The Structure and Regulation of Aldehyde Dehydrogenase Encoding Genes in *Aspergillus niger* and *Aspergillus nidulans*.

by

Matthew J. O'Connell, B.Sc.(Hons.)

Department of Genetics, University of Adelaide.

awarded 12.10.90

A thesis submitted to the University of Adelaide for the degree of Doctor of Philosophy in March, 1990.

Contents.

ABSTRACT.

The primary aim of this research was to determine mechanisms of gene regulation in the filamentous ascomycete fungus *Aspergillus niger*. To this end, a gene encoding an aldehyde dehydrogenase enzyme, designated *ald*A, has been cloned from *A. niger*. This clone and a clone of the *ald*A gene from the related fungus *Aspergillus nidulans* (Pickett *et al.*, 1987) has been used in a comparative study of gene structure and regulation in these organisms.

The complete nucleotide sequence of the *A. niger ald*A gene and its flanking regions, and the structure of its mRNA transcript has been determined. The *ald*A gene of *A. niger* has one major and two minor start points of transcription and 4 sites of polyadenylation. The coding region shows significant similarity to the sequence of the *A. nidulans ald*A gene (Pickett *et al.*, 1987), but contains three introns compared to the two present in the *A. nidulans* gene. Alignment of amino acid sequences of several aldehyde dehydrogenases has implicated two cysteine residues likely to be adjacent at the active site. The promoter region has a general structure similar to other highly expressed fungal genes, including an extensive pyrimidine rich region, TATA and CCAAT box homologies, and several repeated sequences.

Studies of *ald*A mRNA levels in wildtype strains grown under a variety of conditions has revealed that *ald*A expression is subject to induction and repression in both species. However, the relative levels of induction and degree of repression differs markedly between these species in almost all growth conditions tested. Most notably, threonine

induction of *ald*A genes in *A. niger* is insensitive to glucose repression, and in addition to *alc*R mediated control of *ald*A expression in *A. nidulans*, the *amd*R gene (also called *int*A) either directly or indirectly affects *ald*A expression in mycelia grown in fructose or threonine medium.

A strain of *A. niger* carrying a deletion of the *ald*A gene was constructed by gene replacement. In addition to several phenotypes in common with *ald*A mutant strains of *A. nidulans*, it also exhibited some unique phenotypes further suggesting fundamental differences in metabolism.

Heterologous expression of the *A. niger ald*A gene in *A. nidulans* is subject to induction and repression, and induction during growth on ethanol medium requires a functional *alc*R gene product. Expression of the *A. nidulans ald*A in *A. niger* is also controlled by induction and repression, in a similar way, including induction in fructose medium, to that seen for wildtype *A. nidulans*. However, induction in response to growth in threonine medium is, as in wildtype *A. niger*, insensitive to repression by glucose.

The effects of several 5' deletions in the *A. niger ald*A gene have been investigated. Regions required for induction and the high level of expression were identified. Deletion of the putative CCAAT boxes and a sequence common to the promoters of several fungal genes, many of which are subject to carbon catabolite repression, had no observable effects on gene expression. A region containing a silencer of induction during fructose growth was also located, which together with the high level of fructose induced expression of the *A. nidulans ald*A gene in *A. niger*, implies that *A. niger* also has an induction mechanism functionally homolgous to that affected by the *amd*R gene in *A. nidulans*.

Declaration

I declare that this thesis contains no material which has been accepted for the award of any other degree or diploma in any University, and to the best of my knowledge and belief contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

I give my consent for this thesis to be made available for photocopying and loan.

Matthew J. O'Connell

Acknowledgements.

I wish to express my gratitude and respect for my supervisor, Dr. Joan M. Kelly. Her advice and encouragement, both before the start of this candidature, and during the experimental work and preparation of this thesis, together with her insight and enthusiasm for science, was of great benefit to my scientific education. I also appreciate the tolerance and friendship shown by Joan Kelly, which has made it a pleasure to have worked under her supervision.

I wish to acknowledge the significant contributions at both a personal and professional level by several other people:

Celia Dowzer has been a fellow student since my enrolment, and has extended friendship, interest and advice that is greatly appreciated.

I thank Dr. Robin Lockington (Department of Animal Sciences, Waite Agricultural Research Institute, Australia) who has provided many interesting discussions and a great many valuable comments with regard to written and experimental work.

Prof. Michael Hynes (Department of Genetics, University of Melbourne, Australia) has also provided valuable advice during the experimental work and in his critical reading of parts of this thesis.

Dr. Dara Whisson (South Australian Department of Agriculture, Australia), while a Postdoctoral Research Fellow in the Department of Genetics, provided much help with experimental techniques, in particular with DNA sequencing.

Mrs. Michelle Adamou provided much technical help to which I am indebted.

Doug Anderson, Michael Ayliffe, Angela Binns, Steven Cooper, Jo Drake, Chris Hefford, and many other members of the Department of Genetics, both past and present, have contributed to a friendly and scientifically stimulating atmosphere in which to work.

I appreciate the provision of the nucleotide sequences prior to publication of the *A. nidulans acu*D gene by Dr. G. Turner (Department of Microbiology, School of Medical Sciences, University of Bristol, United Kingdom), the *A. nidulans alc*R gene by Dr. B. Felenbok (Institut de Microbiologie, Universite Paris-Sud, Centre d'Orsay, France), and the *A. nidulans acu*E and *fac*A genes by Dr. R. Sandeman and Prof. M. Hynes (Department of Genetics, University of Melbourne, Australia).

Finally I wish to thank my family for their encouragement and support.

I acknowledge the financial support that I have been provided through a Commonwealth Postgraduate Research Award and an Australian Postgraduate Priority Research Award.

Abbreviations.

A: Adenosine; AMV: avian myeloblastosis virus; bp: base pair(s); BSA: bovine serum albumin; C: cytidine; cDNA: DNA complementary to RNA; Ci: curie; dATP: 2'-deoxy-adenosine-5'-triphosphate; dCTP: 2'-deoxy-cytidine-5'-triphosphate; ddATP: 2',3'-dideoxy-adenosine-5'-triphosphate; ddCTP: 2',3'-dideoxy-cytidine-5'-triphosphate; ddGTP: 2',3'-dideoxy-guanosine-5'-triphosphate; ddNTP: 2',3'-dideoxy-nucleoside-5'-triphosphate; ddTTP: 2',3'-dideoxy-thymidine-5'-triphosphate; dGTP: 2'-deoxy-guanosine-5'-triphosphate; dITP: 2'-deoxy-inosine-5'-triphosphate DNA: deoxyribonucleic acid; dNTP: 2'-deoxy-nucleoside-5'-triphosphate; dTTP: 2'-deoxy-thymidine-5'-triphosphate; DTT: dithiothreitol; EDTA: (ethylenedinitrilo)tetraacetic acid; G: guanosine; g: gram/force of gravity; kDa: kilodalton(s); kb: kilobase(s); mg: milligram; ml: millilitre; mRNA: messenger RNA; N: nucleotide; NAD: β-nicotinamide-adenine dinucleotide; NADP: β-nicotinamide-adenine dinucleotide phosphate; p: plasmid; poly(A)$^+$ RNA: polyadenylated RNA; R: purine; RNA: ribonucleic acid; r.p.m.: revolutions per minute; SDS: sodium dodecyl sulphate; T: thymidine; Tris: 2-amino-2-(hydroxymethyl)-1,3-propanediol; U: uridine; uCi: microcurie; ug: microgram; ul: microlitre; v/v: volume per volume; w/v: weight per volume; Y: pyrimidine.

Publications.

The work presented in this thesis has been published in the following papers:

O'Connell, M.J. and Kelly, J.M. (1988). Differences in the regulation of aldehyde dehydrogenase genes in *Aspergillus niger* and *Aspergillus nidulans*. Curr. Genet. 14: 95-104.

O'Connell, M.J. and Kelly, J.M. (1989). Physical characterization of the aldehyde-dehydrogenase-encoding gene of *Aspergillus niger*. Gene 84: 173-180.

O'Connell, M.J. and Kelly, J.M. (1990). The control of aldehyde dehydrogenase (*ald*A) gene expression in *Aspergillus*. Molec. Gen. Genet. Submitted.

Chapter 1:

1.1: Introduction.

The filamentous ascomycete fungus *Aspergillus nidulans* is an ideal organism for the study of gene structure and regulation. In nature, *A. nidulans* is found in temperate and subtropical soils, where it grows on decaying plant material (Raper and Fennell, 1965). In the laboratory, it can be easily cultured in both defined liquid media or on agar solidified plates. It is normally haploid, but stable diploids can be maintained. It possesses both sexual and parasexual life cycles, which can easily be manipulated to construct strains of interest and map genes with respect to many identified loci. Strains are available with markers on each of the eight linkage groups (Clutterbuck, 1974, 1984; Cove, 1977). Classical genetic analysis has been used to obtain much data on structural and regulatory loci involved in many systems, but in particular loci involved in carbon and nitrogen metabolism (reviewed in Smith and Pateman, 1977; Arst and Scazzocchio, 1985).

A. nidulans has a relatively small genome of approximately 3 X 10^7 bp. (Timberlake, 1978; Brody and Carbon, 1989). It can be transformed at significant frequency, with transforming DNA integrating into the host chromosome. There are several systems involving complementation of recessive auxotrophic mutations and dominant markers that allow selection of transformants (reviewed by Gurr *et al.*, 1987 and Fincham, 1989). These two characteristics have been exploited by several groups

for both the isolation of particular genes and the analysis of mutations by reverse genetics. A small genome eases the task of isolating sequences of interest from genomic libraries using either heterologous probes or differential hybridization, and these approaches have been used to isolate several *A. nidulans* genes, for example *amd*S (Hynes *et al*, 1983), *are*A (Caddick *et al*, 1986) and *qut*E (Hawkins *et al*, 1985). Transforming plasmids containing sequences of interest and an antibiotoic resistance marker for *E. coli* can be rescued from the transformant chromosome by digesting the DNA of the transformant with an appropriate restriction enzyme, circularizing DNA fragments by ligation, and selection of an antibiotic resistance marker after transformation into *E. coli*. This approach has been utilized to clone several genes after complementation of mutations, for example *amd*R (Andrianopoulos and Hynes, 1988), *gat*A (Richardson *et al*., 1989), and *cre*A (Dowzer and Kelly, 1989). Reverse genetics has been used to analyse the effects *in vivo* of *in vitro* generated mutations in cloned genes introduced into strains by transformation, for example, mutations affecting the regulation of the *amd*S (Hynes *et al*, 1988) and *trp*C (Hamer and Timberlake, 1987) genes.

The extensive genetic analysis of *A. nidulans* is in contrast to the limited amount of work published on gene structure and regulation in the related ascomycete *Aspergillus niger*. *A. niger* can be found in nature growing in a variety of environments including tropical and subtropical soils, spoiled fruits, grains, and many protein rich substrates (Raper and Fennell, 1965). *A. niger* is used extensively in industry for the production of several enzymes and metabolites as it is capable of secreting large quantities of compounds into the growth media (reviewed by Bennett, 1985). It can be easily grown in the laboratory, but has the great disadvantage for genetic analysis of

lacking a sexual cycle. However, auxotrophic and morphological markers have been used in the identification of seven different linkage groups (Lhoas, 1967; Bos *et al*, 1988; Bos *et al.*, 1989), by the construction and haploidization of diploids. Mitotic recombination during the parasexual cycle is much more frequent in *A. niger* than in *A. nidulans*, and allows markers to be ordered on a chromosome, and indicates mitotic recombination is important for generating genetic variation in *A. niger* (Lhoas, 1967). Thus parasexual genetic analysis can be performed in *A. niger*. Several systems for *A. niger* transformation have also been developed (reviewed by Gurr *et al.*, 1987 and Fincham, 1989), and thus reverse genetics can also be used to investigate gene regulation. Furthermore, mutant strains and cloned genes from other ascomycete fungi, in particular *A. nidulans* and other *Aspergillus* species, can be used to both isolate *A. niger* genes and analyze their regulation via the use of transformation.

This thesis describes experiments performed to investigate the structure and regulation of genes encoding aldehyde dehydrogenase enzymes in *A. niger* and *A. nidulans*. The primary aim of this work was to investigate mechanisms of gene regulation in *A. niger* by undertaking a comparative study of the structure and regulation of an aldehyde dehydrogenase encoding gene (*ald*A) in *A. niger* with that of *A. nidulans*. The *ald*A gene of *A. nidulans* has been subjected to much analysis by several workers, as have the other genes of the ethanol utilization regulon. A review of literature relevant to this research covering the ethanol utilization regulon and other genes of *A. nidulans* that are well understood in terms of their structure and the control of their expression is presented in the following section.

1.2: Literature review.

1.2.1: Alcohol and aldehyde dehydrogenases in *Aspergillus nidulans*.

Early work on the catabolism of ethanol and ethylamine in *A. nidulans* showed that these compounds are utilized via acetate (Page and Cove, 1972). Mutants unable to utilize acetate as a source of carbon are also unable to utilize ethanol or ethylamine. Mutations in strains able to utilize acetate but unable to utilize ethanol or ethylamine were isolated and mapped to four loci. Mutations isolated at the *mau*A and *mau*B loci abolished amine utilization but did not affect ethanol utilization. A monoamine oxidase, which converts amines to their corresponding aldehyde and ammonium, was detected in wildtype strains but was absent from *mau*A and *mau*B mutants. Strains carrying mutations at the *alc* locus were classified into two classes. One class was affected in the utilization of amines as sources of carbon, the other failed to utilize ethylamine as either a source of carbon or nitrogen. Strains of both classes lacked an alcohol dehydrogenase (ADH) that was present in wildtype strains. *ald* mutations affected both ethanol and ethylamine utilization, but an acetaldehyde degrading enzyme was not detected. These results showed that ethylamine and ethanol are utilized independently via acetaldehyde. Futhermore, acetaldehyde produced from ethylamine oxidation is at least in part detoxified by the action of ADH, or acetaldehyde oxidation was affected by *alc* mutations.

Experiments performed by Pateman *et al.*, (1983) identified an aldehyde dehydrogenase enzyme (AldDH) that was absent from strains carrying *ald*A mutations. The *ald*A mutations mapped to linkage group

VIII. AldDH was required for the utilization of both ethanol and threonine, implying that these compounds were utilized via acetaldehdye. Mutations at the *alc* locus were shown to be in two complementation groups, *alc*A and *alc*R, which were adjacent on linkage group VII. Strains carrying non-leakly mutations in either gene failed to grow on ethanol or threonine medium. Mutations isolated at the *alc*A locus were recessive to wildtype and lacked the single major ADH detected in wildtype strains. Mutations at the *alc*R locus lacked both the alcohol dehydrogenase and aldehyde dehydrogenase. One *alc*R allele, *alc*R125, previously known as *alc*A125, was suppressible by allele specific, gene nonspecific suppressors (Roberts *et al.*, 1979), indicating that *alc*R encodes a protein. As *alc*R mutations affected the expression of *alc*A and *ald*A, the *alc*R gene product was hypothesized to be a *trans*-acting regulatory protein required for *alc*A and *ald*A induction.

Enzyme levels for both ADH and AldDH were elevated in mycelia grown in ethanol or threonine medium. This induction was repressed in wildtype strains in the presence of good sources of carbon. Although glycerol and acetate repress expression to some extent, expression was almost completely repressed in the presence of an inducer when 50mM glucose was added to the growth medium. Derepressed levels of ADH expression were seen in strains carrying the *cre*A^{d}1 mutation (Bailey and Arst, 1975; Pateman *et al.*, 1983). The expression of many other genes involved in carbon metabolism is also derepressed in strains carrying this and other recessive *cre*A mutant alleles (Bailey and Arst, 1975; reviewed by Arst and Bailey, 1977). No dominant *cre*A alleles leading to increased repression, or alleles leading to the failure to derepress, have been isolated. These results indicate that the *cre*A gene product is a negatively acting regulatory molecule, in contast to

most regulators of transcription in *A. nidulans* and other eukaryotes, which generally are positively acting. The mechanism by which the *creA* gene product mediates carbon catabolite repression is not known, and thus it can not be ruled out that the *creA* gene product may function indirectly by a positively acting mechanism. For example, the *creA* gene product may activate other molecules that derepress carbon catabolite repression or a general transcription factor that activates expression. However, the the available genetic evidence points towards *creA* having a negative function on gene expression. A clone of the *creA* gene has been obtained and thus the molecular mechanisms involved in carbon catabolite repression in *A. nidulans* should be able to be elucidated soon (Dowzer and Kelly, 1989).

Strains carrying loss of function mutations at the *alcA* locus did not show stronger growth than extreme *alcR* mutants on threonine medium. No AldDH was detectable in these extreme *alcR* mutants, but wildtype levels of AldDH were seen in *alcA* mutant strains. This result implies that reduction of acetaldehyde to ethanol by ADH is required during threonine metabolism to keep acetaldehyde levels below a threshold level toxic to wildtype *A. nidulans* (Pateman *et al.*, 1983).

The molecule acting as the coinducer may either be ethanol or acetaldehyde. The presence of threonine in the growth media fails to result in induced expression of *aldA* and *alcA* in threonine nonutilizing (*tut*) mutants thought to be defective in the metabolism of threonine to acetaldehyde (Pateman *et al.*, 1983). As basal levels of AldDH, ADH I and other ADHs (see below) could interconvert ethanol and acetaldehyde, either of these molecules could be the coinducer of *alcA* and *aldA*.

A second alcohol dehydrogenase (ADH II) has been identified by activity staining. Unlike the *alcA* encoded ADH I, ADH II was repressed

in wildtype strains in the presence of ethanol, but was present in strains with *alc*R mutations, some *alc*A mutations, and deletions of the *alc*R/*alc*A region, when grown in the presence of ethanol. Therefore the lack of a functional *alc*A or *alc*R gene product inhibited, either directly or indirectly, the ethanol repression of ADH II. As ADH II was present in *alc*R and *alc*A mutants which failed to grow on ethanol medium, ADH I is the physiological enzyme for ethanol utilization. The induction of both ADH I and ADH II acted at the level of mRNA accumulation as no protein could be detected in *in vitro* translated mRNA prepared from cultures grown under conditions where no protein was detected by activity staining. Like ADH I, ADH II synthesis was subject to carbon catabolite repression, which also acted at the level of transcription (Sealy-Lewis and Lockington, 1984). No mutations within the ADH II gene have reported, and as there is no known physiological role for ADH II, there is no obvious selection system for mutations within this gene.

A third ADH (ADH III) has been identified in *A. nidulans* by complementation of a mutation in *Saccharomyces cerevisiae*. This gene has been designated *alc*C. No physiological role has been attributed to ADH III. *alc*C mRNA accumulated in the presence of ethanol, and the expression of this gene was not subject to carbon catabolite repression (McKnight *et al.*, 1985). The *alc*C gene has been altered by gene disruption, but no phenotype was detected in a strain carrying an *alc*C disruption. The *alc*C locus has been mapped to linkage group VII, between the *gat*A and *cbx*B loci (Jones and Sealy-Lewis, 1989).

In addition to the classical genetic analysis of mutations, the genes of the ethanol utilization regulon have been characterized at the molecular level. The *alc*R/*alc*A region has been cloned by two

groups. Doy *et al.*, (1985) identified a 3.85 kb. BamHI restriction fragment in genomic Southern blots, using labelled cDNA prepared from threonine induced mRNA, that was present in wildtype strains but absent from a strain thought to carry a deletion of the *alc*R/*alc*A region. This fragment was cloned and used to isolate three bacteriophage lambda clones spanning 24 kb. in the region of *alc*R and *alc*A. The presence of the *alc*A and *alc*R genes within the clones was verified by both the localization of mutations within the clone and cross hybridization with an ADH clone derived from *S. cerevisiae*.

Lockington *et al.*, (1985) used the differential hybridization, to clones of a cDNA library, of labelled mRNA probes prepared from wildtype mycelia grown in both ethanol and threonine induced, noninduced and noninduced/repressed conditions, competed with a large excess of unlabelled ethanol induced RNA prepared from a strain carrying a deletion of *alc*A and *alc*R. The probes were enriched for either *alc*A or *ald*A mRNA as identified by *in vitro* translation by preparative gel electrophoresis. The *ald*A mRNA was approximately 1.8 kb., and that for *alc*A was approximately 1.5 kb.. Partial cDNA clones for both *alc*A and *ald*A were isolated, and their identity confirmed by *in vitro* translation of hybrid selected mRNA isolated with these cDNA clones. The *alc*A cDNA clone was used to isolate a genomic clone from a bacteriophage lambda library that was shown by transformation to contain functional *alc*A and *alc*R genes. These clones were used for northern blot analysis and confirmed earlier results that indicated *alc*A and *ald*A regulation is at the level of mRNA accumulation, most probably via transcription initiation.

Analysis of the expression of the positively acting regulatory gene *alc*R has shown that *alc*R was itself inducible, and that this induction did not occur in strains carrying the *alc*R125 mutation.

Hence, a functional *alc*R gene product is also required to regulate its own expression. Futhermore, the expression of *alc*R was subject to carbon catabolite repression and was derepressed in a *creA*^{d}1 mutant strain (Lockington *et al.*, 1987). Carbon catabolite repression also acts independently on *alc*A and *ald*A as these genes are still subject to carbon catabolite repression in transformants containing a construct of the *alc*R coding region and the promoter of the glyceraldehde-3-phosphate dehydrogenase gene, which is not subject to carbon catbolite repression and hence derepressed for the expression of the *alc*R protein (Felenbok *et al.*, 1989). As carbon catabolite repression acts on both the structural and positively acting regulatory genes, promoter mutations would need to occur in both the *alc*R and either *alc*A or *ald*A genes to stop carbon catabolite repression of either the *alc*A or *ald*A structural genes. A transformant containing approximately 16 copies of *alc*R had approximately 15 fold the basal noninduced/derepressed level of *alc*R expression found in wildtype strains. However, this transformant had *alc*A and *ald*A basal expression levels similar to wildtype, indicating that a coinducer is necessary for induction. As this elevated basal level of *alc*R expression in this transformant was fully repressed by the addition of glucose to the growth media, as was the basal level of *alc*R mRNA in the *alc*R125 strain, carbon catabolite repression does not act by simply preventing the autoregulated induction of *alc*R (Lockington *et al.*, 1987).

The *ald*A cDNA clone isolated by Lockington *et al.*, (1985), was used to isolate a genomic clone of *ald*A (Pickett *et al.*, 1987). Physical characterization of this clone revealed that the *ald*A gene contains two introns and encodes a protein of 497 amino acids with significant amino acid sequence similarity to the aldehyde dehydrogenases of horse and man. The promoter region contains two

start points of transcription and a sequence with similarity to the TATA box consensus sequence. The primary structure of the *alc*A gene has also been determined. The gene encodes a 348 amino acid protein that was similar in amino acid sequence to the ADH III protein of *A. nidulans* and the ADH I and ADH II proteins of *S. cerevisiae*. Like the *ald*A gene, the promoter region of *alc*A contains two start points of transcription and a TATA like sequence (Gwynne *et al.*, 1987a).

The physical analysis of the ethanol utilization regulon has been extended by a study of the biochemical properties of the products of the structural genes. Both the ADH I and AldDH proteins have been purified and characterized. ADH I has a high substrate affinity for acetaldehyde and a specific activity 50 fold higher than AldDH. Hence the oxidation of acetaldehyde would be a rate limiting step in the utilization of ethanol. It has also been shown that the reduction of acetaldehyde to ethanol by ADH I occurs 5 fold faster than the conversion of ethanol to acetaldehyde (Creaser *et al.*, 1985; Creaser *et al.*, 1987). These results support the hypothesis that ADH I is required to detoxify acetaldehyde produced during threonine metabolism and explains why many *alc*A mutants fail to grow on threonine medium. Significant intercellular concentrations of acetaldehyde inhibit enzymes via reactions with essential -SH groups, and thus in *A. nidulans*, ADH I may also be considered as a detoxification enzyme (Creaser *et al.*, 1987).

Comparison of the sequences 5' to the start codons of *ald*A and *alc*A revealed six regions of similar sequence that may play a role in *alc*R mediated control. Three of these regions are located upstream of the start points of transcription, including one that is similar to a sequence required for *ADR1* mediated induction of *ADH2* in *S. cerevisiae*

(Gwynne *et al.*, 1987a). There have been no reports of experiments identifying sequences controlling either *alc*A or *ald*A expression *in vivo*.

Physical analysis of the cloned *alc*R gene showed that it consists of a coding region interupted by a single intron and encodes a deduced protein of 863 amino acids. The original published sequence of the deduced *alc*R protein does not contain structures that resemble known DNA binding domains such as a zinc finger, a helix-turn-helix or a leucine zipper (Felenbok *et al.*, 1988). However, if the first exon is read from a different start codon and in a different reading frame, a zinc finger structure is present (Creaser, cited in Felenbok *et al.*, 1989). Whether this is the correct reading frame, and whether this putative zinc finger structure interacts with the 5' regions of *alc*A, *ald*A and *alc*R is not known. The 5' region contains a single start point of transcription and putative TATA sequence. Comparison of the 5' sequence of the *alc*R gene to that of *ald*A and *alc*A did not reveal any major regions of sequence similarity (Felenbok *et al.*, 1988).

The promoter of the *alc*A gene has been used to direct the expression of a number of heterologous proteins in *A. nidulans*. These include proteins of commercial interest (Gwynne *et al.*, 1987b; Gwynne *et al.*, 1989), proteins controlling conidiation in *A. nidulans* (Adams *et al.*, 1988; Mirabito *et al.*, 1989), and tubulin encoding genes of *A. nidulans* (Waring *et al.*, 1989). Only 320 bp. 5' to the start codon of *alc*A was required to direct regulated expression of a heterologous coding region (Gwynne *et al.*, 1987a). Many workers have used 0.1% fructose in liquid culture as a noninducing/derepressing carbon source in studies of *alc*A and *ald*A regulation. It is noteworthy that in the *alc*A directed expression system using the coding regions of *A. nidulans* tubulin genes, expression gauged by plate tests showed 0.1% fructose to

be as strong a source of carbon catabolite repression as 1% glucose (Waring *et al.*, 1989).

This classical and molecular analysis of the regulation of the genes of the ethanol utilization regulon has identified the structural and regulatory genes involved in the utilization of compounds such as ethanol and threonine. Molecular analysis of clones of these genes has determined the arrangement of the coding and control regions of these genes as well as their protein products. Although much is known about the roles of these genes in metabolism and gene expression, and the conditions that affect their expression, there has been no direct data gained on either the molecular mechanisms of the function of the regulatory gene products or the sites of action of these regulatory proteins in the control regions of the structural genes.

Other systems of regulated gene expression have also been studied in *A. nidulans*. The genes involved in the catabolism of L-proline have also been studied extensively and have generated much data on the control of gene expression in *A. nidulans*. The structural gene for acetamidase (*amd*S) and the regulatory genes which control its expression, and other structural genes under the control of regulatory genes affecting *amd*S expression have also been studied. In these systems, the study of genetic and molecular data has revealed the nature and sites of action of regulatory gene products. These genes are also subject to the wide acting regulatory mechanism of ammonium repression, which is better understood than carbon catabolite repression. In light of the amount of information pertaining to gene structure and regulation in *A. nidulans*, and in fungi in general, relevant literature about these gene systems will be discussed in the following sections.

1.2.2: The L-proline catabolism gene cluster of *Aspergillus nidulans*.

Genes involved in the utilization of L-proline as a source of carbon or nitrogen in *A. nidulans* are clustered as a group of four genes with an internal control region on linkage group VII. The clustering of functionally related and coregulated genes is not consistently seen in *A. nidulans* (reviewed by Arst and Scazzocchio, 1985). The genes of the L-proline catabolism cluster (*prn*) are: *prn*A, which encodes a positive regulatory gene affecting the expression of the other *prn* genes; *prn*B, which encodes a transport protein required for the uptake of exogenous L-proline (formally called L-proline permease); *prn*C, which encodes L- Δ^1-pyrroline-5-carboxylate (P5C) dehydrogenase; and *prn*D, which encodes a proline oxidase. The gene order is *prn*A, *prn*D, *prn*B, *prn*C, with a *cis*-acting control region located between the *prn*D and *prn*B genes (Arst *et al.*, 1980; Jones *et al.*, 1981; Sophianopoulou and Scazzocchio, 1989).

The noninducible phenotype of *prn*A deletions and the phenotypic suppression of *prn*A$^-$ mutations showed that the *prn*A gene product is a positively acting regulatory protein. Induction mediated by the *prn*A gene product of *prn*B and *prn*D in the presence of L-proline and L-ornithine is subject to carbon catabolite repression but shows derepressed levels of expression in *cre*A mutant strains (Arst and MacDonald, 1975; Bailey and Arst, 1975). It is not known whether carbon catabolite repression of the *prn*B gene, which encodes the L-proline transport protein, is sufficent to account for the repression of the L-proline oxidase encoding *prn*D gene. The expression of *prn*C and *prn*D is subject to ammonium repression via the action of the *are*A gene product, but *prn*B expression is not ammonium repressible (Arst and MacDonald, 1975). *prn*^{d} mutations, which mapped to the internal *cis*-acting control

region, lead to derepression of *prn*B, and thus affect carbon catabolite repression (Arst and MacDonald, 1975).

The *are*A gene product has been proposed to be a positively acting regulatory protein that is required for the expression of genes involved in the utilization of nitrogen sources other than ammonium (Arst and Cove, 1973; reviewed by Arst and Scazzocchio, 1985). Due to the heterogeneity and nonhierarchial nature of the phenotypes of different *are*A allelles, it has been suggested that the *are*A gene product is directly involved in regulating the expression of genes involved in nitrogen metabolism, but the structure of the *are*A receptor sites within the many *are*A regulated genes differs (Arst and Scazzocchio, 1985). Further evidence for the direct interaction between the *are*A gene product and the promoters of ammonium repressed genes was gained from the physical analysis of the cloned *are*A gene which has shown that the *are*A gene product has a region of significant similarity to a mouse DNA binding protein in a region surrounding a novel zinc finger structure (Caddick *et al.*, 1986; Arst *et al.*, 1989). There have been no reports of direct evidence, for example gel retardation analysis, that the *are*A gene product binds to DNA sequences in the control regions of ammonium repressible genes. Despite this, the structure of the *are*A gene product and the phenotypes of *are*A mutations and promoter mutations in ammonium repressible genes, indicates that the *are*A gene product is a positively acting regulatory protein required for the expression of many genes involved in nitrogen metabolism.

A system was devised for the selection of induced deletions within the *prn* cluster. The *sas*A-60 mutation results in sensitivity to the semialdehydes L-glutamic γ-semialdehyde, a product of L-arginine

and L-proline metabolism, and succinic semialdehyde, a product of γ-amino butyric acid (GABA) metabolism. Whether the *sas*A locus encodes a semialdehyde dehydrogenase is not known. Strains with mutations at the *ald*A locus are also sensitive to these compounds, although not as severely as *sas*A-60 strains. Thus, aldehyde dehydrogenase must play some role in L-proline and GABA metabolism. It is not known whether regulatory genes that control the expression of genes encoding enzymes in these metabolic pathways also regulate *ald*A expression. Mutations in the *prn*A, *prn*B and *prn*D genes were selected in a *sas*A-60 background via resistance to L-proline toxicity in carbon and ammonium derepressing conditions. Mutations at the *prn*C locus could not be selected using this system because of the extreme toxicity of L-proline to *prn*C mutants (Arst *et al.*, 1981). Some deletions of the *cis*-acting control region and extending into *prn*B but not into *prn*C lead to reduced levels of *prn*C expression. It was suggested that this effect may have been due to transcription of *prn*B and *prn*C as a dicistronic mRNA (Arst and MacDonald, 1978; Arst *et al.*, 1981). However, cloning of the entire *prn* cluster and subsequent northern blot analysis has shown *prn*B and *prn*C to have separate transcripts, and thus the deletions affecting *prn*C expression must have deleted an enhancer element within the *cis*-acting control region that was required for high level *prn*C expression. Northern blot analysis also showed that expression of all the genes within the *prn* cluster was inducible. The *prn*A regulatory gene was hypothesized on the basis of being the only known regulatory gene involved in L-proline metabolism to be, like the *alc*R gene, autoregulatory (Hull *et al.*, 1989). The *prn*B gene has been sequenced and the deduced amino acid sequence of the encoded L-proline transport protein shows significant similarity to the arginine and histidine transport proteins of *Saccharomyces cerevisiae* (Sophianopoulou and Scazzocchio, 1989).

1.2.3: Regulation of the acetamidase encoding *amd*S and coregulated genes in *Aspergillus nidulans*.

The acetamidase encoded by the *amd*S gene in *A. nidulans* is required for growth on acetamide as the sole source of either carbon or nitrogen. The regulation of *amd*S gene expression has been extensively studied by both classical genetic and recombinant techniques, and is controlled by several well understood independent mechanisms. All the regulatory genes affecting *amd*S expression also control the expression of other genes involved in carbon and nitrogen metabolism (Kelly and Hynes, 1977; Hynes, 1978a and references therein; Atkinson *et al.*, 1985, Katz and Hynes, 1989a). The *amd*S gene has been cloned and sequenced, and the sites of action of several regulators have been determined (Hynes *et al.*, 1983; Corrick *et al.*, 1987; Hynes *et al.*, 1988).

Three regulatory genes have been identified that are involved in *amd*S induction. The *amd*R gene (also called *int*A) product acts to induce *amd*S in the presence of omega amino acids. Recessive *amd*R^{-} and semidominant *amd*R^{c} alleles have been isolated that affect *amd*S gene expression (Hynes and Pateman, 1970). *amd*R also regulates the expression of the *gab*A gene which is involved in omega amino acid uptake, the *gat*A gene which encodes an omega amino acid transaminase, and the *lam*A and *lam*B genes which are involved in the metabolism of lactams (Arst, 1976; Arst *et al.*, 1978; Katz and Hynes, 1989a). Due to the coregulation of genes encoding enzymes of unrelated metabolic pathways, Arst (1976) proposed that *amd*R is an integrator gene under the definition of Britten and Davidson, (1969). The *gat*A, *lam*A and *lam*B genes have also been cloned (Richardson *et al*, 1989; Katz and Hynes, 1989b), as has the *amd*R gene (Andrianopoulos and Hynes, 1988).

The *facB* gene product is required for the induction of *amdS* expression in the presence of acetate (Hynes, 1977). The *facB* gene product is also required for the acetate induction of the acetyl-Co A synthetase encoding *facA* gene, the isocitrate lyase encoding *acuD* gene, and the malate synthase encoding *acuE* gene (Armitt *et al.*, 1976; Hynes 1977). Clones of all these genes have been obtained (Ballance and Turner, 1986; Sanderman and Hynes, 1989; Katz and Hynes, 1989c).

The *amdA* gene also acts to induce *amdS* expression in the presence of acetate, but does not induce *amdS* expression to the same extent as *facB* or *amdR* (Arst and Cove, 1973; Hynes 1978a). The *amdA7* mutation results in increased expression of *amdS* (Hynes, 1978a). *amdA* also regulates the *aciA* gene, which is similarly affected by the *amdA7* mutation. This acetate inducible gene was identified by cloning, but no function has been attributed to the gene product (Atkinson *et al.*, 1985). The *amdA* gene has not been cloned.

The wide domain regulatory genes *creA*, affecting carbon catabolite repression, and *areA*, affecting ammonium repression, also control the expression of *amdS* (Kelly and Hynes, 1977; Hynes, 1973, 1975; Arst and Cove, 1973).

One approach used to locate the sites of action of these regulatory proteins within the *amdS* control region has been the determination of the nucleotide sequences of several 5' mutations that affect *amdS* regulation (Hynes *et al.*, 1985, 1988). The *amdI18* mutation results in a general 2 - 3 fold increase in acetamidase synthesis (Hynes, 1978b) and is due to a C -> A substitution at position -118, approximately 77 bp. upstream of the start point of transcription

(Corrick *et al.*, 1987). The *amd*I93 mutation results in the loss of *amd*R mediated omega amino acid induction of *amd*S (Hynes, 1980). This mutation is a deletion of nucleotides -181 to -151 (Hynes *et al.*, 1985, 1988). There are similar sequences to the region within the *amd*I93 deletion in the promoter region of the *amd*R regulated *gat*A gene (Richardson *et al*, 1989). The *amd*I9 mutation results in increased *fac*B mediated acetate induction and is caused by a single T -> C substitution at position -210. The *amd*I66 mutation results in increased *amd*A mediated acetate induction of *amd*S in strains carrying semidominant *amd*A mutations (Hynes, 1982), and is a duplication of nucleotides -107 to -90 (Hynes *et al.*, 1985,1988). Another mutation, *amd*I666, has been isolated in an *amd*A$^+$ *amd*I66 background. This strain had a similar phenotype to the *amd*A7 *amd*I66 double mutant, and resulted from a triplication of the region duplicated in the *amd*I66 mutant. This result showed that a wildtype *amd*A gene product is capable of inducing high levels of *amd*S expression and the increase in expression is likely to be due to an increased probability of the *amd*A gene product occupying one of the three sites (Katz *et al.*, 1990).

Strains transformed with muliple copies of the wildtype *amd*S 5' region showed reduced growth on media containg omega amino acids as sole sources of carbon or nitrogen. Reduced growth on acetate media was also seen in transformants containing multiple copies of the *amd*S 5' region containing the *amd*I9 mutation. These results suggest that these sequences are capable of titrating the *amd*R and *fac*B regulatory proteins, as these transformants showed reduced growth on media on which growth required the expression of *fac*B and *amd*R regulated genes. The requirement for the *amd*I9 mutation for efficient titration of the *fac*B gene product correlated with the genetic evidence that this mutation leads to increased *fac*B mediated induction, presumably

through tighter binding of the *facB* gene product (Kelly and Hynes, 1987).

The investigation of which sequences within the *amdS* promoter caused titration of proteins that regulate *amdS* expression is another approach that has been used to locate the sites of action of these regulatory proteins. Transformation of an *amdI9* strain with subclones of the *amdS* promoter showed that reduced growth on acetate medium and hence titration of *facB* was achieved with constructs containing the region surrounding the *amdI9* mutation. Titration of *facB* was positioned between nucleotides -181 and -219. These experiments also showed that a similar sequence to that surrounding *amdI9* located between nucleotides -558 and -606 did not have a high affinity for *facB* binding. Similar analysis positioned *amdR* titration between nucleotides -151 and -219. Constructs containing the *amdI93* deletion (nucleotides -151 to -181) failed to titrate the *amdR* gene product. Titration of the *amdA7* gene product was also achieved and was positioned in the region of the *amdI66* duplication. Thus the results gained from these titration experiments were consistent with the positioning of the sites of action of *amdA*, *amdR* and *facB* as determined by the physical characterization of 5' mutations (Hynes *et al.*, 1988).

Further confirmation of these results was gained using a third approach. An *amdS* deletion strain was transformed with a construct containing a deletion to nucleotide -111 of the *amdS* promoter and the entire *amdS* coding region. As expected, this deletion abolished *amdR* and *facB* control and was still regulated by *amdA*. This construct also remained under *areA* control, and thus *areA* acts 3' of nucleotide -111. The expression of this construct also indicated that sequences 5' to nucleotide -111 are not required for *amdS* expression even though the up promoter *amdI18* mutation was positioned at nucleotide -118 (Hynes *et*

al., 1988). The titration of *amd*R and *fac*B gene products can be reversed by transformation into the titrating strains of multiple copies of the cloned *amd*R and *fac*B genes (Andrianopoulos and Hynes, 1988; Katz and Hynes, 1990). The site of *cre*A control within the *amd*S promoter, if in fact *cre*A acts directly on promoters to mediate carbon catabolite repression, has not been determined.

The molecular characterization of a relatively large number of genetically characterized mutations that affect *amd*S expression, together with the studies involving titration of regulatory gene products, has lead to a good understanding of how the complex mechanisms controlling *amd*S expression operate. This work represents one of the best understood systems of regulated gene expression in filamentous fungi.

1.2.4: Other studies of regulated gene expression in *Aspergillus nidulans*.

The quinic acid utilization cluster (*qut*) is another example where coregulated genes encoding enzymes of the same metabolic pathway are clustered. Several enzyme encoding structural genes, a gene for quinic acid permease and two regulatory genes are found in a cluster on linkage group VIII. The genetics and molecular biology of quinic acid utilization has also been extensively studied in *Neurospora crassa*, and thus comparisons of gene structure and regulation can be made, and have aided the analysis of the *qut* cluster in *A. nidulans*.

The structural genes encoding quinate dehydrogenase (*qut*B), catabolic 3-dehydroquinase (*qut*E) and dehydratase (*qut*C) are induced in the presence of quinic acid, and repressed in the presence of

glucose. Mutations at any of these loci result in the loss of a single enzyme activity (Bailey and Arst, 1975; Hawkins *et al.*, 1982, 1984; Grant *et al.*, 1988). Carbon catabolite repression of *qut*B is unaffected in *cre*A^{d}-1 mutant strains, but is affected by *cre*B and *cre*C mutations, where quinate dehydrogenase enzyme levels are similar to wildtype repressed levels in both carbon catabolite repressing and derepressing conditions (Bailey and Arst, 1975; Hynes and Kelly, 1977). Mutations at the *qut*D locus result in the loss of all three enzyme activities (Hawkins *et al.*, 1984). This locus was originally thought to encode a positively acting regulatory protein, but has since been shown to encode a permease for quinic acid (Whittington *et al.*, 1987). Mutations at the *qut*R locus result in constitutive synthesis of all three enzymes, but the expression of the structural genes is still subject to carbon catabolite repression in *qut*R^{-} strains (Grant *et al.*, 1988). The *qut* cluster has been cloned and several of the genes have been sequenced (Hawkins *et al.*, 1985; Da Silva *et al.*, 1986; Beri *et al.*, 1987; Whittington *et al.*, 1987; Hawkins *et al.*, 1988). The molecular organization of the region (Hawkins *et al.*, 1988) and fine structure genetic mapping (Grant *et al.*, 1988) has determined the gene order. Molecular analysis also idenified another putative gene, *qut*G, which showed significant sequence similarity to the *qa*-X gene of *Neurospora crassa*. Both *qut*G and *qa*-X are of unknown function. The nucleotide sequence of the other genes within the *qut* cluster are also significantly similar to genes of homologous function in *N. crassa*. Several conserved sequences upstream of genes within the *qut* cluster were hypothesized to be involved in regulation by the control of transcription (Hawkins *et al.*, 1988). The interest in studies of quinic acid utilization in *A. nidulans* is multiplied by the comparisons of gene structure and regulation that can be made with the quinic acid utilization cluster (*qa*) of *N. crassa* (Giles *et al.*, 1987 and

references therein). Furthermore, the quinic acid utilization pathway in *A. nidulans* shares intermediates with polyaromatic amino acid synthesis which is controlled by the pentafunctional AROM polypeptide (Hawkins, 1987).

