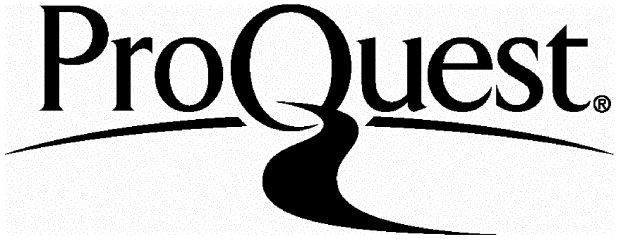
TITLE

GENETIC AND MOLECULAR ANALYSIS OF THE QUINIC ACID UTILIZATION *(QUT)* GENE CLUSTER IN *ASPERGILLUS NIDULANS*

by

ANTONIO JORGE FRANCISCO DA SILVA

THESIS SUBMITTED FOR THE DEGREE OF Ph.D. IN THE DEPARTMENT OF GENETICS, UNIVERSITY OF LEICESTER



ACKNOWLEDGEMENTS

The work of this thesis was carried out in the Department of Genetics of the University of Leicester and I am grateful to all past and present members who have given me help and advice, particularly Dr. Alastair Hawkins and Dr. Clive Roberts for their supervision, encouragement and constructive and patient criticism during the production of this thesis. I would also like to thank the members of Dr. Clive Roberts laboratory for their collaboration in parts of the experimental work.

My special thanks to Professor Norman Giles, Dr. Mary Case and co-workers for their generous gifts of recombinant plasmids, which permitted accelarated progress in the work of this thesis.

And last but not least, my deepest thanks to my parents for their continuous support and encouragement.

TO MY PARENTS

CONTENTS

ABBREVIATIONS

bp	base pairs
BSA	bovine serum albumin
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DTT	dithiothreitol
EDTA	ethylenediaminetetra-acetic acid
kb	kilobases (1000x bp)
mRNA	messenger ribonucleic acid
PEG	polyethylene glycol
RNA	ribonucleic acid
RNase	ribonuclease
SDS	sodium dodecyl sulphate
tris	tris(hydroxymethyl)aminomethane
dNTP	2'-deoxy(N) 5'-triphosphate N=adenosine, cytidine, guanosine or thymidine
ddNTP	2',3'-dideoxy(N) 5'-triphosphate N=adenosine, cytidine, guanosine or thymidine

CHAPTER 1. INTRODUCTION

The mechanisms determining the transient control of gene expression have been very extensively studied in bacteria, and the first system to be comprehensively analysed was the *Lac* operon of *Escherichia coli*, by Jacob & Monod (1961).

They established that the co-ordinate induction of the three enzymes concerned with lactose metabolism, β-galactosidase, galactoside permease and thio-galactoside transacetylase, was directed by a segment of DNA which codes for a single messenger RNA (mRNA) carrying all the necessary genetic information for the synthesis of the three enzymes on the ribosomes. Such a complex mRNA transcript was termed polycistronic, and the complex gene the *Lac* operon. Furthermore, they revealed that a closely linked region controlled the transcription of this polycistronic mRNA. This region codes for a repressor protein that binds to another region called the operator, close to the start of transcription, thus preventing expression of the *Lac* operon. The interaction of the inducer, lactose, with the repressor protein results in the repressor no longer being able to bind to the operator. In this case RNA polymerase can interact with the *Lac* operon at a 5' non-coding region, the promoter region, to intiate synthesis of the polycistronic mRNA.

A third protein involved in the transcription of the *Lac* operon, along with the repressor and the RNA polymerase, is the catabolic activator protein (CAP).

Jacob & Monod showed that the *Lac* operon was repressed when glucose, a more efficient source of carbon than lactose, was present. Glucose is found to repress the concentration of intracellular cyclic AMP, which is required for the activation of CAP (catabolic activator protein), which in turn operates by enhancing the activity of the RNA polymerase.

These observations put forward by Jacob & Monod established the fundamental principles of molecular mechanisms controlling gene expression, which were to form the basis for future models. Subsequent investigations revealed that in bacteria, genes controlling enzymes of catabolic and biosynthetic pathways were frequently found to be organized in operons, similar to the *Lac* operon. In these, the expression of the polycistronic message appears in most cases to be controlled by a repressor protein, although, less commonly, the control can also be exerted by a positively acting, activator protein, as in the *ara* (arabinose) operon (Englesberg *et al.*, 1965).

Advances in the understanding of the mechanisms of the regulation of gene expression in lower eukaryotes, have been largely achieved in fungi, where a comprehensive number of genes that are under physiological control have been identified. In the yeast *Saccharomyces cerevisiae,* examples of such genetic systems that have been more thoroughly analysed include the genes involved in galactose utilization (Douglas & Hawthorn, 1964, 1966; St.John & McDonnel, 1979; Post-Beittenmiller *et al.*, 1984), the glucose repressible alcohol dehydrogenase (Ciriacy, 1979; Denis & Young, 1983), the alkaline and acid phosphatases (Thoh-e *et al.*, 1981) and general

amino-acid control (Jones & Fink,1983; Lucchini *et al*, 1984; Penn *et al.*, 1984). Among the filamentous fungi some of the more thoroughly analysed systems of genetic regulation include the genes controlling the derepression of phosphatases in *Neurospora crassa*, which show remarkable similarities with the equivalent system in yeast (Metzemberg & Chia, 1979) and in *Aspergillus nidulans* the genes encoding the alcohol and aldehyde dehydrogenases (Pateman *et al.*, 1983), the genes for proline utilization (Arst & MacDonald, 1978) and the amide utilization *amd*S gene (Hynes *et al.*, 1983).

In contrast to bacteria where the genes coding for the enzymes in a particular metabolic pathway frequently occur organized in operons, related genes in fungi tend to be scattered throughout the genome and have their expression controlled by more than one, unlinked regulatory gene.

However, exceptions to this general situation in fungi have been found, for example close linkage of the three structural genes, *GAL*1, *GAL*7 and *GAL*10 of *S. cerevisiae*, which code for the three enzymes converting galactose to glucose-1-phosphate (Douglas & Hawthorn, 1964, 1966) and also the genes involved in proline utilization, *prn*A, *prn*B, *prn*C and *prn*D in *A. nidulans* (Arst & MacDonald, 1978).

Another remarkable case of closely linked regulated genes in fungi, is the gene cluster coding for the enzymes catalyzing the first three reactions in the degradation of quinic acid (Figure 1), a product of leaf decay in forests, in at least two filamentous fungi, *Neurospora crassa* (Case & Giles, 1976) and *Aspergillus nidulans*

(Hawkins, Giles & Kinghorn, 1982).

The system has been more extensively investigated in *N. crassa*, where studies have defined a gene cluster consisting of the three enzyme structural genes, together with a closely linked regulatory region which mediates the coincident expression of the three structural genes to achieve very high levels of enzyme induction (Chaleff, 1974). In this respect the quinic acid utilization (*qa*) gene cluster in *N. crassa* appeared to possess striking similarities to the prokaryotic operons, although there was no evidence that the three *qa* genes were transcribed in a single polycistronic mRNA (Case & Giles, 1975).

Initial interest in this gene cluster developed from work on another system to which it is metabolically related (Figure 1), and which concerns the complex gene determining the enzyme activities for five steps in the constitutive biosynthesis of aromatic amino acids and many other compounds (Giles *et al.*, 1967a). Initial speculation that such closely linked genes in fungi constitute a system for gene expression similar to the prokaryotic operons, was discounted after analyis of the *HIS4* genetic region of *S. cerevisiae*, which encodes for three enzymes involved in steps 2, 3 and 10 of histidine biosynthesis. Bigelis *et al.* (1977) showed that this region was transcribed to produce a single mRNA corresponding to a single, multifunctional protein catalyzing all three enzyme fuctions. Similarly, the *arom* complex is now thought to consist of a single gene coding for a single polypeptide, which catalyzes the five enzyme activities for steps two to six in the biosynthetic pathway (Lumsden & Coggins, 1977; Giles, 1978 and Figure 1). Furthermore,

such a complex gene is found not only in *N. crassa*, but also in a number of related fungi: *Ustilago violacea* (Berlyn & Giles, 1972), *S. cerevisiae* (deLeeuw, 1968) and *A. nidulans* (Clive Roberts, unpublished results). In *E. coli*, on the other hand, the five biosynthetic enzymes are coded by five unlinked genes and are readily separated physically and the quinic acid degradative pathway does not exist. A possible explanation for this difference lies in the metabolic competition that exists in these fungi between the biosynthetic and the quinic acid degradative pathways, since both share a common enzymatic step and metabolic intermediates (Figure 1). Such metabolic competition between the two systems may have conferred an advantage for the aggregation and eventually fusion into the *arom* cluster, thus creating a channeling mechanism for intermediates on the enzyme complex.

An indication of the importance of metabolic channeling comes from the methods used to isolate *arom*-9 mutants in *N. crassa* which lack biosynthetic dehydroquinase activity. Because of the high levels of induction of the quinic acid catabolic enzymes and since a lesion in the biosynthetic enzyme would lead to the accumulation of dehydroquinic acid (DHQ), *arom*-9 mutants would actually be able to grow on quinic acid due to endogenous induction of the degradative dehydroquinase. The procedure used to recover *arom*-9 mutants was to first screen for catabolic dehydroquinase mutants (*qa*2$^-$), which are easier to isolate since the constitutive levels of the biosynthetic isoenzyme are not high enough to fully support growth on quinic acid and thus enable the identification of mutants in the catabolic function. Only

then was it possible to isolate the biosynthetic *arom*-9 mutation in a double mutant strain lacking each of the isoenzymes (Giles *et al.*, 1967b).

The analysis of the *qa* mutant strains generated during the screen for *qa*2 mutants revealed quite unexpected results. The first impression was that the mutants identified a system possessing considerable resemblence to the bacterial operons, since not only were the three structural genes clustered, but also their expression was regulated from a closely linked region (Case & Giles, 1975).

However, subsequent genetic and biochemical evidence has shown that although the induction of the three enzymes followed co-incident expression of the three structural genes, these were independently transcribed (Patel *et al.*, 1981). Regulation of expression of the structural genes appeared to be effected by a closely linked control region (Case & Giles, 1975). This region was identified on the basis of a set of pleiotropic, non-inducible mutants (carrying the *arom*-9 mutation) that mapped in the region. Two distinct classes of mutations identified two discrete, non-overlapping arrays within the regulatory region. These mutations were distinguished by contrasting phenotypes based on their ability to relieve an auxotrophic requirement for aromatic amino acids in a *qa*2;*arom*-9 double mutant in heterokaryons grown on quinic acid as carbon source. One group called $qa1^F$ (F for fast complementation) mutants were readily complemented, whereas a second group produced very slow complementation, hence called $qa1^S$. Furthermore, revertants of $qa1^S$ mutant strains, but not $qa1^F$ strains, selected by growth on

quinic acid, were frequently found to express the *QA* structural genes constitutively. These *qa*1^{C} strains, carried additional mutations mapping close to the *qa*1^{S-} mutations (Valone *et al.*, 1971; Partridge *et al.*, 1972). The model developed from these genetic observations was the following: the *QA*1 locus codes for a positively acting regulatory protein that responds to the presence of the inducer molecule, quinic acid. The two types of mutations *qa*1^{S-} and *qa*1^{F-}, identify two domains within this positively acting regulatory protein. The first corresponds to a site binding quinic acid and thus producing allosteric changes in the protein enabling it to bind to the promoter regions of the structural genes and initiate high levels of *QA* mRNA synthesis. The DNA binding region of the *QA*1 protein, was the domain identified by the *qa*1^{F} mutations.

Vapnek *et al.*, 1977, reported the physical isolation of the gene cluster in a plasmid vector on the basis of the expression of the *N. crassa* dehydroquinase gene (*QA*2) in *E. coli*, and thus restoring a mutant lesion in the bacterial biosynthetic gene (*aro*D6) and enabling the selection of transformants by growth in the absence of added aromatic amino acids.

Further analysis of the quinic acid utilization gene cluster and its regulatory mechanisms, in *A. nidulans*, presents some advantages afforded by this organism. *A. nidulans* is an ascomycetous fungus which has been extensively exploited for the investigation of the regulatory processes controlling a variety of metabolic processes, particularly in nitrogen (Kinghorn & Pateman,

1977; Cove, 1979) and carbon (McCullough, Payton & Roberts, 1977; Arst, 1981) assimilation. More recently, the organism has also proved a good tool for the analysis of such cellular processes as mitosis and nuclear movements (Oakley & Morris, 1980; Morris *et al.*, 1982; and Bergen *et al.*, 1984) and development and cell differentiation (Clutterbuck, 1977; Champe *et al.*, 1981; Timberlake *et al.*, 1983).

A. nidulans offers a main advantage over other filamentous fungi in the availability of a powerful system of genetic analysis both through sexual and parasexual processes. Other advantages conferred by this system include, fast growth rates, a colonial growth habit, uninucleate conidiospores and haploid and diploid growth phases (Pontecorvo *et al.*, 1953). *A. nidulans* has also become an important organism in research in general metabolism, since it is considered a good analogue for many filamentous fungi of industrial importance used in the production of organic acids and anti-biotics (Smith & Pateman, 1977).

Preliminary analysis of the quinic acid utilization system in *A. nidulans* by Hawkins, Giles & Kinghorn (1982), revealed a very similar situation with the *qa* gene cluster in *N. crassa*. In their work, quinic acid non-utilizing (*qut*) mutants were isolated and characterized phenotypically. The outcome was the identification of three of the four possible classes of mutants characterized in *N. crassa*, namely mutants lacking the dehydrogenase (*qut*B) and dehydratase (*qut*C) activities and another group in which all three enzyme activities were absent (*qut*A). These, by analogy with the model proposed

for *Neurospora*, were interpreted in the same way; that is to identify a gene coding for a positive regulator. No mutants deficient in dehydroquinase activity were isolated. The *qut*A,B and C mutant genes were shown to be tightly linked and the *qut*A mutants would complement structural mutants, hence proving they had retained their potential enzyme activities for at least two of the structural genes. On the basis of this evidence, it was postulated that *QUT*A codes for a regulatory protein that controls the synthesis of the three enzymes with a positive mode of action. The *qut*A mutants therefore fail to induce the structural genes.

Biochemical studies done on the wild-type have demonstrated further analogies between the *QA* system in *Neuropora* and its *QUT* equivalent in *Aspergillus*. Dehydroquinase enzyme from *A. nidulans* is functionally and structurally very similar to its counterpart in *N. crassa*. It is a heat-stable protein, retaining 98% of its activity after incubation at 71^{o}C for 30 minutes. Likewise, it has the same native molecular weight and sub-unit structure as the enzyme from *N. crassa*, that is a molecular weight of approximately 200,000, composed of some 10 sub-units. The *A. nidulans* enzyme was also shown to exhibit a strong immunological cross-reactivity with anti-body raised against the *N. crassa* enzyme, suggesting a strong amino-acid sequence homology.

AIMS OF THE PROJECT

The general thrust of the project is to further analyse the regulation of the expression of the *QUT* gene

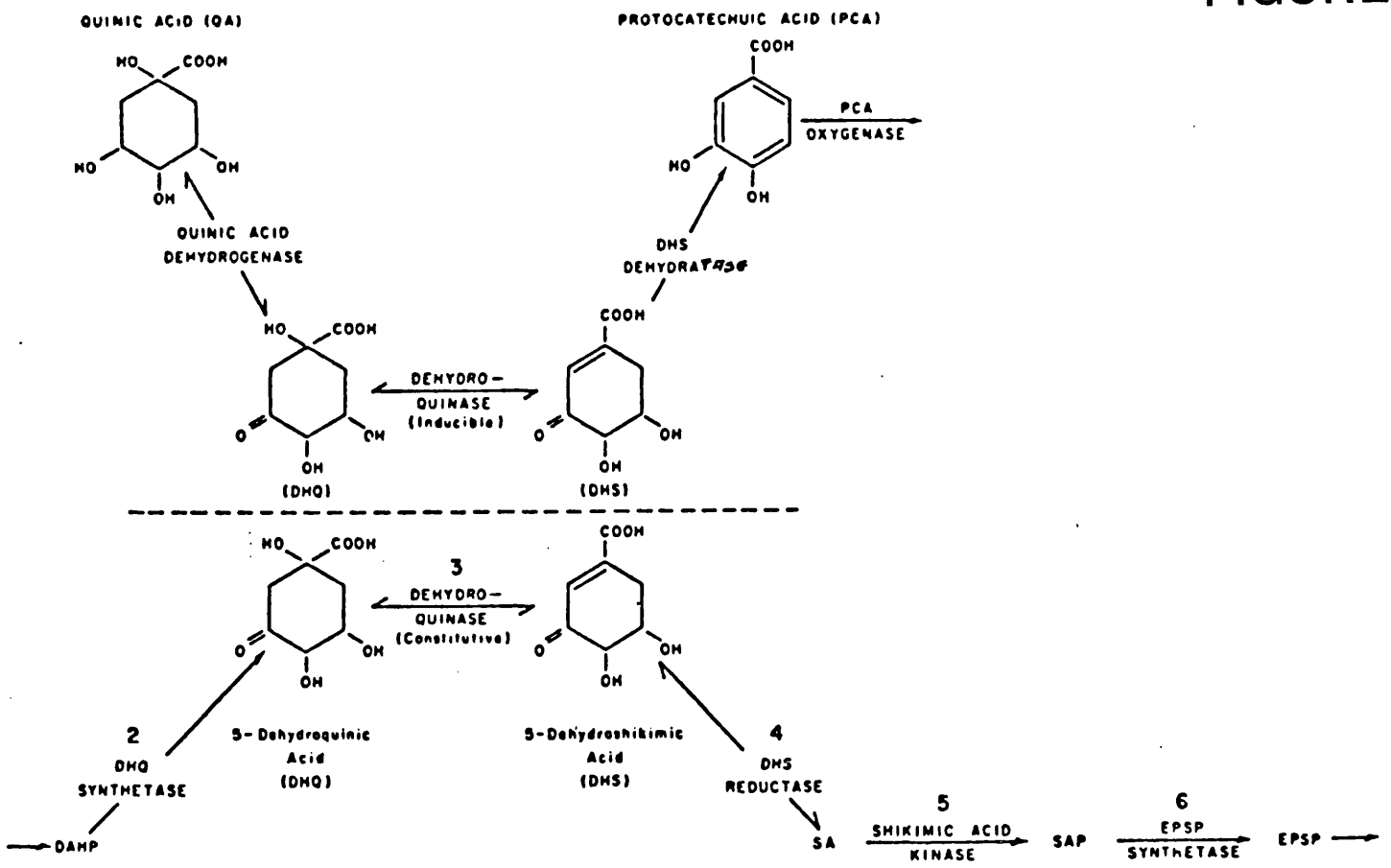
cluster in *A. nidulans*.

