M. (1995). The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 7819-7823.

Chen, H. I., Einbond, A., Kwak, S. J., Linn, H., Koepf, E., Peterson, S., Kelly, J. W. and Sudol, M. (1997). Characterization of the WW domain of human Yes-associated protein and its polyproline-containing ligands. *The Journal of Biological Chemistry* **272**, 17070-17077.

Chen, X. and Fischer, J. A. (2000). *In vivo* structure/function analysis of the Drosophila *fat facets* deubiquitinating enzyme gene. *Genetics* **156**, 1829-1836.

Chen, X., Zhang, B. and Fischer, J. A. (2002). A specific protein substrate for a deubiquitinating enzyme: Liquid facets is the substrate of Fat facets. *Genes and Development* **16**, 289-294.

Chikamatsu, G., Shirai, K., Kato, M., Kobayashi, T. and Tsukagoshi, N. (1999). Structure and expression properties of the endo-β-1,4-glucanase A gene from the filamentous fungus *Aspergillus nidulans. FEMS Microbiology Letters* **175**, 239-245.

Clutterbuck, A. J. (1974). *Aspergillus nidulans* genetics. In *Handbook of Genetics*, pp. 447-510. Edited by R. C. King. New York: Plenum Publishing.

Clutterbuck, A. J. (1993). Fungi: *A. nidulans* (nuclear genes). In *Genetic Maps: Locus Maps of Complex Genomes*. Edited by S. J. Brien: CSH.

Clutterbuck, A. J. (1997). The validity of the *Aspergillus nidulans* linkage map. *Fungal Genetics and Biology* **21**, 267-277.

Cove, D. J. (1966). The induction and repression of nitrate reductase in the fungus *Aspergillus nidulans. Biochimica et Biophysica Acta* **133**, 51-56.

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. The European Molecular Biological Organisation Journal* **13**, 407-415.

Cubero, B., Gomez, D. and Scazzocchio, C. (2000). Metabolite repression and inducer exclusion in the proline utilization gene cluster of *Aspergillus nidulans*. *Journal of Bacteriology* **182**, 233-235.

Cziferszky, A., Mach, R. L. and Kubicek, C. P. (2002). Phosphorylation positively regulates DNA binding of the carbon catabolite repressor Cre1 of *Hypocrea jecorina* (*Trichoderma reesei*). *The Journal of Biological Chemistry* **277**, 14688-14694.

De Vit, M. J. and Johnston, M. (1999). The nuclear exportin Msn5 is required for nuclear export of the Mig1 glucose repressor of *Saccharomyces cerevisiae*. *Current Biology* **9**, 1231-1241.

De Vit, M. J., Waddle, J. A. and Johnston, M. (1997). Regulated nuclear translocation of the Mig1 glucose repressor. *Molecular Biology of the Cell* **8**, 1603-1618.

De Winde, J. H., Crauwels, M., Hohmann, S., Thevelein, J. M. and Winderickx, J. (1996). Differential requirement of the yeast sugar kinases for sugar sensing in establishing the catabolite-repressed state. *European Journal of Biochemistry* **241**, 633-643.

Dowzer, C. E. A. and Kelly, J. M. (1989). Cloning of *creA* from *Aspergillus nidulans*: a gene involved in carbon catabolite repression. *Current Genetics* **15**, 457-459.

Dowzer, C. E. A. and Kelly, J. M. (1991). Analysis of the *creA* gene, a regulator of carbon catabolite repression in *Aspergillus nidulans*. *Molecular and Cellular Biology* **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.

Edmondson, D. G., Mitchell Smith, M. and Roth, S. Y. (1996). Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. *Genes and Development* **10**, 1247-1259.

Errat, J. A., Douglas, P. E., Moranelli, F. and Seligy, V. L. (1984). The induction of α-amylase by starch in *Aspergillus oryzae*: evidence for controlled mRNA expression. *Canadian Journal of Biochemical and Cell Biology* **62**, 678-690.

Fitch, W. M. (1971). Towards defining the course of evolution: Minimum change for a specific tree topology. *Systematic Zoology* **20**, 406-416.