Classical genetic analysis of the genes involved in nitrate assimilation (reviewed by Cove, 1979 and Arst and Scazzocchio, 1985), has shown that three structural genes are clustered on linkage group VIII. The genes encode a nitrate permease (*crn*A), nitrate reductase (*nia*D) and nitrite reductase (*nii*A). Induced expression of *nii*A and *nia*D but not *crn*A is positively controlled by the unlinked *nir*A gene. Mutations at *nir*A fall into three classes. Recessive *nir*A^{-} mutations render the structural genes noninducible. Rare *nir*A^{c} mutations result in constitutive expression. *nir*A^{d} mutations alleviate the need for the *are*A gene product. Double *nir*$^{c/d}$ mutants have been constructed, and thus there are at least two regulatory domains within the *nir*A gene product involved in coinducer recognition and interactions with *are*A protein or the *are*A binding sequence. Mutations at any of several *cnx* loci lead to the absence of a nitrate reductase cofactor. Most *nia*D and *cnx* mutations result in constitutive synthesis of nitrite reductase and inactive nitrate reductase, and therefore it was concluded that nitrate reductase affected the expression of the nitrate reductase encoding *nia*D gene. There have been three models proposed by which the enzyme could affect the expression of its gene: either the *nir*A gene product is inactive when complexed with nitrate reductase, and this is prevented by the presence of coinducers; or, unless nitrate reductase is complexed with a coinducer it prevents *nir*A function; or, each of the proteins interacts with its own genes control region (reviewed by Arst and Scazzocchio, 1985). There have been no reports of molecular investigations of the regulation of these genes, although a clone of

*nia*D has been obtained (Malardier *et al.*, 1989).

Not all studies of gene expression in *A. nidulans* have dealt with the study of metabolism. Several genes controlling asexual development have been cloned and their role in development determined. It is estimated that approximately 1200 genes are expressed selectively during asexual development (Timberlake and Marshall, 1988). A large cluster of developmentally regulated genes (SpoC1) have been identified and cloned (Miller *et al.*, 1987 and references therein), but deletion of this region did not have any associated phenotype (Aramayo *et al.*, 1989). The timing of expression of the three regulatory genes *brl*A, *aba*A, and *wet*A has been shown to be important in controlling development. Mutations at these loci affected the expression of genes that show developmental regulation (Boylan *et al.*, 1987). Changing the timing of *aba*A and *brl*A expression by fusing the coding regions of these genes to the *alc*A promoter also affected development. Physical characterization of the *aba*A and *brl*A genes has shown their protein products contained structures common to putative DNA binding proteins (Adams *et al.*, 1988; Mirabito *et al*, 1989). Many temperature sensitive mutations that affect mitosis have been isolated (Morris, 1976). Molecular analysis of genes involved in controlling cell division isolated using these mutants have identified the products of some of these regulatory genes. These include a protein kinase and a protein phosphatase which are likely to affect the function of other molecules involved in mitosis (Osmani *et al.*, 1988; Doonan and Morris, 1989; reviewed by Morris, 1989).

1.2.5: Studies of gene regulation in *Aspergillus niger*.

The few reports regarding the regulation of gene expression in *A. niger* are mentioned below. It is not known whether general mechanisms controlling the expression of various classes of genes in *A. nidulans* and other filamentous fungi also occur in *A. niger*. A gene encoding an acid phosphatase, the expression of which was repressed in the presence of exogenous phosphate has been cloned (MacRae *et al.*, 1988). Analysis of cDNA and genomic clones of the single glucoamylase gene in *A. niger* revealed that two isozymes of glucoamylase known as glucoamylase G1 and G2 were synthesized from two types of mRNA molecules that differed by differential splicing of an intron (Boel *et al.*, 1984a, 1984b). The inducible synthesis of polygalacturonase and pectinesterase has been shown to subject to carbon catabolite repression. The authors concluded that this repression acted at the level of translation, but the data was unconvincing and experiments performed in the same laboratory showed repression to be acting on transcription (Maldonado *et al.*, 1989). Experiments performed by Kelly and Hynes (1985) showed that *A. niger* grew very poorly on acetamide as a nitrogen source, but transformation of *A. niger* with a cloned *amd*S gene of *A. nidulans* allowed selection of transformants capable of strong growth on acetamide or acrylamide media. High copy number transformants were however impaired for growth on omega amino acids. This result suggested that a gene product of homologous function to that of the *amd*R gene in *A. nidulans*, which induces the expression of genes encoding enzymes required for omega amino acid utilization, was being titrated by the *amd*S sequences.

Despite the large amount of research to improve the production of commercial compounds in *A. niger* for industrial applications, there has been little basic research in *A. niger*. The commercial importance of *A. niger* may mean that a lot of industry based research is not presented in publications. The amount of research into gene expression in *A. niger* is insignificant compared to the amount of published work that has been undertaken in other filamentous fungi such as *A. nidulans* and *N. crassa*.

1.2.6: Gene structure in filamentous fungi.

Analysis of the nucleotide sequences of genes cloned from filamentous fungi and the structure of their mRNA transcripts have revealed some generalities in gene structure (reviewed by Ballance, 1986 and Gurr *et al.*, 1987). Studies of gene regulation ultimately involve the identification of DNA sequences, frequently in the 5' regions of genes, that are involved in the control of gene expression. Therefore it is important to be able to distinguish between promoter elements involved specifically in gene regulation from those involved more generally in transciption. Unlike in *S. cerevisiae*, the genes of filamentous fungi analyzed thus far are, in general, similar in structure to those of higher eukaryotes, in that they have sequences in common to other eukaryotes that fuction as transcription and processing signals. A possible exception to this is polyadenylation signals.

Sequences resembling the TATA box of higher eukaryotes, with the consensus sequence of TATAA/TAA/T (reviewed by Serfling *et al.*, 1985), occur in many but not all promoters of filamentous fungal genes. When present, this sequence is located approximately 30 - 60 bp. upstream of the major transcriptional start points. A second AT rich sequence positioned approximately 275 bp. 5' to the start points of

transcription has been identified by Nussinov *et al.*, (1986). While not all fungal genes have been sequenced this far, AT rich sequences at or near this position are found in some genes, for example the *A. nidulans* genes *pki* (de Graaff and Visser, 1988), *pgk*A (Clements and Roberts, 1986), *pyr*G (Oakley *et al.*, 1987), and *gpd*A (Punt *et al.*, 1988). CAAT box sequences are also found in many promoters but at variable distances 5' to putative TATA boxes. The variable position and similarity to consensus of these motifs in fungal promoters may, in the absence of experimental evidence question their validity in controlling transcription. Another promoter element found in the genes of filamentous fungi is a pyrimidine rich region. This element is particularly evident in highly expressed genes, as it is in some highly expressed genes in *S. cerevisiae* (Dobsen *et al.*, 1982). This sequence is usually positioned directly upstream of the major start point of transcription. A region deleted from the *A. nidulans trp*C promoter which included, albeit small, a pyrimidine rich sequence resulted in transcription initiating from heterogeneous sites (Hamer and Timberlake, 1987). This result, together with the observed positions of these sequences, may implicate the pyrimidine rich sequences in playing a role in directing the site of transcription initiation.

There is considerable similarity in nucleotide sequence around the start points of translation of sequenced genes, with an A at position -3 with respect to the start codon being very prevalent. Codon usage shows a strong bias against A and for C in the third position. The AGN codons of serine and arginine are extremely rare. As in *S. cerevisiae*, codon bias is less marked in genes that exhibit a low level of expression (reviewed by Gurr *et al.*, 1987).

Introns are common in the genes of filamentous fungi, and are found throughout the coding regions of genes. This is in contrast to *S. cerevisiae* where only about 10% of sequenced genes have been shown to contain introns which are generally positioned at the 5' end of the gene. The GT - AG intron/exon junctions are present in almost all cases. A putative intron lariat sequence with consensus of YGCTAACN, with the central T being almost invariant, which resembles the TACTAAC intron lariat sequence of *S. cerevisiae*, is found in the appropriate position of all mapped introns (reviewed by Gurr *et al.*, 1987).

The AAUAAA polyadenylation signal identified by Proudfoot and Brownlee (1976) is not a general feature of the 3' ends of fungal mRNAs, but is sometimes present either in its entirity or in a truncated form. The proposed YGTGTTYY termination siganl found 3' of AAUAAA in mammalian genes (McClauchlan *et al.*, 1985) is also not a general feature (reviewed by Gurr *et al.*, 1987).

Several genes have been cloned from *A. niger* that have also been cloned from *A. nidulans*. Comparison of nucleotide sequences of the *pyr*G and *arg*B genes of *A. niger* and *A. nidulans* have shown a level of sequence similarity of approximately 70% (Buxton *et al.*, 1987; Wilson *et al.*, 1988). This level of sequence similarity represents considerable genetic distance between *A. niger* and *A. nidulans*, and correlates with data based on 5S ribosomal RNA sequences comparisons (Chen *et al.*, 1984) which showed *A. niger* and *A. nidulans* to be less related to each other than to other non-*Aspergillus* fungi such as *Penicillium chrysogenum*. The classification of *A. niger* and *A. nidulans* into the same genus has been on the basis of morphological characteristics.

1.2.7: Properties of purified aldehyde dehydrogenases.

As experiments presented in this thesis in part involve the analysis of the derived primary structure of the *A. niger* aldehyde dehydrogenase enzyme, it is pertinent to discuss properties of purified AldDHs. Compared to the large amount of work published on biochemical properties of many purified alcohol dehydrogenases, few aldeyde dehydrogenases have been purified. A primary reason for this was the comparative difficulty in purifying AldDHs until the 1970s. The properties of the purified fungal and mammalian AldDHs is discussed below.

Two AldDHs have been purified from *S. cerevisiae*. One enzyme, purified by Seegmiller (1953) was activated by divalent cations and reacted only with NADP. This enzyme was shown to exist in the cytosol and was only partly repressed during growth on glucose (Jacobson and Bernofsky, 1974). A mitochondrial enzyme from *S. cerevisae* has also been purified. This enzyme was 99% repressed during growth on glucose and was involved in the oxidative metabolism of ethanol. Enzyme activity requires K^+ and a thiol reagent. This enzyme reacted with both NAD and NADP. The K_m and V_{max} values for a number of aldehydes were determined and showed propionaldehyde, acetaldehyde and succinic semialdehyde to be amoungst the best substrates (Clark and Jakoby, 1970; Jacobson and Bernofsky, 1974; Tamaki *et al.*, 1977; Bostian and Betts, 1978).

Enzymic properties of the *ald*A encoded AldDH of *A. nidulans* suggested it to be very similar to the mitochondrial AldDH of *S. cerevisiae*. This AldDH was also dependent on K^+, a thiol reagent, and active with either NAD or NADP. The preferred substrate of aldehydes

tested was acetaldehyde (Creaser *et al.*, 1987). A second AldDH in *A. nidulans* has not been detected.

Both cytoplasmic and mitochondrial AldDHs have been purified from humans, and their amino acid sequence determined (Hempel *et al.*, 1984, 1985). A partial cDNA clone of the mitochondrial AldDH gene (ALDH2) and a full length cDNA clone of the cytoplasmic AldDH gene (ALDH1) have also been obtained (Hsu *et al.*, 1985; Braun *et al.*, 1987). A cytoplasmic AldDH enzyme has also been isolated and its amino acid sequence determined (Bahr-Lindstrom *et al*, 1984). The cytoplasmic enzyme was inactivated by disulfiram, whereas the mitochondrial enzyme was much less sensitive. A cysteine residue within a hydrophobic segment reactive with iodoacetamide and sensitive to disulfiram inactivation was hypothesized to be adjacent to another cysteine within a hydrophobic cleft at the active site. The sequence of the first 21 amino acids was highly conserved between the mitochondrial enzymes of man and horse but 19 differences were observed in this region between the two isoenzymes in man. This observation lead to the hypothesis that this region was involved in subcellular targeting of the isoenzymes (Hempel *et al.*, 1984, 1985; Bahr-Linstrom *et al.*, 1984).

1.3: Aims of this study.

The primary aim of this research was to investigate mechanisms of gene regulation in *A. niger*. A comparative study of gene structure and regulation with a gene of homologous function in *A. nidulans* was undertaken so that data already gained in *A. nidulans*, as well as mutant strains and clones of genes could be used in this research. Similar approaches have been used elsewhere, for example comparisons of quinic acid utilization and ammonium repression in *A. nidulans* and *N. crassa*, and frequently result in obtaining more data about both systems through defining and explaining differences and similarities between the two systems.

The aldehyde dehydrogenase encoding genes were chosen for several reasons. Compared to *A. nidulans*, *A. niger* shows very poor growth on ethanol and threonine media (chapter 3), which suggested some difference in the utilization of compounds via acetaldehyde. A fully characterized clone of the *ald*A gene and many mutations at both the *ald*A locus and at loci affecting *ald*A expression in *A. nidulans* have been isolated, both of which could be used to isolate and analyze the expression of the *A. niger ald*A gene. Although much is known about alcohol dehydrogenase gene structure and regulation in many systems including humans and several other mammals, Maize, *Drosophila*, Yeast and *A. nidulans*, comparatively little is known about aldehdye dehydrogenase genes. Futhermore, in addition to the metabolism of threonine, and some alcohols and amines, it appears, at least in *A. nidulans*, aldehyde dehydrogenase, but not alcohol dehydrogenase, plays a role in the metabolism of L-proline and GABA. Although much was known

about *ald*A gene expression in *A. nidulans*, many more experiments needed to be performed to allow comparisons to *A. niger*, and this also revealed new data on *ald*A gene regulation in *A. nidulans*.

The initial aim was to isolate a clone of an aldehdye dehydrogenase encoding gene of *A. niger*. Once the clone had been obtained, the project was divided into two parts:

(i) To determine the effects of various growth conditions and regulatory mutations on *ald*A gene regulation in *A. niger* and *A. nidulans*. This involved an analysis of *ald*A gene regulation in both wildtype strains and heterologous expression of *ald*A genes across the two species.

(ii) To physically characterize the *ald*A gene of *A. niger*, including the location of regions of the promoter responsible for the control of gene expression.

Results obtained from these experiments are presented and discussed in the following chapters.

Chapter 2:
Materials and methods.

2.1: Materials.

DNA modification enzymes: Restriction Endonucleases, Polynucleotide Kinase, Klenow fragment of *E. coli* DNA Polymerase I (sequencing grade), Calf Intestinal Phosphatase, Nuclease S1, and Deoxyribonuclease I were purchased from Boehringer-Mannheim GmbH. AMV reverse transcriptase was purchased from International Biotechnologies, Inc.. Bacteriophage T4 DNA Ligase and *E. coli* DNA Polymerase I were purchased from BRESATEC Pty. Ltd.. EcoRI methylase was purchased from New England Bio-Labs.

Nucleotides: All unlabelled deoxyribonucleotides and dideoxyribonucleotides were purchased from Boehringer-Mannheim GmbH. α-^{32}P-dCTP (3000 Ci/mmole) and γ-^{32}P-dATP (4000 Ci/mmole) were purchased from BRESATEC Pty. Ltd.

Other enzymes: Novozyme was purchased from Novo Industries. Proteinase K was purchased from Boehringer-Mannheim GmbH. Lysozyme, β-glucuronidase, Pronase, and Ribonuclease A were purchased from Sigma Chemical Company.

General reagents and media requirements: All general reagents, media requirements and supplements were of laboratory grade and were purchased from Sigma Chemical Company, BDH Ltd., Aldrich Chemical Company Inc., Oxoid Ltd., and United States Biochemical Corp. Ampicillin, isopropylthiogalactoside (IPTG), and 5-bromo-4-chloro-3-

indolyl-β-D-galactoside (X-gal) were purchased from Boehringer-Mannheim GmbH.

2.2: Reaction buffers, solutions, and media.

2.2.1: 1 X Reaction buffers.

Restriction endonucleases: Reaction buffers for Restriction Endonuclease digests were supplied with the enzymes by Boehringer-Mannheim GmbH.

Ligation buffer: 10mM $MgCl_2$, 10mM DTT, 1mM ATP, 1mM Spermidine, 50mM Tris.HCl (pH 7.5).

Polynucleotide kinase buffer: 50mM Tris.HCl (pH 7.5), 10mM $MgCl_2$, 5mM DTT.

Calf intestinal phosphatase buffer: 50mM Tris.HCl (pH 9), 1mM $MgCl_2$, 0.1mM $ZnCl_3$, 1mM Spermidine.

TM buffer (DNA sequencing): 10mM Tris.HCl (pH 8), 10mM $MgCl_2$.

Nick translation buffer: 50mM Tris.HCl (pH 7.5), 7.5mM $Mg(CH_3COO)_2$, 4mM DTT, 100 ug/ml BSA.

2.2.2: Solutions:

1 X SSC: 0.15M NaCl, 0.15M $Na_3C_6H_5O_7.2H_2O$, pH 7.2

1 X SSPE: 0.18M NaCl, 10mM NaH_2PO_4, 1mM EDTA, pH 7.4.

1 X Denhardts: 0.02% (w/v) Ficoll, 0.02% (w/v) Polyvinylpyrrolidone, 0.02% (w/v) BSA (Pentax Fraction V).

1 X TE: 10mM Tris.HCl (pH 8), 1mM EDTA.

SM buffer: 100mM NaCl, 10mM $MgSO_4.7H_2O$, 50mM Tris.HCl (pH 7.5), 0.1% (w/v) Gelatin.

1 X STE: 10mM Tris.HCl (pH 7.5), 100mM NaCl, 1mM EDTA.

1 X TAE: 40mM Tris, 20mM $NaCH_3COO$, 2mM EDTA, pH 7.8.

1 X TBE: 89mM Tris, 89mM H_3BO_4, 2mM EDTA, pH8.4.

Sucrose buffer: 15% (w/v) Sucrose, 50mM Tris.HCl (pH 8.5), 50mM EDTA.

Triton buffer: 5% (v/v) Triton X-100, 50mM Tris.HCl (pH8.5), 50mM EDTA.

Trace elements solution: (Per Litre) 40mg $Na_2B_4O_7$, 400mg $CuSO_4$, 1g $FePO_4$, 600mg $MnSO_4.H_2O$, 800mg $Na_2MoO_4.2H_2O$, 8g $ZnSO_4.7H_2O$, 2ml $CHCl_3$ (preservative).

Vitamin solution: (Per Litre) 40mg p-aminobenzoic acid, 50mg aneurin.HCl, 1mg biotin, 400mg inositol, 100mg nicotinic acid, 200mg Calcium D-pantothenate, 100mg riboflavin, 50mg pyridoxine, 2ml $CHCl_3$ (preservative).

Salt solution: (Per Litre) 26g KCl, 26g $MgSO_4.7H_2O$, 76g KH_2PO_4, 50ml Trace Elements Solution, 2ml $CHCl_3$ (preservative).

2 X M9 salts: (Per Litre) 12g Na_2HPO_4, 6g KH_2PO_4, 1g NaCl, 2g NH_4Cl.

2.2.3: Bacterial growth media.

L-broth: 1% (w/v) NaCl, 0.5% (w/v) Oxoid Yeast Extract, 1% (w/v) Oxoid Tryptone, pH 7.5. Plates solidified with either 1.5% (w/v) Agarose (Sigma) or 1.5% (w/v) Oxoid Class I Agar.

M9 glucose minimal: 50% (v/v) 2 X M9 Salts, 10mM $MgSO_4$, 1mM $CaCl_2$, 10mM Thiamine.HCl, 0.2% (w/v) glucose, 50% (v/v) [3% (w/v) Oxoid Class I Agar] (Molten).

Soc media: L-broth plus: 10mM $MgSO_4.7H_2O$, 10mM $MgCl_2$, 2.5 mM KCl, 1% (w/v) glucose.

2 X YT: 0.5% (w/v) NaCl, 1% (w/v) Oxoid tryptone, 1.6% (w/v) Oxoid yeast extract, pH 7.5.

2.2.4: *Aspergillus* growth media.

Nitrogen free medium: (glucose media) 2% (v/v) salt solution, 1% (w/v) glucose, pH 6.5. Plates solidified with Oxoid Class I Agar added to either 1% (w/v) or 2.2% (w/v). Unless otherwise indicated, nitrogen sources were added to 10mM.

Carbon free medium: 2% (v/v) salt solution, pH 6.5. Plates solidified with Oxoid Class I Agar added to either 1% (w/v) or 2.2% (w/v). Unless otherwise indicated, carbon sources were added to 50mM.

Complete medium: 1% (w/v) glucose, 0.2% (w/v) peptone, 0.15% (w/v) caesein hydrolysate, 0.1% (w/v) yeast extract, 10mM ammonium (+)-tartrate, 2% (v/v) salt solution, 1% (v/v) vitamin solution, 25 ug/ml riboflavin, pH 6.5. Plates solidified with Oxoid Class III Agar added to either 1% (w/v) or 2.2% (w/v).

Protoplast medium: 1M sucrose, 1% (w/v) glucose, 2% (v/v) salt solution, pH 7. Solidifed with Oxoid Class I Agar added to 0.25% (w/v) for overlays or 1% (w/v) for plates.

Supplements: the following supplements were added when required to the indicated final concentration: biotin - 0.01ug/ml; nicotinic acid - 1ug/ml; p-aminobenzoic acid - 50ug/ml; sodium thiosulphate - 0.1% (w/v); uridine - 10mM.

2.3: *Aspergillus strains*.

The genotypes of strains used in this study are shown in Table 2.3.1. For meanings of *A. nidulans* gene symbols see Clutterbuck, (1984). The wildtype strain of *A. niger* used was that of Kelly and Hynes, (1985). Gene symbols used for *A. niger* are the same as those used for homologous loci in *A. nidulans*.

2.4: *E. coli* strains and bacteriophage.

E. coli strain DH1 (Hannahan, 1983) was used for general plasmid maintainance. EMBL 3 lambda bacteriophage were propagated on *E. coli* strain NM538 (Frischauff *et al.*, 1983). M13 bacteriophage were propagated on *E. coli* strain JM101 (Yannisch-Perron *et al.*, 1985). JM101 was also used to screen for recombinant plasmids for cloning experiments using pUC19 as a vector. Recombinant λgt10 bacteriophage were selected on *E. coli* strain C600*Hfl* and propagated on *E. coli* strain C600 (Maniatis *et al.*, 1982 p. 504) which were supplied with packaging extracts by Promega Corp.

Strain referred to as:	Genotype:
A. nidulans,	
wildtype	*bi*A1; *nii*A4
*ald*A67	*bi*A1; *nii*A4 *ald*A67
*alc*R125	*paba*A1; *s*B43; *alc*R125
*alc*R5	*bi*A1; *arg*B2; *alc*R5
*amd*R$^-$	*y*A1; *amd*R$^-$44; *nic*B8
*amd*R^c	*bi*A1; *amd*R^{c}6
*amd*A7	*bi*A1; *amd*A7
A. niger,	
wildtype	not mutant for any known marker
712	*ald*A$^-$712 (gene replacement, see Chapter 5)
216	*pyr*G216; *ald*A$^-$712

Table 2.3.1: Genotypes of *Aspergillus* strains used in this study.

2.5: Plasmids.

A description of plasmids used during this study is shown in Table 2.5.1.

Plasmid	Vector	Insert	Reference
pAN212	pUC12	*A. nidulans aldA*	Pickett *et al.*, (1987)
p3SR2	pBR322	*A. nidulans amdS*	Hynes *et al.*, (1983)
pANA19	pUC19	*A. nidulans amdS*	M. J. O'Connell, unpublished
pM006	pUC19	*A. nidulans argB*	Upshall, (1986)
pANIR1	pBR322	*A. niger* ribosomal repeat unit	M. J. O'Connell, unpublished
pAB4.1	pUC19	*A. niger pyrG*	van Hartingsveldt *et al.*, (1987)
pUC19	(used as general cloning vector)		Yannish-Perron *et al.*, (1985)

Table 2.5.1: Plasmids used in this study.

2.6: Methods.

2.6.1: Isolation of nucleic acids.

DNA was isolated from freeze-dried *Aspergillus* mycelia harvested from liquid cultures grown overnight in glucose medium by the method of Hynes *et al.*, (1983).

For RNA isolation from *Aspergillus*, cultures were grown overnight in glucose media, harvested, washed with a large volume of prewarmed carbon free medium, and transferred to the appropriate prewarmed medium for three to six hours. These cultures were harvested, washed with a large volume of sterile distilled water, and freeze-dried overnight.

Total RNA was isolated from freeze-dried mycelia by the method of Rienert *et al.*, (1981). Poly(A)$^+$ RNA was separated from the poly(A)$^-$ fraction by affinity chromatography through 1ml columns of oligo-dT cellulose (Sigma) by the method of Maniatis *et al.*, (1982) pp. 197-198.

DNA was prepared from lambda bacteriophage by the plate lysate method of Maniatis *et al.*, (1983) pp. 65-66.

Large scale preparation of plasmid DNA was performed by Triton X-100 lysis. *E. coli* cells harboring the desired plasmid were grown to confluence overnight in 500ml Soc medium plus 50ug/ml ampicillin. These cells were harvested by centrifugation at 3500g for 10 minutes, washed in 100ml 1 X STE buffer and repelleted. The cells were resuspended in 40ml of Sucrose buffer plus 2mg/ml lysozyme and incubated on ice for 10 minutes. 4ml of Triton buffer was added and the tubes were incubated at room temperature for a further 10 minutes. Bacterial debris was pelleted by centrifugation at 18000 r.p.m. in a Sorvall SS34 rotor for 60 minutes. The supernatant was collected, extracted twice with an equal volume of tris-buffered phenol/chloroform/isoamyl alcohol (50:48:2), and nucleic acids were precipitated with two volumes of ethanol after the addition of sodium acetate to 0.3M. The precipitate was resuspended in 10ml 1 X TE plus 1mg of heat treated ribonuclease A, and incubated at 37 °C for 60 minutes. Plasmid DNA was seperated from *E. coli* chromosomal DNA on CsCl equilibrium gradients by the method of Maniatis *et al.*, (1982) pp. 93-94.

Small scale preparation of plasmid DNA was performed by the alkali lysis method of Ish-Horowicz and Burke, (1981).

Single plaques of M13 bacteriophage were propagated in 2ml liquid cultures of *E. coli* strain JM101 in 2 X YT medium as described by

Ausubel *et al.*, (1987) p. 7.3.9. These cultures were used for preparation of single stranded M13 templates for DNA sequencing. The cultures were centrifuged at 3500g for 10mins, and 1.5ml of the supernatant was collected in a microcentrifuge tube. This was spun at 12000 r.p.m. in a microcentrifuge for 10 minutes, and the supernatant was decanted into a fresh microcentrifuge tuge. To this, 275ul of [2.5M NaCl, 20% (w/v) polyethylene glycol 8000] was added and the tubes incubated at room temperature for 15 minutes. Bacteriophage particles were pelleted by microcentrifugation at 12000 r.p.m. for 10 minutes. The supernatant was decanted and discarded. The tubes were spun for a further 2 minutes, and the residual supernatant was removed with a drawn-out pasteur pipette. The bacteriophage were resuspended in 200ul of 1 X TE, extracted with 100ul of tris-buffered phenol, with 150ul of the aqueous layer being collected. This was extacted with 75ul of water saturated chloroform, with 100ul of the aqueous layer being collected. M13 DNA was precipitated from this by the addition of 11ul of 4M LiCl and 275ul of ethanol with a 15 minute incubation at -70 °C. The DNA was collected by microcentrifugation at 12000 r.p.m., washed in 70% (v/v) ethanol, dried and resuspended in 12ul of 1 X TE.

2.6.2: General recombinant DNA methods.

General methods for DNA manipulations, propagation of bacteria and bacteriophage, and transformation of *E. coli* as indicated by suppliers of DNA modification enzymes amd published in laboratory manuals (Maniatis *et al.*, 1982; Ausubel *et al.*, 1987). Probes for nucleic acid hybridizations were prepared by nick translation. DNA (0.1-0.5ug) was incubated at room temperature in a total volume of 35ul in 1 X nick translation buffer, 1ug deoxyribonuclease I, 25uM dATP, 25uM dGTP and 25uM dTTP for 15 minutes. This was then incubated at 65 °C for 10

minutes and cooled. 10 units of *E. coli* DNA polymerase I and 20uCi of $-^{32}$P-dCTP was added, and the reaction was continued at 15 °C for a further 15-30 minutes. Unincorporated nucleotides were separated from labelled DNA using Bio-Rad Biogel P60 spun columns by the method of Davis *et al.*, (1980).

2.6.3: Electrophoresis.

Agarose electrophoresis of DNA was carried out in gels made from 1 X TAE buffer and an appropriate concentration of agarose depending on the molecular weight of the DNA under analysis. DNA was recovered from agarose gels by the freeze-squeeze method of Thuring *et al.*, (1975). Bacteriophage lambda DNA digested with HindIII or bacteriophage SPP-1 DNA digested with EcoRI was used as markers of molecular weight.

RNA was electrophoresed in 1.5% (w/v) agarose, 8% (w/v) formaldehyde, 10mM sodium orthophosphate (pH 7) gels with 10mM sodium orthophosphate (pH 7) as the running buffer. Two volumes of [50% (v/v) formamide, 12% (w/v) formaldehyde, 10mM sodium orthophosphate (pH 7)] were added to RNA samples and this was incubated at 65 °C for 10 minutes prior to electrophoresis.

Products of DNA sequencing reactions were electrophoresed in gels containing 4-8% (w/v) polyacrylamide (depending on the molecular weight of DNA fragments to be separated), 8M urea, 1 X TBE. The dimensions of sequencing gels was 450x170x0.2mm, and gels were run at 2000-2200 volts.

2.6.4: Nucleic acid blotting and hybridizations.

Southern blotting of DNA gels, northern blotting of RNA gels, and dot blotting of both DNA and RNA using a Schleicher and Schuell

minifold II apparatus was to Zeta-Probe membrane (Bio-Rad) by the alkali transfer methods recommended by the supplier of the membrane. Plaque lifting to nitrocellulose was performed by the method of Benton and Davis, (1977). Colony blots on nitrocellulose were performed by the SDS/NaOH method of Maniatis *et al.*, (1982) p. 314.

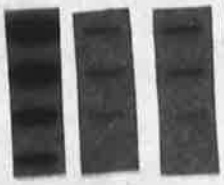
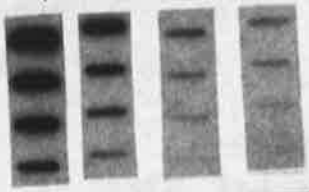
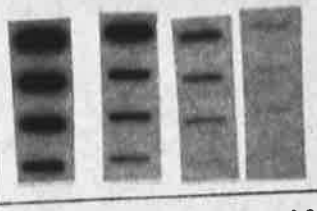
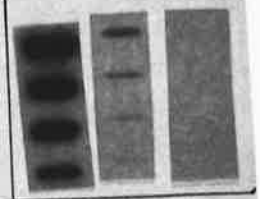
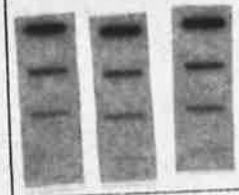
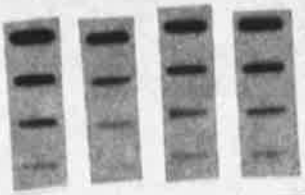
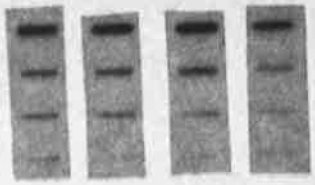
Hybridization of nick translated probes to Zeta-Probe membrane and subsequent washing of filters after hybridization was performed by the method recommended by the membrane supplier. For nitrocellulose membranes, prehybridization of filters for 2 - 4 hours at 42 °C was carried out in 50% (v/v) formamide, 4 X SSPE, 10 X Denhardts, 0.1% (w/v) SDS, and 30ug/ml sheared denatured salmon sperm DNA. Hybridization of denatured probes was under the same conditions except with the addition of 10% (w/v) dextran sulphate. Filters were washed to a final stringency of 0.1 X SSC, 1% (w/v) SDS at 65 °C. For heterologous probing of plaque lifts, the conditions were changed to include 5 X SSPE and 25ug/ml sheared denatured *E. coli* DNA in the prehybridization, and the final washing stringency was 0.5 X SSC, 1% (w/v) SDS at 42 °C.

2.6.5: Determination of mRNA levels and DNA copy number.

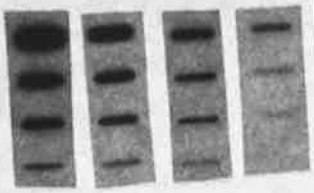
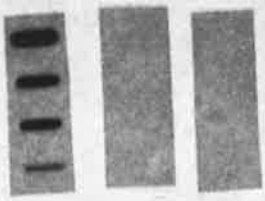
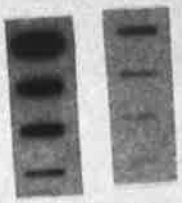
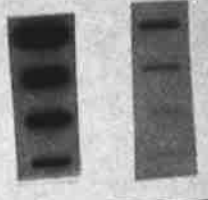
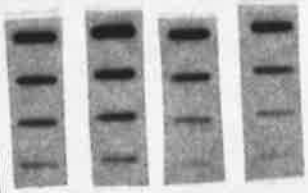
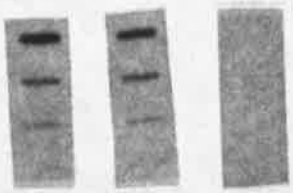
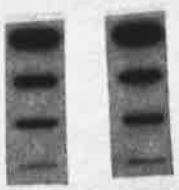
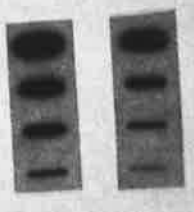
The insert of pNG100 and the insert of pAN212 were used to detect the *A. niger* and *A. nidulans aldA* transcripts respectively. Duplicate dot blots of a dilution series of total RNA extracted from mycelia grown under various conditions were hybridized with both the *aldA* specific probe and another probe to standardize loadings. The probes used to standardize loadings were: for *A. nidulans* samples, the insert of pM006 which detects the *argB* transcript; for *A. niger* samples, the insert of pAB4.1, which detects the *pyrG* transcript or the insert of

Figure 2.6.1: Representative slot blots determining relative *ald*A expression levels. Slots contain a dilution series of (from top to bottom) 2.0, 1.0, 0.5, and 0.1 *u*g of total RNA. Unless otherwise indicated, RNA was extracted from wildtype strains. AU/*u*g: absorbance units per *u*g of total RNA blotted from the linear range as determined by lazer scanning densitometry.

A. nidulans RNAs

	Experiment 1			Experiment 2				Experiment 3				Experiment 4		
	Threonine	Threonine + Glucose	Glucose (prep 1)	Fructose + Ethanol (prep 1)	Fructose (prep 1)	Ethanol (prep 1)	Glucose (prep 1)	Fructose + Ethanol (prep 2)	Fructose (prep 2)	Ethanol (prep 2)	Glucose (prep 2)	Threonine	Glucose (prep 2)	Transfer Buffer
Probe: aldA														
AU/ug	6.0	0.25	0.26	5.5	1.4	0.32	0.18	6.1	1.55	0.5	0.22	7.6	0.35	0
Probe: argB														
AU/ug	0.8	0.8	0.8	1.6	1.1	1.2	1.25	0.8	0.7	0.8	0.7	0.24	0.26	0
Corrected aldA mRNA level	23.1	1.0	1.0	23.8	8.8	1.8	1.0	24.2	7.0	2.0	1.0	23.5	1.0	0

A. niger RNAs

	Experiment 1				Experiment 2			Experiment 3		Experiment 4	
	Ethanol (prep 1)	Threonine	Threonine + Glucose	Glucose (prep 1)	Ethanol (prep 1)	Gene Replacement Strain 712 Ethanol	Transfer Buffer	Ethanol (prep 2)	Glucose (prep 2)	Ethanol (prep 1)	Glucose (prep 2)
Probe: aldA											
AU/ug	3.8	1.5	1.2	0.45	3.0	0	0	3.5	0.41	3.7	0.43
Probe: pyrG											
AU/ug	1.2	1.2	1.1	1.1	0.9	0.9	0	3.5	3.4	3.6	3.5
Corrected aldA mRNA level	7.7	3.0	2.6	1.0				8.3	1.0	8.4	1.0

pANIR1, which hybridizes to ribosomal RNA. Autoradiography was carried out in the absence of intensifying screens. The signal on autoradiograms was quantified using a LKB Ultroscan XL enhanced laser densitometer. The absorbance units detected by densitometry were plotted against the amount of RNA blotted (as determined by spectrophotometry) to obtain data points within a linear range. The arbitary value of absorbance units/ug for the *ald*A specific probes was corrected using the other probes, and values were then normalized against indicated samples (generally glucose grown wildtype samples). The same protocol was used for DNA to determine plasmid copy number in transformants.

2.6.6: DNA sequencing.

Subclones of the insert of pNG100 cloned into the bacteriophage vectors M13mp18 and M13mp19 were constructed for DNA sequencing by standard cloning techniques. DNA sequencing was performed by the dideoxynucleotide chain termination procedure (Sanger *et al.*, 1976). 2.5ng of the Bio-Labs -40 sequencing primer was annealed to 8ul of single stranded template DNA in 1 X TM buffer by heating the mixture to 100 °C and allowing to slowly cool. The annealed DNA was added to a tube containing 20uCi of α-^{32}P-dCTP (dried from ethanolic stock) and 1 unit of the Klenow fragment of *E. coli* DNA polymerase I. This was aliquoted at 2.5ul each into four tubes containing one N° solution (2ul) and 2ul of the corresponding ddNTP. An additional 2.5uCi was added to the C reaction. The reactions were incubated at 37 °C for 15 minutes. 2.5ul of chase (0.5mM dCTP, 0.5mM dGTP, 0.5mM dTTP, 0.5mM dATP, and 0.1 units/ul Klenow) was added to each tube and the reactions were continued for a further 15 minutes. 4ul of loading buffer (10mg/ml bromophenol blue, 10mg/ml xylene cyanol in deionized formamide) was

added to each tube. Samples were denatured at 100 °C for 10 minutes and 2ul was loaded on a sequencing gel (Section 2.6.3). Samples were loaded for a second time 60 minutes after the bromophenol blue had run out of the bottom of the gel. After electrophoresis, the gel was fixed in [20% (v/v) methanol, 10% (v/v) acetic acid] for 30 minutes, dried, and exposed to an X-Ray film overnight. Compressions in the banding pattern were resolved either by substituting dITP for dGTP in the reactions, or carrying out the sequencing reactions at 50 °C. The quantities of dNTPs used in each reaction are shown in Table 2.6.1. The dideoxynucleotides ddATP, ddTTP, ddGTP and ddCTP were aliquoted from stock solutions of 0.3mM, 0.5mM, 0.15mM amd 0.05mM respectively.

	A°	T°	G°	C°
0.5mM dATP	4.3ul	43ul	43ul	29ul
0.5mM dTTP	43ul	4.3ul	43ul	29ul
0.5mM dGTP	43ul	43ul	4.3ul	29ul
0.5mM dCTP	1ul	1ul	1ul	1ul
1 X TE	8.7ul	8.7ul	8.7ul	12ul
Total	100ul	100ul	100ul	100ul

Table 2.6.1 dNTP solutions for DNA sequencing.

Computer analysis of sequence data was carried out using Staden and NIH programs. Database comparisons were to the GenBank nucleic acid sequence database and NBRF protein sequence database.

2.6.7: Synthesis and cloning of cDNA.

cDNA was synthesized from ethyl methyl ketone induced poly(A)$^+$ RNA by the method of Gubler and Hoffman, (1983). After methylation of EcoRI recognition sites within the cDNA, EcoRI linkers were ligated to the cDNA, digested with EcoRI, and ligated into EcoRI digested, dephosphorylated λ gt10 arms (purchased from Promega). The ligation was packaged using a Promega Packagene packaging mix. *E. coli* C600*Hfl* cells were infected with the packaging mix to select recombinants.

2.6.8: S1 nuclease protection analysis

Preparation of M13 probes, annealing of probes to RNA and S1 nuclease digestion was performed by the methods of Burke, (1984). Fragments protected from S1 nuclease digestion were resolved on either sequencing or agarose gels depending on the resolution of bands required.

2.6.9: RNA sequencing and primer extension analysis:

An oligodeoxyribonucleotide of sequence TGGGCTTCTCGTTGCTGG was purchased from the Department of Microbiology and Immunology, University of Adelaide. This was used for both primer extension analysis and RNA sequencing, which were performed by procedure 1 of Geliebter *et al.*, (1986). The oligodeoyribonucleotide was end labelled with γ-^{32}P-dATP and polynucleotide kinase and annealed to the complementary RNA sequence at 53 °C (determined from: annealing temp = 4(G+C)+2(A+T)-5). Primer extension analysis was performed by the same protocol used for sequencing from a RNA template except the ddNTPs were left out of the transcription buffer. The products of the RNA

sequencing reactions were resolved on a 8% (w/v) polyacrylamide, 8M urea, 1 X TBE gel. The products of the primer extension reactions were electrophoresed on a similar gel, with the products of a DNA sequencing reaction being used as a marker of molecular weight.

2.6.10: Transformation of *Aspergillus*.

The production and isolation of protoplasts, and their transformation with plasmid DNA was performed by the methods of Tilburn *et al.*, (1983). Selection of AmdS$^+$ transformants was as Kelly and Hynes, (1985). The selection of PyrG$^+$ transformants of *A. niger* was as described by van Hartingsveldt *et al.*, (1987).

2.6.11: Growth testing and genetic manipulations of *Aspergillus*.

Growth testing and meiotic crosses of *A. nidulans* were performed as in Cove, (1966). Growth testing of *A. niger* was essentially the same, unless otherwise indicated, as that for *A. nidulans*. The formation and haploidization of diploids of *A. niger* was performed by the methods of Bos *et al.*, (1988).