The initial aims were the isolation of dehydroquinase mutants, so that mutations representative of each of the structural genes were available and also to test the predicted linkage to the cluster. Furthermore, because the regulation of particular functions in fungi seem to generally involve two or more regulatory proteins it was reasonable to consider the possibility that this could also be the case for the *QUT* gene cluster in *A. nidulans*. Hence, more mutants in regulatory functions were desirable to test this possibility. *Aspergillus* has an advantage in the opportunity to construct both heterokaryons and diploid strains, the latter not being available in *N. crassa*. Diploid strains have a major advantage over heterokaryons in that the allele ratios in the cell for any given gene can be controlled and therefore the analysis of mRNA transcription and protein activity quantified, and genetic interactions accurately assessed in heterozygous diploid strains. This advantage, it was felt should enable a more detailed and accurate examination of the system *in vivo*, both at the genetic and molecular levels, than in *Neurospora*.

Further insights into the mechanism regulating expression of the *QUT* genes was expected from a study of the sequences flanking the coding regions, with particular emphasis on the promoter regions. This study was prompted by the idea that if the regulatory protein(s) bind to these regions, sequence analysis might reveal features reflecting such interactions.

FIGURE 1 Diagram of the reactions in the quinic acid catabolic (inducible) and polyaromatic synthetic (constitutive) pathways, indicating the metabolic relationships.

FIGURE 1


QUINIC ACID (QA)
PROTOCATECHUIC ACID (PCA)
PCA OXYGENASE
QUINIC ACID DEHYDROGENASE
DHS DEHYDRATASE
DEHYDRO-QUINASE (Inducible)
(DHQ)
(DHS)
3
DEHYDRO-QUINASE (Constitutive)
2
DHQ SYNTHETASE
5-Dehydroquinic Acid (DHQ)
5-Dehydroshikimic Acid (DHS)
4
DHS REDUCTASE
DAHP
SA
5
SHIKIMIC ACID KINASE
SAP
6
EPSP SYNTHETASE
EPSP

CHAPTER 2 . MATERIALS AND METHODS

A. Genetics and Biochemistry of *Aspergillus nidulans*

1. Strains

The strains of *A. nidulans* (Eidam Winter) all derive from the Glasgow stock (Pontecorvo *et al.*,1953). The strains R21 (*paba*A1 *y*A2) and R153 (*w*A3;*pyro*A4), constructed from the translocation free strains of Käfer (1965), were used as standard "wild-type" strains and also for the isolation of mutants. Quinic acid non-utilizing mutant strains *qut*A4, *qut*A7, *qut*A8, *qut*B42 and *qut*C1 were isolated in R21 (Hawkins et al, 1982). Strain R56 (*paba*A1; *w*A3; *orn*B7 *fac*B101 *ribo*B2 *gal*7) was used for meiotic mapping of the *qut* gene cluster.

Strains were cultured on complete medium slants for medium term storage and held on silica gel (Roberts, 1969) for long term storage. Stocks were kept at 4 ^{o}C.

2. Media

(i) Minimal medium

The defined minimal medium (MM) used throughout this work is basically that of Pontecorvo *et al.* (1953), modified by Roberts (1963). This is prepared as a 10X solution of stock salts (see below) to which 1.0M $MgSO_4$ is added to give a final concentration of 10mM.

10X Stock salts solution

Substance	Final amount per litre
$NaNO_3$	6.0g
KCl	0.52g
KH_2PO_4	1.52g
Trace elements solutions	1.0 ml

Trace elements solution

Substance	Concentration in stock solution (g/l)
$FeSO_4.7H_2O$	1.00
$ZnSO_4.7H_2O$	8.80
$CuSO_4.5H_2O$	0.40
$MnSO_4.4H_2O$	0.15
$Na_2B_4O_7.10H_2O$	0.10
$(NH_4)_6Mo_7O_{24}$	0.05

Trace elements were added to distilled water, and undissolved crystals brought into solution by the addition of conc. HCl (2ml) directly onto the crystals, and the solution made to the final volume by the addition of deionised water.

(ii) Culture media

Minimal media (MM) plus $MgSO_4$ for plates was solidified with 1.5% (w/v) agar. Minimal media for liquid culture contained the wetting agent Tween 80, diluted 10^{-5} (v/v), and sterile $MgSO_4$ solution was added to a final concentration of 10mM after autoclaving the medium.

The nutritional supplements for auxotrophic strains were those described previously (Armitt et al, 1976), and were added as required before autoclaving the media:

Supplement	Stock solution (mg/ml)	Amount added (ml/l)
p-Aminobenzoic acid (*paba*)	1.0	1.0
Pyridoxin.HCl (*pyro*)	0.5	1.0
Ornithine (*orn*)	2.0	1.0
Riboflavin (*ribo*)	2.5	1.0

The mutant allele *fac*B101 (Apirion,1965) ≡ *acu*B101, was tested by the inability of *fac*B strains to grow on acetate as sole carbon source.

Malt extract agar (MEA) is a complex medium consisting of:

Substance	Composition (g/l)
Malt extract	20
Bacto-peptone	1
Agar	15

MEA was used when fast growth and good conidiation of strains was required. It was generally supplemented with p-aminobenzoic acid and pyridoxin.

All media were sterilized by autoclaving at 15lb pressure for 15 minutes.

(iii) Carbon sources

Carbon sources were added aseptically to sterile media as follows:

Carbon source	Final concentration
Glucose (solid medium)	40 mM
Glucose (liquid medium)	10 mM
Quinic acid (Growth media)	1% (w/v)
Quinic acid (Induction media)	0.1% (w/v)
Protocatechuic acid (PCA)	1% (w/v)
Acetate	100mM

Quinic acid and PCA were prepared in 10% stock solutions (in distilled H_2O), brought to pH6.5 with NaOH and sterilized by filtration through a 0.45μm membrane filter (Whatman). 1M glucose and 2M acetate (pH6.5 with NaOH) stock solutions, were sterilized by autoclaving at 10lb for 10 min.

(iv) Incubation temperature

All cultures were grown at 37^{o}C unless otherwise stated.

3.Preparation of suspensions of conidiospores

Dilute suspensions of conidiospores were spread on the surface of MEA plates at such a concentration to produce confluent growth and good conidiation. Each plate was then flooded with approximately 20 ml of sterile "Tween-Saline", a solution of 0.08% (w/v) NaCl plus 0.025% (v/v) of the wetting agent Tween 80.

The conidia were harvested by drawing a sterile wire across the surface of the plate. The suspensions were collected in sterile 25 ml disposable universal bottles and vigourously shaken on a "Whirlimixer" to disrupt conidial chains. The suspensions were then washed by repeated centrifugation in neutral phosphate buffer (20mM

KH_2PO_4; 50mM Na_2HPO_4; 50mM NaCl; 0.4mM $MgSO_4.7H_2O$; final pH7.2), resuspended in 10ml of the same buffer and stored at 4°C. The conidiospores retain good viability for at least 4 weeks.

4. Mutagenesis and Filtration Enrichment

Suspensions of recently harvested conidiospores in neutral phosphate buffer at a concentration of 10^8 spores ml^{-1}, were treated with ultra-violet light (UV). A suspension of conidiospores in a petri dish was exposed to UV light at a dose rate of 0.5J m^{-2} for 2.5 min on a rotating platform to ensure uniform illumination. Sufficient exposure was given to produce approximately 50% killing.

The filtration enrichment technique used was based on the method of Woodward *et al.*, (1954). The UV irradiated conidiospores were immediately innoculated into liquid growth medium with quinic acid as carbon source. After incubation for 10 hrs at 30°C, the suspension was filtered through three double layers of sterile cotton gauze to remove growing mycelium and the filtrate incubated for a further 8 hrs. when it was filtered again. This procedure was repeated three times at two hour intervals at which point no further mycelial growth was observed.

The final suspension was stored at 4°C, the viable count determined by plating aliquots on MEA. Surviving conidiospores were poured in glucose MM agar plates to yield 30 - 50 colonies per dish and the colonies screened by replica plating to identify the quinic acid non-utilizing mutants.

5. Standard techniques for genetic analysis

The techniques used were basically those described by Pontecorvo *et al.*, (1953).

Sexual crosses were set up between haploid strains carrying complementary nutritional markers. Dense suspensions of conidiospores of the two strains were mixed on the surface of thick MM plates and "dug" into the medium with a sterile wire-loop. The plates were then sealed with adhesive tape and incubated at 30°C for about 3 weeks.

Fruiting bodies (cleistothecia) were picked from mature crosses and rolled across the surface of 4% (w/v) hard agar plates to remove contaminating mycelia and conidiospores. Individual cleistothecia were then crushed in 1 ml of neutral phosphate buffer, and a sample of the resultant ascospore suspension streaked on MEA to identify hybrid cleistothecia by the presence of recombinant conidial colours. Suitable dilutions of ascospore suspensions from hybrid cleistothecia were plated to produce separate colonies, and the progeny tested on appropriate media for the segregation of markers.

Heterokaryons were synthesized by mixing conidia of two strains carrying complementary nutritional and colour markers on MEA plates and incubated overnight. Small blocks of agar containing mycelium were then transferred to thick MM plates containing no added nutritional supplements and on incubation heterokaryons could be identified after 3 to 4 days as rapidly growing sectors. These heterokaryons were then subcultured on fresh MM plates by repeated transfer to agar blocks containing

mycelium.

Heterozygous diploid strains were selected by plating suspensions of conidiospores harvested from heterokaryons in MM (Roper, 1952) and replated for single cell isolation on the same medium. The diploid strains were verified by plating on MEA containing benomyl (0.7 *mg*/ml) which produces greatly enhanced mitotic segregation (Hastie, 1970).

6. Preparation of cell-free extracts for enzyme assays

Mycelium was grown in liquid culture using 2 litre Erlenmeyer flasks with vertical baffles, formed by indenting the walls of the flask, and treated with water repellent silicone film (Armitt *et al.*, 1976). Each flask contained 250ml liquid MM and was innoculated with a suspension of conidiospores to yield 10^6 spores ml^{-1}. Cultures were grown with glucose as carbon source for 16 hours at 30°C on a gyratory shaker. Mycelium was harvested aseptically by filtration through Whatman No1 filters, washed with salts medium and resuspended in growth medium with either glucose (non-induced culture) or quinic acid (induced culture) as carbon source, and incubated for a further 4 hours. It was then harvested by filtration, washed thoroughly on the filter with buffer (0.1M potassium phosphate,pH7.2; 10mM EDTA; 1mM DTT) and stored at -18°C wrapped in aluminium foil.

A thin pad of frozen mycelium (about 2g) was immersed in liquid nitrogen in a pre-cooled mortar and ground to a fine powder with a cold pestle. The excess liquid nitrogen

was decanted and the frozen powdered mycelium resuspended in 1ml extraction buffer (0.1M potassium phosphate,pH7.2; 10mM EDTA; 1mM DTT; 1mM PMSF -phenylmethylsulphonyl fluoride-). After thawing, the mycelium was extracted by gently shaking the slurry in a polypropylene tube, on ice for 1 hour, after which time cell debris was removed by centrifugation for 15 minutes in an Eppendorf microfuge in a cold-room. The supernatant was decanted, stored on ice and the enzymes assayed within six hours of extraction. A sample of the supernatant was heated at 71°C for 10 minutes to denature biosynthetic dehydroquinase activity (Kinghorn & Hawkins, 1982). The heated sample was cooled on ice and recentrifuged to pellet denatured proteins

7. Enzyme assays

Each of the three enzymes was assayed in a Unicam (Cambridge) SP1800 recording spectrophotometer at 37°C in 1ml reaction mixtures, taking 5 and 10μl samples of the cell-free extracts.

(i) Quinate dehydrogenase (enzyme 1) was assayed by following the reduction of NAD at 340 nm in a reaction mixture containing 10mM Tris.HCl (pH8.6); 1mM quinic acid; 2.5mM NAD.

(ii) Catabolic dehydroquinase (enzyme 2) was assayed by monotoring the increase in absorbance at 240nm in a reaction mixture containing 10mM Tris.HCl (pH7.2); 1mM EDTA; 5μl dehydroquinate solution (see source of materials). The volume of dehydroquinate solution added gives an absorbance at 220nm of 0.2 before addition of cell-free extract. Thermolabile biosynthetic

dehydroquinase had been inactivated by heating (see above).

(iii) 3-Dehydroshikimate dehydratase (enzyme 3) was assayed by measuring the increase in absorbance at 290nm in a reaction mixture containing 0.1M Tris.HCl (pH7.5);2.5mM $MgCl_2$ and 5μl dehydroshikimate solution. The latter was freshly prepared enzymically from dehydroquinate (5μl, A220nm 0.2) using purified *A. nidulans* catabolic dehydroquinase (Hawkins et al, 1982).

8. Protein estimations

The concentration of soluble protein was assayed in each cell-free extract by the method of Lowry et al (1951). Calibration curves were constructed using bovine serum albumin standards each time protein estimations were performed.

9. Source of materials

All reagents were of analar or greater purity unless otherwise indicated. Dehydroquinate and *A. nidulans* catabolic dehydroquinase were supplied by Dr.A Hawkins (Hawkins et al, 1982). Quinic acid and PCA (Procatechuic acid) were supplied by Sigma. Agar was from Davis "New Zealand" source.

B. Recombinant DNA Techniques for *Aspergillus nidulans*

1. Bacterial strains, plasmids and bacteriophage

(i) Strains of *Escherichia coli*

C600 (Appleyard, 1954)

*sup*E, *ton*A, *thr*-1, *leu*B6, *thi*a1, *lac*Y-1

ED8910 (Loenen & Brammar, 1980)

*rec*B21, *rec*C22, *hsd*R, *sup*E44, *met*B, *lac*Y-1

*gal*K2, *gal*T22

Q359 (Karn *et al.*, 1980)

*sup*E44, *ton*A, *hsd*R, P2 lysogen

BHB2688 (Hohn & Murray, 1977; Hohn, 1979)

N205, *rec*A$^-$ (λimm^{434}, C1^{ts}, b2, red$^-$, Eam, Sam)/λ

BHB2690 (Hohn & Murray, 1977; Hohn, 1979)

N205,*rec*A$^-$(λimm^{434}, C1^{ts}, Dam, Sam)/λ

SK3430 (Kushner *et al.*, 1977)

*leu*C, *aro*D6, *thi*a$^-$, *sbc*B15, *hsd*4

JA221 (Clarke & Carbon, 1978)

*hsd*R, *m*K$^+$, *trp*E5D, *leu*B6, *rec*A$^-$, *lac*Y

JM101 (Messing *et al.*, 1981)

Δ (*lac-pro*), *sup*E44, *thi*1.F'*tra*D36, *pro*AB, *lac*I^q, ZΔM15

Strains C600, ED8910, Q359, BHB2688 and BHB2690 were used for the construction and assay of the *A. nidulans* genomic library in λ-phage DB286 (Brammar, 1982). Strains SK3430 and JA221 were used for harbouring recombinant DNA plasmids. JM101 was used for propagation of recombinant M13 phage for DNA sequencing. Reisolation of bacterial strains for new cultures was done by inoculation from

single isolated colonies. Cultures were generally incubated at 37°C.

(ii) **Plasmids**

Plasmid pBR322 (Bolivar *et al.*,1977 a,b) was used in subcloning DNA fragments from recombinant λDNA. Plasmids pVK57 and pVK88 (Kushner *et al.*,1977) and pMSK366 and pMSK375 (Schweizer *et al.*, 1981) containing DNA fragments from the *N. crassa QA* gene cluster were generously supplied by Prof. N. H. Giles and co-workers.

(iii) **Bacteriophages**

The λ replacement vector DB286 (Brammar,1982) was used for the construction of the *A. nidulans* genomic library. M13 sequencing vector M13mp18 and mp19 (Messing and Vieira, 1982) were used in "shotgun" cloning of DNA fragments for M13 sequencing.

2. Media

(i) Luria Broth

Substance	Amount per litre
Difco Bacto Tryptone	10g
Difco Bacto Yeast Extract	5g
NaCl	5g

Luria agar plates were prepared by solidifying liquid media with 15g Difco Bacto agar per litre. Luria broth and agar were used generally for bacterial cultures.

Antibiotics were added to selective media as follows:

Antibiotic	Stock solution (mg/ml)	Final concentration (μg/ml)
Ampicillin (sodium salt)	25	50
Chloramphenicol	34	10
Tetracycline.HCl	12.5	15

Ampicillin was prepared in water, chloramphenicol in ethanol and tetracycline in water/ethanol (50% v/v). Solutions were sterilized by filtration (0.45μm Whatman) and added to molten agar media cooled to 55°C.

(ii) Trypticase Peptone (BBL) Agar

BBL agar was used for phage growth and assays.

Substance	Amount per litre
Trypticase Peptone	10g
NaCl	5g
$MgSO_4.7H_2O$	5g

Media was solidified with 15g/l of Difco Agar. Soft agar overlays were prepared with 6g/l of agar, and supplemented with 10mM $MgSO_4$.

(iii) Glucose MM Agar

This medium was used to maintain *E. coli* strain JM101.

Substance	Amount per 100ml
Difco Agar	1.5g
M9 salts (see below)	10 ml
0.2% B_1 (Thiamine)	0.1 ml
10mM $CaCl_2$/100mM $MgSO_4.7H_2O$	1 ml
20% glucose	2 ml

M9 Salts

Substance	Amount per litre
Na_2HPO_4	60g
KH_2PO_4	30g
NaCl	5g
NH_4Cl	10g

3.Phenol extraction and ethanol precipitation of DNA

Equal volumes of phenol reagent (phenol, 100g; chloroform, 100ml; isoamyl alcohol, 4ml; 8-hydroxyquinoline, 0.1g) saturated with 10mM Tris-HCl,pH7.5 were added for the extraction of proteins from nucleic acid solutions.

DNA solutions were concentrated by precipitating the DNA with two volumes of ethanol after the addition of 0.1 volume of 2M NaAc,pH5.6. The solution was mixed and chilled for 10 minutes in a methanol/dry ice bath for 10 minutes or overnight at -18^{o}C. DNA precipitates were pelleted by centrifugation at 18,000xg for 10 minutes, rinsed in cold 70% (v/v) ethanol, dried under vacuum and resuspended in appropriate volumes of sterile distiled water (DW) or TE buffer (10mM Tris.HCl,ph7.5; 1mM EDTA)

4.Gel Electrophoresis

Horizontal agarose gels (0.8% agarose (w/v)) were prepared and run in electrophoresis buffer (40mM Tris; 5mM NaAcetate; 2mM EDTA; pH8.0 with glacial HAc) containing ethidium bromide at 0.5μg/ml (Aaij and Borst, 1972)

Molecular weight marker DNA standards were provided

by *Hind*III digestion of λDNA (λx*Hind*III) or *Sau*3A digestion of pBR322 (pBR322x*Sau*3A). DNA samples were mixed with 0.1X volume electrophoresis loading buffer (0.25% bromophenol blue, 15% Ficoll400(w/v)), and electrophoresis continued until the bromophenol blue dye travelled to within the last 1/3 of the gel.