Flipphi, M., van de Vondervoort, P. J. I., Ruitjer, G. J. F., Visser, J., Arst, H. N. and Felenbok, B. (2003). Onset of carbon catabolite repression in *Aspergillus nidulans*. *The Journal of Biological Chemistry* **278**, 11849-11857.

Gines, M., Dove, M. J. and Seligy, V. L. (1989). *Aspergillus oryzae* has two nearly identical Taka-amylase genes, each containing eight introns. *Gene* **79**, 107-117.

Gomi, K., Akeno, T., Minetoki, T., Ozeki, K., Kumagai, C., Okazaki, N. and Iimura, Y. (2000). Molecular cloning and characterization of a transcriptional activator gene, *amyR*, involved in the amylolytic gene expression in *Aspergillus oryzae*. *Bioscience, Biotechnology and Biochemistry* **64**, 816-827.

Gough, J. A. and Murray, N. E. (1983). Sequence diversity among related genes for recognition of specific targets in DNA molecules. *Journal of Molecular Biology* **166**, 1-19.

Gurr, S. J., Unkles, S. E. and Kinghorn, J. R. (1987). The structure and organisation of nuclear genes of filamentous fungi. In *Gene structure in eukaryotic microbes*, pp. 93-139. Edited by J. R. Kinghorn. Oxford: IRL press.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* **166**, 557-580.

Herrero, P., Galindez, J., Ruiz, N., Martinez-Campa, C. and Moreno, F. (1995). Transcriptional regulation of the *Saccharomyces cerevisiae HXK1*, *HXK2* and *GLK1* genes. *Yeast* **11**, 137-144.

Herrero, P., Martinez-Campa, C. and Moreno, F. (1998). The hexokinase 2 protein participates in regulatory DNA-protein complexes necessary for glucose repression of the *SUC2* gene in *Saccharomyces cerevisiae. FEBS letters* **434**, 71-76.

Hicks, J., Lockington, R. A., Strauss, J., Dieringer, D., Kubicek, C. P., Kelly, J. and Keller, N. (2001). RcoA has pleiotropic effects on *Aspergillus nidulans* cellular development. *Molecular Microbiology* **39**, 1482-1493.

Hynes, M. J. and Pateman, J. A. (1970). The genetic analysis of regulation of amidase synthesis in *Aspergillus nidulans. Molecular and General Genetics* **108**, 95-106.

Hynes, M. J. and Kelly, J. M. (1977). Pleiotropic mutants of *Aspergillus nidulans* altered in carbon metabolism. *Molecular and General Genetics* **150**, 193-204.

Imhof, M. O. and McDonnell, D. P. (1996). Yeast RSP5 and its human homolog hRPF1 potentiate hormone-dependent activation of transcription by human progesterone and glucocorticoid receptors. *Molecular and Cellular Biology* **16**, 2594-2605.

Innis, M. A. and Gelfand, D. H. (1990). Optimisation of PCRs. In *PCR protocols: A guide to methods and applications.* Edited by M. A. Innis, D. H. Gelfand, J. S. Sninsky and T. J. White. New York: Academic Press Inc.

Jiang, R. and Carlson, M. (1996). Glucose regulates protein interactions within the yeast Snf1 protein kinase complex. *Genes and Development* **10**, 3105-3115.

Jiang, R. and Carlson, M. (1997). The Snf1 protein kinase and its activating subunit, Snf4, interact with distinct domains of the Sip1/Sip2/Gal83 component in the kinase complex. *Molecular and Cellular Biology* **17**, 2099-2106.

Kaneko, A., Sudo, S., Takayasu-Sakamoto, Y., Tamura, G., Ishikawa, T. and Oba, T. (1996). Molecular cloning and determination of the nucleotide sequence of a gene encoding an acid-stable α-amylase from *Aspergillus kawachii. Journal of Fermentation and Bioengineering* **84**, 292-298.

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. *Bioscience, Biotechnology and Biochemistry* **60**, 1776-1779.

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. *Molecular and General Genetics* **254**, 119-126.

Kato, N., Murakoshi, Y., Kato, M., Kobayashi, T. and Tsukagoshi, N. (2002b). Isomaltose formed by α-glucosidases triggers amylase induction in *Aspergillus nidulans*. *Current Genetics* **42**, 43-50.