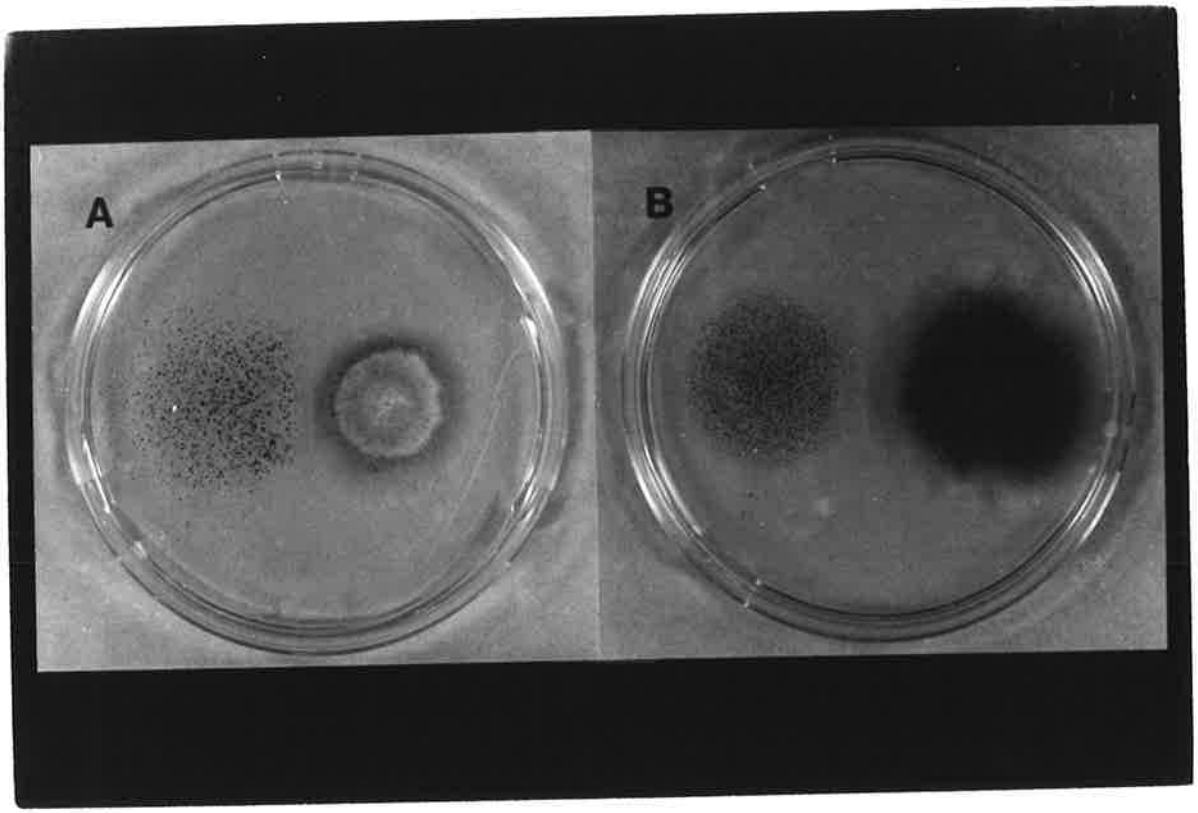
Chapter 3:
Cloning and physical characterization of the *Aspergillus niger aldA* gene.

This chapter describes the cloning of the aldehyde dehydrogenase encoding (*aldA*) gene of *Aspergillus niger* and the physical characterization of this clone. These experiments were performed with the aim of allowing a molecular comparison of *aldA* gene structure between *A. niger* and *A. nidulans* as part of the overall aim of this study, which is to elucidate mechanisms of gene regulation in *A. niger* by carrying out a comparative study of gene structure and regulation between these two species. The physical analysis has involved the determination of the complete nucleotide sequence of *aldA* and the structure of its transcript. Thus, the derived amino acid sequence of the aldehyde dehydrogenase enzyme has been obtained. This amino acid sequence has been compared to several other aldehyde dehydrogenases in an attempt to define functional regions of the enzyme.

3.1: Utilization of ethanol and threonine by *A. niger*.

A. niger shows much weaker growth on ethanol and threonine media than *A. nidulans*. Growth of *A. niger* on these media after 5 days incubation at 37 °C is shown in Figure 3.1.1, and although the mycelial growth was weak, it was fully conidiated. This proved to be a general phenotype of *A. niger* when grown on weak carbon sources which is in contrast to *A. nidulans* which frequently fails to conidiate

Figure 3.1.1: Growth of wildtype *A. niger* (left colony) and *A. nidulans* (right colony) on: A, 1% ethanol, 10mM NH_4Cl; B: 50mM threonine, 10mM NH_4Cl. Plates were incubated at 37 °C for 5 days.


A
B

during growth on poor carbon sources. As discussed in Chapter 5, *A. niger* fails to grow on media lacking added carbon sources, and thus the observed growth provided evidence for aldehyde dehydrogenase activity in *A. niger* as both these compounds are utilized via acetaldehyde.

3.2: Isolation of the *A. niger aldA* gene.

A genomic library of partially MboI digested *A. niger* DNA was constructed in the lambda based EMBL3 vector by S.J.B. Cooper (Dept. of Genetics, Univ. of Adelaide). This library was screened by plaque lifting and hybridization using the cloned *aldA* gene of *A. nidulans* within the insert of pAN212 (Pickett *et al.*, 1987) as a probe. A single hybridizing plaque was identified and isolated. The bacteriophage within this plaque were eluted, plated at low density, and rescreened by plaque lifting and hybridization. Greater than 90% of the plaques hybridized to the *aldA* specific probe (Figure 3.2.1). DNA was isolated from several of these plaques and digested with SalI, to remove the insert from the EMBL3 vector arms. The SalI restriction pattern for all these isolates was identical. Southern analysis showed a single SalI fragment of approximately 6.8 kb. that hybridized to the insert of pAN212 (Figure 3.2.2). This fragment was isolated and ligated into the SalI site of pUC19 to make the plasmid pNG100. It was confirmed that the insert of pNG100 contained a functional aldehyde dehydrogenase gene by its ability to complement the *aldA67* mutation of *A. nidulans* via transformation (Chapter 5). The aldehyde dehydrogenase encoding gene of *A. niger* was designated *aldA*. The insert of pNG100 hybridized to a single mRNA of approximately 1.8 kb. on northern blots (Chapter 4).

Figure 3.2.1: autoradiographs of plaque lifts from the *A. niger* genomic library probed with the insert of pAN212 (Pickett, *et al.*, 1987) under the conditions described in Section 2.6.4.

a): duplicate filters from a primary screen of the library. The solid arrows indicate the position of the hybridizing plaque. The open triangles indicate the position of pin holes used to orientate the filter to the plate.

b): duplicate filters from a plaque lift of bacteriophage eluted from the hybridizing plaque shown above in Figure 3.2.1a. The open triangles indicate the position of pin holes used to orientate the filter to the plate.

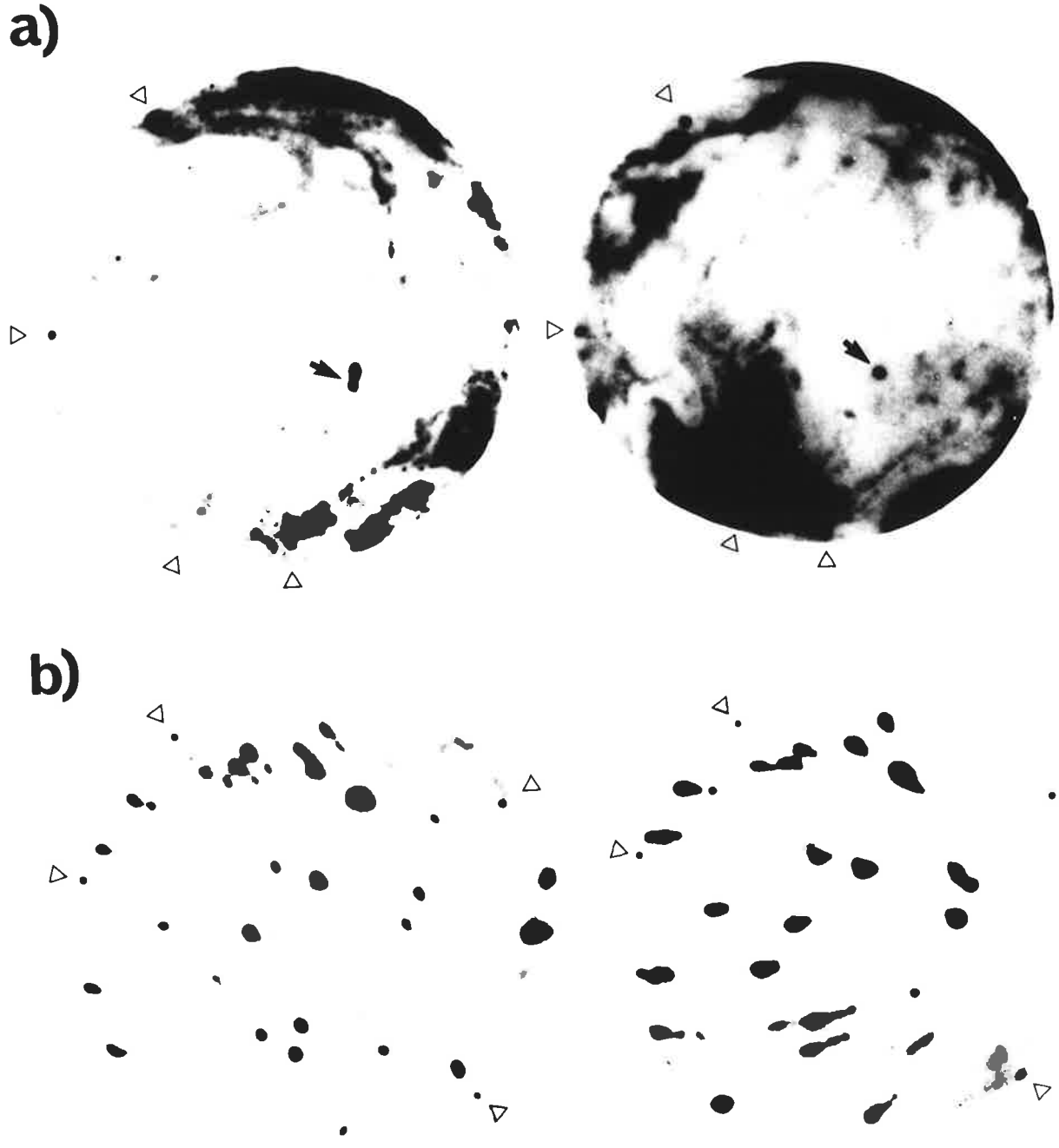
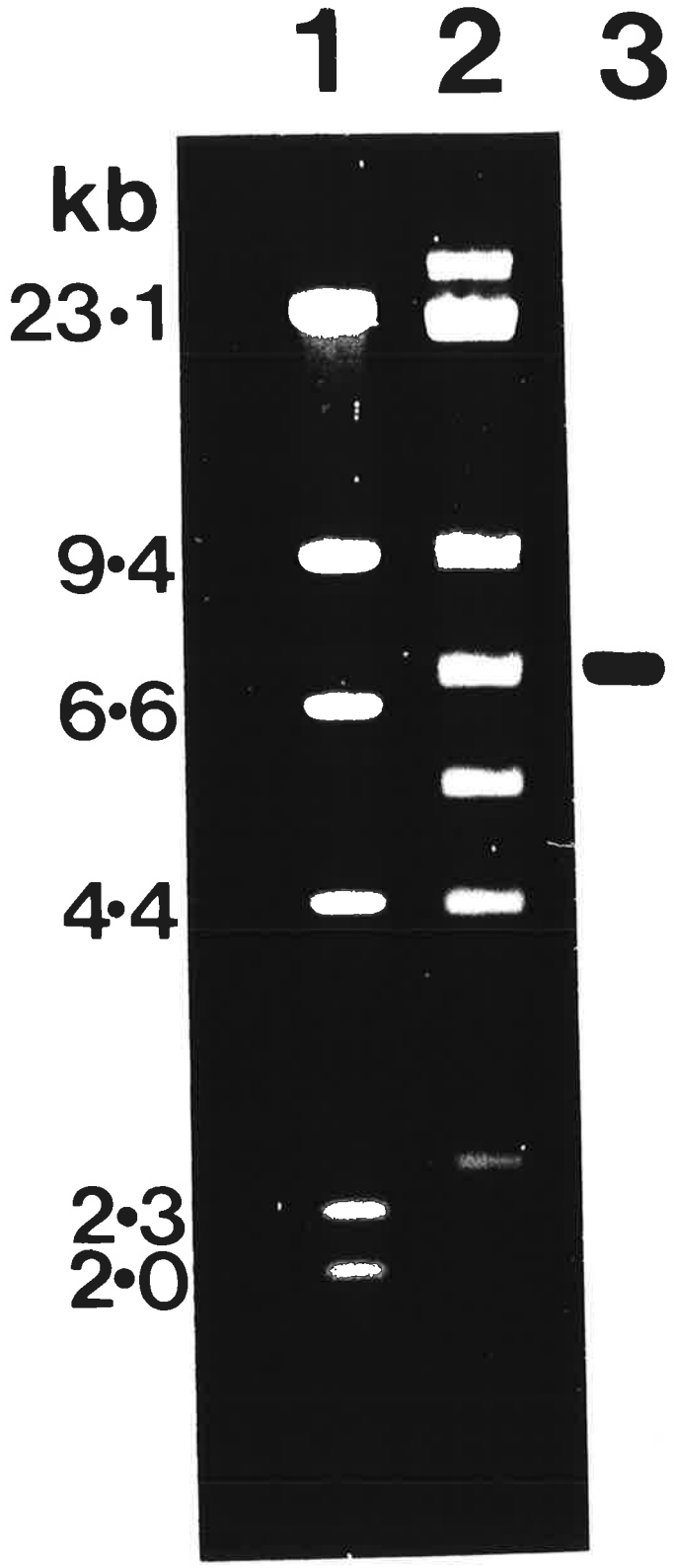

a)
b)

Figure 3.2.2: Southern blot analysis of DNA extracted from bacteriophage of a hybridizing plaque identified from the filter shown in Figure 3.2.1. Tracks are:

1: Wildtype bacteriophage lambda DNA digested with HindIII (marker of molecular weight).

2: DNA extracted from bacteriophage of a plaque that hybridizes to the insert of pAN212 digested with SalI.

3: Autoradiograph of a Southern blot of track 2 hybridized with the insert of pAN212 under the conditions indicated in Section 2.6.4.


1
2
3
kb
23·1
9·4
6·6
4·4
2·3
2·0

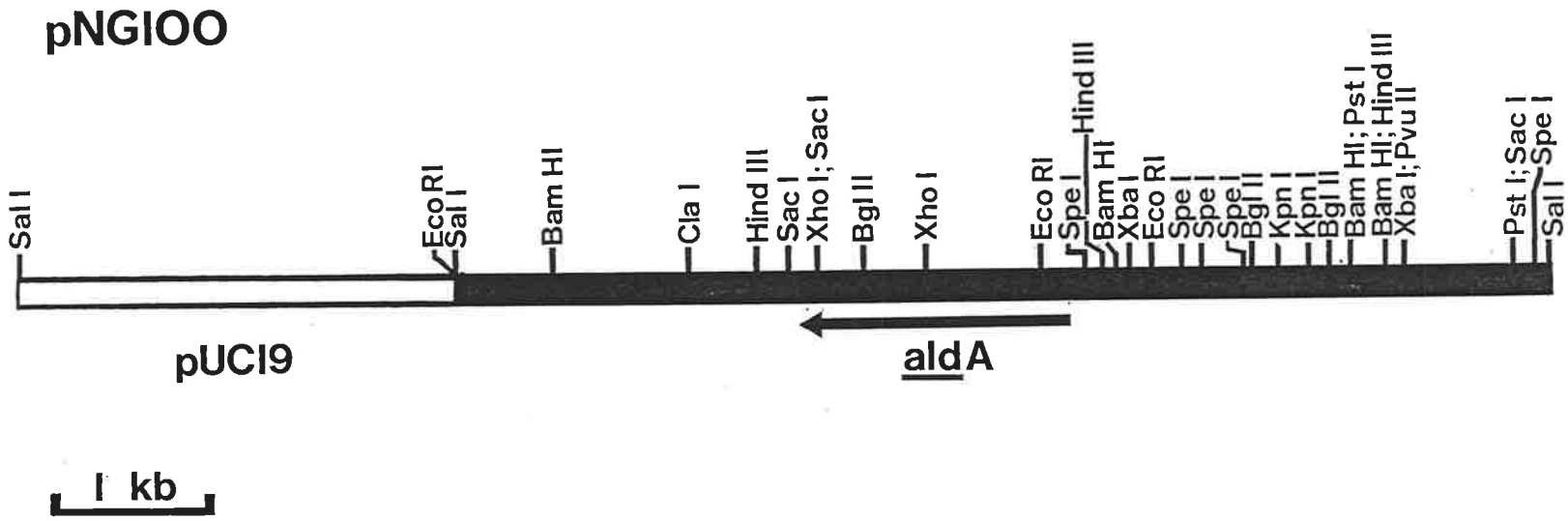
3.3: Positioning of the *ald*A gene within the pNG100 insert.

A preliminary map of restriction endonuclease sites within the pNG100 insert was constructed by conventional methods using combinations of double restriction endonuclease digests of the entire insert and of subclones. This map was confirmed and refined later through DNA sequencing (Figure 3.3.1). Several restriction fragments were used as probes for dot blots of total ethyl methyl ketone induced RNA which contained induced levels of the *ald*A transcript (Chapter 4). This allowed the positioning of the *ald*A coding region within the ClaI and XbaI restriction endonuclease sites (Table 3.3.1). The position and direction of transcription was determined more accurately by comparison of the nucleotide sequence of subclones of this region to the nucleotide sequence of the *A. nidulans ald*A gene, and is shown relative to the map of restriction endonuclease sites in Figure 3.3.1.

Probe	Hybridization to *ald*A mRNA
1.5 kb. SalI -> ClaI	NO
2.2 kb. SalI -> XhoI	YES
3.7 kb. SalI -> EcoRI	YES
4.2 kb. SalI -> XbaI	YES
5.0 kb. SalI -> PstI	YES
2.0 kb. XbaI -> XbaI	NO
1.3 kb. BglII -> SalI	NO
1.5 kb. PstI -> PstI	NO

Table 3.3.1: Positioning of the *ald*A coding region within pNG100. Restriction sites are shown in the left to right order of the map in Figure 3.3.1.

Figure 3.3.1: Map of restriction endonuclease sites within the insert of pNG100. The orientation of the insert in the vector pUC19 is shown relative to the EcoRI site in the pUC19 polylinker. None of the restriction endonucleases tested failed to cleave the plasmid. The position of the *ald*A coding region and the direction of transcription are shown by the arrow below the map. Some additional restriction endonuclease sites surrounding the *ald*A coding region are shown in Figure 3.4.1.


pNG100
Sal I
EcoRI
Sal I
Bam HI
Cla I
Hind III
Sac I
Xho I; Sac I
Bgl II
Xho I
Eco RI
Spe I
Hind III
Bam HI
Xba I
Eco RI
Spe I
Spe I
Spe I
Bgl II
Kpn I
Kpn I
Bgl II
Bam HI; Pst I
Bam HI; Hind III
Xba I; Pvu II
Pst I; Sac I
Spe I
Sal I
pUC19
aldA
1 kb

3.4: Physical characterization of the *A. niger aldA* gene.

3.4.1: The *aldA* coding region.

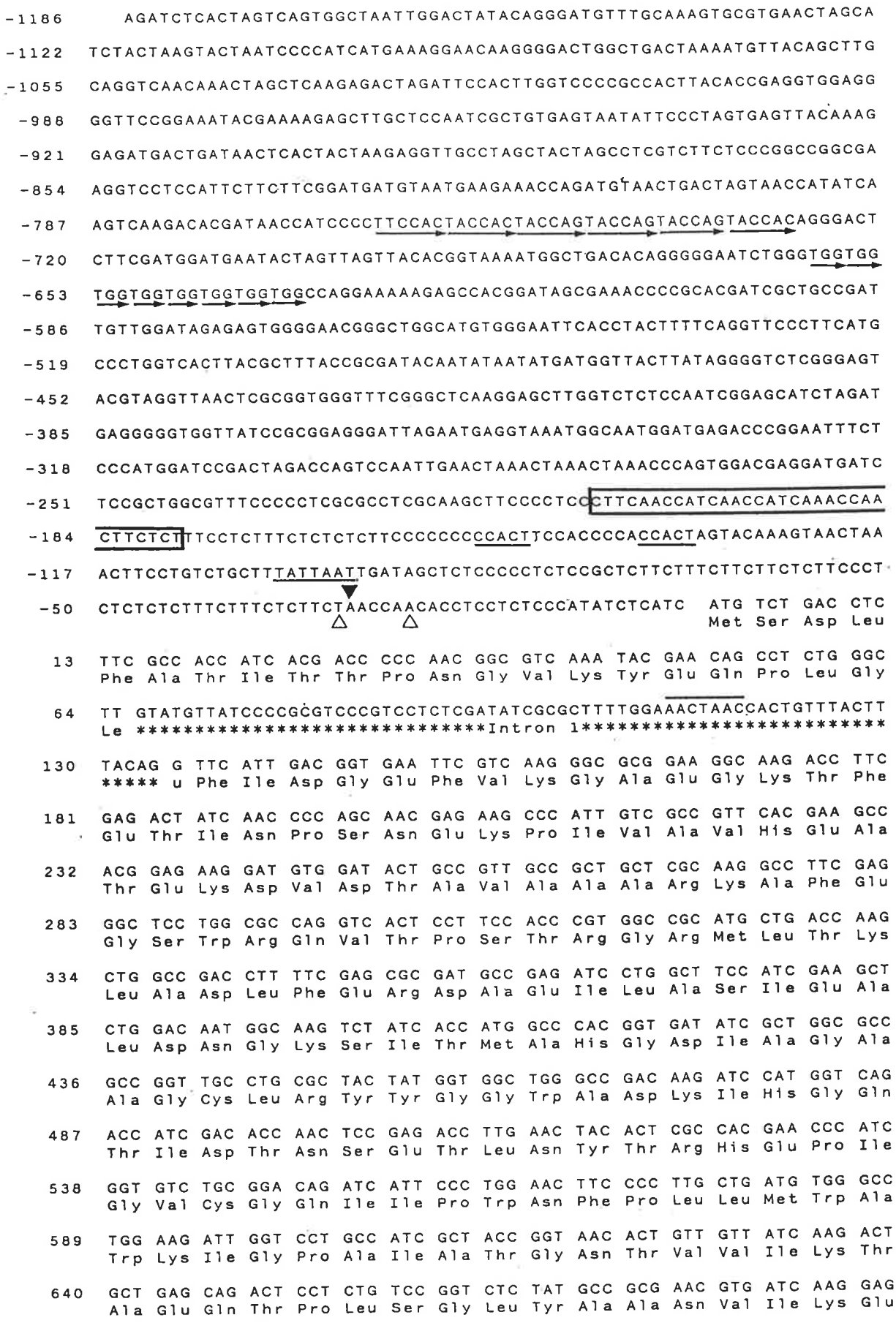
The nucleotide sequence of a 3.5 kb. region spanning 1.2 kb. 5' to 0.7 kb. 3' of the *aldA* coding region was determined from subclones of this region using the strategy shown in Figure 3.4.1. The nucleotide sequence and the amino acid sequence of the deduced 53.8 kDa polypeptide encoded by *aldA* is shown in Figure 3.4.2. This coding region was interrupted by three introns (see below). Excluding these intron sequences, a comparison of the nucleotide sequence of the coding region of the *A. niger aldA* gene to that of *A. nidulans* showed 80% identity of nucleotides, with nucleotide substitutions being scattered throughout the sequence (Figure 3.4.3). This level of nucleotide sequence similarity is slightly higher than that seen for the *argB* and *pyrG* genes of *A. niger* and *A. nidulans* which is approximately 70% (Buxton *et al.*, 1987; Wilson *et al.*, 1988).

Like the *A. nidulans aldA* gene and other highly expressed fungal genes (reviewed by Ballance, 1986 and Gurr *et al.*, 1987) there is a bias against A and for C in the third position of codons, and infrequent use of the AGN codons of arginine and serine. Codon usage within the *A. niger* and *A. nidulans aldA* genes is shown in Table 3.4.1. There are some notable differences in codon usage between these two genes, namely the CTT, GCT, GCC and GAA codons. The relatively frequent use of the GAA codon in the *A. niger* gene (8/51) compared to *A. nidulans* (1/50) and other fungal genes is unusual, as codons with an A in the 3rd position are infrequently used in fungal genes. The few genes from *A. niger* that have been sequenced are not representative of more highly expressed fungal genes and genes of higher eukaryotes both

Figure 3.4.1: Strategy employed for the sequencing of the *A. niger* *aldA* gene and its flanking regions. Extent and direction of sequence gained from subclones in M13mp18 and M13mp19 of the pNG100 insert is indicated by the arrows below the restriction endonuclease map of the 3.5kb. BglII-ClaI restriction fragment of pNG100. Sequence data gained from cDNA clones (after subcloning into the EcoRI site of M13mp19) or directly from the *aldA* mRNA are also indicated. Dotted lines within cDNA and mRNA arrows indicate the position of intron sequences that were absent from these templates. The positions of the start (ATG) and end (TAA) of translation are also indicated.

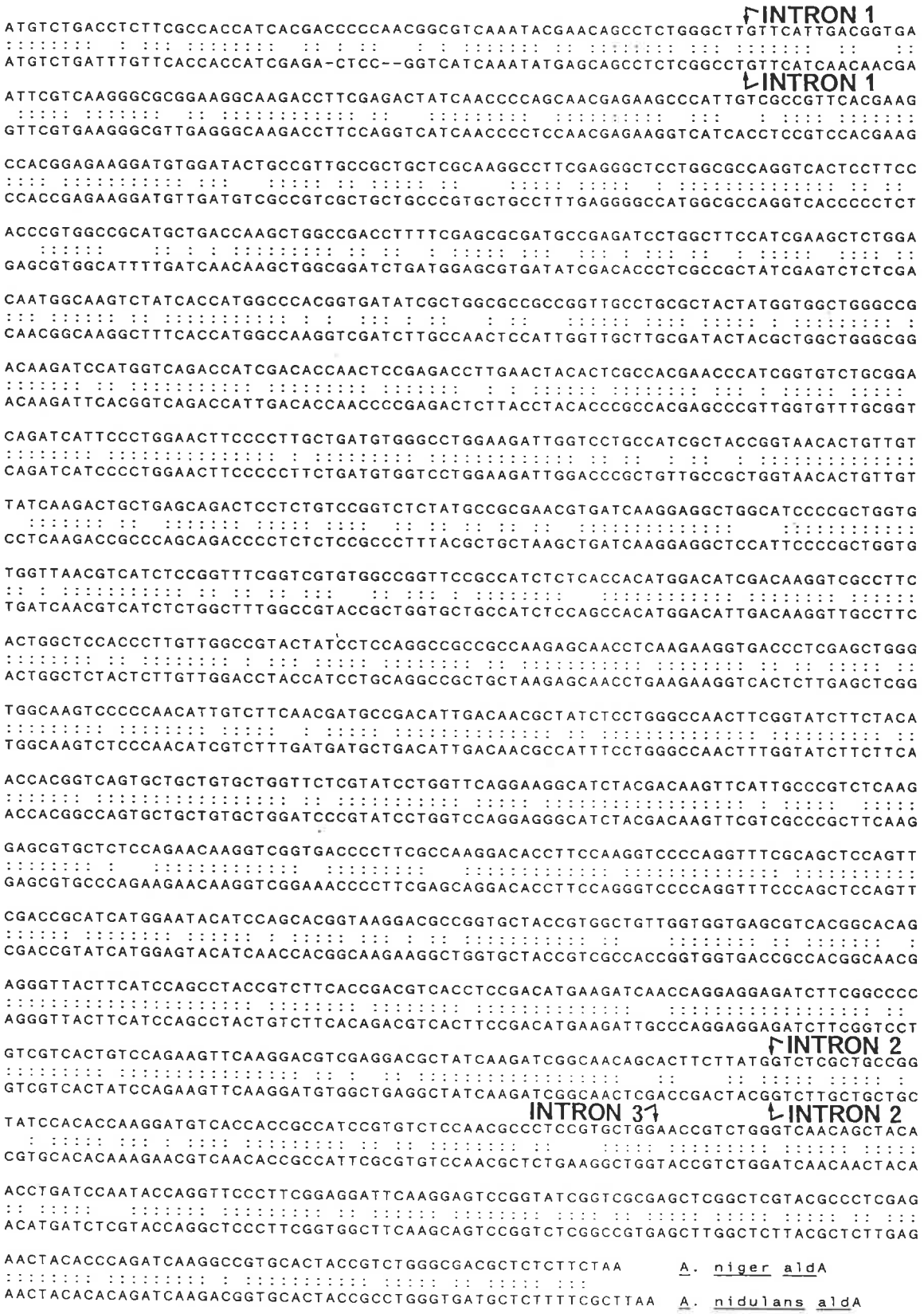
Bgl II
Spe I
Spe I
Spe I
Eco RI
Xba I
Sty I
Bam HI
Hind III
Spe I
Eco RI
Sph I
Sty I
Sty I
Xho I
Sty I
Sty I
Bgl II
Sty I
Sty I
Sac I
Xho I
Sac I
Acc I
Hind III
Sma I
Cla I
ATG
TAA
100 bp
cDNA
mRNA
cDNA

Figure 3.4.2: Nucleotide sequence of the *A. niger aldA* gene and its flanking regions. Numbering of nucleotides begins with the first nucleotide of the start codon as +1 (0 has not been used). Putative TATA and CCAAT box sequences are underlined. Repeated sequences in the 5' region are indicated by arrows. The boxed sequence is discussed in Section 6.1. The positions of the major (solid triangle) and minor (open triangles) start points of transcription are indicated. Sites of transcript polyadenylation are indicated by squares. The derived amino sequence is indicated below the corresponding codons. Intron sequences are indicated by asterisks (*), with lines over putative internal lariat sequences. The sequence is continued overleaf.


-1186 AGATCTCACTAGTCAGTGGCTAATTGGACTATACAGGGATGTTTGCAAAGTGCGTGAACTAGCA
-1122 TCTACTAAGTACTAATCCCCATCATGAAAGGAACAAGGGGACTGGCTGACTAAAATGTTACAGCTTG
-1055 CAGGTCAACAAACTAGCTCAAGAGACTAGATTCCACTTGGTCCCCGCCACTTACACCGAGGTGGAGG
-988 GGTTCCGGAAATACGAAAAGAGCTTGCTCCAATCGCTGTGAGTAATATTCCCTAGTGAGTTACAAAG
-921 GAGATGACTGATAACTCACTACTAAGAGGTTGCCTAGCTACTAGCCTCGTCTTCTCCCGGCCGGCGA
-854 AGGTCCTCCATTCTTCTTCGGATGATGTAATGAAGAAACCAGATGTAACTGACTAGTAACCATATCA
-787 AGTCAAGACACGATAACCATCCCCTTCCACTACCACTACCAGTACCAGTACCAGTACCACAGGGACT
-720 CTTCGATGGATGAATACTAGTTAGTTACACGGTAAAATGGCTGACACAGGGGGAATCTGGGTGGTGG
-653 TGGTGGTGGTGGTGGTGGCCAGGAAAAAGAGCCACGGATAGCGAAACCCCGCACGATCGCTGCCGAT
-586 TGTTGGATAGAGAGTGGGGAACGGGCTGGCATGTGGGAATTCACCTACTTTTCAGGTTCCCTTCATG
-519 CCCTGGTCACTTACGCTTTACCGCGATACAATATAATATGATGGTTACTTATAGGGGTCTCGGGAGT
-452 ACGTAGGTTAACTCGCGGTGGGTTTCGGGCTCAAGGAGCTTGGTCTCTCCAATCGGAGCATCTAGAT
-385 GAGGGGGTGGTTATCCGCGGAGGGATTAGAATGAGGTAAATGGCAATGGATGAGACCCGGAATTTCT
-318 CCCATGGATCCGACTAGACCAGTCCAATTGAACTAAACTAAACTAAACCCAGTGGACGAGGATGATC
-251 TCCGCTGGCGTTTCCCCCTCGCGCCTCGCAAGCTTCCCCTCCCTTCAACCATCAACCATCAAACCAA
-184 CTTCTCTTTCCTCTTTCTCTCTCTTCCCCCCCCCACTTCCACCCCACCACTAGTACAAAGTAACTAA
-117 ACTTCCTGTCTGCTTTATTAATTGATAGCTCTCCCCCTCTCCGCTCTTCTTTCTTCTTCTCTTCCCT
-50 CTCTCTCTTTCTTTCTCTTCTAACCAACACCTCCTCTCCCATATCTCATC ATG TCT GAC CTC
Met Ser Asp Leu
13 TTC GCC ACC ATC ACG ACC CCC AAC GGC GTC AAA TAC GAA CAG CCT CTG GGC
Phe Ala Thr Ile Thr Thr Pro Asn Gly Val Lys Tyr Glu Gln Pro Leu Gly
64 TT GTATGTTATCCCCGCGTCCCGTCCTCTCGATATCGCGCTTTTGGAAACTAACCACTGTTTACTT
Le **********Intron 1**********
130 TACAG G TTC ATT GAC GGT GAA TTC GTC AAG GGC GCG GAA GGC AAG ACC TTC
***** u Phe Ile Asp Gly Glu Phe Val Lys Gly Ala Glu Gly Lys Thr Phe
181 GAG ACT ATC AAC CCC AGC AAC GAG AAG CCC ATT GTC GCC GTT CAC GAA GCC
Glu Thr Ile Asn Pro Ser Asn Glu Lys Pro Ile Val Ala Val His Glu Ala
232 ACG GAG AAG GAT GTG GAT ACT GCC GTT GCC GCT GCT CGC AAG GCC TTC GAG
Thr Glu Lys Asp Val Asp Thr Ala Val Ala Ala Ala Arg Lys Ala Phe Glu
283 GGC TCC TGG CGC CAG GTC ACT CCT TCC ACC CGT GGC CGC ATG CTG ACC AAG
Gly Ser Trp Arg Gln Val Thr Pro Ser Thr Arg Gly Arg Met Leu Thr Lys
334 CTG GCC GAC CTT TTC GAG CGC GAT GCC GAG ATC CTG GCT TCC ATC GAA GCT
Leu Ala Asp Leu Phe Glu Arg Asp Ala Glu Ile Leu Ala Ser Ile Glu Ala
385 CTG GAC AAT GGC AAG TCT ATC ACC ATG GCC CAC GGT GAT ATC GCT GGC GCC
Leu Asp Asn Gly Lys Ser Ile Thr Met Ala His Gly Asp Ile Ala Gly Ala
436 GCC GGT TGC CTG CGC TAC TAT GGT GGC TGG GCC GAC AAG ATC CAT GGT CAG
Ala Gly Cys Leu Arg Tyr Tyr Gly Gly Trp Ala Asp Lys Ile His Gly Gln
487 ACC ATC GAC ACC AAC TCC GAG ACC TTG AAC TAC ACT CGC CAC GAA CCC ATC
Thr Ile Asp Thr Asn Ser Glu Thr Leu Asn Tyr Thr Arg His Glu Pro Ile
538 GGT GTC TGC GGA CAG ATC ATT CCC TGG AAC TTC CCC TTG CTG ATG TGG GCC
Gly Val Cys Gly Gln Ile Ile Pro Trp Asn Phe Pro Leu Leu Met Trp Ala
589 TGG AAG ATT GGT CCT GCC ATC GCT ACC GGT AAC ACT GTT GTT ATC AAG ACT
Trp Lys Ile Gly Pro Ala Ile Ala Thr Gly Asn Thr Val Val Ile Lys Thr
640 GCT GAG CAG ACT CCT CTG TCC GGT CTC TAT GCC GCG AAC GTG ATC AAG GAG
Ala Glu Gln Thr Pro Leu Ser Gly Leu Tyr Ala Ala Asn Val Ile Lys Glu

```
691  GCT GGC ATC CCC GCT GGT GTG GTT AAC GTC ATC TCC GGT TTC GGT CGT GTG
     Ala Gly Ile Pro Ala Gly Val Val Asn Val Ile Ser Gly Phe Gly Arg Val

742  GCC GGT TCC GCC ATC TCT CAC CAC ATG GAC ATC GAC AAG GTC GCC TTC ACT
     Ala Gly Ser Ala Ile Ser His His Met Asp Ile Asp Lys Val Ala Phe Thr

793  GGC TCC ACC CTT GTT GGC CGT ACT ATC CTC CAG GCC GCC GCC AAG AGC AAC
     Gly Ser Thr Leu Val Gly Arg Thr Ile Leu Gln Ala Ala Ala Lys Ser Asn

844  CTC AAG AAG GTG ACC CTC GAG CTG GGT GGC AAG TCC CCC AAC ATT GTC TTC
     Leu Lys Lys Val Thr Leu Glu Leu Gly Gly Lys Ser Pro Asn Ile Val Phe

895  AAC GAT GCC GAC ATT GAC AAC GCT ATC TCC TGG GCC AAC TTC GGT ATC TTC
     Asn Asp Ala Asp Ile Asp Asn Ala Ile Ser Trp Ala Asn Phe Gly Ile Phe

946  TAC AAC CAC GGT CAG TGC TGC TGT GCT GGT TCT CGT ATC CTG GTT CAG GAA
     Tyr Asn His Gly Gln Cys Cys Cys Ala Gly Ser Arg Ile Leu Val Gln Glu

997  GGC ATC TAC GAC AAG TTC ATT GCC CGT CTC AAG GAG CGT GCT CTC CAG AAC
     Gly Ile Tyr Asp Lys Phe Ile Ala Arg Leu Lys Glu Arg Ala Leu Gln Asn

1048 AAG GTC GGT GAC CCC TTC GCC AAG GAC ACC TTC CAA GGT CCC CAG GTT TCG
     Lys Val Gly Asp Pro Phe Ala Lys Asp Thr Phe Gln Gly Pro Gln Val Ser

1099 CAG CTC CAG TTC GAC CGC ATC ATG GAA TAC ATC CAG CAC GGT AAG GAC GCC
     Gln Leu Gln Phe Asp Arg Ile Met Glu Tyr Ile Gln His Gly Lys Asp Ala

1150 GGT GCT ACC GTG GCT GTT GGT GGT GAG CGT CAC GGC ACA GAG GGT TAC TTC
     Gly Ala Thr Val Ala Val Gly Gly Glu Arg His Gly Thr Glu Gly Tyr Phe

1201 ATC CAG CCT ACC GTC TTC ACC GAC GTC ACC TCC GAC ATG AAG ATC AAC CAG
     Ile Gln Pro Thr Val Phe Thr Asp Val Thr Ser Asp Met Lys Ile Asn Gln

1252 GAG GAG ATC TTC GGC CCC GTC GTC ACT GTC CAG AAG TTC AAG GAC GTC GAG
     Glu Glu Ile Phe Gly Pro Val Val Thr Val Gln Lys Phe Lys Asp Val Glu

1303 GAC GCT ATC AAG ATC GGC AAC AGC ACT TCT TAT G GTGAGTTCTTTTTCGCCACTA
     Asp Ala Ile Lys Ile Gly Asn Ser Thr Ser Tyr G **********************

1358 AATTGCAGTGTGTCAAGTGCTAATCTTTCGTCCTAG GT CTC GCT GCC GGT ATC CAC ACC
     ***Intron 2************************** ly Leu Ala Ala Gly Ile His Thr

1417 AAG GAT GTC ACC ACC GCC ATC CGT GTC TCC AAC GCC CTC CGT GCT GG GTAA
     Lys Asp Val Thr Thr Ala Ile Arg Val Ser Asn Ala Leu Arg Ala Gl ****

1468 GCTATACTTAACACTACATTCCTGTATTCATTCACTAACATGCACATAG A ACC GTC TGG GTC
     ********************Intron 3********************** y Thr Val Trp Val

1530 AAC AGC TAC AAC CTG ATC CAA TAC CAG GTT CCC TTC GGA GGA TTC AAG GAG
     Asn Ser Tyr Asn Leu Ile Gln Tyr Gln Val Pro Phe Gly Gly Phe Lys Glu

1581 TCC GGT ATC GGT CGC GAG CTC GGC TCG TAC GCC CTC GAG AAC TAC ACC CAG
     Ser Gly Ile Gly Arg Glu Leu Gly Ser Tyr Ala Leu Glu Asn Tyr Thr Gln

1632 ATC AAG GCC GTG CAC TAC CGT CTG GGC GAC GCT CTC TTC TAA TTTAATATTTG
     Ile Lys Ala Val His Tyr Arg Leu Gly Asp Ala Leu Phe END

1685 AAACCATGAGAAAATGGGAAGAGTTCGCTTCAGCGACGAGCTCGTGGATATTGGGTTACCTTTTTCG

1752 CAGATAGCCATGACTGGTTCATGCAAAACAATCAGGATACGAATTTACGATGATTACAAATTGTTCT

1819 TACATTGCAGTAACTGTAGACCTTTGTAAATATAATCTGTCTCAAGCTTCAACCTCCACATAATCTT

1886 GGTACTGGCTATAATCACTCCCTAGTGGTTAACTAAGGACGAATCACCACCCCTTCCACTCGTTCAA

1953 TTGCAATCGACCTCCATCAGGCCATTGCATATCGTACCCATTGTTACACACCCAAATTCCCGGGACA

2020 ACCACCGGCGAGTTTGAGTAGCCTGGGTCTAGTCTCGGGAGAATTACTACGCCGAGAAGAGAGACCC

2087 CCTCGGGTCGAAGATGGCCCGAAGCGAGGGCATTCATCAGCGAATACACAAGGACCGAATGATTGGT

2154 TATGTCACCGACGTTGATGATGATCAGCCATTCAAAAAGACAGCTAATTTCAATCTTCAGTACTTCA

2221 TAGGTTGATTGATTACTTGCTGGTGCCGAAAATAGGTATTCGGGAACTATAATCTCCTATCTTGTCA

2288 GTTCTGGTACTAGAATGCATCGAT
```

Figure 3.4.3: alignment of the nucleotide sequences of the amino acid coding regions of the *A. niger* (upper) and *A. nidulans* (lower) *aldA* genes. Colons (:) indicate identity between nucleotides of the two sequences. Hyphens (-) indicate spaces placed in the sequence of the *A. nidulans* gene to gain the best alignment. The positions of introns within these sequences are indicated.


INTRON 1
INTRON 1
INTRON 2
INTRON 37
INTRON 2
A. niger aldA
A. nidulans aldA

in terms of promoter structure and the presence of intron sequences, and so it is not possible to gauge the significance of these differences in codon usage between the *ald*A genes of *A. niger* and *A. nidulans* as to whether thay represent general differences in codon usage in these two species.