DNA binding ethidium bromide in agarose gels was visualized using a short wave lengh ultra-violet transilluminator (Ultra-violet Products Inc, California, USA) and photographed using a Polaroid MP-3 land camera and Polaroid 4X5 Type-55 or Type-57 films.

5. Preparations of *A. nidulans* genomic DNA

A. nidulans genomic DNA was prepared from freshly grown mycelia of strain R153, washed in water, rapidly frozen in liquid N_2 and powdered in a mortar and pestel under liquid N_2(see above). The powdered fragments of mycelium were gently resuspended (at 20 ml g^{-1}) in TE ; 4% (w/v) Sodium dodecyl sulphate (SDS); and 25% (w/v) sucrose, for 15 minutes at room temperature and then the solution extracted with phenol. The solvents were separated by centrifugation at 15,000g for 20 minutes, the aqueous supernatant was collected, made 1 molar with potassium acetate and the resulting precipitate removed by centrifugation at 15,000g for 10 minutes. The supernatant was re-extracted twice with phenol reagent. Nucleic acids were precipitated with ethanol, disolved in TE buffer and the DNA further purified by centrifugation at 110,000xg for 2 hours, at 5°C, to pellet contaminating carbohydrates.

Gel electrophoresis of the DNA prepared by this procedure showed a major population of molecules greater than 50Kb in size with very little low molecular weight DNA or RNA contamination.

6. *A. nidulans* RNA preparation

A. nidulans RNA was prepared by John Clements as previously described (Clements & Roberts, 1985).

7.Plasmid DNA preparations

(i) Small scale plasmid DNA preparation

This follows a modification of the method of Birnboim & Doly, (1979).

Overnight cultures of the plasmid carrying *E. coli* strain were grown in Luria broth plus the appropriate antibiotic. A 1.5 ml sample in an Eppendorf tube was centrifuged for 1 minute and the medium removed by aspiration. The bacterial pellet was resuspended in 100μl of a lysis solution (50mM glucose; 10mM EDTA; 25mM Tris-HCl,ph8.0; 4mg/ml of freshly added lysozyme) and left at room temperature for 5 minutes, when 200μl of a freshly prepared, ice cold solution of 0.2N NaOH, 1%SDS (w/v) was added. The solutions were mixed by inverting the tube several times and held on ice for 5 minutes; then 150μl of 3M potassium acetate (pH4.8 with glacial acetic acid) was added and again the mixed solution left for 10 minutes on ice, when it was centrifuged for 5 minutes, in a precooled microfuge in a 4^{o}C cold room. The supernatant was transferred to a new Eppendorf tube, phenol extracted and

the DNA precipitated with ethanol at room temperature for 2 minutes. The sample was resuspended in 50μl TE buffer containing 20μg/ml of DNase-free RNase.

(ii) Large scale plasmid DNA preparation

This was essentially a scaled up version of the same method of Birnboim & Doly, (1979), with an additional purification step by CsCl gradient centrifugation.

Two 400ml volumes of Luria broth (plus the appropriate antibiotic) were innoculated with 1ml of an overnight culture of the *E. coli* strain carrying the required plasmid and incubated with shaking, overnight. The cells were pelleted by centrifugation at 4,200g, for 5 minutes at 4°C, resuspended in 40ml of ice-cold lysis buffer and held on ice for 5 minutes. A volume of 80ml of ice-cold alkaline/SDS was added, mixed and the preparation held on ice for a further 5 minutes. To this solution 60ml of ice-cold 3M potassium acetate,pH4.8 was added and mixed to precipitate the chromosomal DNA and proteins. The precipitate was removed by centrifugation at 6,000g for 10 minutes at 4°C, and the supernatant passed through polyallomer wool to separate the solution from remaining cell debris and the precipitate.Nucleic acids were precipitated by addition of a 0.5 volume of propano-2-ol followed by centrifugation at 4,200g for 10 minutes at 4°C. The pellet was rinsed with cold 70% ethanol, dried under vacuum and dissolved in 10ml TE buffer, containing 10μg/ml DNase free RNase. The volume was adjusted gravimetically to 20ml with TE, and 4ml of a 5mg/ml ethidium bromide solution added together with 23.76g of AR

grade CsCl, to give a final density of ρ=1.392-1.394. This solution was divided between two Beckman polyallomer "Quickseal" tubes and centrifuged at 15°C at 110,000g for 40 hours in a 50Ti rotor or at 270,000g, 15°C overnight in a 75Ti rotor. The plasmid band was removed, and ethidium bromide removed by repeated extractions with butan-2-ol saturated with 10mM Tris-HCl, pH7.5. Salts were removed by dialysis overnight with several changes of TE buffer, and plasmid DNA ethanol precipitated and redissolved in appropriate volumes of water or TE buffer.

The concentration and quality of the plasmid preparation were determined by optical density at 260nm, where an OD of 1 corresponds to approximately 50μg/ml of DNA. Purity of DNA preparation was estimated from ratio of measurements OD_{260}/OD_{280}. If the preparation is pure, the ratio should be ≈1.8 (Maniatis, 1982).

8. Large scale preparation of M13 bacteriophage DNA

An overnight culture in Luria broth of the host *E. coli* strain JM101 was diluted by adding 1 drop (0.05ml) to 20ml of fresh broth and then grown to mid-exponential phase (OD_{550}=0.2 to 0.3) to provide a host cell culture.

Serial dilutions of the M13 phage (in 10mM potassium phosphate,pH7.2; 10mM $MgSO_4$; 0.005% gelatin) were added to 3ml of softened 0.6% BBL agar containing 0.2ml of the host cell culture, 25μl BCIG (25mg/ml in dimethyl formamide) and 25μl IPTG (25mg/ml in water) and poured onto BBL agar plates. The plates were incubated overnight at 37°C.

A well separated blue plaque was picked with a toothpick and used to inoculate 20ml of the host cell

culture. This was incubated for 5-8 hours at 37°C with shaking. The cells were pelleted and the supernatant used to infect a 1 litre of host cells in exponential phase (A_{550}=0.5) and grown with shaking at 37°C.

After 1.5 hours, 1ml of a 250mg/ml chloramphenicol solution was added to arrest the newly synthesized M13 DNA in its double-stranded form and incubation continued for a further 60 minutes. The cells were pelleted, and the RF DNA harvested by the large scale plasmid preparation method described above.

9. Digestion of DNA with restriction endonucleases

DNA samples were routinely incubated with 2 to 5-fold excess of restriction enzymes (manufacturer's definitions) to ensure total digestion of DNA. Typical reaction mixtures (20μl total volume) consisted of 1μg DNA, appropriate restriction enzyme buffer and 1mMDTT. Incubation was for 1 hour at 37°C.

When further manipulation with the DNA was required, one phenol extraction and one ethanol precipitation was carried out.

Restriction enzymes' buffers

*Eco*RI

100mM Tris-HCl.pH7.5

7mM $MgCl_2$

50mM NaCl

7mM 2-Mercaptoethanol

*Sma*I

15mM Tris-HCl,pH8.0

6mM $MgCl_2$

15mM KCl

"Core" (all other enzymes)

50mM Tris-HCl,pH8.0

10mM $MgCl_2$

50mM NaCl

10. Recovery of DNA fragments from agarose gels

DNA was electroeluted onto Whatman DE81 DEAE-cellulose paper and recovered by a modification of the method of Dretzen *et al.*, (1981).

DE81 paper was presoaked in 2.5M NaCl for 15 minutes and rinsed three times in DW. After separating DNA molecules by gel electrophoresis (using Seakem* HGT agarose), DNA bands were visualized under a long wave-lengh UV source and a slice of gel containing the band to be purified, excised. This slice was then wrapped with the DE81 paper and inserted into a slot of equivalent size in an unused part of the gel. Electrophoresis was resumed for 20 to 30 minutes, at 100 volts until the DNA was bound to the paper. The DE81 paper was rinsed in water and blotted dry on Whatman 3MM paper. Excess paper was removed, and the strip with bound DNA placed in an Eppendorf tube. After addition of 450μl of a high salt buffer (1M NaCl, 50mM Tris-HCl,pH7.5, 1mM EDTA,pH7.5), the paper was shredded by dragging the tube on an Eppendorf rack and incubated at 37^{o}C for 1 hour.

The DNA solution was recovered by inverting the tube,

making a small hole at the bottom with a red hot needle, and centrifuging the solution for 1 minute through the hole into a second Eppendorf tube. The DNA solution in the new tube was then centrifuged for 10 minutes to pellet traces of DE81 paper. The solution was transferred to a new tube and phenol extracted, the DNA ethanol precipitated and usually resuspended in 25μl TE buffer. The concentration of DNA was estimated by running 1μl of the solution on a test gel against DNA size markers of known concentration.

11. Construction of an *A. nidulans* genomic library in a λ-phage vector

This *A. nidulans* genomic library was constructed by Dr.A.Hawkins. Only a brief description of the procedure is therefore given here.

The DNA "arms" of the λ-replacement vector DB286 were prepared by conventional methods (Maniatis *et al.*, 1982) following digestion with endonucleases *Bam*HI and *Sal*1. *Aspergillus nidulans* genomic DNA, partially digested with restriction enzyme *Sau*3A, was isolated in the size range 10-15Kb by sucrose gradient centrifugation (Maniatis *et al.*, 1982) and ligated with the λDB286 DNA in the molar ratio of 3 chromosomal : 1 λDNA. Ligation reaction mixtures containing 0.5μg DNA were packed *in vitro* using packaging extracts prepared by the method of Scalenghe *et al.*, (1981) which typically produce 1.5-2.5 x 10^8 p.f.u. μg^{-1} with λ-CI857 DNA.

The final preparation was found to contain 95% recombinant phage particles identified by their *spi*$^-$

phenotype when titrated on an *E. coli* P2 lysogen Q359.

12. Identification and purification of positive recombinant λ-phage

This part of the work was done in collaboration with Dr. A. Hawkins. The method used was a modification of that of Benton and Davis, (1977).

(i) Transfer of λ-Phage DNA onto nitrocellulose filters

A phage suspension containing $4x10^3$ p.f.u., was mixed with 0.2ml of a ED8910 overnight culture in 10mM $MgSO_4$ and left to incubate at room temperature for 15 minutes. 3ml of BBL soft layer agar added and the total mixture poured onto BBL plates. These were incubated overnight inverted at 37°C.

The plates were then held at 4°C for at least 1 hour to harden the agar. A circular nitrocellulose filter (88mm diameter, Schleicher & Schüll BA 85/20) was placed on each agar plate. After 5 minutes at 4°C, the filter was marked relative to the plates by making a hole through the filter into the agar with a needle and then removed and placed on a damp pad (Whatman No1, 90mm diameter filter paper) wetted with 0.5M NaOH for 15 minutes. This step was followed by a transfer to a pad wetted with 1M Tris-HCl,pH7.5 for 2 minutes. This step was repeated once again, after which the nitrocellulose filter was transferred to damp pad wetted with 1M Tris-HCl,pH7.5; 1.5M NaCl, for 15 minutes. The filters were blotted onto 3MM paper, let to dry at room temperature and then baked

at 80°C for 2 hours.

(ii) Screening for recombinant λ-hybrids

Filters were washed in 5X Denhardts solution at 65°C for 4 hours. The $^{\alpha-32}P$ labelled DNA probe (see below) was boiled for 5 minutes to denature the DNA, and added to petri dish containing 5X Denhardts solution and 6X SSC at 65°C. After addition of filters, the petri dish was sealed and placed in a plastic lunch box in a 65°C oven, overnight.

Filters were then washed rapidly in succession 6-7 times until radiation in the washes was at background levels, in 2X SSC, 0.1% SDS (solution degased) at 65°C. After a further 15 minute incubation at 65°C, the filters washed once again, blotted onto 3MM paper, let to dry at room temperature and put on X-ray film with an intensifying screen, at -70°C for 2 days (see below).

Positively hybridizing recombinant plaques were picked and resuspended in 1.0ml of phage buffer (6mM Tris-HCl; 10mM $MgSO_4.7H_2O$, 0.005% gelatin,pH7.5), treated with 10μl of chloroform to kill the bacterial cells, serially diluted, and the same procedure repeated several times until only positively hybridizing plaques were observed on a single plate.

DNA from these positively hybridizing plaques was prepared following the method of Yamamoto *et al.*, (1970).

5X Denhardts solution

Substance	Amount per litre
20X SSC	300ml
Bovine Serum Albumin	1g
Polyvinylpyrrolidone	1g
Ficoll 400	1g

20X SSC - 3M NaCl; 0.3M NaCitrate, pH7.0

13. Construction of hybrid plasmids

Fragments of *A. nidulans* DNA were amplified using plasmid vectors in *E. coli* strains

(i) Ligation of DNA into plasmid vectors

1μg of vector DNA was subjected to a 10-fold over digestion with the appropriate endonuclease(s) in a buffer containing 0.15 units of calf intestinal phosphatase (Worthington Biochemical Corporation). DNA was recovered by one phenol extraction followed by an ethanol precipitation, and disolved 1n 100μl sterile H_2O (final concentration 10ng/μl).

DNA ligations were done using a 10X ligase buffer giving a final, concentration of 50mM Tris-HCl,pH7.5; 10mM $MgCl_2$; 1mM ATP; 20ng of the prepared vector DNA, samples of ½, 1 or 2-fold molar excess of purified *A. nidulans* DNA fragment and one unit (Manufacturer's definitions) of T4 DNA ligase in a final volume of 10μl. Vector DNA alone was also tested to monitor the effectiveness of the phosphatase treatment. Ligation reactions were all incubated at 15°C for 12 or 72 hours depending on the

quality of the ends generated by the restriction endonucleases (cohesive or blunt ends respectively)

(ii) Transformation of *E. coli* strains

Strains of *E. coli* were transformed with plasmid DNA using a modified version of the method of Kushner (1978).

The host bacterial strain was grown overnight in Luria broth, diluted 1/100 in the same medium and incubated with vigorous aeration until reaching an OD_{550} of 0.2 in logarithmic growth.

1.5ml samples of cells in Eppendorf tubes were pelleted by centrifugation for 10 seconds in a MSE Microcentaur bench centrifuge. The cells were resuspended in 500μl 0.1M MOPS (3-[N-Morpholino] propanesulphonic acid) solution,pH7.0 containing 10mM RbCl, pelleted by centrifugation for 10 seconds, resuspended in 500μl 0.1M MOPS solution,pH6.5 containing 10mM RbCl and 50mM $CaCl_2$ and held on ice for 90 minutes.The cells were again pelleted and resuspended in 150μl of the same buffer solution, to which the DNA ligation mixture (see above) plus 3μl of DMSO (Dimethyl Sulphoxide) were added. After at least 30 minutes incubation on ice, cells were heat shocked for 30 seconds at 55^{o}C and then transferred immediately to an ice-water bath for 2 minutes.

The chilled cells were plated directly by spreading on Luria agar plates containing appropriate antibiotics to select transformed cells. Plates were incubated overnight to yield transformant colonies.

14. ^{32}P-labelling of DNA fragments

The method used to label DNA probes was essentially that of Jeffreys *et al.*, (1980).

A sample of 50 to 100ng of DNA were added to the following reaction mixture:

2.5µl 10X nick-translation buffer (500mM Tris-HCl,pH7.5; 50mM $MgCl_2$;100mM 2-mercaptoethanol)
2µl each of 50µM dGTP, dATP and dTTP
1µl 8ng/ml DNase I (freshly diluted from a 1mg/ml stock in 10mM Tris-HCl,pH7.5)
1.5µl α-^{32}P dCTP (10µCi/µl ≈300 Ci/mMol)
1.5µl of *E. coli* DNA Polymerase I (5 units)
H_2O to 25µl

Ater incubation for 60 minutes at 15°C, the reaction was stopped by addition of 25µl of 0.5% SDS; 12.5mM EDTA; 10mM Tris-HCl,pH7.5. The solution was phenol extracted and the phenol reextracted by addition of 50µl of 10mM Tris-HCl,pH7.5 and the aqueous phases combined. 100µl of high molecular weight Salmon Sperm DNA (see below) was added as carrier and the DNA precipitated with ethanol and rinced with 70% ethanol. Due to the high concentration of DNA (mostly carrier) no chilling or centrifugation was necessary. The DNA precipitate was then redissolved in 200µl of Tris-HCl,pH7.5 and a second ethanol precipitation, 70% ethanol wash repeated as above. The final DNA precipitate was redissolved in 500µl of 10mM Tris-HCl,pH7.5.

High molecular weight Salmon sperm DNA was prepared as follows: 200mg of Salmon sperm DNA were dissolved in

200ml of 10mM Tris-HCl,pH7.5 overnight at 4°C. The DNA solution was then phenol extracted, ethanol precipitated, dried under vacuum and redissolved in 70ml of 10mM Tris-HCl,pH7.5. The concentration of DNA was calculated from the optical density at 260nm.

15. Southern blot filter hybridization

(i) Transfering of DNA onto nitrocellulose filters

DNA was bound to nitrocellulose as described by Maniatis *et al.*, (1982)

The DNA samples were separated by electrophoresis and unused areas of the gel were removed. DNA was denatured by immersing the gel in 1.5M NaCl; 0.5M NaOH for 1 hour at room temperature with gentle shaking. This step was repeated with an equal volume of 1M Tris-HCl,pH8.0; 1.5M NaCl, to neutralize the gel. The gel was then inverted and laid on top of a glass plate, which had been wrapped with 3MM paper soaked with 10X SSC, taking care to avoid trapping air bubbles between the gel and the plate. The plate itself lay on top of a perspex block inside a tray filled with 10X SSC, such that the plate was 2 to 3 cm above the surface of the liquid. The nitrocellulose filter and two pieces of 3MM paper cut to a size slightly greater than that of the gel were soaked in 2X SSC. The nitrocellulose filter followed by the two sheets of 3MM paper were laid on top of the gel, again removing any air bubbles by rolling a glass rod over the surface. A stack of paper towels was placed above the 3MM paper, covered with a glass plate compressed with a 1kg weight. Transfer of DNA continued for 12 to 18 hours replacing the stack of

paper towels at intervals.