Kato, N., Suyama, S., Shirokane, M., Kato, M., Kobayashi, T. and Tsukagoshi, N. (2002a). Novel α-glucosidase from *Aspergillus nidulans* with strong transglycosylation activity. *Applied Microbiology and Biotechnology* **68**, 1250-1256.

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*. *Molecular and General Genetics* **257**, 404-411.

Kelly, J. M. (1980). Pleiotropic mutants of *Aspergillus nidulans* affected in carbon metabolism. PhD Thesis, Melbourne: La Trobe University.

Kelly, J. M. (2004). The regulation of carbon metabolism in filamentous fungi. In *The Mycota III*, pp. 385-401. Edited by R. Brambl and G. A. Marzluf. Berlin-Heidelberg: Springer-Verlag.

Kelly, J. M. and Hynes, M. J. (1977). Increased and decreased sensitivity to carbon catabolite repression of enzymes of acetate metabolism in mutants of *Aspergillus nidulans*. *Molecular and General Genetics* **156**, 87-92.

Korman, D. R., Bayliss, F. T., Barnett, C. C., Carmona, C. L., Kodama, K. H., Royer, T. J., Thompson, S. A., Ward, M., Wilson, L. J. and Berka, R. M. (1990). Cloning, characterisation, and expression of two α-amylase genes from *Aspergillus niger* var. *awaromi*. *Current Genetics* **17**, 203-212.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research* **15**, 8125-8148.

Krogh, A., Larrson, B., von Heijne, G. and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* **305**, 567-580.

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

Kumar, S., Tamura, K., Jakobsen, I. B. and Nei, M. (2001). MEGA2: Molecular Evolutionary Genetics Analysis software. Arizona: Arizona State University.

Lee, S. B. and Taylor, J. W. (1990). Isolation of DNA from fungal mycelia and single spores. In *PCR protocols: A guide to methods and applications*. Edited by M. A. Innis, D. H. Gelfand, J. S. Sninsky and T. J. White. New York: Academic Press Inc.

Litzka, O., Then Bergh, K. and Brakhage, A. A. (1996). The *Aspergillus nidulans* penicillin gene *aat* (*penDE*) is controlled by a CCAAT-containing element. *European Journal of Biochemistry* **238**, 675-682.

Lockington, R. A. and Kelly, J. M. (2001). Carbon catabolite repression in *Aspergillus nidulans* involves deubiquitination. *Molecular Microbiology* **40**, 1311-1321.

Lockington, R. A. and Kelly, J. M. (2002). The WD-40 repeat protein CreC interacts with and stabilizes the deubiquitinating enzyme CreB *in vivo* in *Aspergillus nidulans*. *Molecular Microbiology* **43**, 1173-1182.

Ludin, K., Jiang, R. and Carlson, M. (1998). Glucose-regulated interaction of a regulatory subunit of protein phosphatase 1 with the Snf1 protein kinase in *Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences of the United States of America* **95**, 6245-6250.

Mahadevan, M. S., Amemiya, C., Jansen, G., Sabourin, L., Baird, S., Neville, C., Wormskamp, N., Segers, B., Batzer, M., Lamerdin, J. and al., e. (1993). Stucture and genomic sequence of the myotonic dystrophy (DM kinase) gene. *Human and Molecular Genetics* **2**, 299-304.

Martinelli, S. D. (1994). *Aspergillus nidulans* as an experimental organism. In *Aspergillus: 50 years on.* Edited by S. D. Martinelli and J. R. Kinghorn. Amsterdam: Elsevier Science B.V.

Mathieu, M. and Felenbok, B. (1994). The *Aspergillus nidulans* CREA protein mediates glucose repression of the ethanol regulon at various levels through competition with the ALCR-specific trans-activator. *The European Molecular Biological Organisation Journal* **13**, 4022-4027.

Mayordomo, I. and Sanz, P. (2001). Hexokinase PII: structural analysis and glucose signalling in the yeast *Saccharomyces cerevisiae. Yeast* **18**, 923-930.