	NIG.	NID.		NIG.	NID.		NIG.	NID.		NIG.	NID.
F TTT	0.	4.	S TCT	5.	7.	Y TAT	3.	1.	C TGT	1.	1.
F TTC	23.	22.	S TCC	13.	12.	Y TAC	12.	13.	C TGC	4.	4.
L TTA	0.	0.	S TCA	0.	0.	* TAA	1.	1.	* TGA	0.	0.
L TTG	3.	3.	S TCG	2.	2.	* TAG	0.	0.	W TGG	7.	7.
L CTT	2.	10.	P CCT	5.	5.	H CAT	1.	0.	R CGT	10.	8.
L CTC	13.	8.	P CCC	12.	12.	H CAC	10.	9.	R CGC	8.	6.
L CTA	0.	0.	P CCA	0.	2.	Q CAA	2.	0.	R CGA	0.	1.
L CTG	12.	10.	P CCG	0.	1.	Q CAG	18.	22.	R CGG	0.	0.
I ATT	7.	10.	T ACT	11.	9.	N AAT	1.	0.	S AGT	0.	0.
I ATC	34.	29.	T ACC	22.	22.	N AAC	22.	28.	S AGC	4.	2.
I ATA	0.	0.	T ACA	1.	3.	K AAA	1.	1.	R AGA	0.	0.
M ATG	7.	8.	T ACG	2.	1.	K AAG	28.	31.	R AGG	0.	0.
V GTT	10.	9.	A GCT	18.	30.	D GAT	6.	10.	G GGT	28.	20.
V GTC	19.	22.	A GCC	31.	21.	D GAC	20.	15.	G GGC	19.	18.
V GTA	0.	0.	A GCA	0.	0.	E GAA	8.	1.	G GGA	4.	4.
V GTG	7.	6.	A GCG	2.	2.	E GAG	19.	24.	G GGG	0.	1.

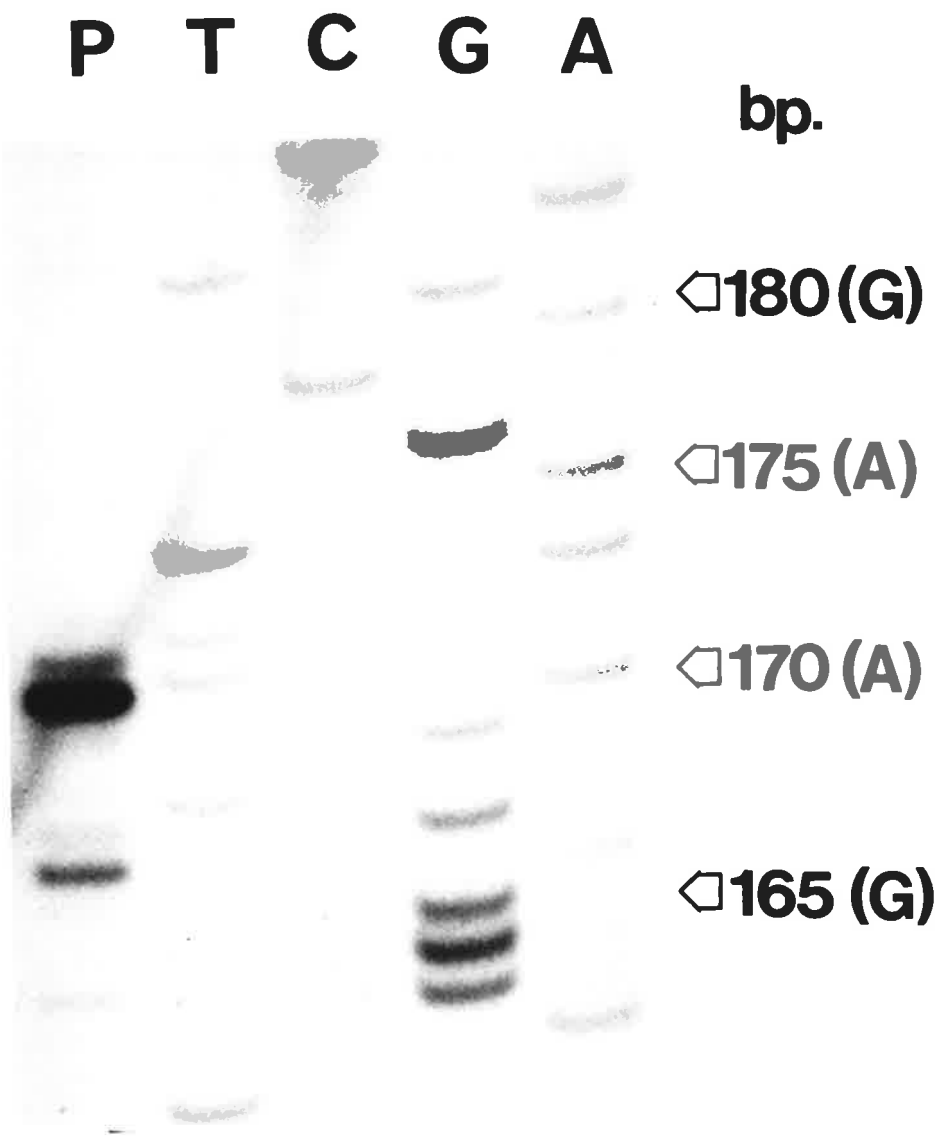
Table 3.4.1: Codon usage in the *ald*A genes of *A. niger* (NIG) and *A. nidulans* (NID). The standard single letter amino acid code is used. *: termination codon.

3.4.2: The 5' region.

The promoter region of the *A. niger ald*A gene includes a TATA like sequence at nucleotide position -102. Two sequences at nucleotide positions -138 and -152 (CCACT) show the closest homology to the CCAAT box consensus sequence. There is one major and two minor start points of transcription, which were determined by primer

extension analysis, and were positioned between nucleotides -30 and -24 (Figure 3.4.4). The sequences surrounding the start point of transcription (CTAACCAAC) are similar to those seen for *A. nidulans alc*A (CATCAACCAAC) and *ald*A (CTACTCCAAC) (Gwynne *et al.*, 1987). The sequence preceeding the start codon resembles that of other fungal genes, with the sequence TCA from nucleotide position -5 to -3 being the preferred bases (Ballance, 1986). There is an extremely pyrimidine rich region (59/60 nucleotides) positioned between the TATA box and the start points of transcription. This is seen in the promoters of other fungal genes and is particularly evident in highly expressed genes of *Saccharomyces cerevisiae* (Dobson *et al.*, 1982). Pyrimidine rich promoter sequences are also present in highly expressed *Aspergillus* genes, such as the *A. nidulans* glyceraldehyde-3-phosphate dehydrogenase encoding gene (Punt *et al.*, 1988), the ATP synthase subunit 9 encoding genes of both *A. nidulans* and *A. niger* (Ward and Turner, 1986; Ward *et al.*, 1988), and to a lesser extent in the 3-phosphoglycerate kinase encoding gene of *A. nidulans* (Clements and Roberts, 1986). The major start point of transcription of the *A. niger ald*A gene is located at the first purine after the pyrimidine rich region. This is also the case for other *Aspergillus* genes containing this element. Deletion of a small pyrimidine rich region of the *A. nidulans trp*C promoter results in transcription initiating from heterogenous positions (Hamer and Timberlake, 1987). Thus, this sequence may play a role in directing transcriptional start points rather than directing high levels of expression. Such a pyrimidine-rich sequence is not as evident in the promoter of the *A. nidulans ald*A gene, although 27 of the 38 nucleotides (71%) between the TATA box and the start point of transcription are pyrimidines (Gwynne *et al.*, 1987a).

Figure 3.4.4: Mapping of the 5' ends of the *A. niger aldA* mRNA transcript. A γ-^{32}P-labelled oligodeoxynuceotide primer of sequence TGGGCTTCTCGTTGCTGG, complementary to nucleotides 195-212 (Figure 3.4.2) was annealed to 2ug of ethanol induced poly(A)$^+$ RNA and extended with AMV reverse transcriptase by the method of Geliebter *et al.*, (1986) as described in Section 2.6.9. The products of the extension (track P) and of a dideoxynucleotide sequencing reaction used as a marker of molecuar weight (tracks T, C, G and A) were electrophoresed on a 6% polyacrylamide sequencing gel (Section 2.6.3). The clone sequenced for the marker of molecular weight was the 2.3kb. SmaI-BamHI fragment of pNG100 subcloned into M13mp19 and primed with the Bio-Labs -40 sequencing primer. The length of the major extended product was 171 nucleotides which positions the major start point of transcription as shown by the solid triangle on Figure 3.4.2. The lengths of the minor bands (166 and 172 nucleotides) positioned the minor start points of transcription as indicated by the open triangles on Figure 3.4.2.

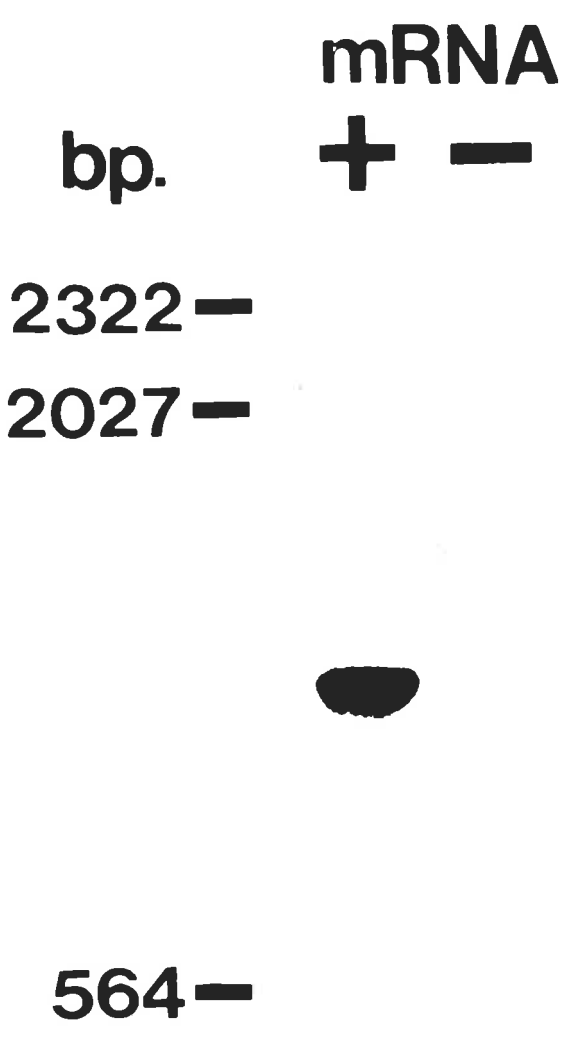

P
T
C
G
A
bp.
180 (G)
175 (A)
170 (A)
165 (G)

The 5' region also contains several repeated sequences, the most marked being the sequence, TGG, which is directly repeated eight times at nucleotide -659 to -636, and the sequence, TA/TCCAC/G, which is repeated six times at nucleotide -763 to -728. There is a also region of sequence similarity in the *A. niger aldA* promoter to several *A. nidulans* and two *N. crassa* genes, many of which are subject to carbon catabolite repression. These regions and a deletion analysis of the promoter region are discussed in Chapter 6.

3.4.3: Introns.

Three introns were found in the *A. niger aldA* gene, compared to the two which interrupt the *A. nidulans aldA* gene (Gwynne *et al.*, 1987a). The position of intron 1 was found by performing dideoxynucleotide sequencing on the mRNA template (Geliebter *et al.*, 1986), using a [γ -^{32}P]-oligodeoxyribonucleotide of sequence TGGGCTTCTCGTTGCTGG, which is complementary to the sequence of nucleotides 195 - 212. This region was 3' of the proposed intron position based on nucleotide and amino acid sequence comparisons to the *A. nidulans aldA* gene. This RNA sequencing analysis positioned intron 1 in the homologous postion with respect to the amino acid sequence as intron 1 of the *A. nidulans aldA* gene. Several partial *aldA* cDNA clones were isolated from a library containing cDNA synthesized from poly(A)$^+$ RNA extracted from ethyl methyl ketone induced mycelia constucted in the bacteriophage vector, λ gt10, by plaque lifting and hybridization using pNG100 as a probe. Comparison of the nucleotide sequence of these clones and that of the genomic clone was used to position introns 2 and 3. Intron 2 is in the same position relative to the amino acid sequence as intron 2 in the *A. nidulans aldA* gene. S1 nuclease protection analysis showed exon 2 to be free of intron

Figure 3.4.5: S1 nuclease protection analysis of exon 2 of the *A. niger aldA* gene. The 1.1kb. EcoRI-BglII fragment of pNG100, which spans almost all of exon 2, was internaly labelled with α $-^{32}$P-dCTP, denatured, and annealed to 5ug of total ethanol induced RNA and digested with S1 nuclease by the method of Burke, (1984). The products of this reaction, and a duplicate reaction performed without the addition of RNA, were electrophoresed on a 1.2% agarose, 1 X TAE gel. The nucleic acids in the gel were transferred to Zeta-Probe membrane by Southern blotting. An autoradiograph of this filter is shown opposite. In the presence but not absence of RNA, a single band corresponding to the full length probe is protected from S1 nuclease digestion, indicating that this region was completely annealed to *aldA* mRNA, and is hence free of intron sequences. Bacteriophage lambda DNA digested with HindIII was used as a marker of molecular weight.

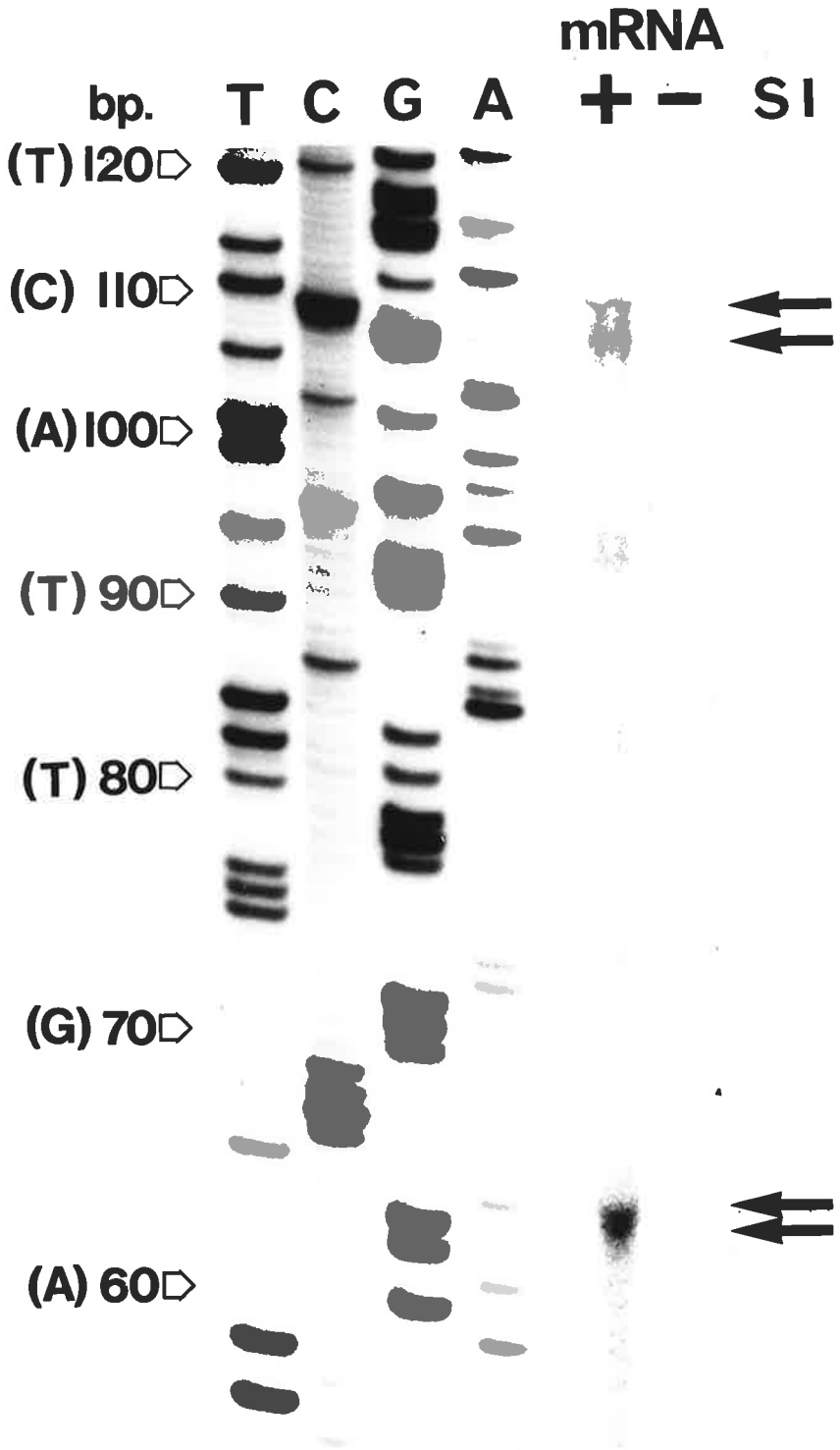

mRNA
bp.
+
−
2322
2027
564

sequences (Figure 3.4.5). In the presence of RNA, a band protected from S1 nuclease digestion corresponding to a full-length probe that spans exon 2 is seen, indicating the absence of introns in this region. All three introns have the GT - AG intron/exon boundaries and sequences resembling the consensus splicing signal for filamentous fungi (reviewed by Ballance, 1986 and Gurr *et al.*, 1987). The introns are similar in size to other fungal introns (Hawkins, 1988), and as expected, show little sequence similarity to the introns of the *A. nidulans aldA* gene. As the loss or addition of an intron to a gene is an all or nothing event, no evolutionary implications of the presence of an extra intron (intron 3) in the *A. niger aldA* gene compared to the *A. nidulans aldA* gene can be deduced.

3.4.4: The 3' Region.

Four major sites of polyadenylation were found by S1 nuclease protection analysis (Figure 3.4.6). They are located between 109 and 157 nucleotides 3' to the stop codon. The AAUAAA polyadenylation signal of higher eukaryotes (Proudfoot and Brownlee, 1976) was not present, but this sequence is not a general feature of mRNA molecules of filamentous fungi. A sequence surrounding the first two polyadenylation sites (CAAAACAATCAG, 103 bp. 3' of TAA codon) has similarity to regions of the 3' end of the *A. nidulans aldA* (TAAAACATATCAG, 104 bp. 3' of TAA codon) and *alcA* (CAAACACAATC, 27 bp. 3'of stop codon) genes, but as the polyadenylation sites of these *A. nidulans* genes have not been determined, it is difficult to say if these sequences could play a role in signalling polyadenylation.

Figure 3.4.6: Mapping of the 3' ends of the *A. niger aldA* mRNA transcript. The 296 bp. SacI-SmaI fragment of pNG100 was internally labelled with α -^{32}P-dCTP, annealed to 5ug of total ethanol induced RNA and digested with S1 nuclease by the method of Burke, (1984). The products of this reaction, and a duplicate reaction carried out in the absence of RNA were electrophored in an 8% polyacrylamide sequencing gel (Section 2.6.4), together with the products of a dideoxynucleotide sequencing reaction of the 296bp. SacI-SmaI fragment in M13mp18 primed with the Bio-Labs -40 sequencing primer (tracks T, C, G and A). In the presence of added RNA, four fragments of 62, 63, 108 and 112 nucleotides were protected from S1 nuclease digestion (arrowed). The corresponding positions of these sites of polyadenylation are shown on Figure 3.4.2. No fragments were observed to be protected from S1 nuclease digestion in the absence of added RNA.


mRNA
bp.
T
C
G
A
+
−
S1
(T) 120
(C) 110
(A) 100
(T) 90
(T) 80
(G) 70
(A) 60

3.4.5: Structure of the deduced AldDH polypeptide.

A comparison of the deduced amino acid sequence of *A. niger* AldDH with that of *A. nidulans*, showed 82% amino acid sequence similarity. This is similar to the levels of similarity seen between the products of the *A. niger* and *A. nidulans pyr*G (75%) and *arg*B (84%) genes (Wilson *et al.*, 1988; Buxton *et al.*, 1987). There is 46% similarity of amino acids of the *A. niger* AldDH to the human and horse polypeptides (Figure 3.4.7).

There is evidence to suggest that the cysteine residue at position 302 (Cys-302) in the human enzymes is the catalytic residue and this is adjacent to another Cys residue within a hydrophobic cleft at the active site (Hempel *et al.*, 1985). Cys-299 in the *A. niger* enzyme is found at the homologous position to the human Cys-302. Apart from Cys-298, the only other Cys common to the five sequences compared is Cys-159 in the *A. niger* enzyme. This result, coupled with the conservation of hydrophobic residues surrounding Cys-159 of the *A. niger* enzyme, implicates this residue as the second Cys at the active site. The coenzyme binding site for the mammalian enzymes has not been determined.

The first 22 amino acids of the *A. niger* enzyme, which are encoded in the first exon, show five amino acid substitutions when compared to the *A. nidulans* enzyme and virtually no similarity to mammalian sequences. This segment is, however, likely to be important for cellular location, rather than enzyme activity, as there is no similarity in amino acid sequence between the human cytoplasmic and the human mitochondrial enzymes in this region, but similarity is seen for these amino acids between human and horse cytoplasmic enzymes (Hempel *et al.*, 1985).

Figure 3.4.7: Comparison of several aldehyde dehydrogenase amino acid sequences. The sequences use the standard single letter code and represent aldehyde dehydrogenases from (top to bottom): *A. niger*; *A. nidulans* (Pickett *et al.*, 1987); human cytoplasm (Hempel *et al.*, 1984); human mitochondria (Hempel *et al.*, 1985); and horse cytoplasm (Bahr-Lindstrom *et al.*, 1984). Numbers refer to the amino acids of the *A. niger* enzyme, with the last digit aligning with the corresponding amino acid. Gaps indicate identity of amino acids to those in the corresponding positions in the *A. niger* enzyme. Hyphens (-) indicate spaces inserted into the sequences to obtain the best alignments. The significance of cysteine residues marked by open triangles is discussed in the text.

```
        10         20         30         40         50         60         70         80
MSDLFATITTPNGVKYEQPLGLFIDGEFVKGAEGKTFETINPSNEKPIVAVHEATEKDVDTAVAAARKAFE-GS-WRQVTPS
     T  E  -VI          NN      V      QV        V TS           V       A     P
S SGTPDLPVLLTDLKI YTKI  NN WHDSVS  K PVF  AT EELCQ E GDKE   K  K   Q  QI  P  TMDA
AAAATQAVPA  QQPEVFCNQI  NN WHDAVSR   P V   TGEV CQ A GDKE   K  K   A  QL  P  RMDA
S SGTPDLPVLLTDLKF YTKI  NN WHDSVS  K PVF  AT EKLCE E GDKED NK       Q  QI  P ETMDA

        90        100        110        120        130        140        150        160
TRGRMLTKLADLFERDAEILASIEALDNGKSITMAH-GDIAGAAGCLRYYGGWADKIHGQTIDTNSETLNYTRHEPIGVCG
E  ILIN     M    IDT  A  S      AF    K V L NSI       A                P   T       V
E   L Y     I    RLL  TM SMNG   LYSN YLN L  CIKT    CA       Q R  PIDGNFFT
H   L NR    I    RTY  AL T      PYVISYLV LDMVLK      A       Y  K  PIDGDFFS        V
E   L Y     V    RL   TM SMNG   LFSN YLM LG CLKT    CA       Q R  PSGDNFFT        V

       170        180        190        200        210        220        230        240
QIIPWNFPLLMWAWKIGPAIATGNTVVIKTAEQTPLSGLYAANVIKEAGIPAGVVNVISGFGRVAGSAISHHMDIDKVAF
            S       V A      L   Q      A     KL     PF     I         T  A   S
          V LI        LSC      V P       TA HV SL      F P     IVP Y PT  A   S
            Q   L     L    V   M V       TA  V  L      F P     IVP    PT  A  AS E V
  L         FL    A   LSC      V P        A HV TL      F P     IVP Y PT  A   S

       250        260        270        280        290        300        310        320
TGSTLVGRTILQAAAKSNLKKVTLELGGKSPNIVFNDADIDNAISWANFGIFYNHGQCCCAGSRILVQEGIYDKFIARLK
       P                                  D                  F                      V  F
    E  KL KE  G      R           C  LA    L    VEF HH V  HQ     I A    F E S    E VR SV
    EI  V QV  GS     R              IMS    M W VEQ H AL F Q          TF    D    E VE SV
    E  KL KE  G      RTV         F    A    LET LEVTHQAL  HQ     V A    F E S    E VR SV

       330        340        350        360        370        380        390        400
ERALQNKVGDPFAKDTFQGPQVSQLQFDRIMEYIQHGKDAGATVAVGGERHGTEGYFIQPTVFTDVTSDMKINQEEIFGP
   QK     N  EQ                       N    K       T  D    N                    A
   KKYIL N LTPGVT     IDKE Y K LDL ES   KE   KLEC   GPW NK    V      SN   DE R AK
A  KSRV  N  DSK E      DET   KK LG   NT   QE   KLLC   GIAADR          G  QDG T AK
   KKYVL N LTPGVS     IDKE Y K LDL ES   KE   KLEC   GPW NK           SN  SDE R AK

       410        420        430        440        450        460        470        480
VVTVQKFKDVEDAIKIGNSTSYGLAAGIHTKDVTTAIRVSNALRAGTVWVNSYNLIQYQVPFGGFKESGIGRELGSYALE
   I       AE         D      AV    N N          K      I N  M S  A       Q  L
 QQIM    SLD V   RA N F    S   VF    IDK   TI S   Q         C GVVSA C       M   N       E GFH
 MQIL    TI EVVGRA NST       AVF    LDK  NYL Q   Q         C DVFGA S     Y M   S       E G Q
 QQIM    SLD V   RA N T    F   SF    LDK   T  A   Q         C GVVSA C       M   N    M E GFH

       490
NYTQIKAVHYRLGDALF
      T           A
E  EV T TVKISQKNS
A  EV T TVKVPQKNS
E  EV T TVKISQKNS
```

3.6 Conclusions.

The *A. niger aldA* gene has been cloned from a lambda genomic library using the cloned *A. nidulans aldA* gene. The complete nucleotide sequence of this gene and its flanking regions and the structure of its mRNA transcript has been determined. The coding region shows 80% nucleotide sequence similarity to that of the *A. nidulans aldA* gene, and contains three intons whereas the *A. nidulans aldA* gene has two. The promoter contains one major and two minor start points of transcription and sequences resembling the TATA and CCAAT box consensus sequences. There are several repeated sequences and a region that shows nucleotide sequence similarity to several fungal genes regulated by carbon catabolite repression. Analysis of these sequences and other regions of the promoter is presented and discussed in Chapter 6. There are four sites of polyadenylation situated between 109 and 157 nucleotides 3' to the stop codon. These sites of polyadenylation are not preceeded by the AAUAAA polyadenylation signal, but there is a sequence similar to regions 3' to the termination codons of the *alcA* and *aldA* genes. It is unknown whether this sequence plays a role in signalling the site of polyadenylation. Comparison of several AldDH amino acid sequences indicated that cysteine residues at homologous postions to Cys-159 and Cys-299 in *A. niger* AldDH are adjacent at the active site.

Chapter 4:
Regulation of *aldA* gene expression in *Aspergillus niger* and *Aspergillus nidulans*.

This chapter describes experiments which were performed with the aim of determining the relative levels of *aldA* expression in both *A. niger* and *A. nidulans* when grown under a variety of conditions. This has been done by using the cloned *aldA* genes of *A. niger* and *A. nidulans* to detect the *aldA* transcript in dot blots of RNA extracted from mycelia that were later quantified by densitometry. These experiments have shown that there are some fundamental differences in the regulation of *aldA* expression between wildtype *A. niger* and *A. nidulans*. Further experiments have also shown that in *A. nidulans*, the product of the *amdR* gene (also called *intA*), in addition to the previously described *alcR* mediated control, plays a role in regulating the expression of the *aldA* gene.

4.1: Detection of the *aldA* mRNA transcripts.

Northern blot analysis of RNA extracted from both *A. niger* and *A. nidulans* mycelia grown in various media using the corresponding cloned *aldA* genes as probes detected transcripts of approximately 1.8kb. (Figures 4.1.1 and 4.1.2). For *A. niger* RNAs, only the 1.8kb. band was present. In *A. nidulans* RNAs extracted from mycelia grown in the absence of glucose and presence of either threonine or ethyl methyl ketone, additional bands of higher molecular weight were also present.

Figure 4.1.1: Northern blot analysis of wildtype *A. niger* RNAs. RNA was isolated from freeze-dried mycelia as described in Section 2.6.1, electrophoresed on a 1.5% agarose, 8% formaldehyde gel, and transferred to Zeta-Probe membrane. The filter was probed with the insert of pNG100. Shown is A): autoradiograph of a short (6 hour) exposure; B): autoradiograph of a long (4 day) exposure. Each track contains 10ug of total RNA from mycelia grown overnight at 37 °C in glucose medium, harvested, washed and transferred to the following media for a further 4 hours: 1, 1% ethanol; 2, 50mM threonine; 3, 0.05% fructose, 50mM ethyl methyl ketone; 4, 0.05% fructose; 5, 1% glucose, 1% ethanol; 6, 1% glucose, 50mM threonine; 7, 1% glucose, 50mM ethyl ammonium chloride; 8, 1% glucose. 10mM NH_4Cl was added to all media as a source to nitrogen.

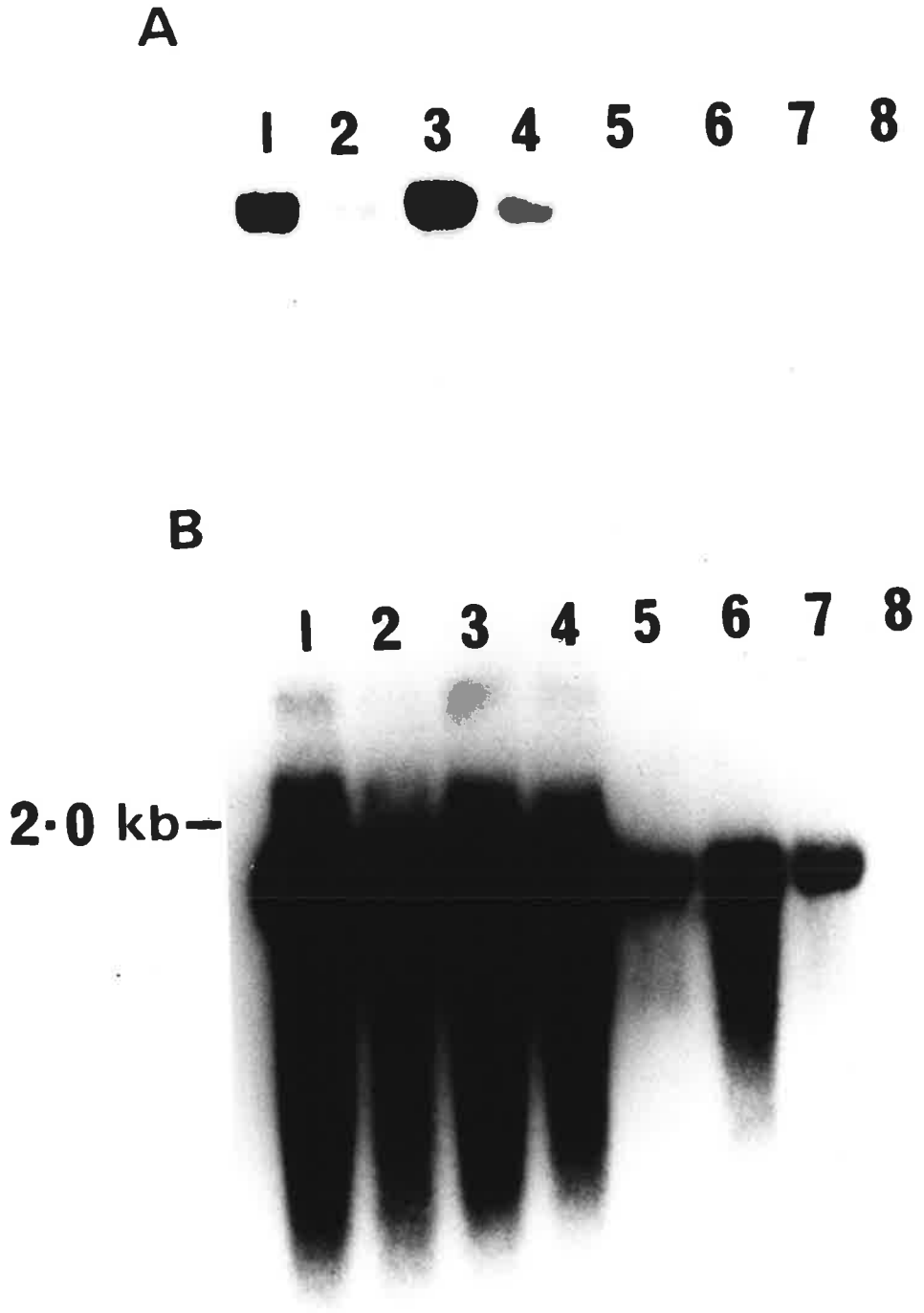
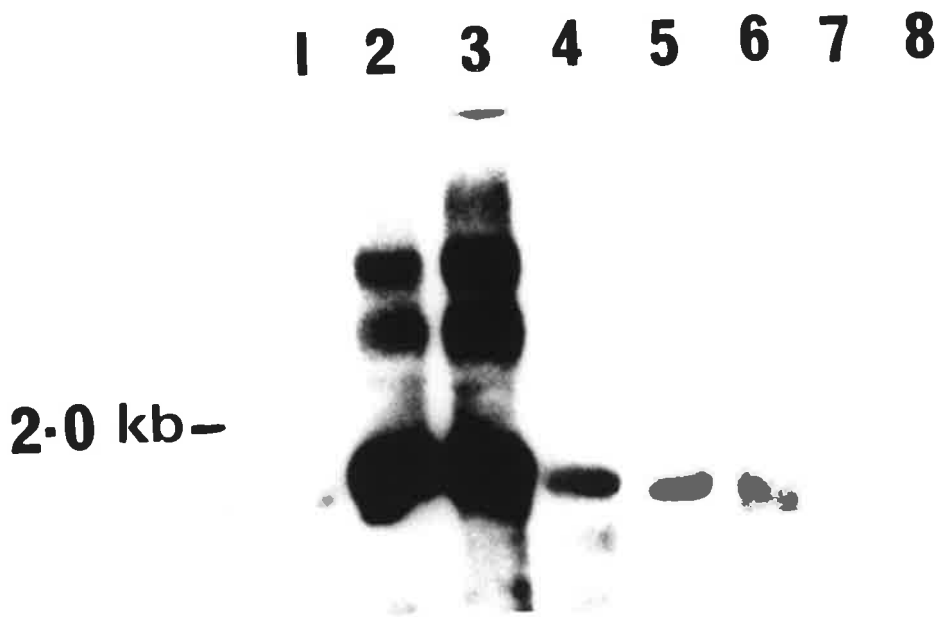

A
1 2 3 4 5 6 7 8
B
1 2 3 4 5 6 7 8
2·0 kb—

Figure 4.1.2: Northern blot analysis of wildtype *A. nidulans* RNAs. RNA was isolated from freeze-dried mycelia as described in Section 2.6.1, electrophoresed on a 1.5% agarose, 8% formaldehyde gel, and transferred to Zeta-Probe membrane. The filter was probed with the insert of pAN212. Shown is an autoradiograph of this filter. Each track contains 10ug of total RNA from mycelia grown overnight at 37 °C in glucose medium, harvested, washed and transferred to the following media for a further 4 hours: 1, 1% ethanol; 2, 50mM threonine; 3, 0.05% fructose, 50mM ethyl methyl ketone; 4, 0.05% fructose; 5, 1% glucose, 1% ethanol; 6, 1% glucose, 50mM threonine; 7, 1% glucose, 50mM ethyl ammonium chloride; 8, 1% glucose. 10mM NH_4Cl was added to all media as a source to nitrogen.


1 2 3 4 5 6 7 8
2·0 kb–

The identity of these bands is unclear, but they did not comigrate with the ribosomal RNAs, nor was there DNA contaminating the samples. Similar higher molecular weight bands were detected in RNA extracted from mycelia grown under inducing conditions by Lockington *et al.*, (1985) in northern blot analysis using the cloned *alc*A gene as a probe, and to a lesser extent in a lighter exposure of an autoradiograph of a northern blot probed with an *ald*A cDNA clone. It is possible that these higher molecular weight bands represent aggregates of mRNA molecules.

4.2: Induction of *ald*A gene expression in wildtype strains.

Relative levels of *ald*A mRNA were determined in RNA extracted from *A. niger* and *A. nidulans* mycelia grown in the presence of a variety of compounds to investigate their effect on *ald*A gene expression (Table 4.2.1). As has been done by many other workers, a low level of fructose (0.05%) was added to media lacking significant carbon sources to prevent starvation of cultures. In both *A. niger* and *A. nidulans*, growth in the presence of several compounds was shown to lead to induced levels of expression, which generally were repressed by the addition of glucose to the growth media. However, in almost all growth conditions tested, the relative levels of *ald*A expression differed between the two species.

One marked difference in the regulation of *ald*A expression between the two species was the effects of growth in the presence of 0.05% fructose. In *A. nidulans*, this growth condition led to *ald*A mRNA levels 8.5 fold higher than those found for glucose grown mycelia. Other workers have used 0.1% fructose as a derepressing/noninducing carbon source when looking at the expression of the ethanol utilization regulon (Lockington, *et al*, 1985; 1987). However, there are conflicting reports of the effects of fructose on the regulation of the ethanol

	A. niger		*A. nidulans*	
Growth Condition (+10mM Ammonium Chloride)	+Glucose	-Glucose	+Glucose	-Glucose
1% Ethanol	1.5	8	1	2
1% Ethanol, 0.05% Fructose	1	7	1	23
50mM Threonine	2.5	3	1	23.5
50mM Ethyl Methyl Ketone, 0.05% Fructose	1	4.5	0.5	20.5
1% Glycerol	1	1	0.5	0.5
50mM Sodium Acetate	1	1	1	1.5
5mM Acetaldehyde, 0.05% Fructose	1	11	1	10
5mM Glyceraldehyde, 0.05% Fructose	1.5	7	1	4
50mM Acetone, 0.05% Fructose	1	6	1	70
50mM Ethyl Ammonium Chloride, 0.05% Fructose	2	21	1	100
50mM Ethyl Ammonium Chloride, 0.05% Fructose (- 10mM Ammonium Chloride)	1	2	1	15
0.05% Fructose	1*	2	1*	8.5

Table 4.2.1. Relative levels of aldA mRNA in wildtype strains. *Values have been normalized against those for glucose grown mycelia. 0.05% fructose was not added to glucose cultures. Glucose was added to 1%. Cultures were grown overnight at 37 °C in glucose medium, harvested, washed and transferred to the media with the indicated growth condition for 4 hours as described in Section 2.6.1.

utilization regulon. Gwynne *et al.*, (1987b), using the *alc*A promoter to direct the expression of the human interferon gene in *A. nidulans*, showed that cultures grown in 1% fructose had a 25 fold higher level of gene expression, as determined by assay of human interferon levels, than glucose grown cultures, and only a 4 fold lower level of gene expression than cultures grown in 0.1% fructose and 100mM threonine. In addition, they found 0.1% fructose to be a derepressing carbon source for threonine induction. In contrast to this, in the expression system developed by Waring *et al.*, (1989) using the *alc*A promoter and the coding regions of *A. nidulans* tubulin genes, 0.1% fructose was found to be as strong a repressing carbon source as 1% glucose in growth tests on agar plates. Several lines of evidence gained from experiments presented in this thesis show 0.05% fructose to be an inducing growth condition for *ald*A gene expression in *A. nidulans*. Significantly lower levels of expression were seen in ethanol compared to fructose grown cultures (Table 4.2.1). Futhermore, although 0.1% arabinose and 0.1% galactose were shown not to repress ethyl methyl ketone induction, growth in media containing these compounds without other sources of carbon led to *ald*A expression levels only 3.5-4 fold higher than glucose grown cultures (Table 4.2.2), which is approximately half the expression level for fructose grown cultures. This observation, together with the level of *ald*A expression in cultures grown in ethanol alone, shows that the level of *ald*A expression in fructose grown cultures represents an induced rather than merely a derepressed/noninduced level of expression.

The higher level of *ald*A expression was not seen in fructose grown cultures of *A. niger*, nor did fructose significantly affect ethanol induction of *ald*A in *A. niger*, which is in contrast to the synergistic effect of fructose on ethanol induction in *A. nidulans* (Table 4.2.1).

As will be discussed in Section 4.3, fructose induction of *ald*A expression in *A. nidulans* appears to be controlled by a separate mechanism to the *alc*R dependent induction mechanism which controls other known sources of induction. Studies of the expression of the *A. nidulans* *ald*A gene in *A. niger*, and the analysis of mutations 5' to the *A. niger* *ald*A gene has shown that this induction mechanism exists in *A. niger* but is silenced by a promoter element (see Chapters 5 and 6).

Acetaldehyde or glyceraldehyde added to the growth medium together with 0.05% fructose induced *ald*A gene expression to a greater level than fructose alone in *A. niger* but not in *A. nidulans*. However, as acetaldehyde can induce *ald*A expression in *A. nidulans* in the absence of fructose induction (see Section 4.3) and other potential sources of acetaldehyde do induce *ald*A, the presence of fructose may be masking the induction of *ald*A in *A. nidulans* during growth in media containing acetaldehyde. The ketones acetone and ethyl methyl ketone induce *ald*A expression in both species, although to a far greater extent in *A. nidulans* than in *A. niger*.

Carbon Source	+EMK	-EMK
1% Glucose	0.5	1*
0.05% Fructose	20.5	8.5
0.1% Arabinose	21	3.5
0.1% Galactose	19.5	4

Table 4.2.2: *ald*A mRNA levels in wildtype *A. nidulans*. *Values have been normalized against those for glucose grown mycelia. EMK: ethyl methyl ketone. Cultures were grown overnight at 37 °C in glucose medium, harvested, washed, and transferred to media containing the indicated carbon source plus 10mM ammonium chloride for 4 hours as indicated in Section 2.6.1.