The position of loading slots in the gel were marked on the filter with a ball point pen, and the filter was then washed in 6X SSC, blotted dry on 3MM paper and heated for 4 hours at 80°C to 'bake' the single stranded DNA onto the filter.

(ii) Hybridization with ^{32}P labelled DNA probe

The conditions used were those commonly practised in our department (Jeffreys *et al.*, 1980).

The filters were cut into convenient strips and pre-hybridized with gentle agitation at 65°C in hybridyzation boxes made of perspex, as follows:

-one wash in 3X SSC for 30 minutes

-one wash in 1X Denhardts for 30 minutes

-one wash in 1X Filter Hybridization mixture (FHM) for one hour

-one wash in 1X FHM plus 9% Dextran Sulphate or alternatively 6% PEG (Polyethylene Glycol - Sigma, grade8000)

The final hybridization solution contained the ^{32}P-labelled DNA which had been denatured by heating to 100°C for 5 minutes prior to addition to the hybridization solution. Hybridization occurred overnight. Dextran Sulphate (Wahl *et al.*, 1979)) or PEG (Renz & Kurz, 1984)) are known to promote the kinetics of hybridization and therefore were used when the filters contained *A. nidulans* genomic DNA.

Unbound labelled DNA was removed from the filters by washing in repeated changes of the filter hybridization solution (15 minutes at 65°C each time) until no

detectable ^{32}P could be measured in the wash solution. The stringency of hybridization and the wash solutions was always 1X SSC, 0.1% SDS. Filters were rinsed in 3X SSC, and dried at room temperature.

(iii) Autoradiography

The nitrocellulose filter strips were assembled into their original order, mounted on 3MM paper, covered with "cling" film and placed in a cassette against X-ray film. If autoradiography required prolonged periods of exposure, an intensifying screen was used, and left at -70°C. Exposure times varied from 12 hours to 10 days.

Autoradiographies were done using Kodak X-ray film (35X40 or 13X18 cm X-Omat RP). The film was developed by immersion in developer (Kodak DX80) for 3-5 minutes followed by a rinse in water (containing a trace of acetic acid) and then 3-5 minute immersion in fixer (Kodak FX40 plus HX-40 Hardener). Films were rinsed in water, and dried at room temperature or in a 37°C hot room.

(iv) Stock solutions

All solutions were degased under vacuum before use.

<u>10X Denhardts Solution</u>

10g Bovine serum albumin (Sigma Fraction V)

10g Polyvinylpyrrolidone

10g Ficoll400

Up to 500ml with 3X SSC

1X Filter Hybridization Mixture:

1X Denhardts

50μg/ml alkali denatured Salmon sperm DNA

0.1% SDS

1X SSC.

Alkali denatured Salmon Sperm DNA (10mg/ml):

1g of Salmon sperm DNA (Sigma, Type III) was added to 500ml of DW containing 20ml of 0.5M EDTA. The DNA was dissolved by heating the solution to 100°C, followed by the addition of 15ml of 10M NaOH and incubation for further 20 minutes to denature the DNA. The solution was cooled on ice, then 20 ml of 1M Tris-HCl,pH7.5 added and pH adjusted to 7-8 with concentrated HCl, added in small volumes and with rapid mixing. The denatured DNA solution was phenol extracted, precipitated with I.M.S. (Industrial Methylated Spirit) and rinsed with 70% ethanol. The DNA was initially dissolved in 50ml of water, and the concentration determined from the optical density at 260nm and adjusted to a final conmcentration of 10mg/ml.

16. Northern blot hybridizations

Northern blot hybridizations were carried out by John Clements in our laboratory, as described by Thomas (1980)

17. Determination of the DNA sequence by the dideoxyribonucleotide chain termination method

(i) Sonication of plasmid DNA

The procedures followed were based on the methods of

Deninger, (1983). These produce randomly sheared DNA fragments, which are end repaired to generate 'blunt'-ended fragments for ligation into the *Sma*I site of the M13 vector DNA.

A 200μl solution of plasmid DNA (0.25μg/μl) was exposed to ultrasonic vibration using a probe sonicator (1-71 MSE-Measuring & Scientific Equipment Ltd, London, England). The sample was treated at an amplitude of 5μm for three 5 second exposures, removing a 60μl sample after each exposure (Deninger, 1983). This procedure generated three samples sonicated for 5 seconds, 10 seconds and 15 seconds respectively.

DNA was recovered following phenol extraction by ethanol precipitation and small aliquots from each sample run on agarose gel to test for efficacy of sonication. The10 and 15 seconds samples were pooled because majority of DNA molecules appeared in the desired 1.2 to 0.6Kb size range (see Chapter 7).

(ii) End-repair of Sonicated DNA and Size-selection

The sonicated DNA was end-repaired by incubation overnight at 15^{o}C in the following reaction mixture:

DNA (in H_2O)	20μl
10x ligase mix	3μl
TM buffer	3μl
Sequence chase mix	2μl
DNA Polymerase I	2μl
(large fragment, Klenow enzyme)	(≡10 units)

Following end repair of the sonicated DNA, the sample was separated by agarose gel electrophoresis and molecules

in the range 600-900 bp (estimated by DNA size markers generated by digesting pBR322 with *Sau*3A) recovered by the use of DE81 paper. An aliquot of the resulting solution was examined by gel electrophoresis to check the efficiency of recovery of the DNA.

Reagent solutions

10X Ligase mixture:

100mM Tris-HCl,pH7.5; 100mM $MgCl_2$; 100mM DTT

TM Buffer:

100mM Tris-HCl,pH7.5; 100mMMgCl_2

Sequence chase mixture:

0.25mM each of dNTP in TM buffer

(iii) Preparation of M13 vector DNA

2μg of M13mp18 or mp19 RF DNA were cleaved at the desired cloning site by a restriction endonuclease(s) (for blunt ended substrates the enzyme used was *Sma*I) and the products treated with calf intestinal phosphatase as described above (section 13 (i)). M13 DNA was extracted once with phenol, precipitated twice with ethanol and dissolved in 200μl of water (10ng/μl).

(iv) Ligation of sonicated DNA into M13 vector DNA

The following six ligation reaction mixes were prepared:

	Reaction mixtures					
	1	2	3	4	5	6
Size selected DNA	1	2	4	-	-	-
M13 vector DNA	2	2	2	-	2	2
RF M13 DNA (1ng/μl)	-	-	-	1	-	-
10mM ATP	1	1	1	1	1	1
10X Ligase buffer	1	1	1	1	1	1
10mM DTT	1	1	1	1	1	1
DW	3	2	-	5	4	5
T4-DNA ligase (400 units/μl)	1	1	1	1	1	-

All volumes are in μl. All samples were incubated for 3 days at 15^{o}C and stored at -20^{o}C until required.

Reaction 4 gives a measure of the efficiency of transformation, whereas samples 5 and 6 monitor the efficiency of the cleavage of vector DNA and of the phosphatase treatment.

(v) Ligation of specific DNA restriction fragments into M13 vector DNA

Ligation of specific DNA fragments was done by the same procedure used for ligation of fragments into plasmid DNA (Section 13(i)). M13 recombinant phage generated in this manner were used for directional DNA sequencing from known restriction sites.

(vi) Transformation of recombinant M13 phage into the *E. coli* strain JM101

Transformations were carried out by a modification of the method of Kushner (1978) and described above for recombinant plasmid transformation (Section 13(ii)) .

Transformed cells were added to 200μl of an overnight culture of JM101, containing 25μl of 25mg/ml BCIG (in dimethyl formamide) and 25 μl of 25mg/ml IPTG (in H_2O). These cells were then mixed with 3ml of BBL soft agar before plating out BBL agar plates and incubated overnight.

BCIG: 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

IPTG: Isopropyl-β-D-galactopyranoside

(vii) Preparation of recombinant M13 DNA sequencing templates

Phage particles from colourless "plaques" (see chapter 7) were transferred with a toothpick and diluted into 1ml of phage buffer. An 100μl alliquot was then used to infect 1.5ml of a JM101 culture in logarithmic growth phase at OD_{550} = 0.2 in a test tube. A total of 24 recombinant M13 phage DNA preparations were usually made on each occasion.

The cultures were incubated with vigorous shaking for 5 to 6 hours. Cultures were transfered to Eppendorf tubes and the cells pelleted by 5 minutes centrifugation. 1ml of the supernatant containing M13 phage was then transferred to a new Eppendorf tube. Phage particles were precipitated by addition of 0.2ml of a PEG solution (10% PEG8000; 2.5M NaCl), incubated at 15°C for 15 minutes and pelleted by centrifugation for 10 minutes. The supernatant was removed by aspiration and remaining traces of PEG removed by spinning tube briefly for a second time and aspiration of remaining PEG solution.

The recombinant phage pellet was resuspended in 100μl of a solution consisting of 1.1M NaAc,pH7.0 followed by a

phenol extraction (80μl phenol). The DNA solution was re extracted with 65μl of chloroform (chloroform/isoamyl alchol, 49:1.v/v) and the final DNA solution was ethanol precipitated with 250μl ethanol, rinsed with 70% ethanol, dried under vacuum and dissolved in 30μl DW.

(viii) Screening M13 clones for recombinant sequences

Recombinant M13 phage generated from the sonicated DNA fragments, were screened for identification of those carrying DNA fragments that were desired for sequencing. After the PEG precipitation step described above, the M13 phagepellet was disolved in 100μl TE buffer and 1μl of this phage suspension was mixed with 1μl of a solution containing 3M NaCl; 0.2M NaOH; 0.1% SDS; trace bromophenol blue. Samples were spotted onto gridded nitrocellulose filter (Schleicher Schüll). After 15 minutes, the filters were washed in 3X SSC, blotted dry and baked at 80°C for 3 hours.

Filter hybridization with specific $^{\alpha-32}P$ labelled DNA fragment followed the procedures described above (Section 15(ii)). After autoradiography, positively hybridizing recombinant M13 phage clones were identified and new DNA prepared from them as described above.

(ix) Preparation and recovery of sequence specific 17-mer synthetic oligonucleotide primers

Specific oligonucleotides for primer extension DNA sequencing reactions were synthesized by Mr. J. Keyte, of the Department of Biochemistry in this University, following the method of Mathes *et al.*,(1984).

The oligonucleotides were provided as a unpurified

mixture of oligomers in the range 17mer, 16mer, 15mer, etc. The specific 17mer oligonucleotides required for sequencing reactions were recovered by the following procedure:

A 20% acrylamide/urea; 1X TBE(see section xi) gel was produced in a gel mould consisting of two siliconized plates. 2μl of formamide dye was added to a 2μl aliquot from the oligonucleotide solution (total volume of 10μl), boiled for 2 minutes and the 4μl sample loaded and run in the gel for 4 hours, until the bromophenol blue indicator dye had run to within 5cm of the bottom of the gel. This should place the 17-mer near the middle of the gel. One plate was removed and cling film spread over the exposed surface of the gel. The plate and the gel were then turned over, the second plate removed and a second sheet of film placed over the other surface of the gel.

The 17-mer oligonucleotide was observed using a fluorescent sheet (indicator tlc plate) and its position marked with a felt tip pen on the surface of the film. Bands were cut out conservatively, and the gel slice placed in an Eppendorf tube with 100μl of water, and left at room temperature overnight. The 17-mer oligonucleotide was purified by passing through a G-25 sephadex column (Maniatis *et al.*, 1982).

(x) M13 recombinant sequencing

Sequencing of M13 recombinant clones was based on the methods of Sanger *et al.*, (1977) for M13 dideoxyribonucleotide chain termination using α-^{35}S dATP. Quantities quoted are for 15 sequencing templates which are a confortable number to process at any one time.

Reactions were performed in 1.5ml Eppendorf tubes and all centrifugations were done in an Eppendorf bench centrifuge.

To ensure that the cloned template DNAs were fully dissolved, they were incubated at 60°C for 10 minutes prior to the annealing reaction.

Each clone was annealed to the 17-mer universal primer by taking 5μl of clone DNA plus 5μl of primer mix (7.2μl of 2μg/ml 17-mer universal primer; 64μl H_2O; 8μl TM buffer).

Alternatively, a 17-mer synthetic oligonucleotide was also used for primer extension sequencing, using only 0.1μl of solution (see above), keeping the same buffer conditions.

Mixed DNA/primer samples were incubated at 60°C for 30 minutes, briefly centrifuged and incubated for a further 30 minutes. The annealed mixture can be held at room temperature until the start of the sequencing reactions. For each clone 4 tubes were prepared containing 1μl of annealed mixture plus 1μl of either a "T", "C", "G" or "A" NTP mixtures.

NTP mixtures for sequencing reactions

	"T"	"C"	"G"	"A"
0.5mM dTTP	12.5	250	250	250
0.5mM dGTP	250	12.5	250	250
0.5mM dCTP	250	250	12.5	250
10mM ddTTP	25	-	-	-
10mM ddCTP	-	4	-	-
10mM ddGTP	-	-	8	-
10mM ddATP	-	-	-	0.5
TE buffer	500	500	500	250

(TE buffer = 10mM Tris-HCl,pH8.0; 0.1mM EDTA).

All volumes are μl. The deoxyNTPs and dideoxyNTPs were prepared in this TE buffer.

To each reaction tube in turn was added 1μl of freshly prepared "Klenow mix" (57μl H_2O; 3.5μl Klenow Polymerase; 5μl α-^{35}S-dATP). The solutions were mixed gently and then incubated at 37^{o}C for 20 minutes. This was followed by addition of 1μl of sequence chase mixture to each reaction and a further 20 minutes incubation at 37^{o}C. At this point the reaction mixtures can be prepared for loading onto a sequencing gel by addition of 2μl of formamide dye (stock solution containing 10ml deionised formamide; 10mg xylene cyanole FF; 10mg bromophenol blue; 0.2ml 0.5M EDTA,pH8.0). Substrates were occasionally stored for several days at -20^{o}C without detriment, provided the formamide dye had not been added.

(xi) Sequencing gels

The standard reagents used, sufficient to prepare one gradient gel are given below:

	0.5X TBE	2.5X (1.5X) TBE
Acrylamide	5.13g	1.14g
Bisacrylamide	0.27g	0.06g
Urea	45.00g	10.00g
Sucrose	-	1.00g
10X TBE	4.5ml	5ml (6ml)
DW final volume	90ml	20ml

10X TBE: 1M Tris-Borate,pH8.3; 20mM EDTA). Figure given in parenthesis refers to mixture used for extension (6 hour) runs.

Preparation and running of 50 X 32 cm 6% (w/v) polyacrylamide gradient gels were as follows:

The solutions were dissolved with stirring and moderate heating, and filtered through two 9cm diameter Whatman Ashless filters using a Buchner funnel and vacuum line. The solutions were degassed before addition of:

	0.5X	2.5X (1.5X)
10% APS	0.7ml	0.175ml
TEMED (prior to pouring)	48μl	12μl

APS: Ammonium persulphate

TEMED: N,N,N',N'-tetramethylethylenediamine

The gradient gels were poured as follows: using a 80ml large syringe, 8ml of the 0.5XTBE solution were drawn, followed by 12ml of the 2.5XTBE solution. The mixture was then poured down one side of the glass plates, carefully avoiding air bubbles, and allowing to settle at the bottom. This mixture was topped up by addition of 80ml of the 0.5XTBE solution, which is poured downn each side

of the gel mould, to keep the gradient even. The glass plates were then laid horizontally, the tooth comb inserted and the gel firmly clamped with "bulldog" clips.

A "sharks" tooth comb was used which enabled 15 X 4 samples to be comfortably run on a single gel. The sequencing reaction mixtures were heated in a boiling water bath for 3 minutes, in order to denature the DNA molecules before loading the gel. After electrophoresis for 3 hours (six hours for extension runs) at 1.4-1.7 Kv the gels were fixed in a 10% methanol, 10% acetic acid solution for 15 minutes and then dried using a Bio-rad gel drier. Gels were autoradiographed at room temperature for 2-4 days.

18. Computing

Various computing facilities were used in the course of this work. The dot-matrix was prepared using a DNA manipulation package written in Fortran 77 by Dr. Z. Nugent and run on the University facilities including a Cyber 73 (Control Data Corporation) mainframe computer and Calcomp 936 graphics plotter. Alignment of sequences was also performed on the Cyber 73 using a basic programme written by Dr. C. Boyd.

M13 clone sequences were aligned against reference sequences and each other using a version of the Staden (1980) programme modified to run on a Digital PDP 11/44 minicomputer.

Finally this thesis was compiled using a word processing package in a Apricot MS-DOS microcomputer.

19. Source of materials

BETHESDA RESEARCH LABORATORIES INC. Rockville, Maryland, USA:

All restriction enzymes and deoxyribonucleoside triphosphates

SIGMA London, England:

Lysozyme, Dextran sulphate (sodium salt), Dithiothreitol (DTT), Ficoll 400, Salmon Sperm DNA (sodium salt), Bovine Serum Albumin, Dimethyldichlorosilane, Isopropyl-β-D-galactopyranoside (IPTG), Ampicillin (sodium salt), Chloramp[henicol and N,N,N',N'-tetramethyl ethylenediamine (TEMED)

WORTHINGTON BIOCHEMICAL CORPORATIONA, Freehold, New Jersey, USA:

DNA Polymerase I (large fragment, Klenow enzyme), Calf Intestinal Phosphatase and Dideoxyribonucleoside Triphosphates.

NEW ENGLAND BIOLABS, Beverly, Massachusetts, USA:

T4-DNA Ligase

FISONS, Loughborough, England:

Polyvinylpyrrolidine and Phenol (AR)

UNISCIENCES LTD., London, England:

Acrylamide

ALDRICH CO LTD, Gillingham, England:

Dimethylsulphoxide

F.M.C. CORPORATION, Rockland, Maine, USA:

Agarose

GLAXO LABORATORIES:

Streptomycin sulphate

RADIOCHEMICAL CENTRE, Amersham, England:

E. coli DNA Polymerase I (for nick translations) and all radionucleotides

BACHEM Inc, Torrance, California, USA:

5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal or BCIG)

The universal 17mer primer (Duckworth *et al.*, 1981) used for M13 sequencing was kindly provided by Dr. J. Almond (Department of Microbiology, Leicester).

All other chemicals were of analytical grade.

20. Containment and Safety

All experiments undertaken in the work described in this thesis were conducted with reference to the Genetic Manipulation Advisory Groups guidelines on safety and containment conditions for such work. All procedures fall with the category of Good Microbiological Practice.

CHAPTER 3 . THE ISOLATION AND CHARACTERIZATION OF FURTHER QUINIC ACID NON-UTILIZING (*qut*) MUTANTS

The expression of fungal genes that are under physiological control, is generally found to be mediated by the function of several distinct regulatory genes and their protein products (Chapter 1).