Minetoki, T., Nunokawa, Y., Gomi, K., Kitamoto, K., Kumagai, C. and Tamura, G. (1996). Deletion analysis of promoter elements of the *Aspergillus oryzae agdA* gene encoding α-glucosidase. *Current Genetics* **30**, 432-438.

Minoda, Y. and Yamada, K. (1963). Acid-stable α-amylase of black *Aspergilli*: Detection and purification of acid-stable dextrinizing amylase. *Agricultural and Biological Chemistry* **27**, 806-811.

Minoda, Y., Arai, M., Torigoe, Y. and Yamada, K. (1968). Acid-stable α-amylase of black *Aspergilli:* Separation of acid-stable α-amylase and acid-unstable α-amylase from the same mold amylase preparation. *Agricultural and Biological Chemistry* **32**, 110-113.

Muratani, M. and Tansey, W. P. (2003). How the ubiquitin-proteasome system controls transcription. *Nature Reviews: Molecular Cell Biology* **4**, 1-10.

Nagamine, K., Murashima, K., Kato, T., Shimoi, H. and Ito, K. (2003). Model of α-amylase production by the shochu koji mold *Aspergillus kawachii. Bioscience, Biotechnology and Biochemistry* **67**, 2194-2202.

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. *Molecular and General Genetics* **237**, 251-260.

Nehlin, J. O. and Ronne, H. (1990). Yeast MIG1 repressor is related to the mammalian early growth response and Wilms' tumour finger proteins. *European Molecular Biological Organisation Journal* **9**, 2891-2898.

Nei, M. and Kumar, S. (2000). *Molecular Evolution and Phylogenetics.* New York: Oxford University Press.

Oakley, C. E., Weil, C. F., Kretz, P. L. and Oakley, B. R. (1987). Cloning of the *riboB* locus of *Aspergillus nidulans. Gene* **53**, 293-298.

Oldham, C. E., Mohney, R. P., Miller, S. L., Hanes, R. N. and O'Bryan, J. P. (2002). The ubiquitin-interacting motifs target the endocytic adaptor protein epsin for ubiquitination. *Current Biology* **12**, 1112-1116.

Palczewski, K. (1994). Structure and functions of arrestins. *Protein Science* **3**, 1355-1361.

Pan, T. and Coleman, J. E. (1990). GAL4 transcription factor is not a "zinc finger" but forms a $Zn(II)_2Cys_6$ binuclear cluster. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 2077-2081.

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. *Molecular and General Genetics* **251**, 412-421.

Pateman, J. A., Rover, B. M. and Cove, D. J. (1967). Genetical and biochemical studies of nitrate assimilation in *Aspergillus nidulans. Biochemical Journal* **104**, 103-111.

Petersen, K. L., Lehmbeck, J. and Christensen, T. (1999). A new transcriptional activator for amylase genes in *Aspergillus. Molecular and General Genetics* **262**, 668-676.

Pokorska, A., Drevet, C. and Scazzocchio, C. (2000). The analysis of the transcriptional activator PrnA reveals a tripartite nuclear localisation sequence. *Journal of Molecular Biology* **298**, 585-596.

Randez-Gil, F., Sanz, P., Entian, K. D. and Prieto, J. A. (1998). Carbon source-dependent phosphorylation of hexokinase PII and its role in the glucose-signaling response in yeast. *Molecular and Cellular Biology* **18**, 2940-2948.

Redd, M. J., Arnaud, M. B. and Johnson, A. D. (1997). A complex composed of Tup1 and Ssn6 represses transcription *in vitro. The Journal of Biological Chemistry* **272**, 11193-11197.

Reece, R. J. and Ptashne, M. (1993). Determinants of binding-site specificity among yeast C6 zinc cluster proteins. *Science* **261**, 909-911.

Roberts, C. F. (1963). The genetic analysis of carbohydrate utilization in *Aspergillus nidulans. Journal of General Microbiology* **31**, 45-58.

Roper, J. A. and Kafer, E. (1957). Acriflavine resistant mutants of *Aspergillus nidulans. Journal of General Microbiology* **16**, 660-667.

Rotin, D., Staub, O. and Haguenauer-Tsapis, R. (2000). Ubiquitination and endocytosis of plasma membrane proteins: role of Nedd4/Rsp5p family of ubiquitin-protein ligases. *Journal of Membrane Biology* **176**, 1-17.