The growth condition leading to the highest level of induced expression of *ald*A in both species was 0.05% fructose, 50mM ethyl ammonium choride, 10mM ammonium chloride (Table 4.2.1). Ethyl ammonium chloride is converted to acetaldehyde and ammonium by monoamine oxidase in *A. nidulans* (Page and Cove, 1972). An amine oxidase has also been isolated and characterized from *A. niger* (Yamada and Adachi, 1971). The induction of *ald*A in the presence of ethyl ammonium chloride is not as significant in *A. nidulans*, and absent in *A. niger*, if an additional source of nitrogen is not added to the growth medium. Both species are, however, capable of strong growth on media containing ethyl ammonium chloride as the sole source of nitrogen. Additional ammonium may stimulate the rate of metabolism of ethyl ammonium chloride to acetaldehyde which could be an inducer for *ald*A, or ammonium repression of amine oxidase encoding genes may lead to an intracellular accumulation of ethyl ammonium chloride that may itself act as an inducer of *ald*A. However, this is no evidence to support either of these hypotheses.

In *A. niger*, growth in the presence of threonine led to induced levels of *ald*A expression that were insensitive to glucose repression. This effect was specific to threonine growth in *A. niger*, but also acted on the *A. nidulans ald*A gene when expressed in *A. niger* (Chapter 5) and was likely to be due to acetaldehyde production from threonine metabolism. The metabolism of threonine to acetaldehyde in *A. nidulans* is most probably via the action of a threonine aldolase (Creaser *et al.*, 1987). Threonine is not an inducer of *alc*A or *ald*A in strains carrying mutations at the *tut*E locus, but these strains are unaffected for induction in the presence of ethyl ammonium chloride (Hynes and Kelly, 1981; Pateman *et al.*, 1983; M.J. Hynes, pers. comm.). Hence the *tut*E locus probably encodes a threonine aldolase. The insensitivity to

carbon catabolite repression of threonine induction of *ald*A expression in *A. niger* could be due to the expression of a gene encoding threonine aldolase that is insensitive to carbon catabolite repression. Thus, it is possible that carbon catabolite repression does not act directly on *ald*A in *A. niger*, but indirectly via inducer formation. The indirect action may be by repressing the expression of genes, such as those encoding alcohol dehydrogenases, which are required to produce aldehydes. If this is so, then genes controlling carbon catabolite repression do not act directly on *ald*A to affect the initiation of transcription. That is, once the intracellular inducer is available, the expression of the gene is induced. However, induction of *ald*A in media containing aldehydes and ketones is subject to carbon catabolite repression, and thus glucose may also affect the uptake of these compounds. Indirect evidence for this came from plate test experiments. Concentrations of acetaldehyde greater than 5mM are toxic to *A. niger* and *A. nidulans* when added to media containing relatively poor sources of carbon such as quinate or glycerol. In contrast, acetaldehyde added to glucose medium to a final concentration of 50mM is not toxic to wildtype strains. Considering these phenotypes, and the carbon catabolite repression of *ald*A induction in media containing acetaldehyde, glucose may affect the uptake of acetaldehyde, and possibly other carbon sources, which are either inducers of *ald*A or the utilization of which produces metabolites that act as inducers of *ald*A expression, but not significantly affect the uptake of threonine. Once inside the cell, metabolism of threonine to acetaldehyde could cause the induction of *ald*A independent of the presence of glucose.

An alternative hypothesis to explain the insensitivity to carbon catabolite repression of *ald*A induction in threonine containing media is that in *A. niger*, induction of *ald*A in threonine media occurs by a

separate mechanism to that affecting *ald*A expression in other growth conditions that lead to induced levels of expression. Under this hypothesis, this separate mechanism must not be affected by carbon catabolite repression, but this does not exclude the possibility of a direct affect of carbon catabolite repression on other induction mechanisms. Experiments have shown that constitutive expression of *alc*R in *A. nidulans* fails to lead to derepression under repressing conditons of *ald*A expression (Felenbok *et al.*, 1989). However, these experiments do not discount the possibility that it is the formation of inducer that is still subject to carbon catabolite repression under these conditions. Lockington *et al.*, (1987) showed that there is an absolute requirement for the coinducer together with the *alc*R gene product for the *alc*R dependent induced expression of *ald*A and *alc*A, and so constitutive expression in *A. nidulans* of *alc*R in the absence of the intracelluar inducer would not lead to high levels of *ald*A and *alc*A expression. The molecular mehanisms controlling carbon catabolite repression in *Aspergillus* are not understood, and may, as is seen in carbon catabolite repression in *Saccharomyces* species (reviewed by Carlson, 1987), involve several separate processes affecting different induction mechanisms including the uptake of different carbon sources. Regardless of the mechanism that is acting during carbon catabolite repression, the induction of *ald*A in threonine plus glucose medium implies that in *A. niger*, the need to detoxify the level of acetaldehdye produced from threonine metabolism causes the induction of *ald*A expression independent of the presence of glucose. Apart from threonine induction, the induced levels of expression in the presence of all other carbon sources were tightly repressed by the addition of 1% glucose.

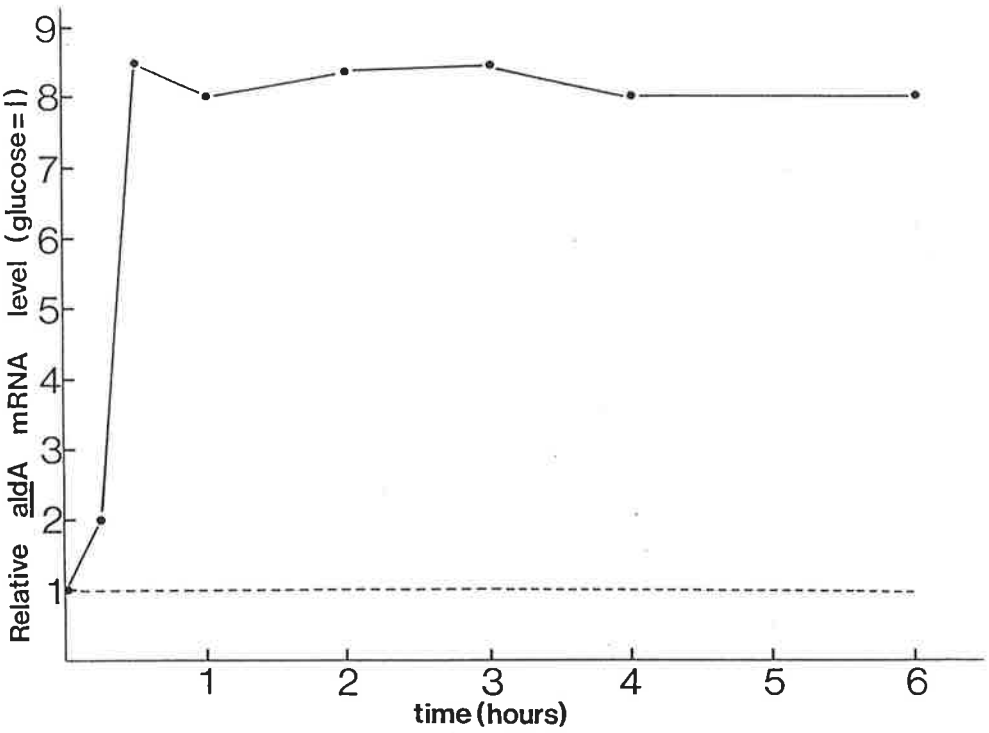
Analysis of the time course of ethanol induction of *ald*A gene expression in *A. niger* showed that the induction observed during ethanol growth is rapid, and *ald*A mRNA levels were found to be at a maximal level within 30 minutes of transfer of cultures to ethanol medium (Figure 4.2.1).

The large number of differences in *ald*A expression between the two species are unlikely to be due to artefacts of the culture conditions and imply that there are some fundamental differences between the metabolic pathways requiring aldehyde dehydrogenase in *A. niger* and *A. nidulans*. Further experiments investigating the regulation of *ald*A gene expression in *A. niger* and *A. nidulans* are discussed in the following two sections.

4.3: The effects of mutations at the *amd*R locus on *ald*A regulation in *A. nidulans*.

In *A. nidulans*, sources of L-glutamic γ-semialdehyde and succinic semialdehyde are toxic to strains carrying mutations at the *ald*A locus (Arst *et al.*, 1981). The *amd*R gene of *A. nidulans* (also called *int*A) regulates the expression of the *lam*A and *lam*B genes, which are responsible for the utilization of 2-pyrrolidone (and other lactams) via GABA (Arst *et al.*, 1978; Katz and Hynes, 1989a), and the *gab*A and *gat*A genes which are responsible for the metabolism of GABA to succinic semialdehyde (Arst, 1976). As the products of these *amd*R regulated genes are involved in a pathway leading to a metabolite toxic to *ald*A$^-$ strains, the effects of *amd*R mutations on *ald*A regulation were investigated. Levels of *ald*A mRNA in an *amd*R$^-$ strain were compared to those in a wildtype strain grown under the same conditions and it was found that this mutation does affect *ald*A expression (Table 4.3.1).

Figure 4.2.1: Time course of *ald*A induction in *A. niger*. Total RNA was extracted from freeze dried mycelia (Section 2.6.1) and assayed for *ald*A mRNA levels as described in Section 2.6.5. A single mycelial culture grown at 37 °C in glucose medium, washed and equal amounts of mycelia were transferred to media containing 1% ethanol, 10mM NH_4Cl for 0, 15, 30, 60, 120, 180, 240, and 360 minutes prior to harvesting and RNA isolation. The broken line indicates the level of *ald*A RNA from a culture transferred to glucose medium for 4 hours against which the other samples have been normalized.


Relative aldA mRNA level (glucose=1)
9
8
7
6
5
4
3
2
1
1
2
3
4
5
6
time (hours)

The levels of *ald*A expression in the *amd*R⁻ strain grown in the presence of threonine, or in the presence of glucose plus threonine was four fold higher than that seen in the wildtype strain. This could have been due to the accumulation of an inducer of *ald*A and would implicate *amd*R as playing a role in threonine metabolism. That is, a block in threonine metabolism caused by the absence of an enzyme encoded by an *amd*R regulated gene could lead to the accumulation of a metabolite that leads to *ald*A induction. *tut*E mutants fail to show threonine induction of *alc*A and *ald*A (Hynes and Kelly, 1981; Pateman *et al.*, 1983; M.J. Hynes, pers. comm.), and thus it is unlikely that an accumulation of threonine itself would increase *alc*R mediated induction of *ald*A. Growth of *amd*R⁻ strains on media containing ethanol or threonine as the sole carbon source is, however, equally as strong as wildtype. Thus, the altered expression of *ald*A but unaltered phenotype of *amd*R⁻ strains grown on threonine as a carbon source may imply that an *amd*R⁻ mutations affect an alternative pathway of threonine metabolism, leading to the accumulation of an inducer for *ald*A derived from threonine, but this alternative pathway is not essential for growth on threonine as a carbon source.

Expression of *ald*A in the *amd*R⁻ strain grown in the presence of 0.05% fructose, both alone and together with ethanol or ethyl ammonium chloride, was significantly lower than that in the wildtype strain. This decreased expression in fructose grown *amd*R⁻ cultures implies that induction in the presence of fructose requires *amd*R and hence *amd*R may also play a role in fructose metabolism. The synergistic effect of fructose on ethanol induction in wildtype *A. nidulans* was absent in the *amd*R⁻ strain, where the *ald*A expression level in mycelia grown in 0.05% fructose, 1% ethanol is the same as that for a wildtype strain grown in 1% ethanol alone. The reduction in expression levels in *amd*R⁻ cultures

Growth Condition (+10mM Ammonium Chloride)	Wildtype	*amd*R⁻	*amd*R^c	*amd*A7
0.05% Fructose, 1% Ethanol	23	2	ND	ND
1% Ethanol	2	ND	ND	ND
1% Glucose, 1% Ethanol	1	1	ND	ND
0.05% Fructose, 50mM EMK	20.5	19	ND	ND
1% Glucose, 50mM EMK	0.5	1	ND	ND
0.05% Fructose, 5mM Acetaldehyde	10	7	ND	ND
1% Glucose, 5mM Acetaldehyde	1	1	ND	ND
0.05% Fructose, 50mM EAC	100	17.5	ND	ND
1% Glucose, 50mM EAC	1	2	ND	ND
0.05% Fructose, 50mM EAC (-10mM Ammonium Chloride)	15	22	ND	ND
1% Glucose, 50mM Ethyl (-10mM Ammonium Chloride)	1	3	ND	ND
50mM Threonine	23.5	80	30	22
1% Glucose, 50mM Threonine	1	4	0.5	1
50mM GABA	2	5.5	1	ND
1% Glucose, 50mM GABA	1.5	2	0.5	ND
50mM Sodium Acetate	1.5	ND	ND	2
1% Glucose, 50mM Sodium Acetate	1	ND	ND	0.5
0.05% Fructose	8.5	1	ND	ND
1% Glucose	1*	1	1	1

Table 4.3.1. aldA mRNA levels in *A. nidulans* strains. *Levels are normalized against those for glucose grown wildtype mycelia. ND: Not Determined. EMK: ethyl methyl ketone. EAC: ethyl ammonium chloride. Cultures were grown overnight at 37 °C in glucose medium, harvested, washed, and transferred to media with the indicated growth conditions for 4 hours as indicated in Section 2.6.1.

grown in 0.05% fructose, 50mM ethyl ammonium chloride, 10mM ammonium chloride implies that the same synergistic effect for fructose on ethanol induction also occurs under this growth condition. As the level of expression is affected less in *amd*R$^-$ cultures grown in 0.05% fructose, 50mM ethyl methyl ketone and 0.05% fructose, 5mM acetaldehyde, it is likely that these conditions lead to increased *ald*A expression that is independent of the effects of fructose, and hence fructose induction is independent of the induction mechanisms acting during growth in the presence of these compounds. These results show that induction of *ald*A expression in *A. nidulans* in ethanol plus fructose and ethyl ammonium chloride plus fructose media is a combination of the separate effects of *alc*R dependent and *amd*R dependent induction mechanisms. The increased *ald*A expression in the *amd*R$^-$ strain grown in the presence of 50mM GABA is likely to be due to starvation as this strain would be lacking *gab*A and *gat*A expression. Although it is clear that *amd*R has an effect on *ald*A regulation, whether the *amd*R product, which is a DNA binding protein (Davis and Hynes, 1989), acts directly on the promoter of *ald*A is not known. It is possible that *amd*R acts on *ald*A indirectly by affecting the expression of genes controlling the levels of metabolites involved in *ald*A induction by other unknown mechanisms.

The semidominant *amd*R^c allele showed no effect on *ald*A expression (Table 4.3.1). The expression of the *amd*S gene is also *amd*R regulated and is increased in *amd*R^c strains (Hynes and Pateman, 1970), but a similar effect may not be as obvious for *ald*A because of the significantly higher levels of expression compared to *amd*S. It is also possible that the *amd*R^c mutation may specifically affect the interaction between the *amd*R gene product and sequences 5' to the *amd*S gene without affecting the putative interaction of the *amd*R encoded

protein and either *ald*A or genes encoding functions that produce molecules involved in *ald*A induction. There is evidence that *A. niger* has a gene of homologous function to the *amd*R gene of *A. nidulans* (Kelly and Hynes, 1985), but any affects of this putative gene on *ald*A expression in *A. niger* are unknown.

The *amd*A7 mutation, which leads to increased expression of the *amd*S and *aci*A genes (Hynes, 1978; Atkinson *et al.*, 1985), did not affect *ald*A expression under the growth conditions tested (Table 4.3.1).

These experiments have shown that in addition to the *alc*R dependent induction of *ald*A gene expression in *A. nidulans*, the product of the *amd*R gene also affects the expression of *ald*A. Whether this effect is caused by a direct interaction between the *amd*R gene product and the *ald*A promoter is not known. Another important finding was the failure of GABA to act as an inducer of *ald*A expression in wildtype *A. nidulans* which implies that it is the basal level of *ald*A expression that is important in GABA metabolism, as GABA is toxic to *ald*A$^-$ strains (Arst *et al.*, 1981).

4.4: Repression of *ald*A expression in *A. niger*

The presence of 1% glucose repressed levels of *ald*A expression in *A. niger* when any of a range of inducing compounds were also present (Table 4.2.1). Several other carbon sources were tested for their ability to repress induction of *ald*A expression in the presence of ethanol in *A. niger*. Apart from 0.05% fructose, all carbon sources tested repressed *ald*A expression to a similar level as 1% glucose (Table 4.4.1). Slightly less repression was observed for quinate and acetate, both of which are weak carbon sources for *A. niger*. Acetate

represses the expression of *alc*A and *ald*A in *A. nidulans*, but not to the same extent as glucose (Pateman *et al.*, 1983). Glycerol was as repressing as glucose for *ald*A induction in *A. niger*, but is a non-carbon catabolite repressing carbon source in *A. nidulans* (Bailey and Arst, 1975). The differences in carbon catabolite repression in response to various carbon sources, and the insensitivity to carbon catabolite repression of threonine induction of *ald*A in *A. niger* shows that there are fundamental differences in carbon catabolite repression, and possibly preference for different carbon sources, between *A. niger* and *A. nidulans*.

Growth Condition (+10mM Ammonium Chloride)	+ 1% Ethanol	- 1% Ethanol
50mM Glucose	1.5	1*
0.05% Fructose	7	2
50mM Fructose	0.5	0.5
50mM Sodium Acetate	2	1
50mM Sucrose	1.5	1
50mM Arabinose	1	1
50mM Quinic Acid	2	1
50mM Maltose	1.5	0.5
50mM Sorbitol	1	1
50mM Glycerol	1	1
No added carbon source	8	ND

Table 4.4.1: Levels of *ald*A mRNA in wildtype *A. niger*.*Levels are normalized against those for glucose grown mycelia. ND: Not Determined. Cultures were grown at 37 °C overnight in glucose medium, harvested, washed, and transferred to media with the indicated growth condition for 4 hours as indicated in Section 2.6.1.

4.5: Conclusions.

Northern blot analysis detected 1.8kb. *ald*A transcripts in *A. niger* and *A. nidulans* RNAs, and additional hybridizing bands of higher molecular weight in *A. nidulans* RNAs containing high levels of the *ald*A transcript. These higher molecular weight bands are probably aggregates of the single transcript.

Analysis of *ald*A expression levels in wildtype *A. niger* and *A. nidulans* showed that the expression of *ald*A is subject to induction and repression in both species at the level of mRNA accumulation. This control is most probably via the initiation of transcription. The highest level of *ald*A expression for the growth conditions tested was detected in cultures grown in a medium containing ethyl ammonium chloride and ammonium chloride. The reason for the requirement of an additional source of nitrogen to obtain the high level of expression is not known, but may be due to either ammmonium repression of amine oxidase encoding genes leading to a build up of intracellular ethyl ammonium chloride that itself acts as an inducer of *ald*A, or added ammonium may increase the rate of ethyl ammonium metabolism and hence the level of *ald*A expression.

For almost all the growth conditions tested there were marked differences in the relative levels of induction and repression of *ald*A gene expression between the two species. Notably, growth in fructose medium leads to induced levels of *ald*A expression in *A. nidulans* but not *A. niger*, and threonine induction is insensitive to carbon

catabolite repression in *A. niger* but not *A. nidulans*. Fructose induction of *ald*A in *A. nidulans* was shown to require a functional *amd*R gene product, and thus is independent of the other induction mechanism acting via *alc*R. Although growth in fructose medium is not an inducing condition for *ald*A expression in *A. niger*, other experiments showed that this induction mechanism does exist in *A. niger* but is silenced by an *ald*A promoter element (Chapters 4 and 5). The increase in threonine induction of *ald*A but unaltered growth characteristics on threonine media of *amd*R$^-$ strains showed that *amd*R probably controls steps in an alternative pathway of threonine metabolism that is not essential for growth but does result in the accumulation of inducers of *ald*A. The failure of GABA to act as an inducer of *ald*A indicates that the basal level of *ald*A expression is important in GABA metabolism as this compound is toxic to *ald*A$^-$ strains.

The existence of an induction mechanism for *ald*A that is likely to be independent of mechanisms controlling the expression of *alc*A is not surprizing in light of the potential requirement in *A. nidulans* for aldehyde dehydogenase to detoxify aldehydes and semialdehydes in metabolic pathways not requiring alcohol dehydrogenase, such as those involved in lactam and L-proline metabolism.

Richardson *et al.*, (1989) identified several sequences in the *gat*A promoter within a region capable of titrating the *amd*R gene product that are similar to the region of the *amd*S promoter which is likely to be the site of action of *amd*R. There are no sequences in the sequenced region of the *ald*A promoter, which extends approximately 310 bp. 5' to the start points of transcription (Gwynne *et al.*, 1987a), that resemble these putative *amd*R binding sites of the *gat*A and *amd*S genes. Although the putative *amd*R binding sites of *gat*A and *amd*S are within 150bp. of the start of transcription, it is possible that such a

sequence could exist further 5' than the sequenced region. Comparison of the nucleotide sequences of genes under common control does not always reveal regions of sequence similarity that could be hypothesized as the binding site of a regulatory protein. For example, comparison of the nucleotide sequences of the promoters of *alc*A, *ald*A and *alc*R genes does not reveal significant regions of nucleotide sequence similarity that could be hypothesized as putative *alc*R binding sites (Gwynne *et al.*, 1987a; Felenbok *et al.*, 1988). Alternatively, it is possible that *amd*R acts indirectly on *ald*A by controlling the expression of genes whose products are responsible for inducer formation from fructose, with an as yet unidentified gene positively regulating the expression of *ald*A in *A. nidulans* in the presence of inducers derived from fructose. However, there is indirect evidence against this which was gained from the analysis of promoter deletions in the *A. niger ald*A gene. There is a promoter element that appears to silence fructose induction of *ald*A in *A. niger* without affecting ethanol induction (Chapter 6), which implies a direct affect of a regulatory gene product on the *A. niger ald*A promoter during fructose growth when this silencing element is deleted. However, it is not known whether this is affecting the binding of a regulatory protein similar to the *amd*R gene product of *A. nidulans*. Although it is clear that the *amd*R$^-$ mutation affected *ald*A expression, further experiments such as the analysis of *ald*A expression levels in strains carrying different *alc*R and *amd*R alleles, and combinations of *alc*R and *amd*R mutations, together with *in vitro* analyses via gel retardation experiments, will need to be performed to elucidate whether the affects of *amd*R on *ald*A expression are via direct or indirect action.

The ethanol induced expression of *ald*A in *A. niger* was shown to be repressed by the addition of several other carbon sources to the growth

medium. Like *A. nidulans*, the repression by weak carbon sources such as acetate and quinate was slightly less severe than the effects of strong carbon sources such as glucose and sucrose. Glycerol was as strong a source of carbon catabolite repression as glucose in *A. niger*, whereas in *A. nidulans*, glycerol is a non-carbon catabolite repressing source of carbon (Bailey and Arst, 1975). Thus, it appears likely that although carbon catabolite repression acts in both species, the level of repression by various carbon sources may differ, and hence may reflect differing preferences for carbon sources between *A. niger* and *A. nidulans*. In addition, positive regulators of *ald*A in *A. niger* must be able to act in the presence of sources of carbon catabolite repression as threonine induction is insensitive to the carbon catabolite repressing effects of 1% glucose. The induced expression of *ald*A in *A. niger* in threonine plus glucose medium implies that carbon catabolite repression does not act directly on *ald*A or genes encoding positively acting regulatory proteins that control *ald*A expression during growth in this medium. It is likely that several different and separate mechanisms, including the control of the uptake of various carbon sources, affect the expression of genes involved in carbon metabolism in *Aspergillus*. *A. niger* contains sequences which hybridize under low stringency to the cloned *cre*A gene of *A. nidulans* as well as sequences of homologous function to *cre*A (C. Dowzer, J.Kelly and K. Heath, pers. comm.). This putative gene would be a good candidate for playing a role in either the direct or indirect control of the carbon catabolite repression of *ald*A expression in *A. niger*.

This analysis of *ald*A gene expression in *A. niger* and *A. nidulans* has revealed that although the expression of *ald*A genes in both species is subject to induction and carbon catabolite repression, there are many differences in reponse to a variety of growth conditions of

relative levels of gene expression. Furthermore, these experiments have shown that a functional *amd*R gene product is required for induced levels of *ald*A gene expression during growth in fructose medium in *A. nidulans*. This induction mechanism probably acts separately to the previously described *alc*R dependent induction mechanism, and may involve a direct interaction of the *amd*R gene product and the *ald*A promoter. The increased threonine induction in the *amd*R$^-$ strain is likely to be due to the accumulation of a metabolite of threonine and is probably independent of the effects of this mutation on fructose induction.

The large number of differences in *ald*A regulation between *A. niger* and *A. nidulans* suggest that there are fundamental differences in metabolic pathways requiring aldehdye dehydrogenase activity between these two fungi. This may be due to the combination of both genetic distance between *A. niger* and *A. nidulans* and differing requirements for aldehyde dehydrogenase activity for growth in the natural environments of the two species.

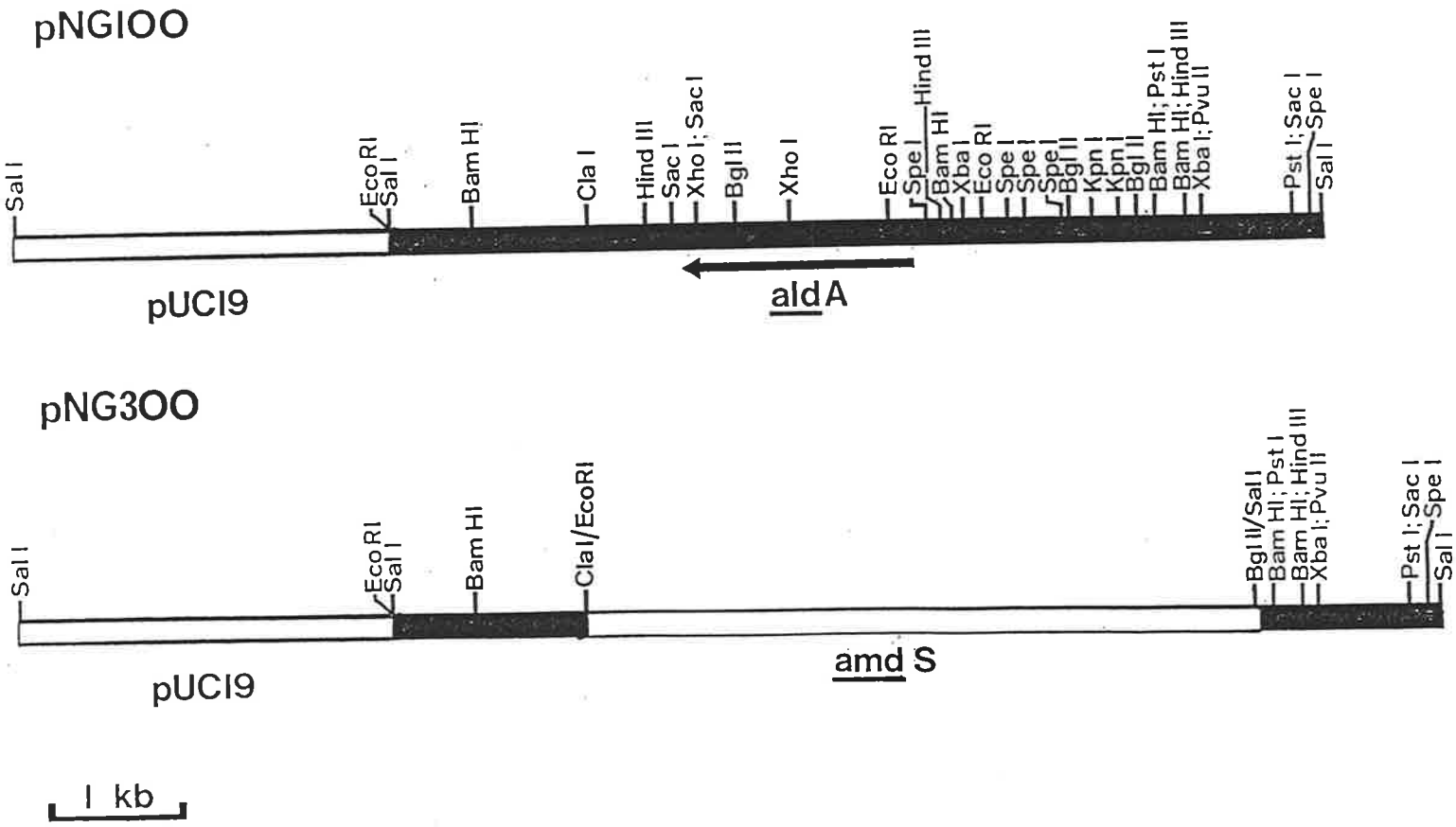
Chapter 5: Heterologous expression of the *aldA* genes of *Aspergillus niger* and *Aspergillus nidulans*.

This chapter describes experiments in which the heterologous expression of the *A. niger* and *A. nidulans aldA* genes was investigated. Expression of the *A. niger aldA* gene in *A. nidulans* has shown that the induction of this gene requires a functional *alcR* gene product. Expression of the *A. nidulans aldA* gene in *A. niger* has given further data with respect to the effects of threonine and fructose on *aldA* expression in both *A. niger* and *A. nidulans*. In addition to these experiments, the effects of increasing *aldA* gene copy number on *aldA* regulation in both species has been investigated. These experiments have extended the data gained from experiments discussed in Chapter 4, and led to a better understanding of *aldA* regulation in *A. niger* and *A. nidulans*.

5.1: Construction of an *A. niger aldA* deletion strain.

A strain of *A. niger* containing a deletion of the *aldA* gene was constructed by gene replacement. A 4 kb region of pNG100 surrounding the *aldA* gene extending 1.7 kb 5' to the start codon and 0.7 kb 3' to the stop codon, was replaced by end filling and blunt end ligation with the insert of p3SR2, which contains the *amdS* gene of *A. nidulans*, to make the plasmid pNG300 (Figure 5.1.1). There is approximately 1.5kb.

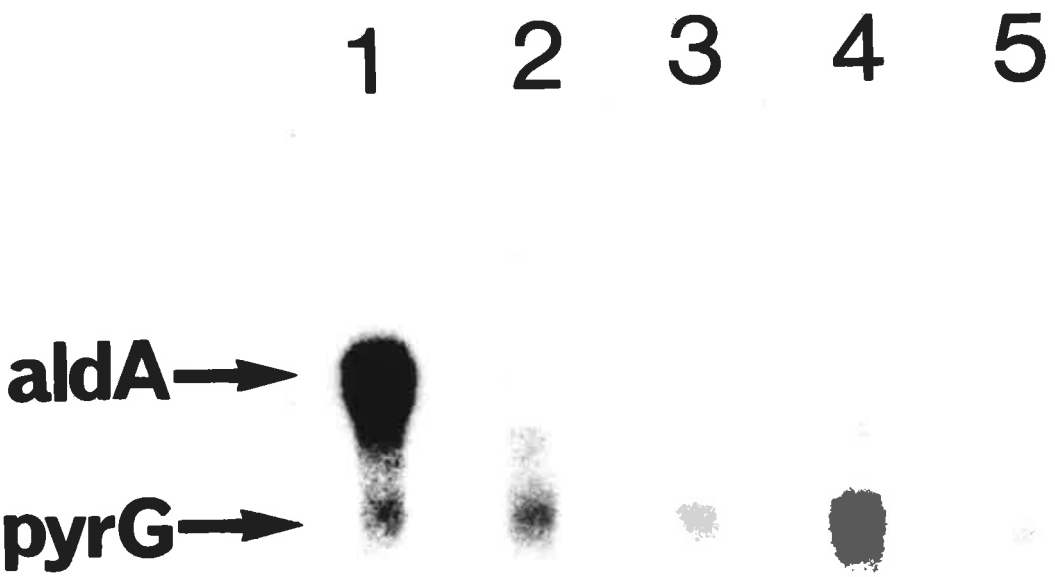
Figure 5.1.1: Structure of the plasmid pNG300 which was used to construct the *A. niger aldA* gene replacement strain. The EcoRI-SalI insert of p3SR2 (Hynes *et al.*, 1983) was end filled and blunt end ligated into end filled, ClaI/BglII digested pNG100. The orientation of the p3SR2 insert in pNG300 is as shown.


pNG100
Sal I
Eco RI
Sal I
Bam HI
Cla I
Hind III
Sac I
Xho I; Sac I
Bgl II
Xho I
Eco RI
Spe I
Hind III
Bam HI
Xba I
Eco RI
Spe I
Spe I
Spe I
Bgl II
Kpn I
Kpn I
Bgl II
Bam HI; Pst I
Bam HI; Hind III
Xba I; Pvu II
Pst I; Sac I
Spe I
Sal I
pUC19
aldA
pNG300
Sal I
Eco RI
Sal I
Bam HI
Cla I/Eco RI
Bgl II/Sal I
Bam HI; Pst I
Bam HI; Hind III
Xba I; Pvu II
Pst I; Sac I
Spe I
Sal I
pUC19
amd S
1 kb

of *ald*A flanking DNA on both sides of the *amd*S sequences in pNG300. This plasmid was used to transform wildtype *A. niger* in an attempt to delete the *ald*A gene by a double homologous recombination event between the *ald*A flanking regions in pNG300 and the chromosome. Transformants were selected by their ability to grow on media containing acetamide as the sole source of nitrogen. AmdS$^+$ transformants were screened by plate testing for possible phenotypes that would suggest deletion of the *ald*A region, including reduced growth on ethanol, threonine and glucose plus acetaldehyde media. Approximately 100 AmdS$^+$ transformants, some of which showed at least slightly different phenotypes in the plate testing and some of which were chosen at random, were screened by Southern hybridization for an altered restriction pattern surrounding *ald*A. One of the transformants, designated 712, which showed reduced growth on ethanol and threonine media, and sensitivity to acetaldehyde, had an altered restriction pattern surrounding the *ald*A locus that indicated integration of pNG300 at the *ald*A locus via the *ald*A flanking regions. Northern blot analysis of RNA extracted from 712 mycelia grown under a variety of conditions showed it lacked detectable *ald*A mRNA which gave further evidence for the loss of the *ald*A region (Figure 5.1.2). The low frequency of homologous integration events would imply that the *ald*A locus in *A. niger* is in a relatively inaccessible region of the genome.

Spontaneous *pyr*G mutations were selected in strain 712 to allow further selection of transformants. Mutations at the *pyr*G locus, which encodes orotidine-5'-phosphate decarboxylase, result in an uridine auxotrophy and can be positively selected by resistance to 5-flouro orotic acid (Boeke *et al.*, 1984). The orotidine-5'-phosphate decarboxylase encoding gene has also been designated *pyr*A by Goosen *et al.*, (1987). Several *pyr*G mutations were selected, and one of these

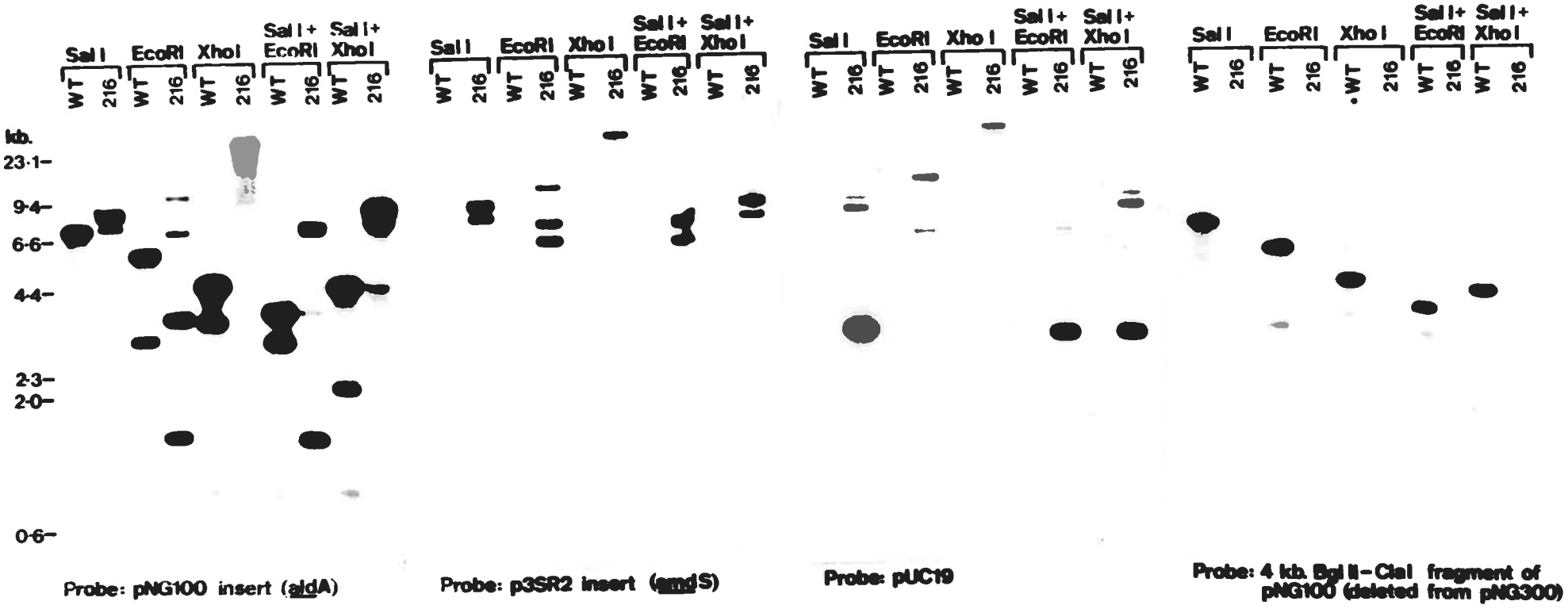
Figure 5.1.2: Northern blot analysis of RNAs extracted from *A. niger* strain 712. Each track contains 10ug of total RNA electrophoresed in a 1.5% agarose, 8% formaldehyde RNA gel (Section 2.6.3) and transferred to Zeta Probe membrane. The filter was probed with the inserts of pNG100, which detects *A. niger aldA* mRNA, and pAB4.1, which detects *A. niger pyrG* mRNA. A 10 fold excess of the *pyrG* probe was used. RNA was extracted from mycelia after overnight growth at 37 °C in glucose medium and transferred to media containing the indicated carbon sources and 10mM NH_4Cl for a further 4 hours. Tracks: 1, wildtype - 1% ethanol; 2, 712 - 1% ethanol; 3, 712 - 0.05% fructose; 4, 712 - 1% ethanol, 1% glucose; 5, 712 - 1% glucose. The identity of the weakly hybridizing transcript of lower molecular weght than the *aldA* transcript in tracks 2 and 4 on the autoradiograph shown opposite is not known.


1
2
3
4
5
aldA
pyrG

strains, designated 216, with the allele being named *pyr*G216, was selected for further analysis. *pyr*G mutations were distinguished from mutations at a second locus, designated *pyr*B in *A. niger* by Goosen *et al.*, (1987), which are also 5-flouro orotic acid resistant, by transformation with the cloned *A. niger pyr*G gene in pAB4.1 (van Hartingsveldt *et al.*, 1987). As positive selection of mutants was available, large numbers of conidia could be plated on plates containing 5-flouro orotic acid to select the mutations, and so no mutagenesis was used. This reduced the risk of strain 712 picking up other undefined mutations which may affect *ald*A expression and so interfere with subsequent analysis of *ald*A expression in *pyr*G/*ald*A cotransformants of strain 216.

Southern analysis of 216 (Figure 5.1.3) indicated that it lacked sequences corresponding to the region that was deleted from pNG100 and replaced with the insert of p3SR2 to form pNG300, and hence this region had been deleted by the integration of pNG300. Strain 216 had *amd*S and pUC19 sequences integrated into its chromosome. pNG300 does not have a site for the restriction endonuclease XhoI. The pNG100 insert, p3SR2 insert and pUC19 hybridized to a single XhoI fragment of greater than 23kb., indicating that all these sequences had integrated at the same site. However, as there were pUC19 sequences present, integration of pNG300 sequences did not occur by a single integration event involving a double crossover between the *ald*A flanking regions in pNG300 and the chromosome. The size of the hybridizing XhoI fragments, and the strength of the hybridization of pUC19 to a 2.8kb. SalI fragment corresponding to the vector sequences of pNG300, indicated that at least two copies of pNG300 had integrated at the *ald*A locus, but not in a straight forward tandem array.

Figure 5.1.3: Southern blot analysis of the *A. niger aldA* gene replacement strain 216. DNA extracted from 216 and wildtype (WT) mycelia was digested with the indicated enzymes prior to electrophoresis in a 1% agarose, 1 X TAE gel. Quadruplicate filters were probed with the indicated probes. Bacteriophage lambda DNA digested with HindIII was used as a marker of molecular weight.


Sal I
EcoRI
Xho I
Sal I + EcoRI
Sal I + Xho I
WT
216
kb.
23·1–
9·4–
6·6–
4·4–
2·3–
2·0–
0·6–
Probe: pNG100 insert (aldA)
Probe: p3SR2 insert (amdS)
Probe: pUC19
Probe: 4 kb. BglII–ClaI fragment of pNG100 (deleted from pNG300)

Attempts were made to form a diploid between strain 216 and the muiltiply marked strains of *A. niger* constructed by Bos *et al.*, (1988) so as to assign the *ald*A locus of *A. niger* to a chromosome. Conventional parasexual analysis techniques and protoplast fusion were used but no stable diploids could be selected. Diploids were sucessfully made between two strains constructed by Bos *et al.*, (1988) and have also been made between derivatives of the wildtype strain used in this study (J. Kelly, pers. comm.). It is likely that the wildtype strains used in this study and by Bos *et al.*, (1988) are of different incompatibility groups. Heterokaryon incompatibility has been observed between different strains of many *Aspergillus species* including *A. niger* (Gossop, *et al.*, 1940; Papa, 1986 and references therein). Hence the *ald*A locus could not be assigned to a linkage group by parasexual analysis.

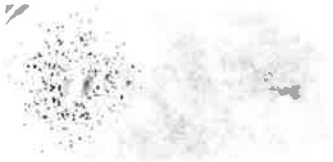
These results showed that strain 216 lacked the *ald*A coding region and flanking regions extending 1.7kb. 5' to the start codon and 0.7kb. 3' to the stop codon, and hence was an *ald*A$^-$ strain for *A. niger*. This strain was used to investigate the regulation of the *A. nidulans* *ald*A gene in *A. niger* (Section 5.3) and for the analysis of the effects of promoter deletions on *ald*A gene expression in *A. niger* (Chapter 6).

5.2: Phenotypes of the gene replacement strain.