The limited number of mutants available in the quinic acid utilization gene cluster in *Aspergillus nidulans* (Hawkins *et al.*, 1982), required the isolation of further quinic acid non-utilizing (*qut*) mutants to fully investigate the number of gene functions specific to the formation of the enzymes metabolizing quinic acid to protocatechuic acid (PCA). Two particular types of mutant were required to test the initial model proposed for regulation of the *qut* gene cluster in *A. nidulans* (Chapter 1) by analogy with the *qa* gene cluster in *Neurospora crassa* (Case & Giles, 1975). The first are mutants deficient in dehydroquinase, required to test linkage of the gene coding for this enzyme to the *qut* gene cluster. The second are additional pleiotropic non-inducible mutants which lack all three enzyme activities. Hawkins *et al.* (1982) isolated three such mutants (*qut*A4, *qut*A7 and *qut*A8) and postulated that these identify a positively acting gene *QUT*A, comparable to the *QA*1 gene of *N. crassa* (Case & Giles, 1975). Additional mutants were required to investigate the possibility that the same types of regulatory mutants occur in *Aspergillus* such as the $qa1^S$ and $qa1^F$ mutant alleles described in *Neurospora* (Case & Giles, 1975). Furthermore, it was necessary to establish

the number of regulatory genes operating in the *qut* gene cluster by genetic analysis of the non-inducible mutants.

1. Mutagenesis and isolation of *qut* mutants

Treatment with ultra-violet light was chosen as a convenient method to generate a variety of mutagenic lesions in the genome of *A. nidulans*. The dose rate was adjusted to produce approximately 50% killing and the filtration method (Chapter 2) used to enrich the proportion of *qut* mutants amongst the surviving conidiospores. Basically, the principle of this technique is to remove germinating conidiospores capable of utilizing quinic acid as sole carbon source, by filtration through cotton gauze, thus enriching the suspension for conidiospores not able to utilize quinic acid.

A suspension of conidiospores of strain R153 treated with UV light was innoculated in liquid MM with quinic acid as carbon source, incubated at 30°C and filtered at intervals over a period of 42 hours. The viable count was reduced from $\simeq 10^6$ ml^{-1} to 10^3 ml^{-1} when aliquots were plated on MEA plus glucose to produce about 30 colonies per plate. Two such mutation and filtration experiments were performed, and ≃2000 colonies screened by replica plating to MM agar plates with either quinic acid (test) or glucose (control) as carbon source. A total of 35 potential *qut* mutants was isolated from the two experiments, *qut*101 to *qut*121, and *qut* 201 to *qut*214 respectively. The mutants were then tested for growth on MM agar containing PCA as carbon source in order to define those with metabolic lesions between quinic acid and this

compound, and seven such new *qut* mutants were recovered. The remaining mutants unable to utilize both quinic acid or PCA are presumed to have lesions in the pathway from PCA to the tricarboxylic acid cycle (Ornston, 1971)). The seven mutants isolated that did not grow on quinic acid but grew normally on PCA and glucose were *qut*113, *qut*118, *qut*119, *qut*203, *qut*204, *qut*208 *and* *qut*214, and were retained for further analysis.

2. Identification of enzyme lesions in the new *qut* mutants

Quinic acid non-utilizing mutants (*qut*$^-$) can be ascribed specific enzyme lesions by assays measuring the activities of the three enzymes : quinate dehydrogenase(enzyme 1); dehydroquinase (enzyme 2) and dehydroquinate dehydratase (enzyme 3). Table 1 shows typical enzyme activities found after induction of a wild-type (*qut*$^+$) strain. When the strain R153 is grown on glucose, very low or non-detectable enzyme levels are observed, indicating that the formation of the three enzymes is not constitutive. However, transfer of mycelium grown on glucose to an induction medium containing quinic acid produces high levels of each of the enzyme activities, which are comparable to those found in mycelium grown on quinic acid. When glucose grown mycelium is transferred to an induction medium containing quinic acid (0.1% v/v) in the presence of glucose (10mM), enzyme induction does not occur (Table 1), indicating catabolite repression by glucose and the preferential utilization of glucose for growth from the mixture.

In order to test the new mutants for enzyme deficiencies, each was grown on glucose medium and the mycelium transferred to quinic acid induction medium. Table 2 shows the results of enzyme assays carried out on the new *qut*$^-$ mutants. As controls, representative mutants with known specific enzyme lesions (Hawkins *et al.*, 1982) were also tested. These were *qut*B42 lacking quinate dehydrogenase (enzyme 1) and *qut*C1 lacking dehydroquinate dehydratase (enzyme 3), and *qut*A4 and *qut*A8 which lack all three enzyme activities (pleiotropic, non-inducible mutants). Mutant *qut*A7 (Hawkins *et al.*, 1982) shows poor growth on all carbon sources including glucose, and therefore was not used for further analysis.

The enzyme assays done on these six new mutants reveal the following:

(i) One mutant strain, *qut*214, is non-inducible, like *qut*A4 and *qut*A8.

(ii) Another mutant, *qut*208, lacks dehydroquinase activity (enzyme 2).

(iii) Two mutants, *qut*203 and *qut*204, lack significant activities of the dehydrogenase enzyme(1), like *qut*B42.

(iv) Two mutants, *qut*113 and *qut*119 lack dehydratase activity(3), like *qut*C1.

The remaining mutant, *qut*118, raised initial interest in that it showed reduced activity for all three enzymes after induction. At first it was thought possible that the mutant might be defective in the uptake of quinic acid, since such a leaky phenotype could identify permease mutants. Genes coding for permeases have been identified in fungi, for example the *ure*A gene of *A. nidulans*

(Kinghorn & Pateman, 1977) and the *gal2* gene of *S. cerevisiae* (Douglas & Hawthorn, 1964, 1966). However, subsequent genetic tests carried out by Dr. A. Hawkins, revealed that *qut*118 was in fact a leaky *qut*C mutant, since it did not complement the *qut*C1 mutant allele.

Mutant strain *qut*208 was interpreted to contain a mutant allele of the gene coding for the quinate dehydroquinase (2), and is therefore designated *qut*E208; the wild-type allele is thus *QUT*E. This strain was used for the linkage studies described below, and further complementation analysis (see Chapter 4).

Mutant strain *qut*214 was used for complementation analysis against the known non-inducible mutant strains.

3. Carbon catabolite repression of the *QUT* gene cluster

It is shown in the previous section that induction of the three enzymes for quinic acid utilization are subject to catabolite repression by glucose.

Further analysis into this aspect of metabolic regulation was done by Dr. Susan Grant. Figure 2 shows the effects on enzyme induction of increasing concentrations of four carbon sources added to growth media containing quinic acid. These results reveal that glucose and sucrose strongly repress induction of the three enzymes. The effects of glycerol and acetate are less marked.

4. Close linkage of *qut*E208 to the *qut* gene cluster

The initial work by Hawkins *et al.* (1982) did not include analysis of a mutant strain deficient in catabolic dehydroquinate hydrolase (dehydroquinase (2)) activity. Acquisition of the mutant *qut*E208, deficient in this enzyme activity, enables a test of the hypothesis that *QUT*E falls within the *qut* gene cluster. First, it was necessary to establish that *qut*E208 phenotype identified a single mutant gene. In a cross of *qut*E208 to the wild-type strain R21, approximately one half of the progeny tested (84 of 158 =53%) were not able to utilize quinic acid for growth, demonstrating that *qut*E208 segregates as a single gene mutation.

To test for linkage of the *qut*E208 allele to the *qut* gene cluster, the mutant strain was crossed to three of the previously characterized, closely linked *qut*$^-$ mutants *qut*A4, *qut*A8 and *qut*B42 (Hawkins *et al.*, 1982). The results of these crosses are given as the frequency of wild-type recombinant progeny estimated as the fraction of ascospores from a single hybrid cleistosthecia growing on quinic acid selective medium (Table 3). These results clearly demonstrate tight linkage of *qut*E208 to these mutants, and that the *QUT*E gene therefore also lies within the cluster.

5. Location of the *QUT* gene cluster in linkage group VIII

The *qut* gene cluster has been allocated to linkage group VIII (Hawkins *et al.*, 1982) upon the basis of

mitotic segregation (Pontecorvo & Käfer, 1958). To map the *qut* locus on this linkage group, the strategy adopted was to cross a *qut*A8 (*pyro*A4) strain with the mapping strain R56, *paba*A1; *w*A3; *orn*B7 *fac*B1 *ribo*B2 *gal*7 (Clutterbuck & Cove, 1976), to locate the region of the chromosome within which the *qut* gene cluster lies. A portion of the genetic map of *A. nidulans* linkage group VIII is shown in Figure 3

The original *qut*A8 strain was isolated by Dr. M. Hynes (Hawkins *et al.*, 1982) and carries auxotrophic requirements for biotin (*bi*$^-$) and NH_4^+ (*cnx*$^-$). In order to utilize this *qut* mutation for mapping purposes, it was necessary to remove these two undesired auxotrophic markers by crossing the original *qut*A8 strain of Hynes to standard strains R21 and R153 and isolating *qut*A8 progeny with the R21 and R153 markers. The recombinant strains were also required for the construction of heterozygous diploids, as described in the next chapter.

Analysis of progeny from a cross between the mapping strain R56 and *qut*A8 in the R153 background revealed linkage of *qut*A8 only to *orn*B7 (RF=12.5%) and to none of the other markers tested (data not shown)

Subsequent meiotic analysis carried out by Hayley Whittington determined the precise orientation of the *qut* locus in relation to the *orn*B7 marker.

A strain carrying the *qut*E208 allele, which is closely linked to the *qut*A8 allele (see above) was crossed to a strain G191 (Balance *et al.*, 1983), which carries the *fw*A1 (fawn coloured spores) marker mapping 25 units from the *orn*B7 marker on linkage group VIII (Figure 3). The recombination frequency between the *qut*E208 and *fw*A1 alleles, was 11.5%. This result shows that the *qut* cluster

is located between the *ORN*B and the *FW*A genes (Figure 3), since the sum of the recombination frequencies between the *qut* cluster alleles and the *orn*B7 and *fw*A1 alleles is 24%, and is in good agreement with the published map distance of 25 units between *orn*B7 and *fw*A1 (Clutterbuck & Cove, 1977).

6. Discussion

Progress in the analysis of the quinic acid utilization (*qut*) gene cluster in *A. nidulans*, required the isolation of more *qut*$^-$ mutants.

The filtration enrichment experiment, yielded one particular mutant that was sought, namely the *qut*E208 strain lacking the dehydroquinase (2) activity. Crosses of this strain to other *qut*$^-$ mutants established linkage of the *QUT*E gene to the cluster, as the initial model suggests, and in agreement with the *QA* gene cluster data in *N. crassa* (Chapter 1). This result is important in that it predicts the possibility of isolating the three structural genes and the regulatory locus coding for the activator protein, as a contiguous sequence of DNA (see Chapters 5 and 6).

Mapping of the *qut* gene cluster between the *orn*B and *fw*A gene loci provides some progress in formal genetic analysis and confirms the location of the gene cluster in linkage group VIII by Hawkins *et al.* (1982). Absence of a closely linked (≈1 to 2 map units) marker to the *qut* locus, means that fine structure mapping of the various *QUT* genes is not possible by conventional genetic methods, since the recombination frequencies to *orn*B or *fw*A would

be too large for accuracy.

Induction by quinic acid of the three enzymes specific to the quinic acid metabolism is severely decreased when alternative carbon sources such as glucose or sucrose are present in the medium and to a lesser extent in the presence of glycerol or acetate. This result is expected in the case of the metabolism of alternative carbon sources.

Preferential carbon utilization is now a well established phenomenon in fungi, but its precise molecular mechanisms of action are far less well understood than in prokaryotes. In *E. coli* for instance, cyclic AMP levels, which are depressed by glucose and other carbon sources, regulate the activation of a variety of glucose repressible functions. In fungi cAMP appears to be involved in a wide variety of apparently unrelated functions, but does not seem to regulate expression of carbon catabolite repressible genes (Pall, 1981).

Bayley & Arst (1975) investigated the mechanisms of carbon catabolite repression in *A. nidulans* and they provided an argument for the existence of wide domain, negatively acting regulatory genes, *CRE* which are concerned in preventing induction of many particular catabolitic enzymes when alternative carbon sources such as glucose are available. However, the exact mechanisms by which these genes operate have remained unresolved.

TABLE 1 ***QUT*** **enzyme specific activities in a wild-type strain**

Enzyme specific activities of the wild-type strain of *A. nidulans* R153, grown in liquid culture on a variety of carbon sources. Enzymes are: dehydrogenase (1), dehydroquinase (2) and dehydratase (3). Mycelium was grown on glucose (10mM) for 16 hours, or quinic acid (1% w/v) for 20 hours. Glucose grown mycelium was washed aseptically on the filter and the pad transferred to the carbon sources indicated for further 4 hour incubation. Concentration of glucose was 10mM, and quinic acid was 0.1% (w/v). Units of specific activity are: 1. dehydrogenase - μmol product min^{-1} mg^{-1}; 2. dehydroquinase - A_{240} min^{-1} mg^{-1} (1 unit is ΔA_{240} of 0.1 min^{-1} at 37°C); 3. dehydratase - μmol product min^{-1} mg^{-1}.

TABLE 2 Enzyme specific activities in the new quinic acid non-utilizing ***qut*** **mutants.**

The mutant strains *qut*A4, *qut*A8, *qut*B42 and *qut*C1 are included as representative examples of mutant strains, previously identified to carry *QUT* gene lesions. Activities of dehydrogenase (1), dehydroquinase (2) and dehydratase (3) were assayed following enzyme induction, as in above, and are presented as the ratio to the wild-type control strain, R153, which was included in each experiment.

TABLE 1

Growth medium	Transfer of mycelium to	Enzyme specific activities 1	2	3
Glucose	Glucose	0	0	0
Glucose	Glucose + Quinic acid	0,19	0,8	0,06
Glucose	Quinic acid	0,95	3,11	0,43
Quinic acid	——	0,83	2,77	0,11

TABLE 2

	Strains	Relative Enzyme Activities 1	2	3
Control	*qut*$^+$	1	1	1
Known Mutant Strains	*qut* A4	0,22	0	0
	qut A8	0,09	0,07	0
	qut B42	0,14	0,84	1,08
	qut C1	1,58	1,08	0,23
New Mutant Strains	*qut* 113	1,04	0,6	0
	qut 118	0,41	0,36	0,58
	qut 119	0,73	1,01	0
	qut 203	0	0,20	0,50
	qut 204	0,10	0,81	1,30
	qut 208	0,71	0	1,04
	qut 214	0	0	0

TABLE 3 Observed frequencies of wild-type recombinant progeny in meiotic crosses between the *qut*E208 mutant strain and representative mutants from the *qut* gene cluster.

TABLE 3

Cross	No. of viable conidiospores plated	No. of colonies growing on quinic acid	Frequency
*qut*E208×*qut*A4	$3,9\times10^4$	58	0,15%
*qut*E208×*qut*A8	$5,2\times10^3$	14	0,27%
*qut*E208×*qut*B42	$1,4\times10^4$	28	0,20%

FIGURE 2 Carbon catabolite repression of enzyme induction by quinic acid in *A. nidulans*.

Mycelium grown on glucose (10mM) for 16 hours, was transferred to a medium containing 0.1% (w/v) quinic acid plus the carbon sources indicated for a further 4 hour incubation, when mycelia was harvested for enzyme assays. Values are given as percentage of dehydroquinase activity observed in mycelium from a sample transferred to quinic acid in the absence of any added alternative carbon sources.