Saitou, N. and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406-425.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual,* 2nd Edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sanz, P., Alms, G. R., Haystead, T. A. J. and Carlson, M. (2000). Regulatory interactions between the Reg1-Glc7 protein phosphatase and the Snf1 protein kinase. *Molecular and Cellular Biology* **20**, 1231-1328.

Schild, L., Lu, Y., Gautschi, I., Schneeberger, E., Lifton, R. P. and Rossier, B. C. (1996). Identification of a PY motif in the epithelial Na channel subunits as a target sequence for mutations causing channel activation found in Liddle syndrome. *European Molecular Biological Organisation Journal* **15**, 2381-2387.

Schuller, H. J. (2003). Transcriptional control of nonfermentative metabolism in the yeast *Saccharomyces cerevisiae. Current Genetics* **43**, 139-160.

Schultz, J., Milpetz, F., Bork, P. and Ponting, C. P. (1998). SMART, a simple modular architecture research tool: identification of signalling domains. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 5857-5864.

Shaw, D. J., McCurrach, M., Rundle, S. A., Harley, H. G., Crow, S. R., Shon, R., Thirion, J. P., Hamshere, M. G., Buckler, A. J., Harper, P. S., Housman, D. E. and Brook, J. D. (1993). Genomic organisation and transcriptional units at the myotonic dystrophy locus. *Genomics* **18**, 673-679.

Shroff, R. (1997). Mutational analysis of *creA*, the mediator of carbon catabolite repression in *Aspergillus nidulans.* PhD Thesis, Adelaide: University of Adelaide.

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. Canadian Journal of Microbiology* **42**, 950-959.

Shroff, R. A., O'Conner, S. M., Hynes, M. J., Lockington, R. A. and Kelly, J. M. (1997). Null alles of *creA*, the regulator of carbon catabolite repression in *Aspergillus nidulans. Fungal Genetics and Biology* **22**, 28-38.

Springael, J. Y. and Andre, B. (1998). Nitrogen-regulated ubiquitination of the Gap1 permease of *Saccharomyces cerevisiae. Molecular and Cellular Biology* **9**, 1253-1263.

Steidl, S., Papagiannopoulos, P., Litzka, O., 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. Molecular and Cellular Biology* **19**, 99-106.

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

Stryer, L. (1988). *Biochemistry*, 3rd Edn. New York: W.H. Freeman.

Sudo, S., Ishikawa, T., Takayasu-Sakamoto, Y., Sato, K. and Oba, T. (1993). Characteristics of acid-stable α-amylase production by submerged culture of *Aspergillus kawachii. Journal of Fermentation and Bioengineering* **76**, 105-110.

Suganuma, T., Noda, N., Honbo, H. and Kitahara, K. (1997). Distinguishable action between acid-stable and neutral α-amylases from *shochu koji (Aspergillus kawachii). Bioscience, Biotechnology and Biochemistry* **61**, 1617-1619.

Tada, S., Iimura, Y., Gomi, K., Takahashi, K., Hara, S. and Yoshizawa, K. (1989). Cloning and nucleotide sequence of the genomic Taka-amylase A gene of *Aspergillus oryzae. Biological Chemistry* **53**, 593-599.

Tanaka, A., Kato, M., Nagase, T., Kobayashi, T. and Tsukagoshi, N. (2002). Isolation of genes encoding novel transcription factors which interact with the Hap complex from *Aspergillus* species. *Biochimica et Biophysica Acta* **1576**, 176-182.

Tanaka, A., Kato, M., Hashimoto, H., Kamei, K., Naruse, F., Papagiannopoulos, P., Davis, M. A., Hynes, M. J., Kobayashi, T. and Tsukagoshi, N. (2000). An *Aspergillus oryzae* CCAAT-binding protein, AoCP, is involved in the high-level expression of the Taka-amylase A gene. *Current Genetics* **37**, 380-387.