Several phenotypes for the *ald*A$^-$ strain were found by plate testing (Figure 5.2.1). Like the strains of *A. nidulans* with the *ald*A67 mutation, the *ald*A$^-$ strain 712 showed no growth on either ethanol or threonine media, and GABA and 2-pyrrolidone were toxic to it. However, toxicity of L-proline could not be clearly identified as it is also

toxic to wildtype *A. niger*. Acetaldehyde added to glucose media was toxic to 712, which is also a phenotype of *A. nidulans aldA*67. 712 showed only slightly inhibited growth on ethyl ammonium chloride as a nitrogen source, which indicated that acetaldehyde was not accumulated to significantly toxic levels under this growth condition. Putrescine, which can theoretically be converted to GABA via 4-amino butyraldehyde, inhibits growth of 712 when added as a source of nitrogen. This effect was not seen for *A. nidulans aldA*67. Whether this pathway of putrescine metabolism exists in *A. niger* is not known, but would explain the toxicity of putrescine to 712. Glycerol or aspartic acid as a carbon source was also toxic to 712, but not to *A. nidulans aldA*67. Glycerol toxicity may be due to its conversion to glyceraldehyde. Aspartic acid can be converted to beta-alanine by the action of a decarboxylase. Beta-alanine is very toxic to *A. nidulans aldA*67 when used as a carbon source, probably due to the production of malonic-semialdehye by the action of the *gatA* encoded omega amino acid transferase. This phenotype is not as marked for *A. niger* strain 712 as beta-alanine is a very poor carbon source for wildtype *A. niger*. Whether these pathways of glycerol and aspartic acid metabolism occur in *A. niger* is not known, but these results show that aldehyde dehydrogenase activity is involved. Strains of *A. nidulans* containing the *aldA*67 are sensitive to allyl alcohol when added to glucose medium, but this phenotype is difficult to gauge for *A. niger* strain 712 as wildtype *A. niger* is as sensitive to allyl alcohol as the *aldA*67 mutant of *A. nidulans*. These phenotypes give a further indication that there are some metabolic differences involving aldehyde dehydrogenase between *A. niger* and *A. nidulans*. As these phenotypes resulted from the *aldA* deletion, it is unlikely that other aldehyde dehydrogenase encoding genes exist in *A. niger*. Pathways of the metabolism of the compounds discussed above are shown in Figure 5.2.2.

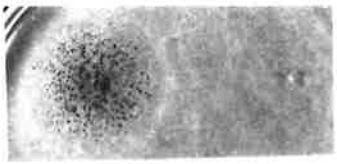
Figure 5.2.1: Phenotypes of the *A. niger aldA* gene replacement strain 712. Plates were incubated at 37 °C for either 2 days, for plates containing glucose, or 5 days, for other carbon sources. Strain 712 is on the right of wildtype *A. niger*.



1% Ethanol
10mM NH_4Cl

50mM GABA
10mM NH_4Cl

1% Glucose
50mM Putrescine

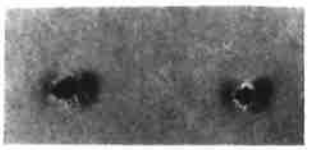


50mM Threonine
10mM NH_4Cl

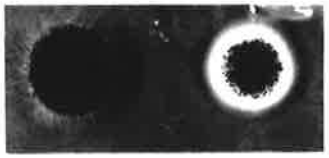
1% Glucose
50mM GABA

1% Glucose
10mM NH_4Cl
50mM Acetaldehyde

1% Glycerol
10mM NH_4Cl

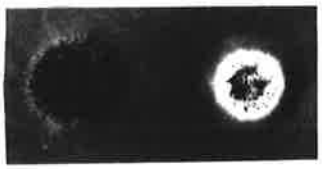


50mM Proline
10mM NH_4Cl



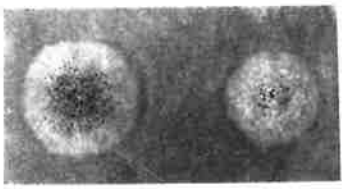
1% Glucose
50mM Ethyl-NH_4Cl

50mM Aspartic acid
10mM NH_4Cl



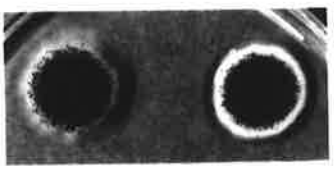
1% Glucose
50mM Proline

1% Glucose



50mM Sodium Acetate
10mM NH_4Cl

50mM 2-pyrrolidone
10mM NH_4Cl

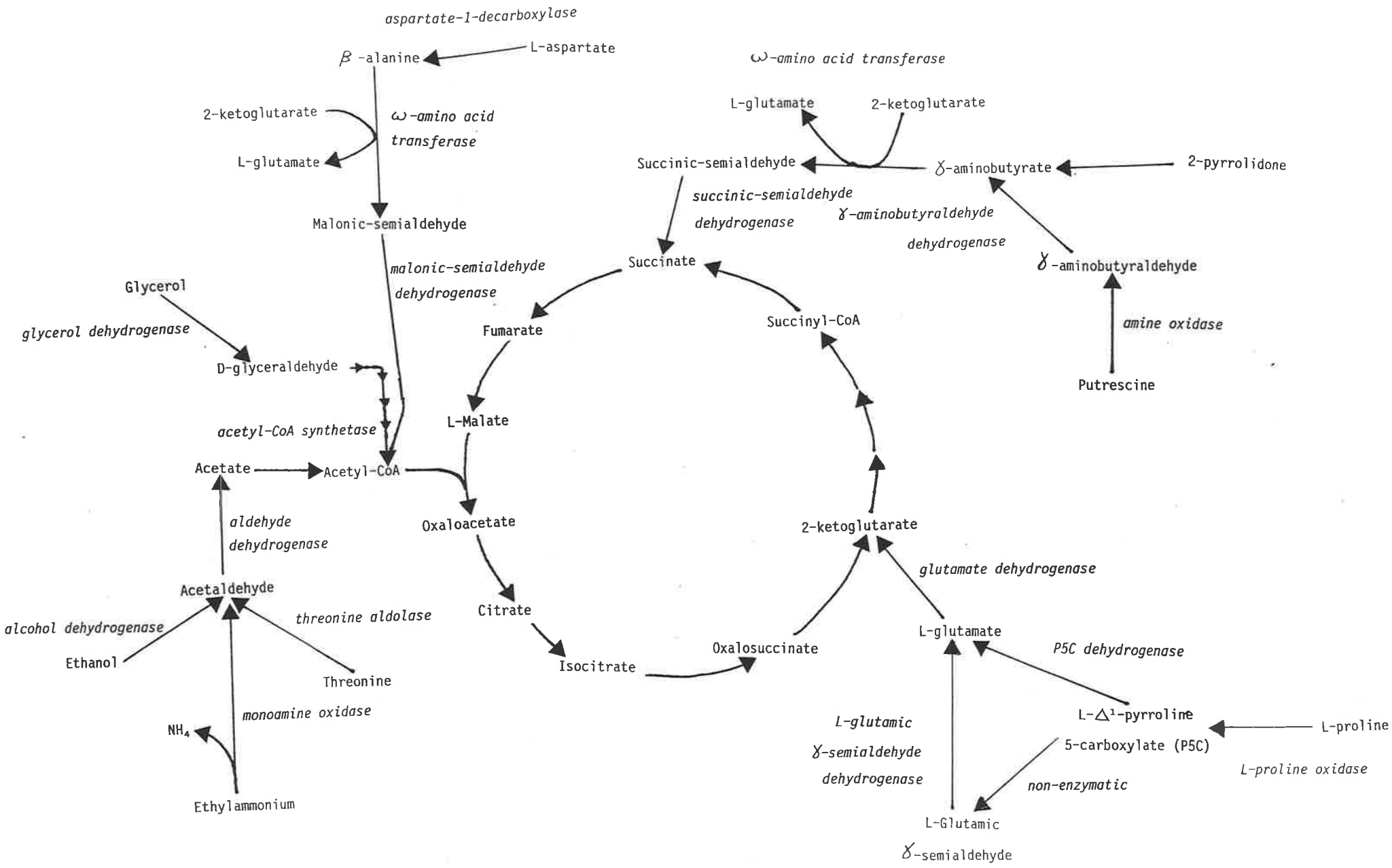


1% Glucose
10mM NH_4Cl

10mM NH_4Cl

1% Glucose
50mM 2-pyrrolidone

Figure 5.2.2: Possible metabolic pathways for the metabolism of compounds discussed in Section 5.2. Enzyme names are given in italics.


aspartate-1-decarboxylase
β -alanine
L-aspartate
ω-amino acid transferase
2-ketoglutarate
L-glutamate
ω-amino acid transferase
L-glutamate
2-ketoglutarate
Succinic-semialdehyde
γ-aminobutyrate
2-pyrrolidone
succinic-semialdehyde dehydrogenase
γ-aminobutyraldehyde dehydrogenase
Malonic-semialdehyde
malonic-semialdehyde dehydrogenase
Succinate
γ-aminobutyraldehyde
Glycerol
glycerol dehydrogenase
Fumarate
Succinyl-CoA
amine oxidase
D-glyceraldehyde
Putrescine
acetyl-CoA synthetase
L-Malate
Acetate
Acetyl-CoA
Oxaloacetate
2-ketoglutarate
aldehyde dehydrogenase
glutamate dehydrogenase
Acetaldehyde
Citrate
alcohol dehydrogenase
threonine aldolase
Oxalosuccinate
L-glutamate
Ethanol
Threonine
Isocitrate
P5C dehydrogenase
L-Δ¹-pyrroline
5-carboxylate (P5C)
L-proline
monoamine oxidase
L-glutamic γ-semialdehyde dehydrogenase
L-proline oxidase
NH4
non-enzymatic
Ethylammonium
L-Glutamic γ-semialdehyde

5.3: Expression of the *A. nidulans aldA* gene in *A. niger*.

The cloned *A. nidulans aldA* gene in plasmid pAN212 was cotransformed into *A. niger* strain 216 with pAB4.1 as the selectable marker, with transformants being selected by growth in the absence of uridine. Several PyrG$^+$ transformants were screened by Southern hybridization for the presence of pAN212. Two of these transformants, designated T212.6 and T212.7, with estimated pAN212 copy numbers as determined by Southern and dot blot analysis of 10 and 2 respectively were chosen for further analysis (Figure 5.3.1). Expression of the *A. nidulans aldA* gene in these transformants restored growth on media containing either ethanol or threonine as the sole source of carbon (Figure 5.3.2), which in the case of the higher copy number transformant T212.6, was stronger than wildtype *A. niger*. The phenotype of this transformant correlates with the observed high levels of expression of the *A. nidulans aldA* gene (see below). Similar increases in threonine and ethanol growth were seen in all heterologous and homologous multicopy *aldA* transformants of *A. niger* and *A. nidulans*, and this is in keeping with the findings of Creaser *et al.*, (1987) who showed that in *A. nidulans*, aldehyde dehydrogenase activity was the rate limiting step in the metabolism of these compounds.

Analysis of *A. nidulans aldA* mRNA levels in these transformants (Table 5.3.1) showed that, as in wildtype *A. nidulans*, relatively high levels of expression were seen in fructose grown cultures and that growth in the presence of 1% ethanol alone does not significantly induce *aldA* expression. However, unlike wildtype *A. nidulans*, no increase in expression level results by the addition of 0.05% fructose to 1% ethanol. Induction of *aldA* during fructose growth in *A. nidulans*

Figure 5.3.1: Southern blot analysis of *A. niger* strain 216 PyrG$^+$ cotransformants containing pAN212. DNA extracted from T212.6 and T212.7 mycelia was digested with EcoRI, which cuts pAN212 once, prior to electrophoresis in a 0.7% agarose, 1 X TAE gel and transfer to Zeta Probe membrane. The filter was probed with the insert of pAN212 and the autoradiograph is shown opposite. Bacteriophage lambda DNA digested with HindIII was used as a marker of molecular weight.

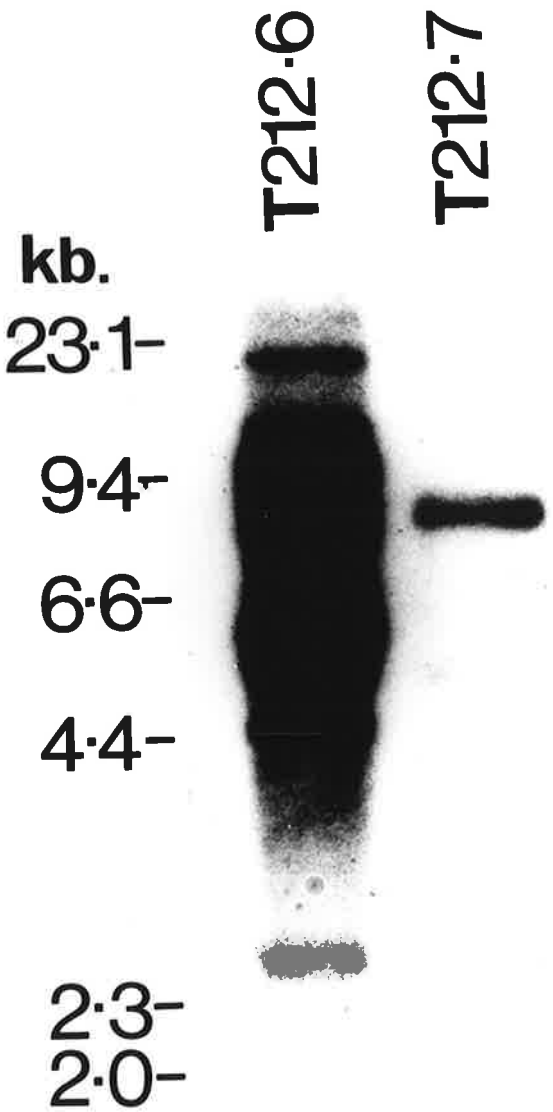
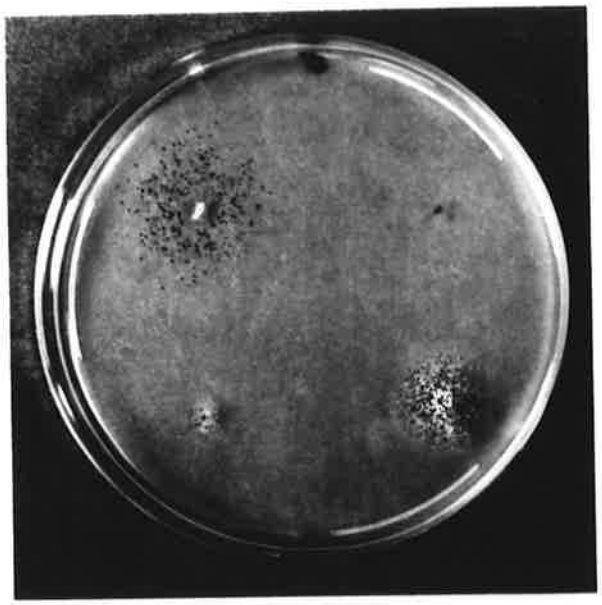
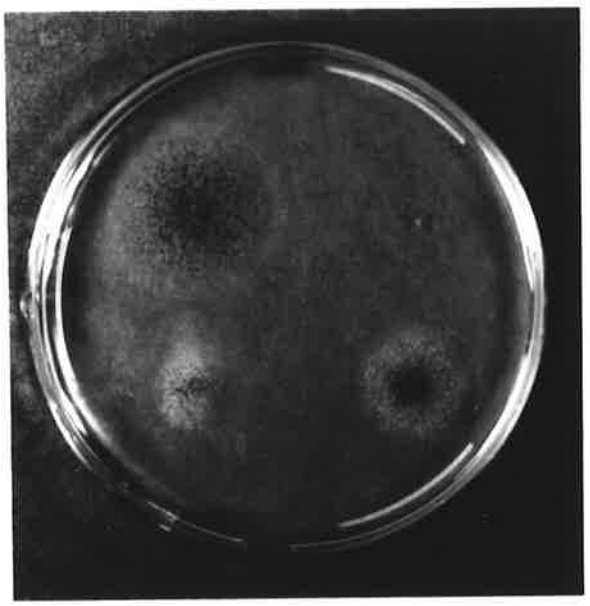

T212·6
T212·7
kb.
23·1–
9·4–
6·6–
4·4–
2·3–
2·0–

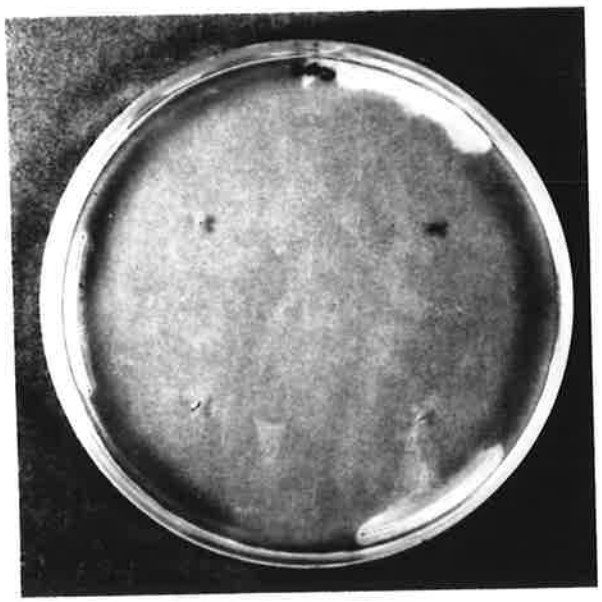
Figure 5.3.2: Phenotypes of PyrG$^+$ cotransformants of *A. niger* strain 216 containing pAN212. Plates were incubated at 37 °C for either 2 days, for the glucose plate, or 5 days, for the plates containing the other carbon sources. Strains are (from top left): wildtype, 712, T212.7, T212.6.



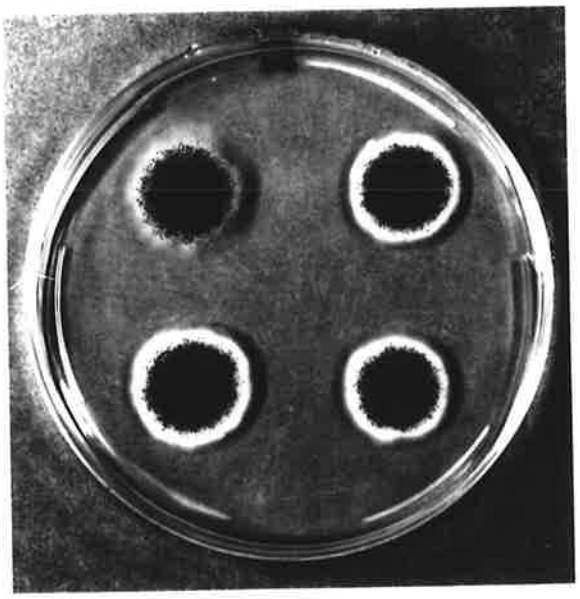
1% Ethanol, 10mM NH_4Cl



50mM Threonine, 10mM NH_4Cl



10mM NH_4Cl



1% Glucose, 10mM NH_4Cl

Growth Condition (+10mM Ammonium Chloride)	T212.6	T212.7	*A. niger*	*A. nidulans*
1% Ethanol	2	2.5	8	2
1% Ethanol, 0.05% Fructose	1.5	2	7	23
1% Ethanol, 1% Glucose	0.5	1	1.5	1
50mM Threonine	1	2.5	3	23.5
50mM Threonine, 1% Glucose	1	2.5	2.5	1
0.05% Fructose	10	5	2	8.5
1% Glucose*	1	1	1	1

Table 5.3.1: Levels of *A. nidulans aldA* mRNA in pAN212 transformants of *A. niger* strain 216. *Levels are normalized against those for glucose grown mycelia for each strain. The *aldA* mRNA levels in T212.6 are 6.5 fold that of T212.7. Cultures were grown overnight at 37 °C in glucose medium, harvested, washed, and transferred with the indicated growth condition for 4 hours as described in Section 2.6.1.

is likely to be controlled by *amdR* by a separate induction mechanism to that controlled by *alcR* (Chapter 4). As induction during growth on 0.05% fructose occurs in both species but only on the *A. nidulans* gene, then factors controlling the mechanism of induction during fructose growth must exist in *A. niger*, but fail to act on its own *aldA* gene. A promoter element responsive to induction during fructose growth may be present within the 5' region of the *A. nidulans aldA* gene, but not 5' to the *A. niger aldA* gene. However during the analysis of some 5' deletions within the *A. niger aldA* promoter (Chapter 6), one deletion was found to result in increased expression during growth on 0.05% fructose and hence an element responding to fructose induction may be present 5' to the *A. niger aldA* gene but be masked by a silencing sequence which renders it cryptic.

Growth of T212.7 in the presence of threonine led to induced levels of *ald*A expression, and like wildtype strains of *A. niger*, this was unaffected by the addition of 1% glucose. As it is unlikely that threonine itself can induce the expression of *ald*A expression in *A. niger*, since it fails to in *A. nidulans* (Pateman *et al.*, 1983; M. Hynes, pers. comm.), acetaldehyde produced from threonine metabolism, regardless of the presence of glucose, leads to the induction of *ald*A expression . These results, as discussed in Chapter 4, imply that in *A. niger* during growth on threonine in the presence of glucose, carbon catabolite repression does not act directly on *ald*A. This effect is specific to *A. niger*, but is independent of whether the *A. niger* or *A. nidulans ald*A gene is being expressed, which implies that in *A. niger*, induction of *ald*A expression occurs whenever significant levels of the intracellular inducer are present, and hence carbon catabolite repression may act in other growth conditions by preventing the formation of the intracellular inducer from other carbon sources. That is, it is differences in the expression of genes other than *ald*A during threonine metabolism that leads to this difference in *ald*A gene regulation between the two species. In addition to structural genes involved in threonine metabolism, if the product of a positively acting regulatory gene is required for *ald*A induction in *A. niger*, then during growth in threonine medium, the expression of this gene is also not directly subject to carbon catabolite repression. The lack of this effect in T212.6 may be an artefact of its higher basal levels of expression which are 6.5 fold that of T212.7, in that the basal level of expression in T212.6 may be sufficient to detoxify acetaldehyde produced from threonine metabolism, and so in this transformant, the need to induce *ald*A expression does not exist. Induction in the presence of 5mM acetaldehyde is subject to carbon catabolite repression

and thus it is possible that glucose affects the uptake of exogenously added acetaldehyde into the cell, as it is likely that it is the acetaldehyde produced from threonine metabolism that leads to the carbon catabolite repression insensitive induction of *ald*A in threonine plus glucose medium. Allyl alcohol is metabolized to the toxic compound acrolein by alcohol dehydrogenase, and has been used to select mutations in alcohol dehydrogenase genes in many species. *A. nidulans* is resistant to allyl alcohol when added to glucose medium due to carbon catabolite repression of the *alc*A gene. In contrast, *A. niger* is sensitive to allyl alcohol on glucose medium and thus an alcohol dehydrogenase is expressed in *A. niger* in the presence of glucose. Although this could also produce acetaldehyde in the presence of glucose, it is not known whether this enzyme has ethanol as a substrate *in vivo*, and so may not affect *ald*A expression.

Regulatory proteins in *A. niger* must recognise elements present in the promoter of the *A. nidulans ald*A gene and result in controlling the observed regulated expression. There is probably a protein homologous to the *alc*R gene product of *A. nidulans* (see Section 5.4), and separate mechanisms of induction involving a homologue of the *amd*R gene of *A. nidulans* affecting fructose and threonine induction may also exist in *A. niger*.

5.4: Expression of the *A. niger ald*A gene in *A. nidulans*.

A strain of *A. nidulans* carrying the *ald*A67 mutation was cotransformed with the cloned *A. niger ald*A gene in pNG100 and pANA19, which carries the *A. nidulans amd*S gene. Transformants were selected by increased growth in the presence of CsCl on media containing acetamide as the sole source of nitrogen. Approximately 50% of the AmdS$^+$

transformants showed an AldA$^+$ phenotype in that they grew on media containing ethanol or threonine as the sole source of carbon, and were resistant to allyl alcohol in glucose medium. Several transformants were screened by Southern and dot blot analysis to determine copy numbers of integrated pNG100 sequences. Two of these transformants, designated T1/H and T1/L, with pNG100 copy numbers of 15 and 3 respectively, were chosen for further analysis (Figure 5.4.1). Growth properties on relevant media of these transformants are shown in Figure 5.4.2.

RNA was prepared from T1/H and T1/L mycelia grown under a variety of conditions and the relative levels of *A. niger aldA* mRNA in these samples was determined (Table 5.4.1). The expression of the *A. niger aldA* gene in *A. nidulans* is induced in the presence of ethanol and repressed in the presence of glucose. The level of expression in glucose grown T1/H mycelia was approximately five fold that of T1/L, and thus there was an approximately linear relationship between pNG100 copy number and *A. niger aldA* expression levels. This lower but easily detectable level of *A. niger aldA* mRNA in T1/L correlated with its poor growth on ethanol and threonine media. There was at least partial complementation of the *aldA67* mutation in this transformant as gauged by growth, albeit weaker than T1/H, on glucose plus allyl alcohol medium. T1/L had the lowest pNG100 copy number of all AmdS$^+$ transformants with an AldA$^+$ phenotype tested. Whether inefficiency of gene expression or lower enzyme activity can explain the requirement for relatively high pNG100 copy number to give an AldA$^+$ is not known. It is also possible given the differences in amino acid sequence of the region encoded by the first exons of the two genes that inefficient subcellular localization of the *A. niger* AldDH may occur in *A. nidulans* as this region is thought to be important for targeting of enzyme

Figure 5.4.1: southern blot analysis of AmdS$^+$ cotransformants containing pNG100. Each track contains 2ug of DNA digested with the indicated enzymes prior to electrophoresis in a 0.7% agarose, 1 X TAE gel and transfer to Zeta Probe membrane. The probe used was the insert of pNG100. Tracks: 1, *A. niger* - SalI; 2, *A. nidulans* - SalI; 3, *A. niger* - ClaI; 4, *A. nidulans* - ClaI; 5, *A. nidulans aldA67* - ClaI; 6, T1/H - ClaI; 7, T1/L - ClaI; 8, T2/L - ClaI; 9, T2/H - ClaI; 10, T1/H2.5 - ClaI. The arrow in track 3 of the autoradiograph shown opposite indicates a faint band of hybridization. Under these hybridization conditions (Section 2.6.4), longer autoradiographic exposure is required to show hybridization of the pNG100 insert to *A. nidulans* DNAs. ClaI cuts pNG100 once. Bacteriophage lambda DNA digested with HindIII was used as a marker of molecular weight.

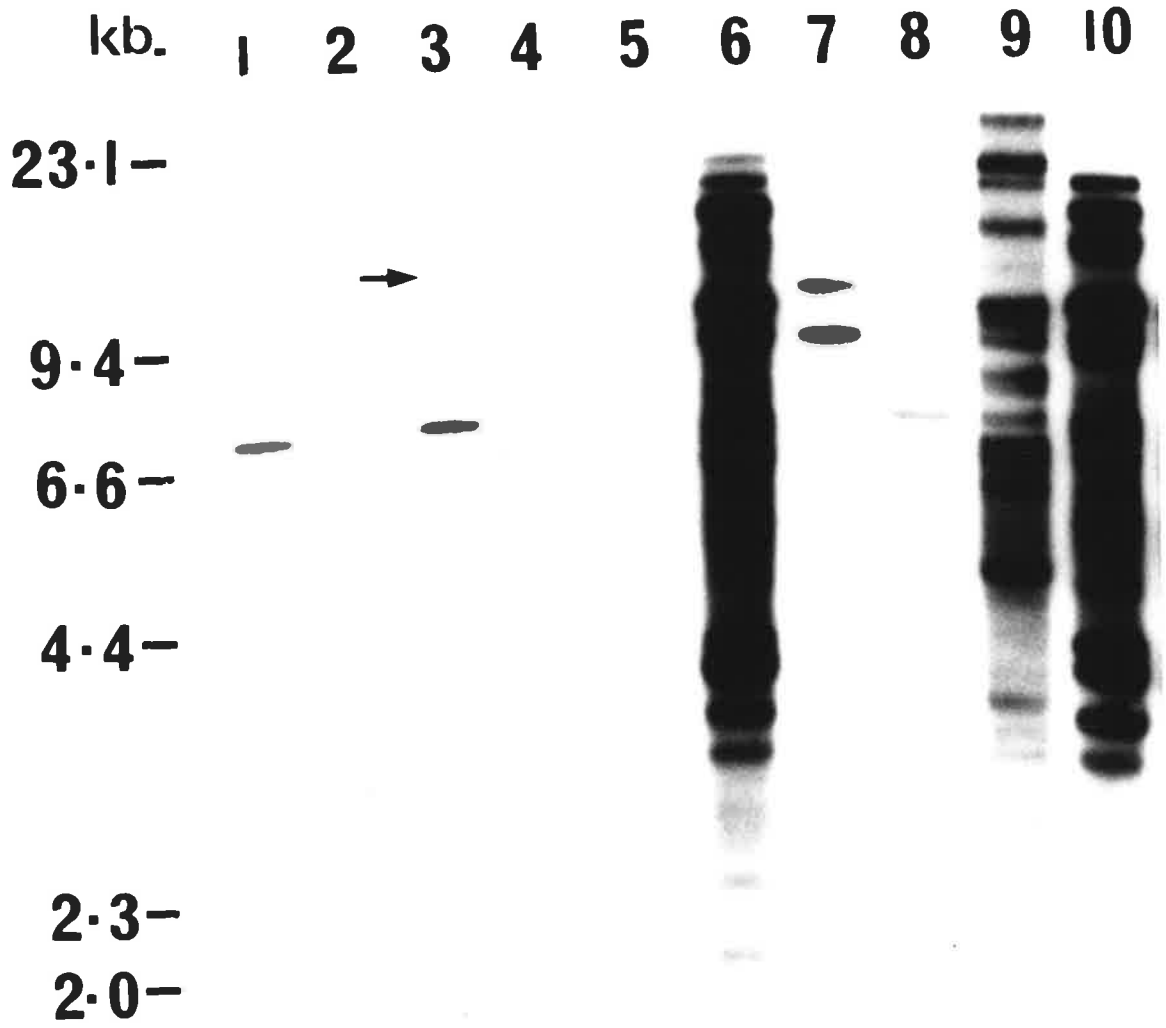
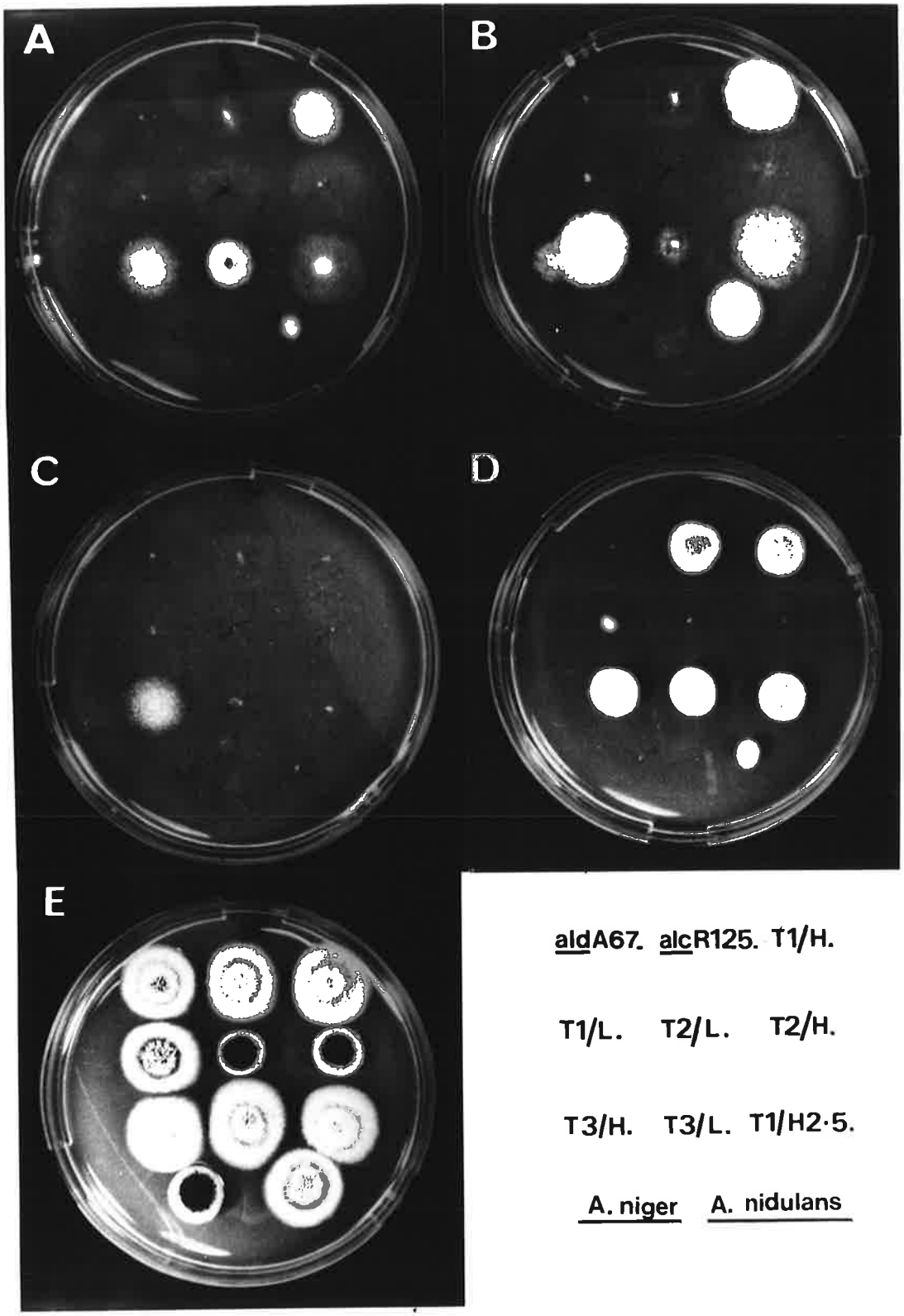

kb.
1
2
3
4
5
6
7
8
9
10
23·1
9·4
6·6
4·4
2·3
2·0

Figure 5.4.2: Growth properties of strains discussed in Sections 5.4 and 5.5. Media: A: 1% ethanol, 10mM NH_4Cl; B: 50mM threonine, 10mM NH_4Cl; C: 10mM NH_4Cl; D: 1% glucose, 50mM allyl alcohol, 10mM NH_4Cl; E: 1% glucose, 10mM NH_4Cl. The plates were incubated at 37 °C for 2 days. The positions of each strain are indicated.



alcA67. alcR125. T1/H.

T1/L. T2/L. T2/H.

T3/H. T3/L. T1/H2·5.

A. niger A. nidulans

location in mammals (Hempel *et al.*, 1985). It is also possible that the *ald*A sequences had integrated in such a way as to interupt the coding or promoter regions preventing the production of active enzyme. However, northern analysis of these transformants only detected full length *ald*A mRNA molecules. Further experiments using hybrids of the *A. nidulans* and *A. niger ald*A coding regions in *A. nidulans* may resolve this. These effects were not seen in *A. niger* transformants of *A. nidulans ald*A.

Growth Condition (+ 10mM Ammonium Chloride)	T1/H	T1/L	T1/H2.5
1% Ethanol	11	3	6.5
0.05% Fructose	11	1.5	3
1% Ethanol, 1% Glucose	4.5	1.5	6.5
1% Glucose	5	1*	6.5

Table 5.4.1: Levels of *A. niger ald*A mRNA in pNG100 transformants of *A. nidulans ald*A67. *Levels have been normalized against those for glucose grown T1/L mycelia. Cultures were grown overnight at 37 °C in glucose medium, harvested, washed, and transferred to media with the indicated growth conditions for a further 4 hours as indicated in Section 2.6.1.

The higher copy number transformant T1/H was crossed to a strain carrying the *alc*R125 mutation to test whether the expression of *A. niger ald*A required a functional *alc*R gene product. Progeny of the cross were screened by plate testing for resistance to allyl alcohol added to glycerol medium to a final concentration of 2.5mM. This phenotype indicates the lack of ADHI activity due to the lack of *alc*A

gene expression, which is caused by the absence of a functional *alc*R gene product in strains with the *alc*R125 allele. Progeny showing the phenotype of the *alc*R125 strain were screened by plate testing on media containing acetamide and acrylamide as sole sources of nitrogen to indicate whether they contained the *amd*S sequences used to select the T1/H transformant. *alc*R125 strains showing similar growth characteristics to T1/H on acetamide and acrylamide as a nitrogen source were screened by Southern hybridization to detect the presence of the pNG100 sequences (Figure 5.4.1). One of these progeny, designated T1/H2.5, had a very similar pattern of bands hybridizing to the insert of pNG100 as T1/H, and so had not suffered major loss or rearrangement of pNG100 sequences during the cross. The T1/H transformant was used for further analysis.

The induced levels of *A. niger ald*A expression in T1/H were absent in T1/H2.5 (Table 5.4.1). T1/H2.5 also showed poorer growth on ethanol and threonine as carbon sources (Figure 5.4.2). The reduction in growth and induced expression in a derivative of T1/H containing a mutation at the *alc*R locus indicated that a functional *alc*R gene product was required to induce the expression of the *A. niger ald*A gene in *A. nidulans*. There is evidence to suggest that *A. niger* has a homologous gene to *alc*R. At low stringency, the cloned *alc*R gene hybridizes to Southern blots of *A. niger* DNA. Furthermore, in an attempt to gain a clone of a homologue of *alc*R from *A. niger*, the *alc*R5 mutation of *A. nidulans* (Pateman *et al.*, 1983), which is non-revertible, was complemented by sequences from an *A. niger* gene bank constructed in pM006 using the *arg*B gene as a selectable marker. Attempts to rescue this sequence were unsuccessful, although many plasmids, nearly all containing *arg*B sequences, were isolated from this transformant, and thus it is likely that the complementing sequences had integrated in

such a way that the putative *alc*R sequence and the ampicillin resistance gene used to select rescued plasmids in *E. coli* could not easily be isolated from the chromosome. It is interesting to note in light of the insensitivity to carbon catabolite repression of threonine induction in *A. niger*, that although this transformant had an AlcR$^+$ phenotype in terms of growth on ethanol and threonine as carbon sources, it was sensitive to allyl alcohol added to either glucose or glycerol medium. Thus, this transformant may have contained an *alc*R homologue from *A. niger* that was not subject to carbon catabolite repression in *A. nidulans*, and in the presence of glucose could direct expression levels of *alc*A that were sufficent to produce toxic levels of acrolein. Alternative hypotheses to explain the phenotype of this transformant could be that it was transformed with a suppressor of *alc*R14 or a constitutively expressed alcohol dehydrogenase gene. No further analysis was performed on this transformant.

5.5: The effects of an increase in *ald*A gene copy number in *A. niger* and *A. nidulans*.

Wildtype *A. niger* was cotransformed with pNG100 and pANA19, with transformants being selected by growth on media containing acetamide as the sole source of nitrogen. Transformants were screened by plate testing for altered phenotypes on relevant media, but no AmdS$^+$ transformants were shown to exhibit increased growth on media containing either ethanol or threonine as the sole carbon source (Figure 5.4.2). Southern analysis of several transformants confirmed however that most AmdS$^+$ transformants had extra copies of pNG100 inserted in their chromosomes. Two transformants used in this Southern blot analysis (Figure 5.4.1) designated T2/L and T2/H, with *ald*A copy numbers of 2 and 7 respectively, were chosen for further analysis.

Analysis of *ald*A mRNA levels in these transformants (Table 5.5.1) showed that they had altered *ald*A expression, but this did not correlate with gene copy number. The expression levels were in some cases lower than wildtype. This is in keeping with the unaltered growth characteristics of these transformants (Figure 5.4.2). This was in contrast to similar experiments in which the *ald*A gene copy number in *A. niger* was increased by cotransformation into a *pyr*G$^-$ strain with pAB4.1, in which levels of *ald*A expression were elevated but still regulated in a similar fashion to wildtype (Chapter 6). Thus, either regulatory molecules required for *ald*A expression are titrated by the *amd*S sequences or acetamidase activity interfered with induction. This is an interesting result in light of the experiments of Kelly and Hynes, (1985) which showed the presence of an *amd*R like induction mechanism for the expression of the *A. nidulans amd*S gene in transformants of *A. niger*. A homologue of the *amd*R gene in *A. niger*

Growth condition (+ 10mM Ammonium Chloride)	T2/L	T2/H	Wildtype
1% Ethanol	11	4	8
0.05% Fructose	0.5	12	2
1% Ethanol, 1% Glucose	0.2	1.5	1.5
1% Glucose	0.2	1	1*

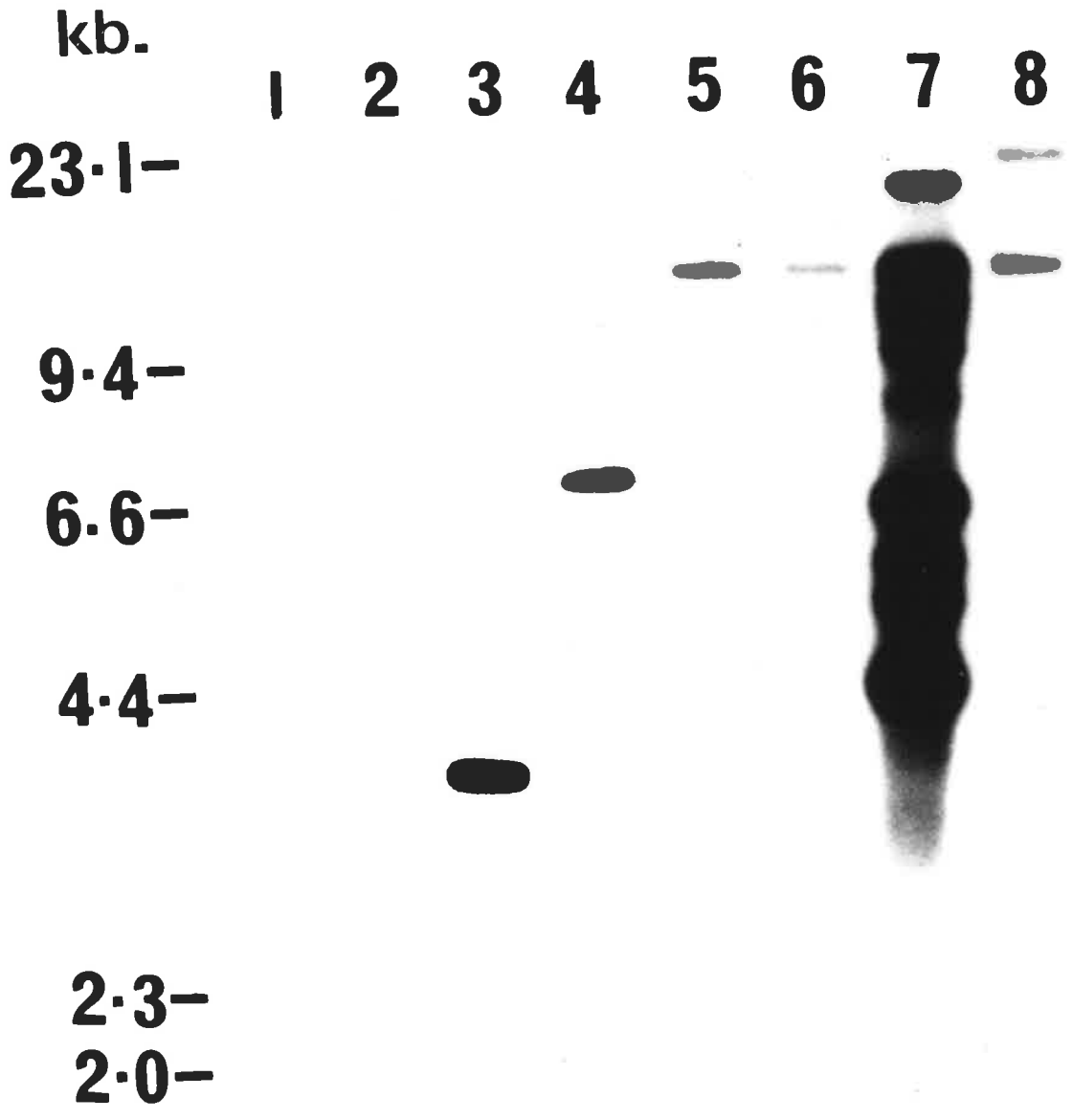
Table 5.5.1: Levels of *ald*A mRNA levels in pNG100 transformants of *A. niger*. *Levels are normalized against those for glucose grown wildtype mycelia. Cultures were grown overnight at 37 °C in glucose media, harvested, washed and transferred into media with the indicated growth condition as described in Section 2.6.1.

may, as in *A. nidulans*, play a role in *ald*A expression. As the sites of action of regulatory molecules controlling *amd*S expression have been identified, further analysis using the various 5' constructs of the *amd*S gene used to determine these sites of regulatory control by Hynes *et al.*, (1988) may identify which type of molecule, if any, is being titrated by the *A. nidulans amd*S sequences.