FIGURE 2

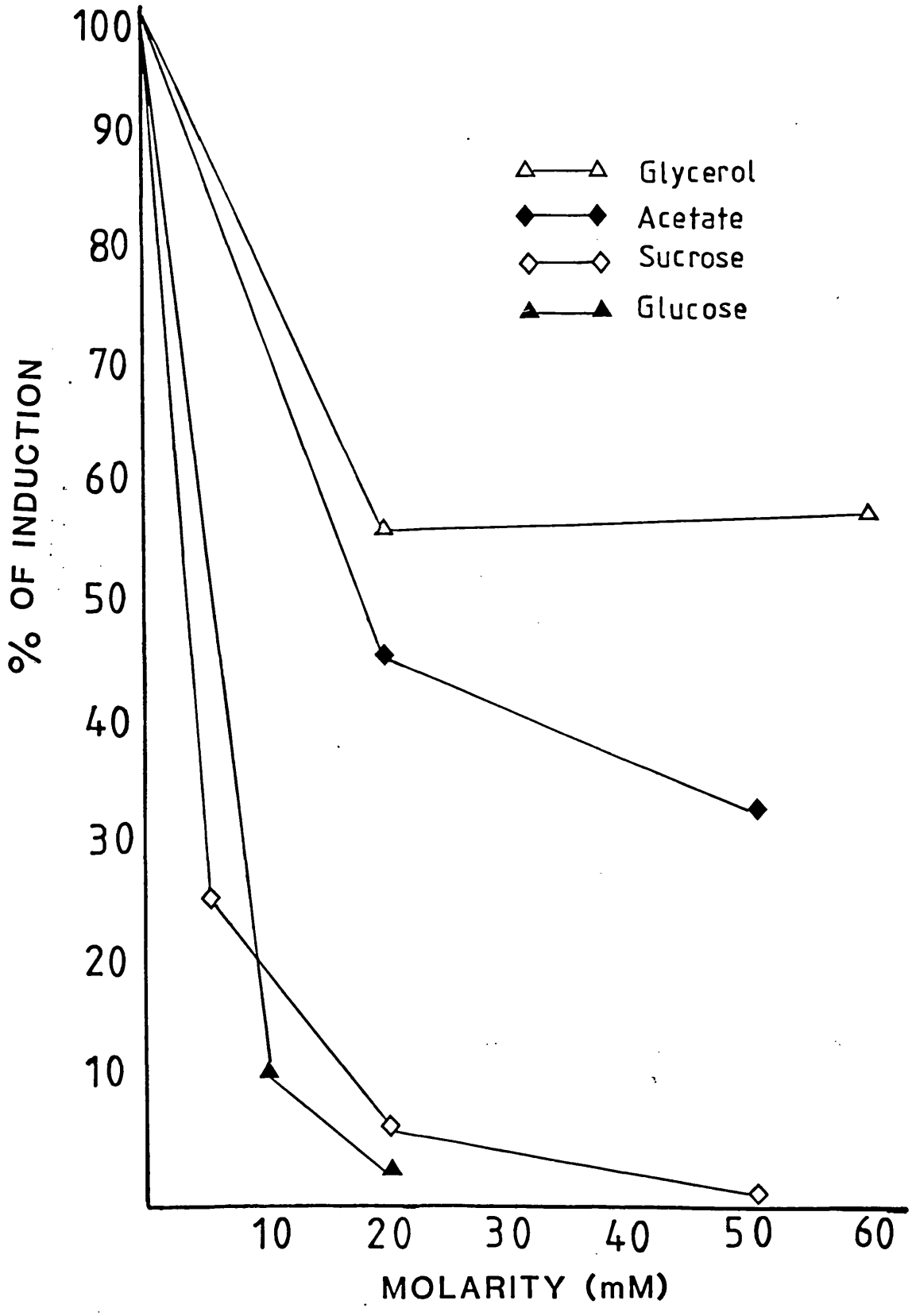
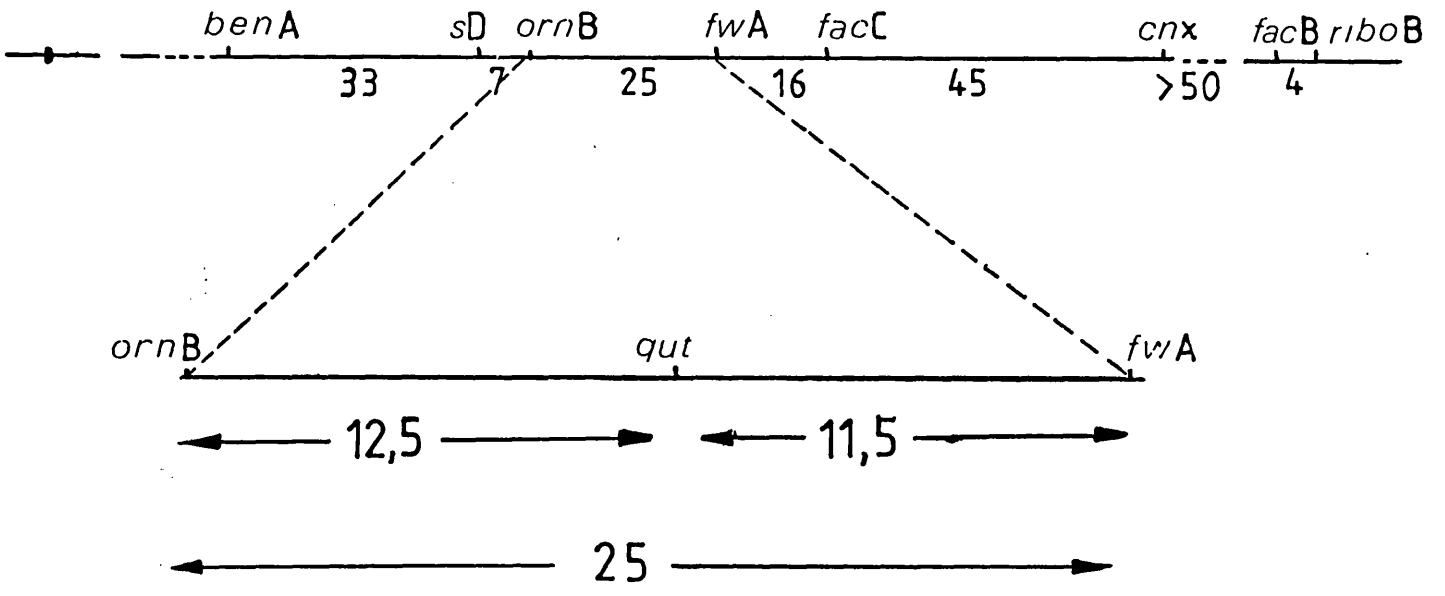

100
90
80
70
60
50
40
30
20
10
% OF INDUCTION
Glycerol
Acetate
Sucrose
Glucose
10
20
30
40
50
60
MOLARITY (mM)

FIGURE 3 Location of the *qut* gene cluster in linkage group VIII.

The figures refer to distances in map units.

FIGURE 3


benA
sD
ornB
fwA
facC
cnx
facB
riboB
33
7
25
16
45
>50
4
ornB
qut
fwA
12,5
11,5
25

CHAPTER 4 . GENETIC AND BIOCHEMICAL ANALYSIS OF THE EXPRESSION OF THE *QUT* GENE CLUSTER

The analysis of the non-inducible *qut*$^-$ mutants in *A. nidulans* was undertaken in order to establish two major points: (1) the number of genes which are identified by this mutant phenotype and involved in the regulation of the expression of the *QUT* enzyme structural genes; and (2) whether in *Aspergillus* one would also find the two types of non-inducible mutant found in *Neurospora*: *qa*1^F and *qa*1^S (Case & Giles, 1975 and Chapter 1).

This chapter describes the analysis of three non-inducible *qut* mutants in *A. nidulans* when combined in heterokaryons or, principally, in heterozygous diploid strains. The three mutant alleles were separated into two complementation groups on the basis of growth tests on quinic acid and this result confirmed by measuring enzyme induction in diploid strains.

These results have led to the construction of a new model for the regulation of the *QUT* gene cluster in *A. nidulans*.

1. Complementation analysis of the *qut* mutant strain

Four *qut*$^-$ alleles, *qut*A4, *qut*A8, *qut*214 and *qut*B42, each combined into the R21 and R153 strains were used for the construction of heterokaryons and diploid strains. The non-inducible mutant strain *qut*A7 (Hawkins *et al.*, 1982) was not used as it was found to grow very poorly on all media and therefore was not amenable to analysis. The work

with the *qut*214 mutant strain was done by Dr. Susan Grant and Michael Stout.

The mutants *qut*A4, *qut*A8 and *qut*214 are non-inducible and were initially thought to be alleles identifying a single genetic regulatory function in the expression of the *QUT* enzyme structural genes. The *qut*B42 mutant strain which lacks dehydroquinase (1) activitywas used as a representative mutant allele of an enzyme structural gene.

Heterokaryons and diploid strains were constructed by conventional methods (Chapter 2). Basically, the heterokaryons were synthesized by mixing conidiospores from two strains carrying complementary nutritional requirements, and selecting for growth on MM plates with glucose as carbon source and no added supplements. Diploid strains were selected from the corresponding heterokaryons by plating uninucleated conidiospores in MM. The strains are prototrophic and produce wild-type green conidiospores as the result of complementation between the component yellow (yA^-; wA^+) and white (y^+; wA^-) mutant genomes in the heterozygous diploids (yA^-/yA^+; wA^-/wA^+). Properties of the *qut* mutant alleles in heterokaryons or in diploid strains were deduced in growth tests on MM with quinic acid as sole carbon source. The same medium with glucose as carbon source served as a control.

The different heterokaryon construction in all cases yielded growth on glucose media identical to that of the wild-type control.

The results for the heterokaryons tested for growth on quinic acid are shown in Table 4a. Combination of strains containing the same *qut* mutation yielded

diminished or no growth, which is expected of the mutant phenotype. Combinations of different mutant strains in reciprocal arrangement in all cases yielded identical results. The heterokaryon between the wild-type and the mutant strain *qut*B42 grew normally on quinic acid demonstrating that mutant allele *qut*B42 is fully recessive to wild-type, as is expected for such an enzyme structural gene loss mutation. On the other hand each of the heterokaryons between wild-type and the non-inducible mutants *qut*A4, *qut*A8 and *qut*214 did not grow on quinic acid, indicating dominance of these mutant alleles over wild-type. Similarly the mutants combined one-with-another produced heterokaryons unable to grow on quinic acid.

The rather striking result found for the non-inducible mutants combined in heterokaryons, suggested that further analysis by investigation of the corresponding set of diploid strains (Table 4b). In contrast, the growth responses of the various heterozygous diploid strains revealed a different range of genetic relationships. All diploid strains grew normally on glucose with the wild phenotype. All of the homozygous mutant diploid strains showed typical poor growth or no growth on quinic acid, expected from the phenotypes of the component mutant strain.

There was good agreement for the results of reciprocal pairs of heterozygous diploid strains and analysis of the growth responses of the strains on quinic acid reveals the following points:

(i) When the *qut*B42 mutant (lacking dehydrogenase(1) activity) is combined with the wild-type *qut*B^{+} allele,

growth on quinic acid is fully restored, indicating that the mutant allele is fully recessive, as expected of a structural gene loss mutation.

(ii) Similarly the *qut*A8 allele is fully recessive to the wild-type allele.

(iii) The *qut*A4 mutation is partially dominant over the wild-type allele, since the *qut*A4/*qut*$^+$ heterozygous diploid gives intermediate growth response to quinic acid, between that of the wild-type control and the qutA4/qutA4 homozygous diploid.

(iv) The *qut*214 mutation is almost fully dominant over the wild-type since in the growth tests, the response to quinic acid of the *qut*214/*qut*$^+$ diploid is only marginally above that of the *qut*214/*qut*214 homozygote.

(v) Combination of *qut*A4/*qut*A8 yields a heterozygous diploid strain that shows ability to grow on quinic acid to a lesser extent than the wild-type, but nevertheless far better than that of each of the homozygous mutant diploids (*qut*A4/*qut*A4 or *qut*A8/*qut*A8) and equally as well as the heterozygous *qut*A4/*qut*$^+$ strain. Taken together, these results show that *qut*A4 and *qut*A8 mutant genomes complement in heterozygous diploids, indicating that each mutation affects different gene functions, and that at least two regulatory genes control expression of the *QUT* structural genes. Therefore, the *qut*A8 allele is now renamed *qut*D8 to identify a second regulatory gene *QUT*D, an addition to *QUT*A.

Complementation analysis by growth tests with the *qut*214 strain, shows that this allele is fully dominant over the wild-type allele, and it is thus impossible to allocate it to a complementation group. However *qut*214

shares with *qut*A4 the characteristic of dominance. Moreover, enzyme assays (done by Michael Stout) of heterozygous diploids in which *qut*214 is combined with *qut*A4 and *qut*D8 exhibits similar properties to those of *qut*A4 (see below). No enzyme induction is detectable when *qut*214 is combined with *qut*A4, but when *qut*214 is combined with the *qut*D8 mutant, significant, albeit reduced levels when compared with the measurements from the *qut*A4/*qut*D8 construct, can be measured. These observations together strongly suggest that *qut*214 is an allele of the *QUT*A gene.

2. Analysis of enzyme induction in heterozygous diploid strains

The specific activities of each of the three enzymes for the initial steps in the metabolism of quinic acid were measured in glucose grown non-induced and quinic acid induced mycelium of the diploid strains described above. The procedures followed are described in Chapter 2.

No significant levels of enzyme activities are detected in non-inducible mycelium from the diploids. The observed levels of enzyme induction (Table 5) show the following:

(i) The homozygous *qut*B42 diploid strain totally lacks dehydrogenase(1) activity but forms normal levels of the other two enzymes. The non-inducible *qut*A4 and *qut*D8 homozygous diploids each lack all three enzymes. These results are those expected for homozygous mutant diploid strains from the properties of the haploid mutant strains.

(ii) Data for all reciprocal pairs of heterozygous

diploids are in good agreement, the only major descrepancy being the dehydrogenase(1) activity in the *qut*A4/*qut*B42 diploid strains.

(iii) The mutant *qut*B42 which lacks dehydrogenase(1) activity and is recessive to the wild-type in growth of heterozygous diploid strains on quinic acid (Table 4b) shows gene dosage in respect to enzyme activity, producing some 60-70% of the wild-type dehydrogenase activity in the heterozygote.

(iv) The mutant *qut*D8, which lacks all three enzyme activities and is fully recessive to the wild-type *qut*D$^+$ allele in heterozygotes growth tests, is also fully recessive with respect to enzyme induction, which is the wild-type level. Furthermore, the *qut*D8 mutant allele fully complements *in trans* the *qut*B42 dehydrogenase mutant and the levels of enzyme activity in the *qut*B$^+$*qut*D8/*qut*B42*qut*D$^+$ heterozygous diploid show no difference from these in the *qut*B42/*qut*B$^+$ diploid.

(v) By contrast, the second non-inducible mutant, *qut*A4 which is semi-dominant in respect to the wild-type in growth tests, exhibits a similar phenotype in producing levels of enzyme activity that are some 60% of those of the wild-type control. When combined with the *qut*B42 mutant, the levels of enzyme induction show some variation, perhaps depending on the haploid strain used in the construction of the heterozygotes. To investigate this possibility, the two new reciprocal diploid strains were constructed from the original haploid strains and the levels of enzyme induction again measured. The results (Table 5a), show consistency with the original values obtained, again showing in one heterozygous diploid strain

that induction of the dehydrogenase enzyme(1) appears to suffer a further reduction in activity as compared to the *qut*A4*qut*B$^+$/*qut*A$^+$*qut*B$^+$ diploid. However, when other recombinant haploid strains are used to construct equivalent diploid combinations, there is no detectable variation in dehydrogenase activity between *qut*A4*qut*B$^+$/*qut*A$^+$*qut*B42 and *qut*A*qut*B42/*qut*A$^+$*qut*B$^+$. It is thus possible that this difference may be due to modifying genes segregating in the crosses done in the construction of the haploid strains.

(vi) Finally, complementation of the *qut*A4 and *qut*D8 mutant genomes *in trans* is comfirmed by the levels of enzyme activity found for all three enzymes. Thus, the *qut*A4*qut*D$^+$/*qut*A$^+$*qut*D8 diploid shows the same response to induction by quinic acid as the *qut*A4/*qut*$^+$ diploid strain, whereas no enzyme induction was observed in either of the *qut*A4 or *qut*D8 homozygous diploids. This result comfirms that the *qut*A4 and *qut*D8 mutations define two separate regulatory functions, and therefore at least two regulatory genes control expression of the structural genes.

3. A new model for the regulation of the quinic acid utilization *qut* gene cluster in *A. nidulans*

The properties of the heterozygous diploid strains, and particularly analysis of the data generated by assaying the three enzyme activities induced by growth in the presence of quinic acid suggests a series of points concerning the characteristics of the products of the genes identified by the *qut* mutant alleles.

(i) The *qut*B42 allele has all the characteristics of a recessive loss mutation, rendering the haploid strain silent for dehydrogenase activity. In the *qut*B42/*qut*$^+$ heterozygous diploid a reduction to some 60% of wild-type enzyme level is a typical gene dosage response, and the normal growth of the diploid on quinic acid shows that this amount of enzyme is not limiting growth.

(ii) The pleiotropic *qut*D8 mutant allele is deficient in the induction of all three enzyme activities, and is fully recessive to the wild-type allele both in terms of growth and enzyme induction. This suggests that the *qut*D8 mutation results in the loss of function of a gene product normally required for enzyme induction and confers a non-inducible phenotype on the *qut*D8 haploid strain and the homozygous diploid strains. Several observations suggest that a single functional copy of the *qut*D$^+$ allele is able to support full enzyme induction in heterozygous diploid strains, and therefore that the concentration of the *QUT*D gene product is not rate limiting for enzyme induction.

First, all three enzymes are fully induced in *qut*D8/*qut*D$^+$ heterozygous diploid strains. Second, dehydrogenase (1) activity measured for *qut*B42 *qut*D$^+$/*qut*B$^+$ *qut*D8 shows no reduction in level compared to that of the *qut*B42 *qut*D$^+$/*qut*B$^+$ *qut*D$^+$ diploid. Similarly, in the *qut*A$^+$ *qut*D8/*qut*A4 *qut*D$^+$ double mutant heterozygote, in which one might expect a two-fold gene-dosage effect, there is no detectable difference in the levels of enzyme induction when compared to the *qut*A4 *qut*D$^+$/*qut*A$^+$ *qut*D$^+$ diploid.

The observation that *qut*D8 is fully recessive and therefore that the product of the wild-type *QUT*D gene is

not rate limiting, indicates that this gene codes for a positively acting, activator protein.

(iii) In contrast to *qut*D8, the *qut*A4 allele exhibits semi-dominance over the wild-type allele, as expressed by both reduced growth and enzyme induction. A possible interpretation therefore is that the product of the *QUT*A gene is rate limiting and that the *qut*A4 allele codes for an altered protein which no longer recognizes the inducer quinic acid. Thus, *QUT*A could code for a negatively acting, repressor protein.

Based on these interpretations of the rôles of the products of the two regulatory genes *QUT*A and *QUT*D, the model for the regulation of expression of the three *QUT* structural genes has been revised. Figure 4 shows a schematic representation of this new basic model, and is explained as follows:

In the absence of the inducer, quinic acid, the product of the *QUT*A gene, which is constitutively transcribed, binds to the promoter regions of the *QUT*D positive regulatory gene and possibly also to the promoter regions of the three structural genes, *QUTB*, *QUTE* and *QUTC*. As a consequence, gene expression is repressed, and no mRNA is transcribed from these genes.

In the presence of quinic acid, induction requires the recognition and binding of quinic acid by the *QUT*A repressor protein causing it to dissociate from the promoter regions of *QUT*D and of the structural genes. The *QUT*D gene is then strongly transcribed, and the activator protein binds to the promoter regions of the enzyme structural genes to promote high levels of mRNA transcription.

Upon the depletion of quinic acid from the medium, the *QUT*D gene and also the structural genes, would not be transcribed due to repression by the *QUT*A gene. In this event, expression of the enzyme structural genes would cease.

4. Discussion

The discovery from the analysis of heterozygous diploid strains that two genes control the expression of the *QUT* structural genes in *A. nidulans*, has led to a significant revision of the model which was initially proposed (Hawkins *et al.*, 1982). This model was closely similar to that for the *qa* system in *Neurospora crassa*, where one single gene, coding for a positively acting protein was thought to effect control over the structural genes (Case & Giles, 1975 and Chapter 1).

The discovery that two regulatory genes could be identified by complementation was made possible by the diploid stage in the life cycle of *A. nidulans*, a feature that is not available in *Neurospora*. Heterokaryon formation in *Aspergillus* is unstable, requiring constant selection for prototrophic growth to maintain a low proportion of heterokaryotic hyphal-tip cells against segregation of homokaryotic hyphae (Clutterbuck and Roper, 1966). Thus the ratio of component nuclei in the bulk of the heterokaryon is subject to wide fluctuations, making analysis of data from this material unreliable. In diploid strains however, the 1:1 allele ratio is precisely defined, thus providing accurate quantitative data. Comparision of the results for the complementation growth

tests done in the heterokaryons and in the heterozygous diploid strains clearly demonstrates this point.