Tanaka, A., Kamei, K., Tanoue, S., Papagiannopoulos, P., Steidl, S., Brakhage, A. A., Davis, M. A., Hynes, M. J., Kato, M., Kobayashi, T. and Tsukagoshi, N. (2001). AoHapB, AoHapC and AoHapE, subunits of the *Aspergillus oryzae* CCAAT-binding complex, are functionally interchangeable with the corresponding subunits in *Aspergillus nidulans.* *Current Genetics* **37**, 175-182.

Tani, S., Kawaguchi, T., Kato, M., Kobayashi, T. and Tsukagoshi, N. (2000). A novel nuclear factor, SREB, binds to a cis-acting element, SRE, required for inducible expression of the A*spergillus oryzae* Taka-amylase A gene in *A. nidulans. Molecular and General Genetics* **263**, 232-238.

Tani, S., Itoh, T., Kato, M., Kobayashi, T. and Tsukagoshi, N. (2001b). *In vivo* and *in vitro* analyses of the AmyR binding site of the *Aspergillus nidulans agdA* promoter; requirement of the CGG direct repeat for induction and high affinity binding of AmyR. *Bioscience, Biotechnology and Biochemistry* **65**, 1568-1574.

Tani, S., Katsuyama, Y., Hayashi, T., Suzuki, H., Kato, M., Gomi, K., Kobayashi, T. and Tsukagoshi, N. (2001a). Characterization of the *amyR* gene encoding a transcriptional activator for the amylase genes in *Aspergillus nidulans. Current Genetics* **39**, 10-15.

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

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680.

Tilburn, J., Scazzocchio, C., Taylor, G. T., Zabicky-Zissman, J. H., Lockington, R. A. and Davies, R. W. (1983). Transformation by integration in *Aspergillus nidulans. Gene* **26**, 205-221.

Toda, H., Kondo, K. and Narita, K. (1982). The complete amino acid sequence of Taka-amylase A. *Proceedings of the Japanese Academy* **58**, 208-212.

Todd, R. B., Lockington, R. A. and Kelly, J. M. (2000). The *Aspergillus nidulans creC* gene involved in carbon catabolite repression encodes a WD40 repeat protein. *Molecular and General Genetics* **263**, 561-570.

Treitel, M. A. and Carlson, M. (1995). Repression by Ssn6-Tup1 is directed by Mig1, a repressor/activator protein. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 3132-3136.

Tsukagoshi, N., Kobayashi, T. and Kato, M. (2001). Regulation of the amylolytic and (hemi-) cellulolytic genes in aspergilli. *Journal of General and Applied Microbiology* **47**, 1-19.

Tsukagoshi, N., Furukawa, M., Nagaba, H., Kirita, N., Tsuboi, A. and Udaka, S. (1989). Isolation of a cDNA encoding *Aspergillus oryzae* Taka-amylase A: evidence for multiple related genes. *Gene* **84**, 319-327.

Tzamarias, D. and Struhl, K. (1995). Distinct TPR motifs of Cyc8 are involved in recruiting the Cyc8-Tup1 co-repressor complex to differentiallly regulated promoters. *Genes and Development* **9**, 821-831.

van Heeswijck, R. and Hynes, M. J. (1991). The *amdR* product and a CCAAT-binding factor bind to adjacent, possibly overlapping DNA sequences in the promoter region of the *Aspergillus nidulans amdS* gene. *Nucleic Acids Research* **19**, 2655-2660.

Varanasi, U. S., Klis, M., Mikesell, P. B. and Trumbly, R. J. (1996). The Cyc8 (Ssn6)-Tup1 corepressor complex is composed of one Cyc8 and four Tup1 subunits. *Molecular and Cellular Biology* **16**, 6707-6714.

Vautard-Mey, G. and Fevre, M. (2000). Mutation of a putative AMPK phosphorylation site abolishes the repressor activity but not the nuclear targeting of the fungal glucose regulator CRE1. *Current Genetics* **37**, 328-332.

Vautard-Mey, G., Cotton, P. and Fevre, M. (1999). Expression and compartmentation of the glucose repressor CRE1 from the phytopathogenic fungus *Sclerotinia sclerotiorum*. *European Journal of Biochemistry* **266**, 252-259.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Research* **14**, 4683-4690.