The *ald*A gene copy number was also increased in *A. nidulans* by cotransformation of the cloned *A. nidulans ald*A and *amd*S genes. This transformation was into a strain carrying the *ald*A67 mutation. AmdS$^+$ transformants were screened by plate testing for complementation of *ald*A67. Several AldA$^+$ transformants were subjected to Southern and dot blot hybridization analysis to determine pAN212 copy number. Several of these transformants showed significant growth on media lacking an added carbon source. Payton *et al.*, (1976) showed that the weak growth of wildtype *A. nidulans* on media lacking an added carbon source required C_2 metabolism via acetate. Wildtype *A. niger* does not show any growth on media lacking an added source of carbon. The phenotype of growth on carbon free media of some of the pAN212 transformants shows that this growth is also via acetaldehyde. This phenotype was the same on plates solidified with electrophoresis grade agarose. Two transformants designated T3/L and T3/H, with pAN212 copy numbers of 1 and 6 respectively, were chosen from the Southern analysis (Figure 5.5.1) for further analysis. Growth characteristics of these transformants are shown on Figure 5.4.2.

Analysis of *ald*A mRNA levels in these transformants (Table 5.5.2) showed that, unlike the *A. niger* pANA19/pNG100 cotransformants, an increase in gene copy number led to an increase in mRNA levels, and so factors required for induced levels of *ald*A expression were not

Figure 5.5.1: southern blot analysis of AmdS$^+$ cotransformants of *A. nidulans aldA67* containing pAN212. Each track contains 2ug of DNA digested with the indicated enzymes prior to electrophoresis in a 0.7% agarose, 1 X TAE gel and transfer to Zeta Probe membrane. The probe used was the insert of pAN212. Tracks: 1, *A. niger* - SalI; 2, *A. niger* - ClaI; 3, *A. nidulans* - SalI; 4, *A. nidulans* - ClaI; 5, *A. nidulans* - EcoRI; 6, *A. nidulans aldA67* - EcoRI; 7, T3/H - EcoRI; 8, T3/L - EcoRI. Under these hybridization conditions (Section 2.6.4), longer autoradiographic exposure is required to show hybridization of the pAN212 insert to *A. niger* DNA. EcoRI cuts pAN212 once. Bacteriophage lambda DNA digested with HindIII was used as a marker of molecular weight.


kb.
1
2
3
4
5
6
7
8
23·1
9·4
6·6
4·4
2·3
2·0

limiting in these transformants. The levels of expression were not, however, correlated with gene copy number. T3/L has only one functional *ald*A gene but shows very high levels of expression. The site of integration of this single copy of pAN212 in T3/L must in some way affect its expression. For example, it may have integrated near a general enhancer sequence. It is also possible that the transforming *ald*A sequence in T3/L integrated into a region of the genome where chromatin structure allows easier access to the *ald*A promoter for RNA polymerase than the native *ald*A locus. The levels of *ald*A mRNA in the *ald*A67 mutant were only slightly different to wildtype, indicating that aldehyde dehydrogenase enzyme activity does not significantly affect *ald*A expression.

Growth Condition (+10mM Ammonium Chloride)	T3/L	T3/H	*ald*A67	Wildtype
1% Ethanol	87.5	37	5.5	2
0.05% Fructose	19.5	33.5	5	8.5
1% Ethanol, 1% Glucose	5.5	26	0.5	1
1% Glucose	7.0	17.5	0.5	1*

Table 5.5.2: Levels of *ald*A mRNA in pAN212 transformants of *A. nidulans ald*A67. *Levels have been normalized against those for glucose grown wildtype mycelia. Cultures were grown overnight at 37 °C in glucose medium, harvested, washed and transferred to media with the indicated growth condition for a further 4 hours as describes in Section 2.6.1.

5.6: Conclusions.

The construction of the *A. niger aldA* deletion strain by gene replacement has allowed the investigation of both metabolic pathways requiring aldehyde dehydrogenase activity in *A. niger*, and the expression of *A. nidulans aldA* gene in *A. niger* without any affects being caused by a resident functional aldehyde dehydrogenase gene. Phenotypes of this deletion mutant show that in addition to differences in *aldA* gene expression between *A. niger* and *A. nidulans*, there are differences in the requirements for the aldehdye dehydrogenase enzyme for growth on various media.

The expression of the *aldA* genes of *A. niger* and *A. nidulans* in homologous and heterologous transformants was investigated to determine whether induction mechanisms acting in each species operate independently of which *aldA* promoter sequence is present. These experiments also allowed the investigation of the effects of increased *aldA* gene copy number on *aldA* expression.

Expression of the *A. niger aldA* gene in *A. nidulans* showed that the *alcR* gene product is capable of regulating the *A. niger aldA* gene, and gives evidence that a similar induction mechanism operates in *A. niger*. *A. niger* has DNA sequences that hybridize to the cloned *alcR* gene, and a sequence capable of complementing an *alcR* mutation of *A. nidulans*. This putative gene may play a role in *aldA* gene regulation in *A. niger*. The control of gene expression by regulatory genes of other species has been reported previously in *Aspergillus*. Although *A. niger* lacks an acetamidase encoding gene, *A. niger* strains transformed with the *amdS* gene of *A. nidulans* show growth properties that indicate the

presence of a gene homologous to the *amd*R gene of *A. nidulans* that regulates *amd*S expression (Kelly and Hynes, 1985). Loss of function mutations at the *are*A locus in *A. nidulans* are complemented by the *nit-2* gene of *N. crassa* (Davis and Hynes, 1987). Is is noteworthy that although the *A. niger* and *A. nidulans ald*A genes are regulated when expressed heterologously in transformants, there are no significant regions of nucleotide sequence similarity specific to the *A. niger ald*A gene and genes of the ethanol utilization regulon of *A. nidulans* that may indicate an *alc*R binding site, nor are there significant regions of nucleotide sequence similarity of greater than 5bp. in length between the *A. nidulans alc*A, *ald*A and *alc*R genes that may be involved in *alc*R regulation of these genes (Felenbok, *et al.*, 1988). These observations indicate that if the *alc*R gene product is a DNA binding protein, then it must be able to bind to sites that are heterogeneous or small in their specificity.

Studies of the effects of increased *ald*A gene copy number in *A. niger* and *A. nidulans* (including data from Chapter 6) showed that although expression levels are affected by the position of transforming sequences, there is an increase in expression with an increase in gene copy number. Lockington *et al.*, (1987) showed that increased levels of *alc*R expression only give increased levels of *alc*A and *ald*A gene expression in the presence of a coinducer. The increase in expression in the higher copy number *ald*A transformants shows that the factors required for induction are not limiting. Furthermore, although gene position may be at least in part involved, there is an increase in basal levels of expression, presumably without increased levels of *alc*R gene product, and thus basal expression is independent of *alc*R function. That is, in multicopy *ald*A transformants, there is an increase in the basal level of *ald*A expression that could be explained,

at least in part, by the additive effects of wildtype levels of basal expression from each gene. The reliance on coinducer for induction of the structural genes, given that the regulatory gene product is not limiting, implies that the varying levels of gene expression in different growth conditions reflect varying levels of coinducer formation.

The study of both heterologous and homologous *ald*A gene expression in transformants has been fruitful in gaining regulatory data. Expression of the *A. nidulans ald*A gene in *A. niger* gave further evidence that the metabolism of threonine to acetaldehyde in *A. niger* is not subject to carbon catabolite repression. This result showed that carbon catabolite repression does not act directly on *ald*A, or genes encoding positively acting regulatory proteins required for *ald*A induction, to affect *ald*A expression in this growth condition. It is likely that carbon catabolite repression insensitive expression of genes other than *ald*A involved in threonine metabolism in *A. niger* but not in *A. nidulans* result in this difference in *ald*A regulation between the two species. This does not exclude the possibility of a direct interaction of carbon catabolite repression on *ald*A under other growth conditions. In addition to this, the induction in the presence of fructose of the *A. nidulans ald*A gene in *A. niger* indicates that this induction mechanism exists in *A. niger* but fails to operate on its own *ald*A gene. This point is discussed further in the next chapter.

Chapter 6:

Analysis of the 5' region of the *Aspergillus niger aldA* gene.

This chapter describes experiments performed with the aim of determining which regions of the *A. niger aldA* promoter are required to control the regulated expression of this gene. This has involved a reverse genetic approach. A number of plasmids containing deletions within the *aldA* promoter were constructed, and the regulatory effects of these deletions were determined *in vivo*. Also discussed in this chapter are results of computer comparisons between the nucleotide sequences of the *A. niger aldA* promoter and other fungal gene promoters in an attempt to define common sequences that may be involved in the control gene expression.

6.1: Similar sequences within the *Aspergillus niger aldA* promoter and other fungal genes.

Studies of the expression of the *A. niger aldA* gene in *A. nidulans* (Chapter 5) showed that its expression in this heterologous system is controlled, at least in part, by the *alcR* gene product. As the product of the *alcR* gene also controls the expression of *alcA*, *aldA* and *alcR* itself, a comparison of the 5' regions of these genes was performed to look for common sequences that may play a role in this *alcR* mediated control. A region of the *A. niger aldA* promoter spanning from nucleotides -204 to -178 (where the first base of the start codon is

+1), which contains a number of direct and invert repeats, is similar to sequences in the 5' regions of *alc*A (Gwynne *et al.*, 1987), *ald*A (Pickett *et al.*, 1987) and *alc*R (Felenbok *et al.*, 1988). The regions of similarity between the *alc*A and *ald*A genes was identified as two regions of similar sequence (called regions D and E) by Gwynne *et al.*, (1987). However, further nucleotide sequence comparisons to the GenBank database and other published and unpublished sequences revealed that similar sequences also existed in the 5' regions of several *A. nidulans* genes. No other regions of sequence similarity between this region of the *A. niger ald*A promoter and the 5' regions of genes whose sequences are held on the plant and fungal section of the September 1989 GenBank database were detected. All of these genes from *A. nidulans* showing the common promoter sequence are subject to carbon catabolite repression and have derepressed levels of expression in strains carrying *cre*A mutations, except of the *trp*C gene which is not known to be under *cre*A control and encodes a trifunctional polypeptide involved in tryptophan biosynthesis These sequence alignments and the arrangement of repeats within the *A. niger* sequence are shown in Figure 6.1.1. It may be significant that this region is rich in repeated sequences, as direct and inverted repeats are often found in recognition sequences for DNA binding proteins.

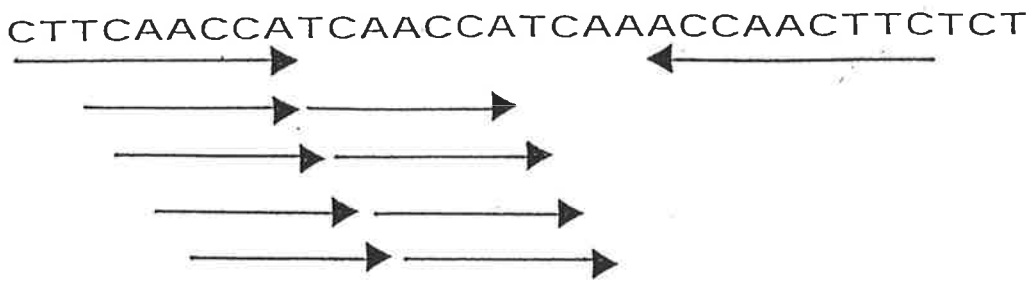
As greatest identity of nucleotides was seen in the region of the sequence ATCAACC, further GenBank database searches and sequence comparisons were performed using this sequence. In this analysis, only sequences 5' to the start codon and having at least 6/7 nucleotides in common to ATCAACC were recorded (Figure 6.1.2). In addition to the exact matches found for *alc*A (Pickett *et al.*, 1987), *alc*R (Felenbok *et al.*, 1988), the isocitrate lyase encoding *acu*D (G. Turner, pers. comm.) and malate synthase encoding *acu*E (R. Sanderman and M. Hynes, pers.

Figure 6.1.1: Nucleotide sequence similarities in *Aspergillus* promoters.

a) A sequence from the *A. niger aldA* promoter region (nucleotides -209 to -178) which has sequence similarity with promoter sequences from several *Aspergillus* genes. The arrows indicate direct and inverted repeat sequences within this region. This sequence is boxed on the complete *A. niger aldA* nucleotide sequence on Figure 3.4.2.

b) Alignment of nucleotide sequences from *Aspergillus* promoter regions showing similarity to the region shown above. Numbering of nucleotides begins with the first base of the start codon as +1 (0 is not used). The position of sequences similar to the TATA box consensus are shown. Hyphens (-) indicate spaces inserted in the sequence to produce the best alignment. An asterisk (*) below a nucleotide indicates identity of that nucleotide with that in the corresponding sequence from the *A. niger aldA* promoter. Sources of sequence data are indicated in the text.

a)



b)

		TATA	MATCHES
A. niger aldA	−204 ACCATCAACCATCAAACCA--ACTTCTCT −178	−102	
A. nidulans alcA	−73 ATCATCAACCAACAATCAA-CAGTTCTCT −46 * ******** *** * * * *****	−110	22/28
A. nidulans aldA	−73 GTCCTC--CCACCGAGCCAGCATTTCTCC −47 * ** *** * * *** * *****	−88	17/29
A. nidulans alcR (1)	−404 GAGATGACCCGTCAAAGGA--TCTTCATT −377 ** * ** ***** * **** *	−192	16/27
A. nidulans alcR (2)	−130 ATCAGCCCTAAAATGA −115 **** ** * *** *	−192	11/16
A. nidulans prnB	−384 AGGATCAACACTGAAACCA--A-TTCTCT −356 (T insertion) * ****** * ****** * ******	−63	21/28
A. nidulans acuD	−655 CCATCAACCATC −644 ************	−159	12/12
A. nidulans acuE	−298 ACGATCAACCGTC −286 ** ****** **	−116	11/13
A. nidulans amdS	−70 TACAT-AACCCTCATGCCA--ACTCCCAG −45 *** **** *** *** *** *	−70	17/26
A. nidulans facA	−1434 G insertion ATCAAAC-TCAAAC −1421 ***** * *****	−460	12/15
A. nidulans trpC	−310 A insertion TCAACCATCAA −299 ***********	–	11/12

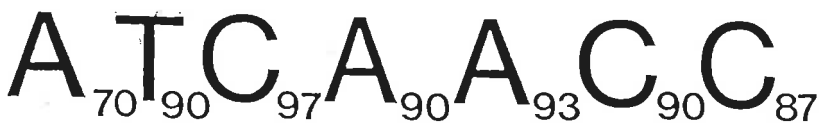
comm.) sequences (Figure 6.1.1), several other significantly similar regions were found in the promoters of carbon catabolite repressed genes. Four regions in the promoter of the *amd*S gene show 6/7 matches with ATCAACC. One of these regions, spanning from -135 to -128, is close to the sites of action of the positively acting regulatory proteins known to be involved in *amd*S induction (Corrick *et al.*, 1987; Hynes *et al.*, 1988). A sequence, ACCAACC, in the promoter of the acetyl-Co A synthetase encoding *fac*A gene at nucleotide position -576 is considerably further 3' than the the sequence similarity identified using the larger sequence discussed above. Although this sequence is a long way from the ATG start codon compared to the other sequences identified, transcription begins at nucleotide -424 and there is a large intron between the start point of transcription and the ATG codon (R. Sanderman and M. Hynes, pers. comm.).

There were also regions of sequence similarity in the promoters of the *A. nidulans* triosephosphate isomerase encoding *tpi*A (McKnight *et al.*, 1986) and orotidine-5'-phosphate decarboxylase encoding *pyr*G (Oakley *et al.*, 1987) genes, and the *trp*C genes of *A. niger* (Kos *et al.*, 1988) and *A. nidulans* (Mullaney *et al.*, 1985). There is no evidence for the expression of these genes being subject to carbon catabolite repression. There were 6/7 matches for each of these genes, but the variant base was frequently one of the most conserved nucleotides of the sequences from the promoters of the genes subject to carbon catabolite repression.

The promoter of the *A. nidulans ben*A gene which encodes a β-tubulin involved in vegetative growth (May *et al.*, 1987), shows several matched sequences, as does the promoter of a β-tubulin gene of *N. crassa* (Orbach *et al.*, 1986). While these genes have not been shown to

Figure 6.1.2: Fungal genes with promoter sequences showing at least 6/7 matches to the sequence ATCAACC. The positions of the matching sequences and, where present, sequences similar to the TATA box consensus sequence are shown relative to the first nucleotide of the start codon (+1, 0 is not used). Solid circles indicate genes that exhibit regulated expression that is subject to carbon catabolite repression. The percentages with which each of the nucleotides of the ATCAACC sequence occurs in these sequences is shown at the bottom of the figure. Sources of nucleotide sequence data are indicated in the text.

		POSITION	TATA
●*A. niger aldA*	ATCAACC	-201	-102
●*A. nidulans alcA*	ATCAACC	-70	-110
●*A. nidulans acuD*	ATCAACC	-653	-159
●*A. nidulans acuE*	ATCAACC	-295	-116
●*A. nidulans alcR*	ATCAgCC	-130	-192
●*A. nidulans amdS*	ATCAACa	-135	-70
	AaCAACC	-499	-70
	ATCAAgC	-805	-70
	ATCAACa	-966	-70
●*A. nidulans facA*	AcCAACC	-576	-460
●*A. nidulans prnB*	ATCAACa	-381	-63
	cTCAACC	-260	-63
	ATCAcCC	-203	-63
●*A. nidulans qutE*	ATCtACC	-53	-
A. nidulans benA	ATCAACC	-489	-
	ATCAACC	-1105	-
	cTCAACC	-505	-
A. nidulans tpiA	ATgAACC	-230	-
	ATCtACC	-212	-
A. nidulans pyrG	ATCcACC	-886	-
A. nidulans trpC	ATCAgCC	-218	-
	ATCAACt	-738	-
	cTCAACC	-827	-
	ATCAAtC	-846	-
A. niger trpC	ATCcACC	-193	-
	AcCAACC	-220	-
●*N. crassa qa-2*	tTCAACC	-116	-
	cTCAACC	-425	-
	cTCAACC	-433	-
N. crassa btub	tTCAACC	-46	-
	ATCAAaC	-142	-



be subject to carbon catabolite repression, the control of vegetative growth would be expected to respond to the carbon status of the cell, and so may share control mechanisms with genes involved in carbon metabolism.

The other *N. crassa* promoter to have a similar sequence is that of the *qa*-2 gene. Like the *qut*E gene of *A. nidulans*, *qa*-2 encodes the catabolic dehydroquinase involved in quinic acid utilization. Both *qut*E and *qa*-2 gene expression is subject to carbon catabolite repression. The two sequences in the *qa*-2 promoter at -425 and -433 are approximately 700 bp. upstream of the divergently transcribed *qa*-X gene of unknown function, which is also subject to carbon catabolite repression. These sequences in the *N. crassa qa* genes are close to transcription start points (Tyler *et al.*, 1984; Giles *et al.*, 1987; Geever *et al.*, 1989). *A. nidulans* shows a similar arrangement of genes with *qut*E and *qut*G being divergently transcribed (Hawkins *et al.*, 1988).

The presence of these sequences in the promoters of genes sharing the common regulatory mechanism of carbon catabolite repression, and the absence of this sequences from genes not subject to this form of control, implies that this sequence may play a role in mediating carbon catabolite repression. However, experimental evidence for a regultory function of these regions needs to be gained to test this hypothesis. The objective in performing these sequence comparisons was to locate conserved sequences within the *A. niger aldA* promoter that may, on the basis of similarity of sequence with genes sharing common regulation, be involved in controlling the expression of the *A. niger aldA* gene, and to correlate any sequence similarities with the effects on expression of promoter mutations. The construction and analysis of deletions within the *A. niger aldA* promoter is discussed in the

Figure 6.2.1: construction of plasmids containing deletions within the *A. niger aldA* promoter region.

a): outline of the method used to construct plasmids containing deletions 5' to the *A. niger aldA* start codon, and derivatives of pUC19 with altered restriction endonuclease sites which were required to make these deletions.

b): position of promoter deletions within plasmids described above relative to restriction endonuclease map. ATG represents the postion of the start of translation.

a)

VECTOR CONSTRUCTION

pUC19 ----> Digest XbaI, End Fill, Ligate ----> Digest EcoRI, End Fill, Ligate ----> pUCXE1

pUCXE1 ----> Digest SalI ----> Insert 6.8kb Insert of pNG100 ----> **pNG101**

pUC19 ----> Digest SacI & HindIII, End Fill, Ligate --> pUCSH1 --> Insert 0.7kb EcoRI Fragment of pNG100 ----> **pNG400**

pUC19 ----> Digest BamHI & EcoRI, End Fill, Ligate --> pUCBE1 --> Insert 2.0kb XbaI Fragment of pNG100 ----> **pNG500**

DELETION SERIES

pNG101 ----> Digest XbaI ----> 7.6kb Fragment, Ligate ----> **pNG106**

pNG400 ----> Digest BamHI & HindIII End Fill, Ligate ----> Ligate EcoRI Fragment into pNG101 (correct orientation) ----> **pNG102**

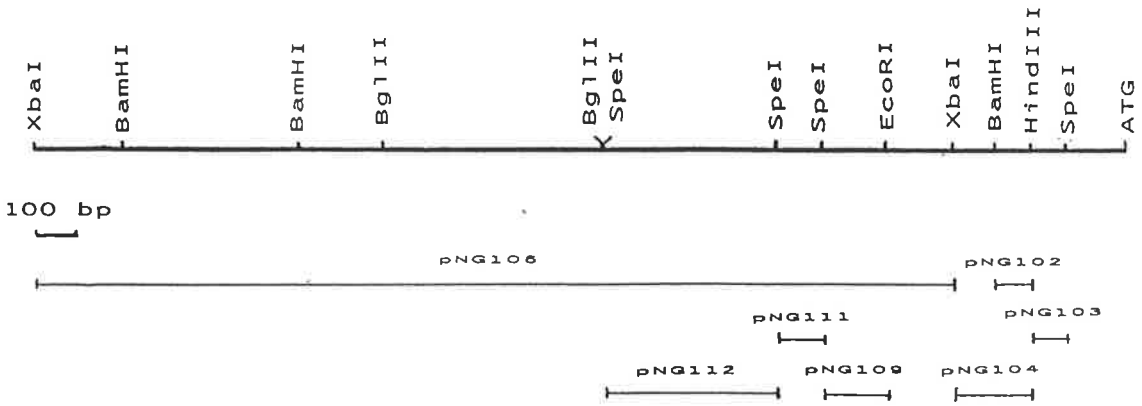
pNG400 ----> Digest SpeI & HindIII End Fill, Ligate ----> Ligate EcoRI Fagment into pNG101 (correct orientation) ----> **pNG103**

pNG400 ----> Digest XbaI & HindIII End Fill, Ligate ----> Ligate EcoRI Fragment into pNG101 (correct orientation) ----> **pNG104**

pNG500 ----> Digest EcoRI & Partial SpeI, End Fill, Ligate ----> Ligate XbaI Fragment into pNG101 (correct orientation) ----> **pNG109**

pNG500 ----> Partial Digest SpeI Ligate ----> Ligate XbaI Fragments into pNG101 (correct orientation) ----> **pNG111** **pNG112**

b)



following section.

6.2: Deletion analysis of the *Aspergillus niger aldA* promoter.

6.2.1: Construction of clones containing deletions in the *aldA* promoter.

Several plasmids containing clones of the *A. niger aldA* gene with regions 5' to the start codon deleted were constructed and transformed into the *A. niger aldA* deletion strain 216 to investigate the effect on *aldA* regulation of these deletions *in vivo*. An outline of the method used to construct these deletions is shown in Figure 6.2.1. Three derivatives of pUC19 containing altered polylinker sites were first constructed so that restriction sites within these regions would not interfere with the isolation and deletion of regions within subclones of pNG100. The orientation of derivatives of the inserts of pNG400 and pNG500 once inserted into pNG101 was determined by the size of restriction fragments resulting from digestion with enzymes that have recognition sites which are asymmetric within either the EcoRI or XbaI fragment containing the deletion. The relative positions of these deletions is shown in Figure 6.2.1.

6.2.2: Transformation of plasmids of the deletion series into *A. niger* strain 216.

In addition to carrying the gene replacement generated deletion of the *aldA* gene, strain 216 also has a mutation at the *pyrG* locus which results in an uridine auxotrophy (see Chapter 5). Transformation of strain 216 with the plasmids of the deletion series was achieved by cotransformation with pAB4.1 (van Hartingsveldt *et al.*, 1987), with

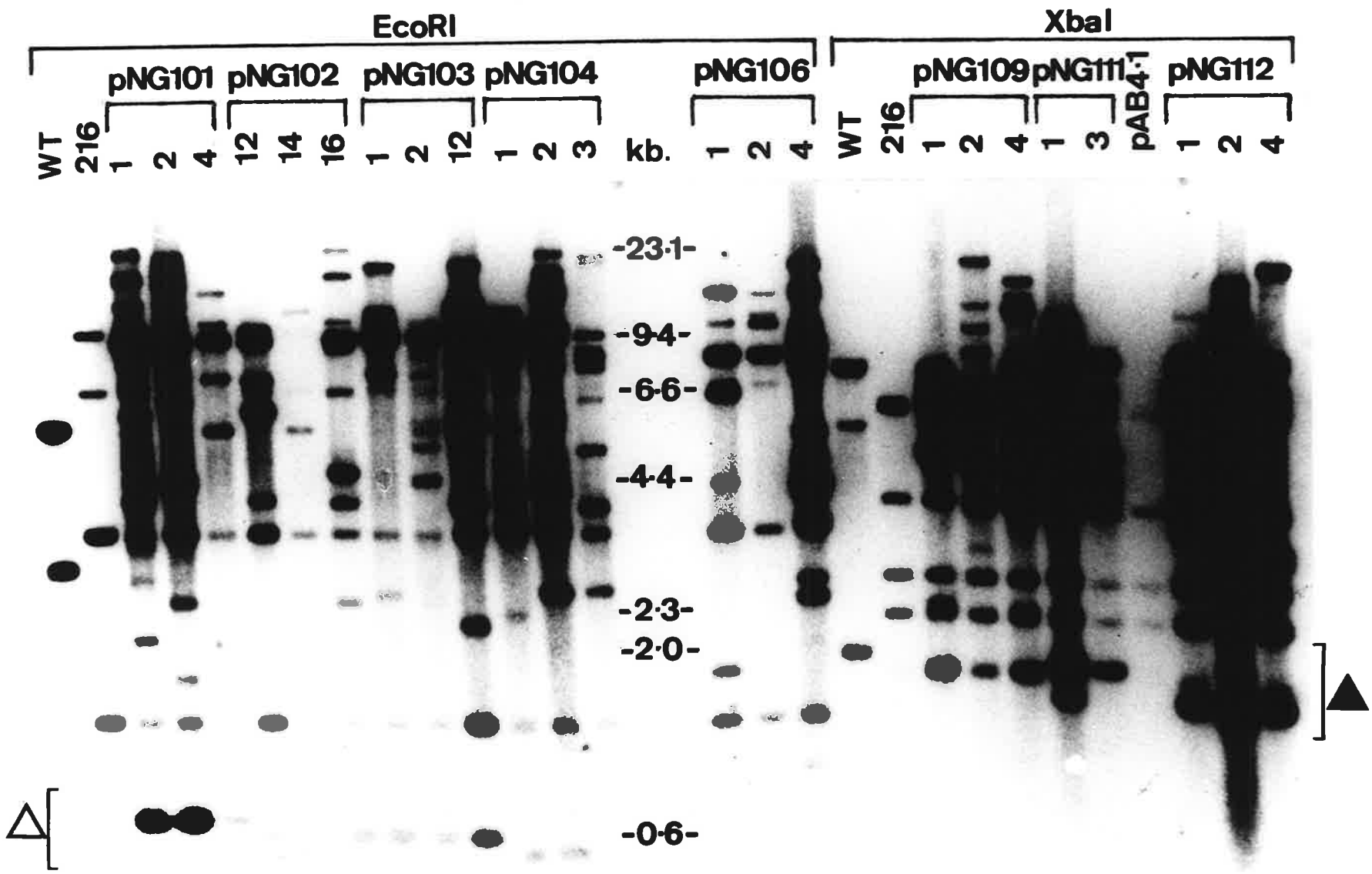
transformants being selected on media lacking added uridine. PyrG$^+$ transformants were screened by Southern hybridization to identify those that also had copies of the *in vitro* mutated *ald*A genes in their chromosomes. Autoradiographs of southern blots of DNA extracted from the transformants used for further analysis, probed with the insert of pNG100, are shown in Figure 6.2.2. Integration of deletion containing plasmids was confirmed by the absence of either the 700bp. EcoRI fragment (for pNG400 derivatives) or the 2.0kb. XbaI fragment (for pNG500 derivatives) and, apart from pNG106 transformants, the appearance of lower molecular weight bands due to the deletion of internal regions of these two fragments. The number of plasmids of the deletion series integrated into each transformants chromosomes, shown in Table 6.2.1, was estimated from Southern and dot blot analysis.

6.2.3: *ald*A regulation in transformants containing plasmids of the deletion series.

The nucleotide sequence of the *ald*A promoter has been determined from the second of the two BglII sites (nucleotide -1187) shown in Figure 6.2.1 (see Chapter 3). Based on data locating regulatory regions of several other *A. nidulans* genes including *amd*S (Hynes *et al.*, 1987), *alc*A (Gwynne *et al*, 1987), *arg*B (Goc and Weglenski, 1988) and *trp*C (Hamer and Timberlake, 1987), it is unlikely that regulatory sequences occur further 5' than this position and so the analysis was concentrated within the sequenced region. Because of the low frequency of integration events at the *ald*A locus (see Chapter 5), only transformants carrying copies of these plasmids integrated at heterologous positions could be analyzed.

*ald*A mRNA levels in the transformants containing plasmids of the deletion series that were grown under various conditions were

Figure 6.2.2: Southern blot analysis of DNA extracted from cotransformants of *A. niger* strain 216 containing plasmids of the deletion series and pAB4.1. The plasmids present in each transformant, and the restriction endonuclease with which the DNA has been digested are indicated. The insert of pNG100 was used as the probe. The open triangle indicates the region showing hybridization to deleted derivatives of the 0.7kb. EcoRI fragment. The solid triangle indicates the region showing hybridization to deleted derivatives of the 2.0kb. XbaI fragment. Bacteriophage lambda DNA digested with HindIII was used as a marker of molecular weight. The autoradiograph shown opposite has been heavily exposed to highlight the small molecular weight bands. WT: wildtype.


EcoRI
XbaI
pNG101
pNG102
pNG103
pNG104
pNG106
pNG109
pNG111
pNG112
pAB4·1
WT
216
1
2
4
12
14
16
3
kb.
-23·1-
-9·4-
-6·6-
-4·4-
-2·3-
-2·0-
-0·6-

determined to investigate the *in vivo* effects of the deletions on *ald*A gene regulation in terms of induction, repression, and the overall level of gene expression. These mRNA levels are shown in Table 6.2.1.

Transformants containing plasmid pNG101, which contains an intact *ald*A promoter region, showed expression similar to wildtype, although there was some increase in the derepressed/uninduced (0.05% fructose) level of expression in the higher copy number transformants. This may have been caused by the additive effects of a small amount of derepressed/uninduced expression from each of the integrated genes. Furthermore, the levels of expression were, particularly in the lower copy number transformant, reasonably closely correlated with gene copy number. This is in contrast to other cotransformants gained using the *amd*S gene of *A. nidulans* that had several copies of the wildtype gene integrated into their chromosomes along with multiple copies of the *amd*S gene (see Chapter 4). This observation implies that either copies of the 5' region of the *A. nidulans amd*S gene titrate factors required for the induction of *ald*A in *A. niger*, or the presence of the acetamidase enzyme affects *ald*A expression. This effect in not seen in AmdS$^+$ transformants of *A. nidulans* (see Chapter 4). This is an interesting result in light of the role of *amd*R in *ald*A regulation in *A. nidulans* and possibly *A. niger* (Chapters 4 and 5) and the experiments of Kelly and Hynes (1985) that showed titration of an *amd*R like protein in multicopy *amd*S transformants of *A. niger*.

Normal expression was also seen for transformants containing plasmids pNG102 and pNG103. The HindIII-SpeI deletion of pNG103 deletes the ATCAACC containing region discussed in Section 6.1. As no alteration in carbon catabolite repression was detected for pNG103 transformants, it appears this region does not play a major role in

216 Transformed with pAB4.1 and:	pNG101			pNG102			pNG103			pNG104			pNG106			pNG109			pNG111		pNG112			pAB4.1	WT
Transformant Number	1	2	4	12	14	16	1	2	12	1	2	3	1	2	4	1	2	4	1	3	1	2	4	ONLY	
Carbon Source																									
1% Ethanol	11.5	13	9	10	11.5	10	14.5	5	5	7	3	5	1	1	1	6	5	13	18	11.5	8	9	15	ND	8
0.05% Fructose	8.5	9.5	4	7	2	0.5	7	1	4	14	7	5	1.5	1	1	1	1	5	5	4.5	3	1	2	ND	2
1% Glucose, 1% Ethanol	1.5	2	1.5	2	2	2.5	0.5	1.5	0.5	1	1	1	1	1	0.5	1	1	1	0.5	1	1	0.5	1	ND	1.5
1% Glucose	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	ND	1
Fold WT Expression	22.5	23.5	4	1	1	7.5	5	23	3	7.5	14	13	22	14	13	1	1.5	1	2.5	5	1.5	7	9	0	1
Gene Copy Number	15	11	4	5	1	5	5	4	10	4	15	5	7	4	10	12	7	6	5	7	5	16	7	0	1
% Expression/Gene	150	210	100	20	100	150	100	580	30	190	90	260	310	350	130	8	21	16	50	71	30	43	130	0	100

Table 6.2.1: Levels of *ald*A mRNA in transformants of *A. niger* strain 216. Levels are normalized against those for glucose grown mycelia for each transformant. 10mM Ammonium chloride was added to the indicated carbon source. ND: none detected. Cultures were grown overnight at 37 ^{o}C, harvested, washed, and transferred to media containing the indicated carbon source for a further 4 hours as indicated in Section 2.6.1. The % Expression/Gene was calculated by dividing the fold wildtype expression by the gene copy number and mutiplying by 100. WT: wildtype.

carbon catabolite repression, such as the binding site for a repressor protein. However, this does not rule out the possibility of these sequences having some function in carbon catabolite repression that could not be detected from these experiments. For example, in light of the similarity to promoter elements of β-tubulin genes involved in vegetative growth in *A. nidulans* and *N. crassa*, this element may in some way be responsive to the carbon status of the cell. Furthermore, as it is possible that there are several separate control mechanisms affecting carbon catabolite repression, a change in gene expression may require mutations at more than one site. As discussed in Chapter 4, it is also possible that there is no direct interaction between molecules controlling carbon catabolite repression and the *A. niger aldA* promoter. Although this sequence may control some other function in the expression of genes involved in carbon metabolism, further analysis of the effects under various growth conditions of mutagenizing this sequence in the *A. niger aldA* promoter, and in other genes containing this sequence, will need to be performed to elucidate whether it controls any regulatory function. The HindIII-SpeI deletion of pNG103 also deleted the two putative CCAAT boxes, and thus it is unlikely that these sequences affect expression, which may be a result of their limited similarity to the CCAAT box consensus sequence.

Transformants containing the plasmid pNG104 showed elevated levels of expression in fructose grown cultures without drastically affecting *aldA* expression in cultures grown under other growth conditions. As discussed in Chapter 5, this deletion results in the *A. niger aldA* gene showing regulated expression that is more characteristic of the *A. nidulans aldA* gene, irrespective of whether the *A. nidulans aldA* gene is in *A. niger* or *A. nidulans*. As this effect is not seen in transformants of pNG102, then a promoter element situated between the

XbaI and BamHI sites, positioned between nucleotides -387 and -312, must be involved in reducing induction during fructose growth, thus giving this promoter element the characteristics of a silencer.

Transformants containing the plasmid pNG106 showed constitutive expression of *ald*A. Sequences required for induction of *ald*A are thus located within the XbaI - XbaI deletion in pNG106, which spans from position -394 to approximately 2.5kb. 5' of the start codon. It is not known whether this deletion removes sequences controlling carbon catabolite repression in addition to sequences controlling induction. However, as discussed in Chapters 4 and 5, the insensitivity to carbon catabolite repression of threonine induction of *ald*A expression in *A. niger* implies that, at least for threonine induction, carbon catabolite acts indirectly to affect *ald*A expression by controlling the formation of the intracellular inducer molecules. If this mechanism of carbon catabolite repression of *ald*A in *A. niger* is the only mechanism of repression that operates, then no mutations in the *ald*A promoter would be expected to specifically affect carbon catabolite repression. However, the molecular mechanisms of carbon catabolite repression in *Aspergillus* are unknown. It is possible that carbon catabolite repression may act directly on *ald*A in some growth conditions. Proteins encoded by genes of homologous function to the *cre*A gene of *A. nidulans* and other genes involved in carbon catabolite repression may, for example, be protein modifying enzymes, such as a kinase or phosphatase, that activate and inactivate positively acting molecules required for induction. These carbon catabolite repression regulatory genes may also encode DNA binding proteins that control the expression of the genes encoding the positive regulators without directly interacting with the promoters of the structural genes. Both of these mechanisms could also repress the expression of genes required for inducer formation and

hence phenotypically repress *ald*A. If either of these is the case, then promoter mutations affecting carbon catabolite repression would also not be identified. However, if these carbon catabolite repression regulatory genes encode DNA binding proteins that do bind to the structural genes' promoters, mutations affecting carbon catabolite repression should be able to be identified. A model of repression of gene expression by DNA binding proteins that is frequently used in eukaryotic systems (reviewed by Levine and Manley, 1989) involves competition for binding sites between repressors and the positively acting molecules involved in induction. If this mechanism operates in *A. niger* to control carbon catabolite repression of *ald*A expression, this type of analysis using deletions spanning relatively large regions compared to the size of known recognition sequences for DNA binding proteins, would be unlikely to separate regions involved in induction and repression.

In an attempt to more accurately position regulatory sequences, the effects of three smaller deletions within this 2kb. XbaI fragment were also studied. The EcoRI-SpeI deletion of pNG109 resulted in very low levels of expression per integrated copy, but regulation similar to the wildtype gene. This region is likely to contain an enhancer element that is required for the normally high level of *ald*A expression. Within this fragment, which extends from position -549 to -701, the sequence TGG is directly repeated eight times in a linear array (see Chapter 3). Although the TGG repeat would be a good candidate for this enhancer element, more thorough analysis would be required to assign the exact sequence involved. Thus is in addition to the specific positive and negative control, enhancers of transcription, which may act on other highly expressed genes which are regulated in a different manner, are also involved in the regulation of *ald*A

expression in *A. niger*.

Deletions in plasmids pNG111 and pNG112 had no significant effect on regulation. Some transformants did have reduced levels of expression, but this effect was not as consistent or significant as that caused by the deletion within pNG109. Therefore, none of these smaller deletions had the same effect as that for pNG106, that is, resulting in constitutive expression. As it is unlikely that regulatory sequences would be found further 5' than the region covered by the deletion within pNG112, a sequence required for the induction of *ald*A is probably located between the EcoRI and XbaI sites within the *A. niger ald*A promoter (position -549 to -391).

It is interesting that none of these smaller deletions specifically abolished glucose repression of ethanol induction. This was to be expected given the evidence that carbon catabolite repression may not act directly on *ald*A. However, from this analysis it is still not known whether a promoter element involved in the carbon catabolite repression of *ald*A is positioned close to sequences required for *ald*A induction.

Due to the absence of restriction endonuclease recognition sequences, no deletions were generated closer to the start codon than that in pNG103 (extending to position -135), and thus it is possible that elements controlling expression could exist 3' to this postion. However, shortly after this position is the putative TATA box, followed immediately by the pyrimidine rich sequence and the start points of transcription. Hence, apart from the pyrimidine rich sequence, there is a small region upstream from the transcriptional start points within this region that could potentially contain regulatory sequences.

6.3: Conclusions.