More recently, Huiet (1984) has shown in *N. crassa* by the analysis of transcription products and by sequencing the *N. crassa QA1* locus DNA, that the *qa1*^{S} and the *qa1*^{F} alleles identify two separate genes. This is in contrast to the original interpretation of one gene coding for two distinct domains in an activator protein (Case & Giles, 1975). The revised model proposed in *Neurospora* is remarkably similar to the new version of the model in *Aspergillus*. It postulates that the *QA1*^{S} gene (≡ *QUTA*) codes for a repressor proten which inhibits expression of the *QA1*^{F} gene (≡*QUTD*). In the presence of quinic acid, the repressor protein coded by the *QA1*^{S} gene dissociates from the 5' control region of the *QA1*^{F} gene, allowing high level expression of the activator protein, which in turn promotes expression of the *QA* structural genes and enzyme synthesis.

Several observations in *N. crassa* arising from the analysis of transcription of the *QA1*^{S} and *QA1*^{F} regulatory genes, exhibit remarkable similarities to the proposals based upon analysis of the biochemical data from the heterozygous diploids in *A. nidulans*. Patel *et al.* (1981) and Giles *et al.* (1985) have shown that expression of *QA1*^{F} requires the presence of the inducer, and that low level constitutive expression increases some 50-fold upon growth in the presence of quinic acid. Furthermore, it has also been shown (Giles *et al.*, 1985) that the *QA1*^{F} gene product is not rate limiting upon induction, as suggested by the data for enzyme induction in the *qut*D8/*qut*$^{+}$; *qut*D8/*qut*A4 and *qut*D8/*qut*B42 diploid strains (see above). The data in

Neurospora is based on an experiment in which a plasmid carrying the *QA1*^{F} gene deleted of its 5' control region, was used to transform a *qa1*^{F} deletion mutant. The results of this transformation experiment showed that some of the transformed strains had regained the ability to grow on quinic acid, presumably due to a fusion of the coding sequence of the *QA1*^{F} gene to an unrelated promoter. Transcrption studies showed that the *QA1*^{F} gene was no longer under quinic acid control, and titration of specific mRNA showed that the gene was poorly expressed even in the presence of quinic acid. Nevertheless, expression of the three structural genes achieved levels of mRNA transcription comparable to those in the wild-type strain.

The *QA1*^{S} gene in *N. crassa*, apparently equivalent to the *QUT*A gene of *A. nidulans*, has a low level of constitutive expression (Patel *et al.*, 1981). It is autogeneously regulated since growth in the presence of quinic acid causes a 2 to 3-fold increase in *QA1*^{S} mRNA. It has been postulated that the expected increase in the concentration of repressor protein induced by quinic acid, reflects a requirement for titration of the synthesis of the activator protein, coded by the *QA1*^{F} gene, in order to respond to the depletion of quinic acid (Giles *et al.*, 1985).

In *Aspergillus* partial dominance of the *qut*A4 allele over the wild-type indicates competition between mutant and wild-type gene products and suggests that the *QUT*A product is rate limiting for enzyme induction. Hence, it appears that the protein coded by the *QUT*A gene, and also that of its putative counterpart in *N. crassa*, the *QA1*^{S},

has a rôle in determining the fine tuning of the control of the expression of the *QUT* cluster structural genes.

There is some evidence from the biochemical data in *Aspergillus* that an interaction between the *QUT*A and *QUT*D gene products may also occur. It has been shown that the concentration of the *QUT*D product does not limit induction; There is no detectable difference in the induction levels measured from the *qut*A4/*qut*D8 and *qut*A4/*qut*D^{+} heterozygous diploids, although one might expect a further reduction in enzyme activity in the first diploid compared to the second as a consequence of the presence of the two mutant alleles. This observation appears to exclude the interpretation that the *QUT*A protein represses only expression of the *QUT*D gene. This follows, since if the *QUT*A protein only exerted its function by repressing expression of the *QUT*D gene, and since the product of the *QUT*D gene is not rate-limiting, there should still be enough activator protein to give wild-type levels of induction in the *qut*A4/*qut*D8 heterozygous diploid. However, as noted above, this is not the case and enzyme induction is reduced to some 60% of fully induced levels in both diploid strains. Therefore, if the product of the *qut*A4 mutant allele, is interpreted as producing a super-repressor no longer sensitive to quinic acid, it must also be exerting its effect directly upon the transcription of the structural genes rather than on *QUT*D alone.

It is possible that the effect of the *QUT*A repressor protein reflects an interaction between the two regulatory proteins not on DNA, in which the *QUT*A protein inactivates the *QUT*D protein, preventing it from binding to the enzyme

structural genes and thus promoting their expression. Alternatively, the repressor protein coded by the *QUT*A gene may bind directly to the 5' control regions of the structural genes, hence directly repressing their expression, until quinic acid causes it to dissiciate from these regions

There are very striking homologies between the mechanisms of regulation of expression of the quinic acid utilization gene cluster in *A. nidulans* and *N. crassa*, to those of the *GAL* system in *Saccharomyces cerevisiae* (see chapter 1). Two regulatory genes, *GAL*4 and *GAL*80 control the expression of five genes coding for galactose inducible functions, involved in the uptake and metabolism of galactose (Douglas & Hawthorne, 1966; Hopper & Rowe, 1978; Matsumoto *et al.*, 1978; and St.John & Davis, 1981). The initial model proposed by Douglas & Hawthorne (1966) is analogous to the simplest version model for the regulation of the *QUT* cluster. Hence, *GAL*80 gene which is constitutively expressed, codes for a repressor protein that binds to the 5' control region of the *GAL*4 gene, inhibiting its expression. However, in the presence of the inducer, galactose, the *GAL*80 repressor protein dissociates from the *GAL*4 gene allowing for the initiation of transcription. The product of the *GAL*4 gene in turn binds to the 5' control regions of the structural genes to promote transcription of the mRNAs. Hemce, the *GAL*80 and *GAL*4 genes seem to perform the same rôles as the *QUT*A and *QUT*D genes respectively.

Matsumoto *et al.* (1978) and Laughon & Gestland (1982) however, found that the expression of the activator gene, *GAL*4 was constitutive, and that therefore it was not under

the control of the *GAL*80 gene. One possible model then, was that the repressor protein coded by the *GAL*80 gene effected its rôle by binding to the activator protein which is coded by the *GAL*4 gene, inactivating it rather than preventing its expression at the level of mRNA transcription. More recent evidence however (Johnston & Hopper, ,1982; Hashimoto *et al.*, 1983; and Nogi *et al.*, 1984) seems to indicate that perhaps the antagonistic effects of the *GAL*80 protein derive from it directly binding to the DNA sequences in the structural genes, preventing the initiation of transcription.

Furthermore, the most common mode of action for regulatory proteins, both repressor and activator, in the lower organisms appears to be by direct binding to the DNA sequence just upstream from the start of mRNA transcription. Hence, a similar mode of action is proposed for the *QUT*A repressor protein.

The model for the regulation of expression of the quinic acid utilization *(QUT)* gene cluster makes predictions that are amenable to experimental test. For example, a mutation abolishing the production of the putative repressor protein might be expected to lead to the constitutive production of enzyme activities in the absence of quinic acid. However, attempts to recover mutants of this class by selection of revertants capable of growth on quinic acid in the mutant strains *qut*A4 and *qut*A214 have failed to yield any quinic acid utilizing strains which are constitutive for enzyme formation. This is in sharp contrast to the situation in *N. crassa*, where *qa*1^{S-} mutants frequently revert to yield constitutive strains (Valone *et al.*, 1971; Partridge *et al.*, 1972).

This observation indicates that there might be significant differences in the nature and modes of action of the *QUT*A and *QA*1^{S} proteins.

A possible alternative model therefore, is one in which *qut*A4 is interpreted not as mutant allele coding for a functionally active superrepressor protein that is not sensitive to the inducer quinic acid, but as a mutation that has lost the ability to regulate expression of the *QUT* enzyme structural genes, in response to quinic acid. In this case, the *QUT*A gene could be interpreted as coding for a co-activator required, together with the *QUT*D activator protein for the full expression of the enzyme structural genes. Moreover, the *QUT*A co-activator would seem to be rate limiting in induction, in contrast to the *QUT*D product and in a *qut*A^{-} heterozygous diploid the reduction in the levels of *QUT*A protein, would lead to a reduction in induction of gene expression, having as a consequence the observed gene dosage effects.

The model can be put to test by measuring mRNA transcription in certain heterozygous diploid strains containing particular regulatory and structural gene mutants, for example *qut*A4/*qut*$^{+}$ and *qut*A4/*qut*D8. This approach, not available is *Neurospora,* would enable one to measure interaction between the products of *trans*-acting genes by quantifying both mRNA transcription and enzyme specific activities. Hence, it should be possible to test whether the *q*utA4 product directly represses transcription of the structural genes, by measuring their mRNA levels.

Secondly, the model proposes that the two regulatory proteins bind to common DNA sequences in the 5' control regions of the structural genes. Determination of the DNA

sequence of the *qut* gene cluster and an analysis of the promoter regions seeking putative sites for the binding of the regulatory proteins, become important strategies in testing the model.

The following chapters describe the procedures followed for the physical isolation of the *QUT* gene cluster and sequencing of the *QUT*E enzyme structural gene.

TABLE 4 Growth of heterokaryons and heterozygous diploid strains on quinic acid.

The extent of growth on quinic acid MM agar was scored visually on a scale from 0 (no growth) to 5 (normal growth) in comparision to the wild-type combination (qut^+/qut^+). All heterokaryons and diploid strains grow normally on glucose MM agar.

TABLE 4

a) HETEROKARYONS

R153 \ R21	*qut*+	*qut*A4	*qut*A8	*qut*B42	*qut*214
qut+	5	0	1	4	0
*qut*A4	2	0	1	1	0
*qut*A8	2	0	1	2	0
*qut*B42	4	0	2	1	0

b) DIPLOID STRAINS

R153 \ R21	*qut*+	*qut*A4	*qut*A8	*qut*B42	*qut*214
qut+	5	3	5	5	1
*qut*A4	3	0	3	3	0
*qut*A8	5	3	0	5	1
*qut*B42	5	3	5	0	1

TABLE 5 Enzyme activities in diploid strains relative to the wild-type (*qut*$^+$/*qut*$^+$) diploid strain.

Specific activities of the enzymes in each of the diploid strains induced by growth in the presence of quinic acid are shown relative to the activity of each of the three enzymes found in the wild-type (*qut*$^+$/*qut*$^+$) diploid strain. For each set of enzyme assays (horizontal rows), a wild-type diploid strain was included as a reference (right hand side block in the table). Mycelium was grown on glucose (10mM) for 16 hours, washed aseptically and a filtered pad transferred to 0.1% (w/v) quinic acid for a further 4 hour incubation. Units of specific activity are: 1. dehydrogenase - μmol product min^{-1} mg^{-1}; 2. dehydroquinase - A_{240} min^{-1} mg^{-1} (1 unit is ΔA_{240} of 0.1 min^{-1} at 37°C); 3. dehydratase - μmol product min^{-1} mg^{-1}. (a) indicates that diploid strains were reconstructed from the same haploid strains. (b) indicate that diploid strains were constructed from newly isolated haploid strains (see text in Chapter 4)

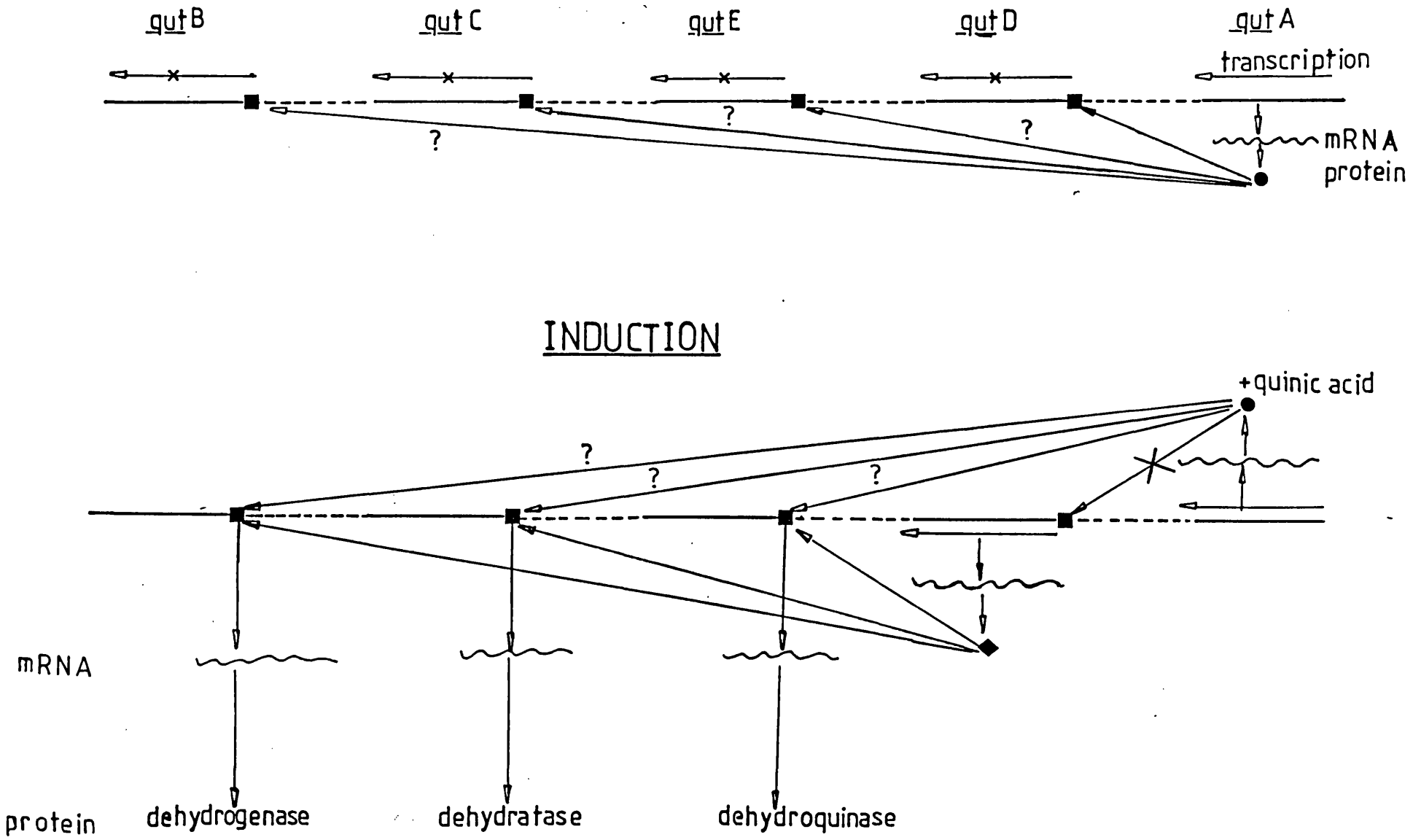
TABLE 5

	qut$^+$			*qut*A4			*qut*D8			*qut*B42			Enzyme specific activities in *qut*$^+$/*qut*$^+$ controls		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
qut$^+$	1	1	1	0,58	0,63	0,58	1,08	1,11	1,0	0,6	1,35	0,94	0,70	3,0	0,45
*qut*A4	0,6	0,76	0,88	0	0	0	0,66	0,69	0,79	0,75	0,9	0,88	0,96	2,47	0,73
										(a) 0,60	0,71	1,11	0,55	2,67	0,27
										(b) 0,65	0,77	0,94	0,20	3,00	0,36
*qut*D8	1,02	0,95	0,87	0,63	0,70	0,86	0	0	0	0,6	1,34	1,32	1,27	3,99	0,84
*qut*B42	0,77	1,26	1,02	0,36	0,67	0,84	0,75	1,25	1,10	0	0,90	0,94	1,51	3,45	0,86
				(a) 0,27	0,60	1,09							0,55	2,67	0,27
				(b) 0,65	0,53	0,58							0,20	3,00	0,36

FIGURE 4 Model for the regulation of expression of the *QUT* gene cluster of *A. nidulans*.

This diagram represents a revision of the model. It is proposed that the *QUT*A gene codes for a repressor protein that binds to the 5' control region of the *QUT*D and possibly of, *QUT*E, *QUT*C and *QUT*B genes, preventing their expression. Upon induction, the repressor protein interacts with quinic acid, which induces an allosteric change in the *QUT*A protein eliminating its ability to prevent transcription of the activator gene *QUT*D and the structural genes. Transcription of *QUT*D produces an activator protein which, either on its own or by binding *QUT*A protein and quinic acid, stimulates transcription of the *QUT* enzyme structural genes, leading to enzyme induction. See the text for more detailed discussion and alternative interpretations of the model.

FIGURE 4


NON-INDUCTION
gutB
gutC
gutE
gutD
gutA
transcription
?
?
?
mRNA
protein
INDUCTION
+quinic acid
?
?
?
mRNA
protein
dehydrogenase
dehydratase
dehydroquinase

CHAPTER 5. ISOLATION AND PHYSICAL CHARACTERIZATION OF THE *QUT* ENZYME STRUCTURAL GENES

The new model proposed for the regulation of the *qut* gene cluster in *A. nidulans* makes predictions that are amenable to experimental test. One such prediction is that the two regulatory proteins, coded by the *QUT*A and *QUT*D genes, interact with one another at the 5' non-coding control regions of the structural genes, to regulate the expression of these genes. Ultimately one would seek to test this proposition using *in vitro* systems. A more immediate aim however, is to study gene expression by qualitative and quantitative analysis of mRNA transcription of haploid strains and particularly from a set of diploids heterozygous for regulatory and structural genes mutant alleles. These studies require the availability of the DNA sequences from the various *QUT* genes to use as hybridization probes to detect the specific mRNA populations. The physical isolation of the *QUT* gene cluster was undertaken for these reasons.

1. Isolation of a λ recombinant phage containing the *QUT* genes from a phage λ genomic library.

Hawkins *et al.* (1982) found that dehydroquinase purified from *A. nidulans* reacted strongly with rabbit anti-serum raised against the equivalent enzyme from *N. crassa*, thus showing that the two enzymes possess similar amino acid sequences. This observation suggests that strong DNA homology might also be found between the two

genes.

The plasmid pVK57 (Kushner *et al.*, 1977) contains a *N. crassa* DNA fragment from the *qa* gene cluster (Figure 5), including the whole of the *N. crassa QA2* gene, which codes for the catabolic dehydroquinase (≡ *QUTE* gene in *A. nidulans*). The 3.5 Kb *Hind*III fragment was isolated from pVK57, radioactively labelled with $^{\alpha-32}$P-dCTP and used as a probe in an hybridization experiment against *Aspergillus* genomic DNA digested with *Hind*III, *Bam*HI or *Eco*RI, and transferred to nitrocellulose filters. The results of this experiment are shown in Figure 6 and demonstrate that *N. crassa* DNA sequences from the *QA2* gene hybridizes in each case to a single band in *A. nidulans* genomic DNA and thus could be used for the identification and isolation of the *A. nidulans qut* gene cluster.