Wirsel, S., Lachmund, A., Wildhardt, G. and Ruttkowski, E. (1989). Three α-amylase genes of *Aspergillus oryzae* exhibit identical intron-exon organisation. *Molecular Microbiology* **3**, 3-14.

Wu, A., Hallstrom, T. and Moye-Towley, W. S. (1996). *ROD1*, a novel gene conferring multiple resistance phenotypes in *Saccharomyces cerevisiae*. *The Journal of Biological Chemistry* **271**, 2914-2920.

Wu, Z., Li, Q., Fortini, M. E. and Fischer, J. A. (1999). Genetic analysis of the role of the *Drosophila fat facets* gene in the ubiquitin pathway. *Developmental Genetics* **24**, 312-320.

Zhu, Y., Carroll, M., Papa, F., Hochstrasser, M. and D'Andrea, A. D. (1996). DUB-1, a deubiquitinating enzyme with growth-suppressing activity. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 3275-3279.

Zhu, H., Choi, S., Johnstone, A. K., Wing, R. A. and Dean, R. A. (1997). A large-insert (130 kbp) bacterial artificial chromosome library of the rice blast fungus *Magnaporthe grisea*: genome analysis, contig assembly, and gene cloning. *Fungal Genetics and Biology* **21**, 337-347.

Addendum

Corrections:

Page 38 Section 3.2.3 Line 23 "suppressed" should replace "complemented".

Additions:

Page 7, Section 1.3.1 In *S. cerevisiae,* Rgr1p is a component of the yeast mediator complex required for the transcriptional regulation by RNA polymerase II. [**Li, Y., Bjorklund, S., Jiang, Y.W., Kim, Y.J., Lane, W.S., Stillman, D.J., and Kornbery, R.D.** (1995) Yeast global transcriptional regulators Sin4 and Rgr1 are components of mediator complex/RNA polymerase II holoenzyme. *Proceedings of the National Academy of Sciences of the United States of America* **92:**10864-10868].

Page 12, Section 1.3.5 PEST sequences are hydrophilic regions rich in proline (P) glutamic acid (E) serine (S) and threonine (T) flanked by positively charged residues that are thought to target sequences for degradation via the 26S proteasome.

Page 13 Figure 1.3 PrnB is the proline permease encoded by the *prnB* gene in *A. nidulans*. The genetic evidence suggests that the CreB/CreC complex also affects permeases (eg. PrnB).

Page 53 Section 3.6.2 The C2 domain is a Ca^{2+}-dependent membrane-targeting module found in many cellular proteins involved in signal transduction or membrane trafficking, and is involved in membrane targeting processes such as subcellular location and calcium-dependent phospholipid binding [**Daveletov, B. A and Sudhof, T. C** (1993). A single C2 domain from synaptotagmin I is sufficient for high affinity Ca2+/phospholipid binding. *Journal of Biological Chemistry* **296**, 26386-26390].

Page 64 Section 4.3 The weak suppression of the *creA204* mutation by the *creD34* mutation was not able to be detected on plate tests, but was determined by specific enzyme assays (Kelly and Hynes, 1977).

Chapter 6 The model for carbon catabolite repression proposed by Lockington and Kelly (2002, see Figure 1.3) can now be expanded to include CreD, ApyA, HulA and AcrB. It has been proposed that the CreB/CreC deubiquitination complex acts on ubiquitinated CreA to stabilize or alter the function of CreA as well as on other targets. CreD and ApyA have been shown to interact with the HECT ubiquitin ligase HulA, and these proteins likely form the "ubiquitinating complex" that is responsible for the ubiquitination of CreA. The AcrB protein is also implicated in the ubiquitination aspect of this regulatory network, and the three transmembrane domains imply a role in signalling. Thus mutants in both the deubiquitination and ubiquitination aspects of this network have been isolated as a result of their mutant phenotypes. To test this model, the ability of the CreD/ApyA/HulA complex to ubiquitinate CreA and other substrates should be tested. This would require the isolation of a specific antibody to CreA (which so far has been unsuccessful) to identify CreA. This would be used to determine the presence or absence of ubiquitinated CreA in *creD*, *apyA* and *hulA* mutant backgrounds, thus testing the model.