This preliminary deletion analysis has located three regions within the *ald*A promoter affecting gene expression. A silencer of induction during fructose growth was positioned between -387 and -312. A region probably affecting ethanol induction is located between -391 and -545. An enhancer element required for the high level of *ald*A expression was found between positions -549 and -701. Further analysis, involving the construction of mutations covering smaller sequences, is required to more accurately position regulatory sequences within the *A. niger ald*A promoter. No sequences within these regions show significant similarity to promoter sequences of other *Aspergillus* genes including the *A. nidulans ald*A gene. However, similar sequences may exist 5' to the *A. nidulans ald*A and *alc*A genes as the regions identified that affect *A. niger ald*A expression are positioned further 5' to the start codon than the sequence data available for the *A. nidulans* genes (Gwynne *et al.*, 1987a). Although these regions are more 5' than known regulatory regions for some *A. nidulans* genes, their distance from the start codon is not unusually large compared to many eukaryotic promoter elements.

Chapter 7:
Summary and concluding discussion.

At the outset of this study in March 1987, there were few publications describing experiments investigating gene structure and regulation in *A. niger*, but there was an ever increasing amount of data appearing in the literature pertaining to these topics in *A. nidulans*. As no mutant strains, clones of genes or regulatory data from *A. niger* was available, it was decided to choose a relatively well understood system of regulated gene expression from *A. nidulans* and undertake a comparative study of gene structure and regulation. This would allow available clones, strains and regulatory data from *A. nidulans* to be used in experiments to elucidate mechanisms of gene regulation in *A. niger*. Many of the initial experiments involved a basic description of the system. For a variety of reasons which are outined in Section 1.3, genes encoding aldehyde dehydrogenase (*ald*A) were chosen for the subject of this analysis.

The cloning of the *A. niger ald*A gene made use of the cloned *A. nidulans* gene to screen a genomic library. Obtaining and physically characterizing this clone was a prerequisite for the subsequent experiments. Physical analysis of this clone revealed that nucleotide sequence similarity between the two genes did not extend outside the protein coding regions and that the regions of dissimilarity were scattered throughout the coding regions. The promoter region of the *A. niger ald*A gene had a general structure resembling that of other highly

expressed fungal genes, with a sequence resembling the TATA box consensus and an extended pyrimidine rich region. There were also several repeated sequences within this region of the gene. The determination of the nucleotide sequence allowed the amino acid sequence of the *A. niger* aldehyde dehydrogenase to be compared to amino acid sequences of the aldehyde dehydrogenases of *A. nidulans*, man, and horse. These comparisons, together with biochemical data gained from the analysis of these mammalian enzymes (Bahr-Lindstrom *et al.*, 1984; Hempel *et al.*, 1984, 1985), allowed the identification of the cysteine residues that are likely to form a disulphide bridge at the active site.

The physical analysis of the *A. niger aldA* gene was performed with the eventual aim of allowing the identification of promoter sequences controlling the expression of this gene. Before this type of analysis could be performed, data pertaining to the conditions under which this gene is expressed needed to be obtained. A comparative approach was undertaken with the primary aim of gaining data on *aldA* gene regulation in *A. niger*, but it also produced a considerable amount of new information about the regulation of *aldA* gene expression in *A. nidulans*. The use of studies of heterologous expression of *aldA* genes was also useful in gaining regulatory data.

The expression of both genes was subject to induction and carbon catabolite repression, but the levels of induction and the extent of repression in different growth conditions showed several differences in the regulation of *aldA* gene expression in the two species. The two most notable differences were found in cultures grown in threonine containing media and in media containing fructose as the sole carbon source. It was impossible to directly compare the basal level of *aldA*

expression in *A. nidulans* and *A. niger* since there exists no probe to detect a second transcript with equal levels of expression, and homology, for the two species. Determination of enzyme activity would not have allowed such a comparison, since mRNA levels are not directly related to enzyme activity, as translational control and protein stability are also involved.

Induction of *ald*A in *A. niger* in response to growth in the presence of threonine was not sensitive to carbon catabolite repression. This was independent of whether the *A. niger* or *A. nidulans* *ald*A genes were being expressed in *A. niger*, but did not occur in wildtype *A. nidulans*. These results showed that threonine metabolism in *A. niger* occurs in the presence of glucose, and under these growth conditions, carbon catabolite repression does not act directly on *ald*A, or genes encoding positively acting regulatory proteins required for *ald*A induction in the presence of threonine. It is possible that this induction mechanism is different to that controlling induction in the presence of other inducing compounds such as ethanol or ethyl ammonium. It is assumed that threonine itself is not an inducer of *ald*A in *A. niger* by inference from published results regarding *ald*A regulation in *A. nidulans* (Pateman *et al.*, 1983), and that induction of *ald*A in *A. nidulans* is sensitive to carbon catabolite repression due to, at least in part, the repression of other genes involved in threonine metabolism, such as the threonine aldolase encoding gene and genes encoding uptake functions, which thus prevents the formation of the intracellular inducer of *ald*A, which is probably acetaldehyde. Other experiments looking at the toxicity and effects on *ald*A expression of the addition to various growth media of acetaldehyde gave further evidence of indirect effects of carbon catabolite repression through

affecting the uptake of carbon sources. Although these results imply that carbon catabolite repression of *ald*A expression in *A. niger* occurs by preventing inducer formation, they do not rule out the possibility of direct effects also occurring under different growth conditions.

Growth in fructose containing media was shown to lead to the induction of *ald*A expression in wildtype *A. nidulans* but not in wildtype *A. niger*. This induction in the presence of fructose in *A. nidulans* required a functional *amd*R gene product, and thus in addition to the control of *ald*A expression by the product of the *alc*R gene, either the *amd*R gene product or the product of an *amd*R regulated gene controls the expression of *ald*A. An alternative hypothesis to explain the increase in *ald*A expression in fructose grown cultures is that of extreme derepression. If this is the case, then the level of *ald*A expression in glucose grown cultures of strains containing mutant *cre*A alleles would be similar to that of fructose grown wildtype cultures due to their derepressed phenotype. This effect will vary with different *cre*A alleles since they are derepressed to different extents. Since little is known of the molecular mechanism of carbon catabolite repression in *A. nidulans*, the interpretation of experiments such as these would not be straightforeward. However, strongest support for *amd*R regulation being independent of the action of the *alc*R gene, and being truly induction rather than extreme derepression, comes from the synergistic effect of fructose on ethanol induction in *A. nidulans*. This synergistic effect was absent in the *amd*R mutant and, by analogy to studies in Yeast and other organisms, implies independent positive regulatory mechanisms. In order to test this further, the analysis of *ald*A expression in an *alc*R mutant strain would be of value. An absence of fructose induction in an *alc*R mutant strain would imply that *amd*R dependent induction of *ald*A in fructose grown cultures operates via the

*alc*R gene product. Although fructose induction of *ald*A expression was not seen in wildtype *A. niger*, it was seen for the *A. nidulans ald*A gene when expressed in *A. niger*. Hence the mechanism leading to fructose induction of *ald*A in *A. nidulans* exists in *A. niger*, but fails to act on the wildtype *A. niger ald*A gene. This was shown to be due to the presence of a silencer element in the *A. niger ald*A promoter, that once deleted led to fructose induction of this gene in *A. niger*. This reverse genetic approach in *A. niger* of analyzing the effects of *in vitro* generated mutations *in vivo* also identified other regions that were required for the induction of *ald*A and the enhancement of transcription to achieve the observed high level of expression. These experiments bought together the data gained from the physical analysis of the *A. niger ald*A gene and the analysis of its regulation.

The construction, by gene replacement, of the strain of *A. niger* containing a deletion of *ald*A was necessary for the analysis of the *A. nidulans ald*A gene in *A. niger* and the identification of regulatory elements in the promoter, but also gave some indication of the involvement of aldehyde dehydrogenase in carbon metabolism in *A. niger*. The enzyme is required for the utilization of a number of carbon sources that are utilized via aldehydes and semialdehydes. While many of the phenotypes of this strain are the same as those for an *ald*A mutant of *A. nidulans*, there were some differences in phenotype observed which, in addition to the differences in *ald*A gene regulation, possibly indicate fundamental differences in metabolism between the two species.

While the primary physical analysis of the *ald*A gene of *A. niger* is complete, further experiments involving the generation of mutations

covering smaller regions of the promoter are required to identify exactly which sequences within the identified fragments are involved in *ald*A gene expression. One good approach would be to fuse mutated derivatives of these fragments to a reporter gene, such as the *lac*Z gene of *E. coli*, so that each of the functions can be looked at separately. The analysis of the 5' deletions presented here involves heterologous integration of transforming DNA. The original design of the experiments outlined in this thesis to examine the *ald*A 5' region included replacement of the pNG300 sequences bearing *amd*S of *ald*A with the inserts of the plasmids of their deletion series. However, the very low frequency of integration events at the *A. niger ald*A locus made this approach infeasible. The integration of pNG300 at the *ald*A locus in the gene replacement strain 712 was the only *ald*A homologous integration event detected in any *A. niger* transformant during this study. While heterologous transformation introduces possible position effects, the analysis of several transformants can allow trends to be detected for each deletion. Heterologously integrated transforming sequences have frequently been used in systems ranging from simple eukaryotes to mammalian cells to define promoter elements. There have been no published reports of plasmids that are self replicative in *A. nidulans* or *A. niger*.

There are a considerable number of experiments that were not performed as part of this study that would be required to gain a better understanding of some of the data gained from the analysis of the regulation of *ald*A gene expression in *A. niger* and *A. nidulans*. Although experiments presented in this thesis clearly show that the product of the *amd*R gene affects *ald*A regulation in *A. nidulans*, and that a similar mechanism probably exists in *A. niger*, the exact role that the *amd*R gene product plays in *ald*A gene regulation has not been

determined. It would be of value to analyze *ald*A expression in an *alc*R125*amd*R$^-$44 double mutant so as to investigate whether the effects of *amd*R on threonine induction were independent of *alc*R. That is, if the effects of *alc*R125 are epistatic to the effects of *amd*R$^-$44, then it would be predicted that the increase in threonine induction in the *amd*R$^-$44 mutant strain would be absent from the double mutant. Valuable data as to whether the *A. niger ald*A gene is also controlled by a gene of homologous function to *amd*R may be gained from the analysis of the heterologous expression of the *A. niger ald*A gene in an *amd*R$^-$ strain of *A. nidulans*. It is not known whether multiple copies of the 5' region of the *ald*A genes are capable of titrating regulatory gene products, but if so, then it would be interesting to see if transformation with cloned *alc*R and *amd*R genes could reverse this titration.

The molecular mechanisms of carbon catabolite repression of *ald*A are also not known. However, experiments presented in this thesis show that carbon catabolite repression of *ald*A in *A. niger* is, at least in part, indirectly acting via the prevention of inducer formation. Therefore, the study of *ald*A gene regulation is unlikely to be as beneficial in determining these mechanisms as the analysis of the regulatory genes controlling carbon catabolite repression. This approach is underway in *A. nidulans*, as the *cre*A gene has been cloned (Dowzer and Kelly, 1989), and is being physically analyzed (C. Dowzer and J. Kelly, pers. comm.).

This analysis of the *A. niger ald*A gene represents the best understood system of gene structure and regulation for this organism. This type of work needs to be performed before regulatory genes controlling the expression of the structural gene of interest can be

identified and analyzed. Rather than trying to identify genes affecting *ald*A expression by isolating mutant alleles in these genes, a similar approach to that used in this work could result in gaining clones of genes involved in *ald*A regulation. These experiments may involve the use of cloned genes from *A. nidulans* to screen *A. niger* libraries. Alternatively, the complementation of mutations in *A. nidulans* or other fungi by transformation, and subsequent rescue of the complementing sequences may also allow the cloning of regulatory genes from *A. niger*. This approach was used during this study, and although not rescued from the transformant, an *A. niger* sequence complemented an *A. nidulans alc*R mutation. Furthermore, this reverse genetic approach, both within and across species, need not be confined to the study of fungal gene regulation, and could be applied to many other organisms that are not easily amenable to genetic analysis.

References:

Adams, T.H., Boylan, M.T. and Timberlake, W.E. (1988). *brlA* is necessary and sufficient to direct conidiophore development in *Aspergillus nidulans*. Cell 54: 353-362.

Andrianopoulos, A. and Hynes, M.J. (1988). Cloning and analysis of the positively acting regulatory gene *amdR* from *Aspergillus nidulans*. Mol. Cell. Biol. 8: 3532-3541.

Aramayo, R., Adams, T.H. and Timberlake, W.E. (1989). A large cluster of highly expressed genes is dispensable for growth and development in *Aspergillus nidulans*. Genetics 122: 65-71.

Armitt, S., McCullough, W. and Roberts, C.R. (1976). Analysis of acetate non-utilizing (*acu*) mutants in *Aspergillus nidulans* J. Gen. Microbiol. 92: 263-282.

Arst, H.N. (1976). Integrator gene in *Aspergillus nidulans*. Nature 262: 231-234.

Arst, H.N. and Bailey, C.R. (1977). The regulation of carbon metabolism in *Aspergillus nidulans*. pp. 131-146. In J.E. Smith and J.A. Pateman (eds.) Genetics and Physiology of *Aspergillus*. Academic Press, Inc. N.Y.

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

Arst,H.N., Jones, S.A. and Bailey, C.R. (1981). A method for the selection of deletion mutations in the L-proline catabolism gene cluster of *Aspergillus nidulans*. Genet. Res. Camb. 38: 171-195.

Arst,H.N., Kudla,D. Martinez-Rossi, N., Caddick, M.X., Sibely, S. and Davies, R.W. (1989). *Aspergillus* and mouse share a new class of 'zinc finger' protein. TIG. 5: 291.

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

Arst, H.N. and MacDonald, D.W. (1978). Reduced expression of a distal gene of the *prn* gene cluster in deletion mutants of *Aspergillus nidulans*: genetic evidence for a dicistronic messenger in an eukaryote. Molec. Gen. Genet. 163: 17-22.

Arst, H.N., MacDonald, D.W. and Jones, S.A. (1980) Regulation of proline transport in *Aspergillus nidulans*. J. Gen. Microbiol. 116: 285-294.

Arst, H.N., Penfold, A., and Bailey, C.R. (1978). Lactam utilization in *Aspergillus nidulans*: evidence for a fourth gene under the control of the integrator gene *int*A. Molec. Gen. Genet. 166: 321-327.

Arst, H.N. and Scazzocchio, C. (1985). Formal genetics and molecular biology of the control of gene expression in *Aspergillus nidulans*. pp. 309-343. In J.W. Bennett and L.L. Lasure (eds.), Gene manipulations in fungi. Academic Press, Inc. N.Y.

Atkinson, P.W., King, J.A. and Hynes, M.J. (1985). Identification of the *aci*A gene controlled by the *amd*A regulatory gene in *Aspergillus nidulans*. Curr. Genet. 10: 133-138.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Smith, J.A., Seidman, J.G. and Struhl, K. (eds.) (1987). Current Protocols in molecular biology 1987-1988. John Wiley and Sons, N.Y.

Bahr-Lindstrom, H., Hempel, J. and Jornvall, H. (1984). The cytoplasmic isoenzyme of horse liver aldehdye dehydrogenase. Relationship to the corresponding human isoenzyme. Eur. J. Biochem. 141: 37-42.

Bailey, C. and Arst, H.N. (1975). Carbon catabolite repression in *Aspergillus nidulans*. Eur. J. Biochem. 51: 573-577.

Ballance, D.J. (1986). Sequences important for gene expression in filamentous fungi. Yeast 2: 229-236.

Ballance, D.J. and Turner, G. (1986). Gene cloning in *Aspergillus nidulans*: isolation of the isocitrate lyase gene (*acu*D). Molec. Gen. Genet. 202: 271-275.

Bennett, J.W. (1985). Molds, manufacturing and molecular genetics. pp. 345-366. In W.E. Timberlake (ed.), Molecular genetics of filamentous fungi. Alan R. Liss, Inc. N.Y.

Benton, W.D. and Davis, R.W. (1977). Screening λgt recombinant clones by hybridization to single plaques *in situ*. Science 196: 180-185.

Beri, R.K., Whittington, H., Roberts, C.F. and Hawkins, A.R. (1987). Isolation and characterization of the positively acting regulatory gene *qut*A from *Aspergillus nidulans*. Nucleic Acids Res. 15: 7991-8001.

Boeke, J.D., LaCroute, F. and Fink, G.R. (1984). A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-flouro-orotic acid resistance. Molec. Gen. Genet. 197: 345-346.

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

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

Bos, C.J., Debets, A.J.M., Swart, K., Huybers, A., Kobus, G. and Slakhorst, S.M. (1988). Genetic analysis and the construction of master strains for the assignment of genes to six linkage groups in *Aspergillus niger*. Curr. Genet. 14: 437-443.

Bos, C.J., Debets, A.J.M., Kobus, G., Slakhorst, S.M. and Swart, K. (1989). Adenine and Pyrimidine genes of *Aspergillus niger* and evidence for a seventh linkage group. Curr. Genet. 16: 307-310.

Bostian, K.A. and Betts, G.F. (1978). Rapid purification and properties of potassium-activated aldehyde dehydrogenase from *Saccharomyces cerevisiae*. Biochem. J. 173: 773-786.

Boylan, M.T., Mirabito, P.M., Willett, C.E., Zimmerman, C.R. and Timberlake, W.E. (1987). Isolation and physical characterization of three essential conidiation genes from *Aspergillus nidulans*. Mol. Cell. Biol. 7: 3113-3118.

Braun, T., Bober, E., Singh, S., Agarwal, D.P. and Goedde, H.W. (1987). Evidence for a signal peptide at the amino-terminal end of human mitochondrial aldehyde dehydrogenase. Febs. Lett. 215: 233-236.

Britten, R.J. and Davidson, E.H. (1969). Gene regulation for higher cells: a theory. Science 165: 349-357.

Brody, H. and Carbon, J. (1989). Electrophoretic karyotype of *Aspergillus nidulans*. Proc. Natl. Acad. Sci. U.S.A. 86: 6260-6263.

Burke, J.F. (1984). High-sensitivity S1 mapping with single stranded [^{32}P]DNA probes synthesized from bacteriophage M13mp templates. Gene 30: 63-68.

Buxton, F.P., Gwynne, D.I., Garven, S., Sibely, S. and Davies, R.W. (1987). Cloning and molecular analysis of the ornithine carbomyl transferase gene of *Aspergillus niger*. Gene 60: 255-266.

Caddick, M.X., Arst, H.N., Taylor, L.H., Johnson, R.I. and Brownlee, A.G. (1986). Cloning of the regulatory gene *are*A mediating nitrogen metabolite repression in *Aspergillus nidulans*. EMBO J. 5: 1087-1090.

Carlson, M. (1987). Regulation of sugar utilization in *Saccharomyces* species. J. Bacteriol. 109: 4873-4877.

Chen, M.W., Anne, J., Volckaert,G., Haysmans, E., van den Burghe, A., and de Wachter, R. (1984). The nucleotide sequences of the 5S rRNAs of seven molds and a yeast and their use in studying ascomycete phylogeny. Nucleic Acids Res. 12: 4881-4892.

Clark, J.F. and Jakoby, W.B. (1970). Yeast aldehyde dehydrogenase. J. Biol. Chem. 245: 6065-6071.

Clements, J.M. and Roberts, C.F. (1986). Transcription and processing signals in the 3-phosphoglycerate kinase (PGK) gene from *Aspergillus nidulans*. Gene 44: 97-105.

Clutterbuck, A.J. (1974). *Aspergillus nidulans* genetics. pp. 447-510. In R.C. King (ed.), Handbook of genetics, vol. 1. Plenum Publishing Corp., N.Y.

Clutterbuck, A.J. (1984). Loci and linkage map of the filamentous fungus *Aspergillus nidulans*. pp. 265-273. In S.J. O'Brien (ed.), Genetic maps, vol 3. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Corrick, C.M., Twomey, A.P. and Hynes, M.J. (1987). The nucleotide sequence of the *amd*S gene of *Aspergillus nidulans* and the molecular characterization of 5' mutations. Gene 53: 63-71.

Cove, D.J. (1966). The induction and repression of nitrate reductase in the fungus *Aspergillus nidulans*. Biochem. Biophys. Acta. 113: 51-56.

Cove, D.J. (1977). The genetics of *Aspergillus nidulans*. pp. 81-95. In J.E. Smith and J.A. Pateman (eds.), Genetics and physiology of *Aspergillus nidulans*. Academic Press, N.Y.

Cove, D.J. (1979). Genetic studies of nitrate assimilation in *Aspergillus nidulans*. Biol. Rev. 54: 291-327.

Creaser, E.H., Porter, R.L., Britt, K.A., Pateman, J.A. amd Doy, C.H. (1985). Purification and preliminary characterization of alcohol dehydrogenase from *Aspergillus nidulans*. Biochem. J. 225: 449-454.

Creaser, E.H., Porter, R.L. and Pateman, J.A. (1987). Purification and properties of aldehyde dehydrogenase from *Aspergillus nidulans*. Int. J. Biochem. 19: 1009-1012.

Da Silva, A.J.S., Whittington, H., Clements, J., Roberts, C.F. and Hawkins, A.R. (1986). Sequence analysis and transformation by the catabolic 3-dehydroquinase (*qut*E) gene from *Aspergillus nidulans*. Biochem. J. 240: 481-488.

Davis, R.W., Bolstein, D. and Roth, J.R. (1980). Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Davis, M.A. and Hynes, M.J. (1987). Complementation of *are*A$^-$ regulatory gene mutations of *Aspergillus nidulans* by the heterologous regulatory gene *nit-2* of *Neurospora crassa*. Proc. Natl. Acad. Sci. U.S.A. 84: 3753-3757.

Davis, M.A. and Hynes, M.J. (1989). Regulatory genes in *Aspergillus nidulans*. TIG 5: 14-19.

De Graaff, L. and Visser, J. (1988). Structure of the *Aspergillus nidulans* pyruvate kinase gene. Curr. Genet. 14: 553-560.

Dobson, M.J., Tuite, M.F., Roberts, N.A., Kingsman, A.J., Kingsman, S.M., Perkins, R.E., Conroy, S.C., Dunbar, B. and Fothergill, L.A. (1982). Conservation of high efficiency promoter sequences in *Saccharomyces cerevisiae*. Nucleic Acids Res. 8: 2625-2637.

Doonan, J.H. and Morris, N.R. (1989). The *bim*G gene of *Aspergillus nidulans*, required for completion of anaphase, encodes a homolog of mammalian phosphoprotein phosphatase 1. Cell 57: 987-996.

Doy, C.H., Pateman, J.A., Olsen, J.E., Kane, H.J. and Creaser, E.H. (1985). Genomic clones of *A. nidulans* containing *alc*A, the structural gene for ADH, and *alc*R, the regulatory gene for ethanol metabolism. DNA 4: 105-114.

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

Felenbok, B., Sequeval, D., Mathieu, M., Sibely, S., Gwynne, D.I. and Davis, R.W. (1988). The ethanol regulon in *Aspergillus nidulans*: characterization and sequence of the positive regulatory gene *alc*R. Gene 73: 385-396.

Felenbok, B., Sophianopoulou, V., Mathieu, M., Sequeval, D., Kulmberg, P., Diallinas, G. and Scazzocchio, C. (1989). Regulation of genes involved in the utilization of carbon sources in *Aspergillus nidulans*. In H. Nevalainen and M. Penttila (eds.), Proceedings of the EMBO-Alko workshop on molecular biology of filamentous fungi, Helsinki, 1989. Foundation for biotechnical and industrial fermentation research. 6: 73-83.

Fincham, J.R.S. (1989). Transformation in fungi. Microbiological Reviews 53: 148-170.

Frischauff, A.M., Lehrach,H., Poustka, A. and Murray, N. (1983). Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. 170: 827-842.

Geever, R.F., Huiet, L., Baum, J.A., Tyler, B.M., Patel, V.B., Rutledge, B.J., Case, M.E. and Giles, N.H. (1989). DNA sequence, organization and regulation of the *qa* gene cluster of *Neurospora crassa*. J. Mol. Biol. 207: 15-34.

Geliebter, J., Zeff, R.A., Melvold, R.W. and Nathenson, S.G. (1986). Mitotic recombination in germ cells generated two major histocompatibility complex mutant genes shown to be identical by RNA sequence analysis: K^{bm9} and K^{bm6}. Proc. Natl. Acad. Sci. U.S.A. 83: 3371-3375.

Giles, N.H., Case, M.E., Baum, J., Geever, R. and Patel, V. (1987). Mechanisms of positive and negative regulation in the *qa* gene cluster of *Neurospora crassa*. pp 13-32. In W.F. Loomis (ed.), Genetic Regulation of Development. Alan R. Liss, Inc. N.Y.

Goc, A. and Weglenski, P. (1988). Regulatory region of the *Aspergillus nidulans arg*B gene. Curr. Genet. 14: 425-429.

Goosen, T., Bloemheuval, G., Gysler, C., de Bie, D.A., van den Broek, H.W.J. and Swart, K. (1987). Transformation of *Aspergillus niger* using the homologous orotidine-5'-phosphate decarboxylase gene. Curr. Genet. 11: 499-503.

Gossop, G.H., Yuill, E. and Yuill, J.H. (1940). Heterogeneous fructifications in species of *Aspergillus*. Trans. Brit. Mycol. Soc. 24: 337-344.

Grant, S., Roberts, C.F., Lamb, H., Stout, M. and Hawkins, A.R. (1988). Genetic regulation of the quinic acid utilization (*qut*) cluster in *Aspergillus nidulans*. J. Gen. Microbiol. 134: 347-358.

Gubler, U. and Hoffman, B.J. (1983). A simple and very efficient method for generating cDNA libraries. Gene 25: 263-269.

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

Gwynne, D.I., Buxton, F.P., Sibely, S., Davies, R.W., Lockington, R.A, Scazzocchio, C. and Sealy-Lewis, H.M. (1987a). Comparison of the *cis*-acting control regions of two coordinately controlled genes involved in ethanol utilization in *Aspergillus nidulans*. Gene 51: 205-216.

Gwynne, D.I., Buxton, F.P., Williams, S.A., Garven, S. and Davies, R.W. (1987b). Genetically engineered secretion of active human interferon and a bacterial endoglucanase from *Aspergillus nidulans*. Biotechnology. 5: 713-719.

Gwynne, D.I., Buxton, F.P., Williams, S.A., Sills, A.M., Johnstone, J.A., Buch, J.K., Guo, Z., Drake, D., Westphal, M. and Davies, R.W. (1989). Development of an expression system in *Aspergillus nidulans*. Biochem. Soc. Trans. 17: 338-340.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166: 557-580.

Hamer, J.E. and Timberlake, W.E. (1987). Functional organization of the *Aspergillus nidulans trp*C promoter. Mol. Cell. Biol. 7: 2352-2359.

Hawkins, A.R. (1987). The complex *Arom* locus of *Aspergillus nidulans*: evidence for multiple gene fusions and convergent evolution. Curr. Genet. 11: 491-498.

Hawkins, A.R., Giles, N.H. and Kinghorn, J.R. (1982). Genetical and biochemical aspects of quinate breakdown in the filamentous fungus *Aspergillus nidulans*. Biochem. Genet. 20: 271-286.

Hawkins, A.R., Da Silva, A.J.S. and Roberts, C.F. (1984). Evidence for two control genes regulating expression of the quinic acid utilization (*qut*) gene cluster in *Aspergillus nidulans*. J. Gen. Microbiol. 130: 567-574.

Hawkins, A.R., Da Silva, A.J.S. and Roberts, C.F. (1985). Cloning and characterization of three enzyme structural genes *qut*B, *qut*C and *qut*E from the quinic acid utilization gene cluster in *Aspergillus nidulans*. Curr. Genet. 9: 305-311.

Hawkins, J.D. (1988). A survey on intron and exon lengths. Nucleic Acids Res. 16: 9893-9908.

Hempel, J., Bahr-Lindstrom, H. and Jornvall, H. (1984). Aldehyde dehydrogenase from human liver. Primary structure of the cytoplasmic isoenzyme. Eur. J. Biochem. 141: 21-35.

Hempel, J., Kaiser, R. and Jornvall, H. (1985). Mitochondrial aldehyde dehydrogenase from human liver. Primary structure, differences in relation to the cystolic enzyme, and functional correlations. Eur. J. Biochem. 153: 13-28.

Hsu, L.C., Tani, K., Fujiyoshi, Y., Kurachi, K. and Yoshida, A. (1985). Cloning of cDNAs for human aldehyde dehydrogenases 1 and 2. Proc. Natl. Acad. Sci. U.S.A. 82: 3771-3775.

Hynes, M.J. (1977). Induction of the acetamidase of *Aspergillus nidulans* by acetate metabolism. J. Bacteriol. 131: 770-775.

Hynes, M.J. (1978a). Multiple independent control mechanisms affecting the acetamidase of *Aspergillus nidulans*. Molec. Gen. Genet. 161: 59-65.

Hynes, M.J. (1978b). An "up promoter" mutation affecting the acetamidase of *Aspergillus nidulans*. Molec. Gen. Genet. 166: 31-36.

Hynes, M.J. (1982). A *cis*-dominant mutation in *Aspergillus nidulans* affecting the expression of the *amd*S gene in the presence of mutations in the unlinked gene *amd*A. Genetics 102: 139-147.

Hynes, M.J., Corrick, C.M. and King, J.A. (1983). Isolation of genomic clones containing the *amd*S gene of *Aspergillus nidulans* and their use in the analysis of structural and regulatory mutations. Mol. Cell. Biol. 3: 1430-1439.

Hynes, M.J., Corrick, C.M., Kelly, J.M. and Littlejohn, T.G. (1988). Identification of the sites of action for regulatory genes controlling the *amd*S gene of *Aspergillus nidulans*. Mol. Cell. Biol. 8: 2589-2596.

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

Hynes, M.J. and Kelly, J.M. (1981). Threonine, quinate and glucuronic acid utilization mutants of *Aspergillus nidulans*. Asp. Newslett. 15: 18-21.

Hynes, M.J., Kelly, J.M., Corrick, C.M. and Littlejohn, T.G. (1985). Structure and expression of the *Aspergillus amd*S gene. pp. 157-169. In W.E. Timberlake (ed.), Molecular genetics of filamentous fungi. Alan R. Liss, Inc. N.Y.

Hynes, M.J. and Pateman, J.A.J. (1970). The genetic analysis of regulation of amidase synthesis in *Aspergillus nidulans*. Molec. Gen. Genet. 108: 97- 106.

Hull, E.P., Green, P.M., Arst, H.N. and Scazzocchio, C. (1989). Cloning and physical characterization of the L-proline catabolism gene cluster of *Aspergillus nidulans*. Molec. Microbiol. 3: 553-559.

Ish-Horowicz, D. and Burke, J.F. (1981). Rapid and efficient cosmid cloning. Nucleic Acids Res. 9: 2989-2998.

Jacobson, M.H. and Bernofsky,C. (1974). Mitochondrial acetaldehyde dehydrogenase from *Saccharomyces cerevisiae*. Biochim. Biophys. Acta. 350: 277-291.

Jones, I.G. and Sealy-Lewis, H.M. (1989). Chromosomal mapping and gene disruption of the ADHIII gene in *Aspergillus nidulans*. Curr. Genet. 15: 135-142.

Jones, S.A., Arst, H.N. and MacDonald, D.W. (1981). Gene roles in the *prn* cluster of *Aspergillus nidulans*. Curr. Genet. 3: 49-56.

Katz, M.E. and Hynes, M.J. (1989a). Gene function by interspecific transformation. Gene 78: 167-171.

Katz, M.E. and Hynes, M.J. (1989b). Characterization of the *amd*R-controlled *lam*A and *lam*B genes of *Aspergillus nidulans*. Genetics. 122: 331-339.

Katz, M.E. and Hynes, M.J. (1989c). Isolation and analysis of the acetate regulatory gene, *fac*B, from *Aspergillus nidulans*. Mol. Cell. Biol. 9: 5696-5701.

Katz, M.E., Saleeba, J.A., Sapats, S.I. and Hynes, M.J. (1990). A mutation affecting *amd*S expression in *Aspergillus nidulans* contains a triplication of a *cis*-acting regulatory sequence. Molec. Gen. Genet. In Press.

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*. Molec. Gen. Genet. 156: 87-92.

Kelly, J.M. and Hynes, M.J. (1985). Transformation of *Aspergillus niger* by the *amd*S gene of *Aspergillus nidulans*, EMBO J. 4: 475-479.

Kelly, J.M. and Hynes, M.J. (1987). Multiple copies of the *amd*S gene of *Aspergillus nidulans* causes titration of *trans*-acting regulatory proteins. Curr. Genet. 11: 21-31.

Kos, A., Kuijvenhoven,J., Werners,K., Bos, C.J., van den Broek, H.W.J., Pouwels, P.H. and van den Hondel, C.A.M.J.J. (1985). Isolation of the *Aspergillus niger trp*C gene. Gene 39: 231-238.

Lhoas, P. (1967). Genetic analysis by means of the parasexual cycle in *Aspergillus niger*. Genet. Res. Camb. 10: 45-61.

Lockington, R.A., Sealy-Lewis, H.M., Scazzocchio, C. and Davies, R.W. (1985). Cloning and characterization of the ethanol utilization regulon in *Aspergillus nidulans*. Gene 33: 137-149.

Lockington, R.A., Scazzocchio, C., Sequeval, D., Matieu, M. and Felenbok, B. (1987). Regulation of *alc*R, the positive regulatory gene of the ethanol utilization regulon of *Aspergillus nidulans*. Molec. Microbiol. 1: 275-281.

MacRae, W.D., Buxton,F.P., Sibley,S., Garven,S., Gwynne,D.I., Davies, R.W. and Arst, H.N. (1988). A phosphate-repressible acid phosphatase gene from *Aspergillus niger*: its cloning, sequencing and transcriptional analysis. Gene 71: 339-348.

Malardier, L., Daboussi, M.J., Julien,J., Scazzocchio, C. and Brygoo, Y. (1989). Cloning of the nitrate reductase gene (*nia*D) of *Aspergillus nidulans* and its use for transformation of *Fusarium oxysporum*. Gene 78: 147-156.

Maldonado, M.C., de Saad, A.M.S. and Callieri, D. (1989). Catabolite repression of the synthesis of inducible polygalacturonase and pectinesterase by *Aspergillus niger sp*. Curr. Microbiol. 18: 303-306.

May, G.S., Tsang, M.L.S., Smith, H., Fidel, S. and Morris, N.R. (1987). *Aspergillus nidulans* β -tubulin genes are usually divergent. Gene 55: 231-243.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

McKnight, G.L., Kato, H., Upshall, A., Parker, M.D., Saari, G. and O'Hara, P.J. (1985). Identification and molecular analysis of a third *Aspergillus nidulans* alcohol dehydrogenase gene. EMBO J. 4: 2093-2099.

McKnight, G.L., O'Hara, P.J. and Parker, M.L. (1986). Nucleotide sequence of the triosephosphate isomerase gene from *Aspergillus nidulans*: implications for a differential loss of introns. Cell 46: 143-147.

McLauchlan, J., Gaffney, D., Whitten, J.L. and Clements, J.B. (1985). The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNAs 3' termini. Nucleic Acids Res. 13: 1347-1368.

Miller, B.L., Miller, K.Y., Roberti, K.A. and Timberlake, W.E. (1987). Position dependent and independent mechanisms regulate cell specific expression of the SpoC1 gene cluster of *Aspergillus nidulans*. Mol. Cell. Biol. 7: 427-434.

Mirabito, P.M., Adams, T.H. and Timberlake, W.E. (1989). Interactions of three sequentially expressed genes control temporal and spatial specificity in *Aspergillus nidulans*. Cell 57: 859-868.

Morris, N.R. (1976). Mitotic mutants of *Aspergillus nidulans*. Genet. Res. Camb. 26: 237-254.

Morris, N.R. (1989). The study of cytoskeletal proteins and mitosis using *Aspergillus* molecular genetics. Cell Motility and the Cytoskeleton. 14: 58-61.

Mullaney, E.J., Hamer, J.E., Roberti, K.A., Yelton, M.M. and Timberlake, W.E. (1985). Primary structure of the *trp*C gene from *Aspergillus nidulans*. Molec. Gen. Genet. 199:37-45.

Nussinov,R., Owens,J. and Maizel, J.B. (1986). Sequence signals in eukaryotic upstream regions. Biochim. Biophys. Acta. 866: 109-119.

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

Orbach, M.J., Porro, E.B. and Yanofsky, C. (1986). Cloning and characterization of the gene for beta-tubulin from a benomyl resistant mutant of *Neurospora crassa* and its use as a dominant selectable marker. Mol. Cell. Biol. 6: 2452-2461.

Osmani, S.A., Pu, R.T. and Morris, N.R. (1988). Mitotic induction and maintenance by overexpression of a cell cycle specific protein kinase homologue (*nim*A) in *Aspergillus*. Cell 53: 237-244.

Page, M.M. and Cove, D.J. (1972). Alcohol and amine catabolism in the fungus *Aspergillus nidulans*. Biochem. J. 127: 17P

Papa, K.E. (1986). Heterokaryon incompatibility in *Aspergillus flavus*. Mycologica. 78: 98-101.

Pateman, J.A., Doy, C.H., Olsen, J.E., Norris, U., Creaser, E.H. and Hynes, M.J. (1983). Regulation of alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (AldDH) in *Aspergillus nidulans*. Proc. R. Soc. Lond. B217: 243-264.

Pickett, M., Gwynne, D.I., Buxton, F.P., Elliot, R., Davies, R.W., Lockington, R.A., Scazzocchio, C. and Sealy-Lewis, H.M. (1987). Cloning and characterization of the *ald*A gene of *Aspergillus nidulans*. Gene 51: 217-226.

Proudfoot, N.J. and Brownlee, G.G. (1976). 3' non-coding region sequences in eukaryotic messenger RNA. Nature 263: 211-214.

Punt, P.J., Dingemanse, M.A., Jacobs-Meijsing, B.J.M., Pouwels, P.H. and van den Hondel, C.A.M.J.J. (1988). Isolation and characterization of the glyceraldehyde-3-phoshate dehydrogenase gene of *Aspergillus nidulans*. Gene 69: 49-57.

Raper, K.B. and Fennell, D.I. (1965). The genus *Aspergillus*. Williams and Wilkins Co. Baltimore.

Reinert, W.R., Patel, V.B. and Giles, N.H. (1981). Genetic regulation of the *qa* gene cluster of *Neurospora crassa*: induction of *qa* messenger ribonucleic acid and dependency on *qa*-1 function. Mol. Cell. Biol. 1: 829-835.

Richardson, I.B., Hurley, S.K. and Hynes, M.J. (1989). Cloning and molecular characterisation of the *amd*R controlled *gat*A gene of *Aspergillus nidulans*. Molec. Gen. Genet. 217: 118-125.

Roberts, T.J., Martinelli, S.D. and Scazzocchio, C. (1979). Allele specific, gene unspecific suppressors in *Aspergillus nidulans*. Molec. Gen. Genet. 177: 57-64.

Sanderman, R.A. and Hynes, M.J. (1989). Isolation of the *fac*A (acetyl-Coenzyme A synthetase) and *acu*E (malate synthase) genes of *Aspergillus nidulans*. Molec. Gen. Genet. 218: 87-92.

Sanger, R., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. 74: 5463-5467.

Seegmiller, J.E. (1953). Triphosphopyridine nucleotide-linked aldehyde dehydrogenase from yeast. J. Biol. Chem. 201: 629-637.

Sealy-Lewis, H.M. and Lockington, R.A. (1984). Regulation of two alcohol dehydrogenases in *Aspergillus nidulans*. Curr. Genet. 8: 253-259.

Serfling, E., Jasin, M. and Schaffner, W. (1985). Enhancers and eukaryotic gene transcription. TIG 1: 224-230.

Smith, J.E. and Pateman, J.A. (1977). (eds.) Genetics and physiology of *Aspergillus*. Academic Press, N.Y.

Sophianopoulou, V. and Scazzocchio, C. (1989). The proline transport protein of *Aspergillus nidulans* is very similar to amino acid transporters of *Saccharomyces cerevisiae*. Molec. Microbiol. 3: 705-714.

Tamaki, N., Nakamura, M., Kimura, K. and Hama, T. (1977). Purification and properties of aldehdye dehydrogenase from *Saccharomyces cerevisiae*. J. Biochem. 82: 73-79.

Thuring, R.W.J., Sanders, J.P.M. and Borst, P. (1975). A freeze-squeeze method for the recovery of long DNA from agarose gels. Anal. Biochem. 66: 213-220.

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

Timberlake, W.E. (1978). Low repetitive DNA content in *Aspergillus nidulans*. Science 202: 973-975.

Timberlake, W.E. and Marshall, M.A. (1988). Genetic regulation of development in *Aspergillus nidulans*. TIG 4: 162-169.

Tyler, B.M., Geever, R.F., Case, M.E. and Giles, N.H. (1984). *Cis*-acting and *trans*-acting regulatory mutations define two types of promoters controlled by the *qa*-1F gene of *Neurospora*. Cell 36: 493-502.

Ward, M. and Turner, G. (1986). The ATP synthase subunit 9 gene of *Aspergillus nidulans*: sequence and transcription. Molec. Gen. Genet. 205: 331-338.

Ward, M., Wilson, L.J., Carmona, C.L. and Turner, G. (1988). The *oli*C3 gene of *Aspergillus niger*: isolation, sequence, and use as a selectable marker for transformation. Curr. Genet. 14: 37-42.

Waring, R.B., May, G.S. and Morris, N.R. (1989). Characterization of an inducible expression system in *Aspergillus nidulans* using *alc*A and tubulin-coding genes. Gene 79: 119-130.

Whittington, H.A., Grant, S., Roberts, C.F., Lamb, H. and Hawkins, A.R. (1987). Identification and isolation of a putative permease gene in the quinic acid utilization (*qut*) gene cluster of *Aspergillus nidulans*. Curr. Genet. 12: 135-139.

Wilson, L.J., Carmona, C.L. and Ward, M. (1988). Sequence of the *Aspergillus niger pyr*G gene. Nucleic Acids Res. 16: 2339.

van Hartingsveldt, W., Mattern, I.E., van Zeijl, C.M.J., Pouwels, P. and van den Hondel, C.A.M.J.J. (1987). Development of an homologous transformation system for *Aspergillus niger* based on the *pyr*G gene. Molec. Gen. Genet. 206: 71-75.

Yamada, H. and Adachi, O. (1971). Amine Oxidase (*Aspergillus niger*). pp. 705-709. In. H. Tabor and C.W. Tabor (eds.), Methods in Enzymology, vol. 17B, Adademic Press, N.Y.

Yannisch-Perron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33: 103-119.