The same 3.5 *Hind*III fragment of *N. crassa* DNA from pVK57 was then used to screen an *A. nidulans* genomic library constructed in the phage λ replacement vector DB286 (Brammar, 1982). One positive hybridization signal was found corresponding to a single plaque. Phage growth from this area was picked, replated and retested through six cycles of purification, until all plaques were observed to give strong positive signals. This recombinant phage was isolated and named λQ1.

A restriction map (Figure 7a) of the DNA prepared from λQ1 was generated using the five restriction enzymes *Hind*III, *Bam*HI, *Eco*RI, *Bgl*II and *Xba*I, and λQ1 was found to contain an *A. nidulans* DNA insert of 13.2 Kb. Since only one recombinant phage had been isolated and to guard against the possibility that λQ1 may have undergone rearrangement during the cloning procedures, a second

screen of the library was done, using as hybridization probe the 3.2 Kb *Bam*HI fragment from λQ1 which was found to cross hybridize to the 3.5 Kb *Hind*III fragment in pVK57 (This result is not shown since it is confirmed by results presented below and data presented in Figure 9). A second recombinant phage was isolated, called λQ2. The restriction map of the 8.9 Kb of *A. nidulans* insert in this new phage clone (Figure 7a), shows that it falls wholly within the cloned fragment in λQ1 and that there are no differences in the restriction patterns.

In order to further verify that λQ1 did not possess any major DNA rearrangements, a Southern blot hybridization was carried out using as DNA probe the whole of λQ1 DNA against *Aspergillus* genomic DNA digested with the same five restriction enzymes as above . The results (Figure 8) show that the restriction patterns exhibited in the autoradiograph are consistant with those found in λQ1, thus proving that the cloned *Aspergillus* DNA had the same sequence organization in genomic DNA.

2. Identification of the *A. nidulans QUT* enzyme structural genes in genomic and λQ1 DNA.

In order to identify the *QUT* genes in *A. nidulans* genomic and cloned DNA, the equivalent genes from *N. crassa* were used as hybridization probes.

A. nidulans genomic DNA was digested with *Hind*III, *Bam*HI, *Eco*RI, *Bgl*II and *Xba*II and then tested by Southern blot hybridization using the *N. crassa* heterologous probes. The fragments were purified from plasmids pVK57 or pVK88, according to the established restriction maps

(Schweizer *et al.*, 1981) shown in Figure 5. Each of the fragments is largely contained within the structural genes for *N.crassa* dehydrogenase (*QA*3≡*QUT*B), dehydroquinase (*QA*2≡*QUT*E) or dehydratase (*QA*4≡*QUT*C). The results of these hybridization experiments are presented in Figure 7b. The approximate location of each of the three putative *QUT* structural genes is indicated, based on a comparision of the size of the bands on the genomic Southern blots and the the restriction map of λQ1.

Confirmation of the localization of DNA sequences within λQ1 that show homology with the *N. crassa QA* genes was obtained from hybridization studies against the cloned DNA. In these, λQ1 DNA was digested with each of the five listed restriction enzymes singly and in all possible pairwise combinations, the products separated by gel electrophoresis and bound to nitrocellulose filters which were hybridized to the same *Neurospora* probes used for the genomic Southern blot hybridizations. The results (Figure 9) show that all three structural genes appear to fall within the 3.2 Kb *Bam*HI fragment in λQ1, confirming the results indicated in Figure 7.

3. Verification of the isolation of the *QUT*E gene: complementation by transformation of dehydroquinase deficient strains of *A. nidulans* and *E. coli*

Kushner *et al.* (1977) have shown that the catabolic dehydroquinase (*QA*2) gene from *N. crassa* can replace the enzyme deficiency ("complement") of an *aro*D6 biosynthetic dehydroquinase mutation of *Escherichia coli*, enabling growth of the bacterial mutant strain on minimal medium in

the absence of added aromatic amino acids. It was therefore of interest to discover whether the cloned *A. nidulans* catabolic dehydroquinase *QUTE* gene would also express in *E. coli* and complement the *aro*D6 mutation, and thus confirm the isolation of the functional *QUTE* gene.

For this purpose, the 6.5 Kb *Hind*III fragment from λQ1, which contains all of the *QUT* structural genes, was subcloned into the *Hind*III site of pBR322 (work done in collaboration with Peter Estribeiro). This plasmid was called pEH1 (see Figure 13, in Chapter 7). The plasmid was used in transformation experiments using the *aro*D6 mutant strain SK3430 of *E. coli* (Kushner *et al.*, 1977) and transformants were selected by growth on complete medium containing ampicillin. Colonies were then replica plated onto minimal medium without aromatic supplements and the plates incubated at 30°C and 37°C.

Slow growth representing weak enzyme formation was observed only on the plates at 30°C after incubation for 4 to 6 days. This result will be fully discussed below.

In collaboration with Hayley Whittington, protoplasts of the *qut*E208 mutant strain of *A. nidulans* which lacks catabolic dehydroquinase (Chapter 3) were transformed using plasmid pEH1 and following a protocol developed by Ballance *et al.* (1983). Several transformed strains were recovered (at a rate of 7 recombinants / μg of plasmid DNA / 10^6 viable protoplasts), by their growth on quinic acid as sole carbon source. No colonies were observed on the control plates where no DNA had been added to the protoplasts, indicating that *qut*E208 is a stable mutant strain (reversion rate less than 10^{-9}). Furthermore, subsequent Southern blot hybridizations of genomic DNA

from the transformed strains revealed the presence of vector pBR322 sequences in $^{1}/_{3}$ of the strains tested. This result shows that pEH1 had integrated into the genome, and that the wild-type *qut*E^{+} allele it carries was capable of repairing the functional lesion in the *qut*E208 mutant strain.

4. Identification of the dehydroquinase gene (*QUT*E) mRNA

The model proposed for the regulation of the expression the quinic acid utilization (*qut*) structural genes implies that regulation operates at the level of transcription. Proof of this was provided from work done in collaboration with John Clements using the technique of Northern blot hybridization. In this, total cellular RNA was separated by electrophoresis on denaturing gel, and transferred to a nitrocellulose filter. This was then tested by hybridization, using as DNA probe the 1.1 Kb *Eco*RI-*Bgl*II fragment from pEH1 (Figure 13, Chapter 7), which contains the *QUT*E gene. The result of this experiment is presented in Figure 10, and it clearly shows that specific *QUT*E mRNA was readly detected in RNA isolated from quinic acid grown mycelium, whereas in glucose grown mycelium no hybridization was detected. The size of the *QUT*E mRNA is estimated to be ≈0.8Kb. These results show that regulation of expression of the *QUT*E gene, and by inference *QUT*B and *QUT*C structural genes, is at the transcriptional level and that no constitutive mRNA production could be detected from the glucose grown mycelium, even after prolonged exposure of the

autoradiograph to the filters.

5. Discussion

The isolation of DNA sequences from in *A. nidulans* containing the *QUT* enzyme structural genes was made possible by DNA sequence homology between the *Aspergillus* and *Neurospora* genes. Two hybrid λ-phage clones were isolated in two separate screening experiments. Since on each occasion the library plated corresponded to about 10 genomes' equivalent one would have expected to find a similar number of copies of *QUT* sequences. However, only one λ-recombinant was isolated on each occasion, and it thus appears that the *QUT* gene cluster is under represented in the library. The reason for this is not known.

The three structural *QUT* genes have been physically mapped in the cloned DNA by hybridization, using as probes DNA fragments largely within the coding sequences of the *Neurospora QA* genes. The *Aspergillus* genes are contained within a 3.2 Kb *Bam*HI fragment, whereas in *N. crassa* the *QA* enzyme structural genes are defined within a 5.4 Kb fragment (Figure 5). The relative orientation of the equivalent enzyme structural genes in the two organisms however, is identical: *QUT*B - *QUT*C - *QUT*E in *A. nidulans* and *QA*3 - *QA*4 - *QA*2 in *N. crassa*, that is the sequence: enzymes (1), (3) and (2) (see Figures 5 and 7).

Analysis of transcription of the *QA* gene cluter in *N. crassa* has revealed that two additional presumptive structural genes, not previously identified by mutations, are also induced by growth with quinic acid (Patel *et al.*,

1981). These genes are called *QAX* and *QAY*, and it has been postulated that they may code for two additional functions : a permease involved in quinic acid uptake and an enzyme for the conversion of chlorogenic acid to quinic acid (Giles *et al*, 1985). Preliminary DNA hybridization experiments, yet to be repeated suggest that these two gene identify DNA sequences in *Aspergillus* probably in genes controlling equivalent functions, that are not only adjacent to the characterized *QUT* genes, but also appear in the same relative orientation (Figures 5 and 7). The *qut* gene cluster in *A. nidulans* appears therefore to possess a remarkably similar organization compared to that of *N. crassa*.

The plasmid pEH1, which contains the 6.5 Kb *Hind*III fragment from λQ1 and the putative *QUT*E gene, weakly restores growth at 30°C, of an *E. coli aro*D6 mutant, lacking the biosynthetic dehydroquinase activity. However, no complementation is observed at 37°C. This is in contrast to a similar experiment in *N. crassa* using *QA*2 gene, which readly complements the *aro*D6 lesion at 37°C and was the procedure used for the isolation of the *QA* gene cluster (Kushner *et al.*, 1977). Unsuccessful attemps by Hawkins, Kinghorn & Giles (umpublished work) to isolate the *QUT* cluster in *Aspergillus* by the same procedure, are now explicable since selection was done at 37°C, at which temperature no complementation is observed. The expression of the *Neurospora QA*2 gene in *E. coli* probably reflects the presence in the 5' non-coding region of the canonical sequence for the *E. coli* ribosome binding site (Shine & Dalgrano, 1975). The weak complementation exhibited by the *A. nidulans* gene, suggests that this sequence may not

present in the *QUTE* gene. This speculation was later proved to be correct by analysis of the DNA sequence (Chapter 7).

Complementation of the *qut*E208 lesion of *Aspergillus* by the pEH1 plasmid provides good evidence that the functional *QUTE* gene has been cloned, and is expressed upon transformation and integration into the genome.

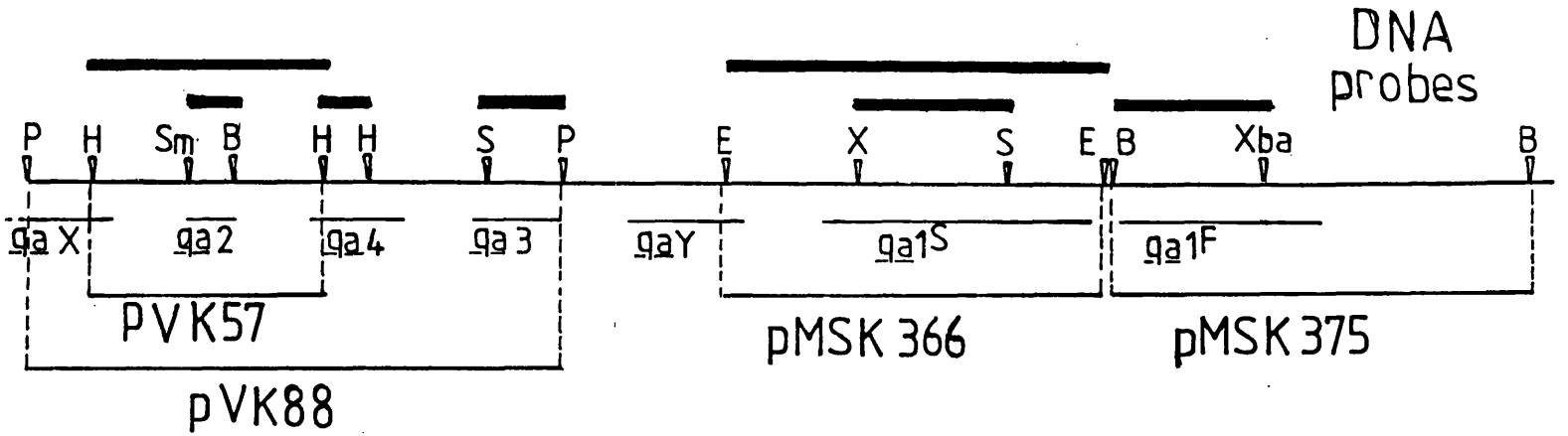
Finally, it has been shown that the *QUTE* gene is under transcriptional control in *A. nidulans*. The size of the mRNA identified by Northern blot hybridization appears to be around 0.8 Kb, which is marginally less than that of the mRNA in *N. crassa*, where multiple *QA*2 gene transcripts are found in a range 1.0 to 1.2 Kb (Patel *et al.*, 1981). However, the experiment was done with total cellular RNA, and the data is still very preliminary. Determination of the start and termination positions of transcription is therefore necessary for a precise assessment the nature of the *QUTE* mRNA.

It has shown that the *QUTE* gene is strongly transcribed upon induction, reaching levels comparable to those of the highly expressed phosphoglycerate kinase gene (*PGK*) and several orders of magnitude greater than the *AROM* complex gene (John Clements, PhD Thesis).

FIGURE 5 Genetic and physical organization of the *N. crassa* quinic acid (*QA*) gene cluster.

Solid bars represent DNA fragments used as probes in Southern blot hybridization experiments described in this chapter. References: pVK57 (Kushner *et al.*, 1977); pVK88 and pMSK375 (Schweizer *et al.*, 1981); pMSK366 (Hiuet, 1984).

FIGURE 5



1Kb

P=PstI S=SalI

H=Hind III E=EcoRI

Sm=SmaI X=XhoI

B= BamHI Xba=XbaI

FIGURE 6 Southern blot hybridization of the 3.5 Kb *Hind*III *N. crassa* DNA fragment from pVK57 to *A. nidulans* genomic DNA.

Restriction enzymes used to digest *A. nidulans* DNA are indicated above the tracks in which the digests were loaded for electrophoresis. The figures record the positions of standard size markers.

FIGURE 6

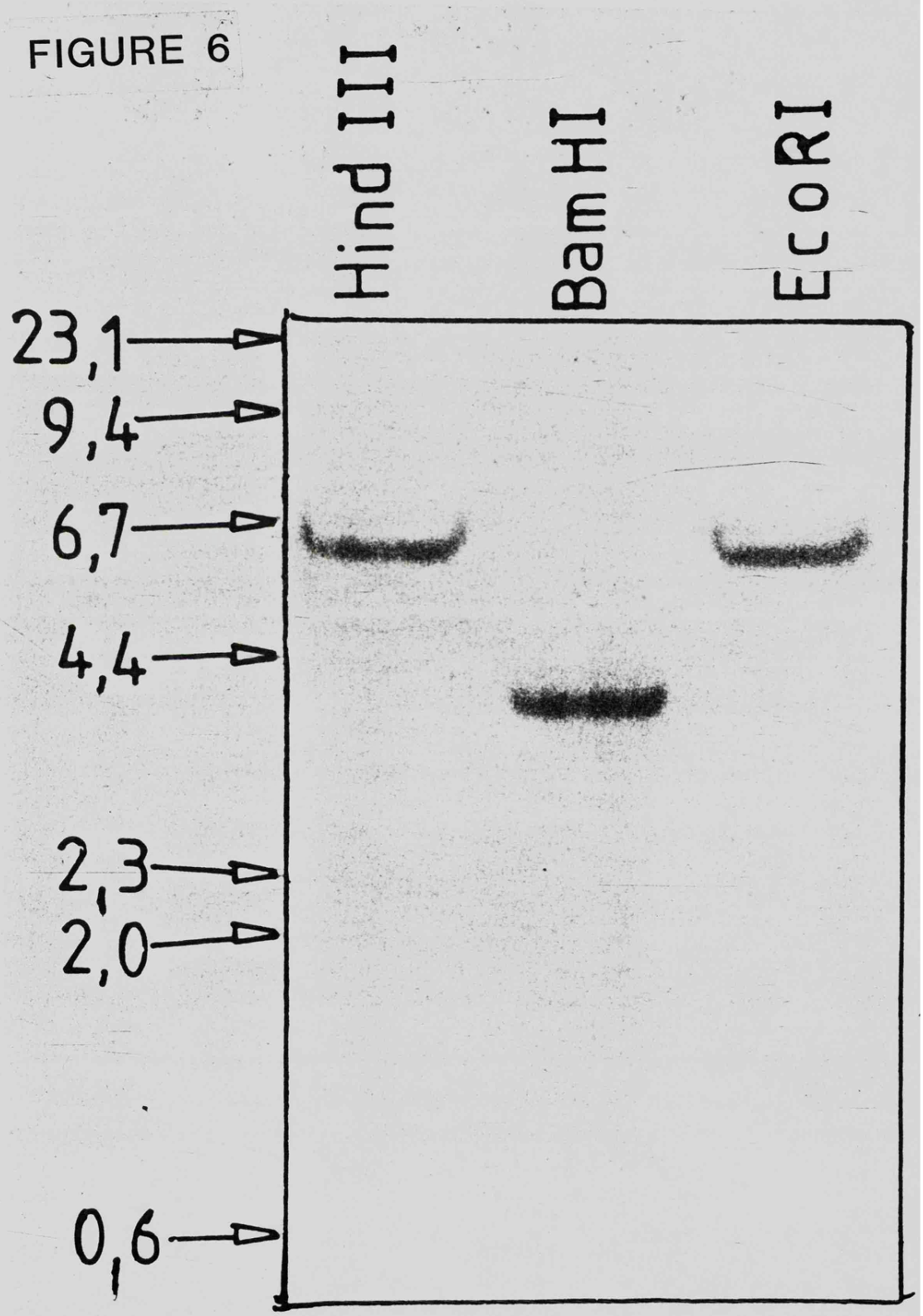


FIGURE 7

7a. Restriction map of the two recombinant phage: λQ1 and λQ2. 7b. Identification of sequences in *A. nidulans* genomic DNA hybridizing with the *N. crassa* structural genes and approximate location in cloned DNA of the putative genes *QUT*B(1), *QUT*E(2) and *QUT*C(3) in *A. nidulans*.

7a. The lines represent *A. nidulans* DNA and the boxes phage vector arms. The positions of restriction enzyme sites are indicated.

7b. The five tracks on each autoradiograph represent *A. nidulans* genomic DNA digested with the restriction enzymes indicated, the products separated by electrophoresis and transferred to nitrocellulose filters. The three genes from *N. crassa* used as hybridization probes are the fragments indicated in Figure 5 as follows:

*QA*3, 1Kb *Sal*I-*Pst*I fragment

*QA*4, 0.5 Kb *Hind*III fragment

*QA*2, 0.6 Kb *Sma*I-*Bam*HI fragment

FIGURE 7

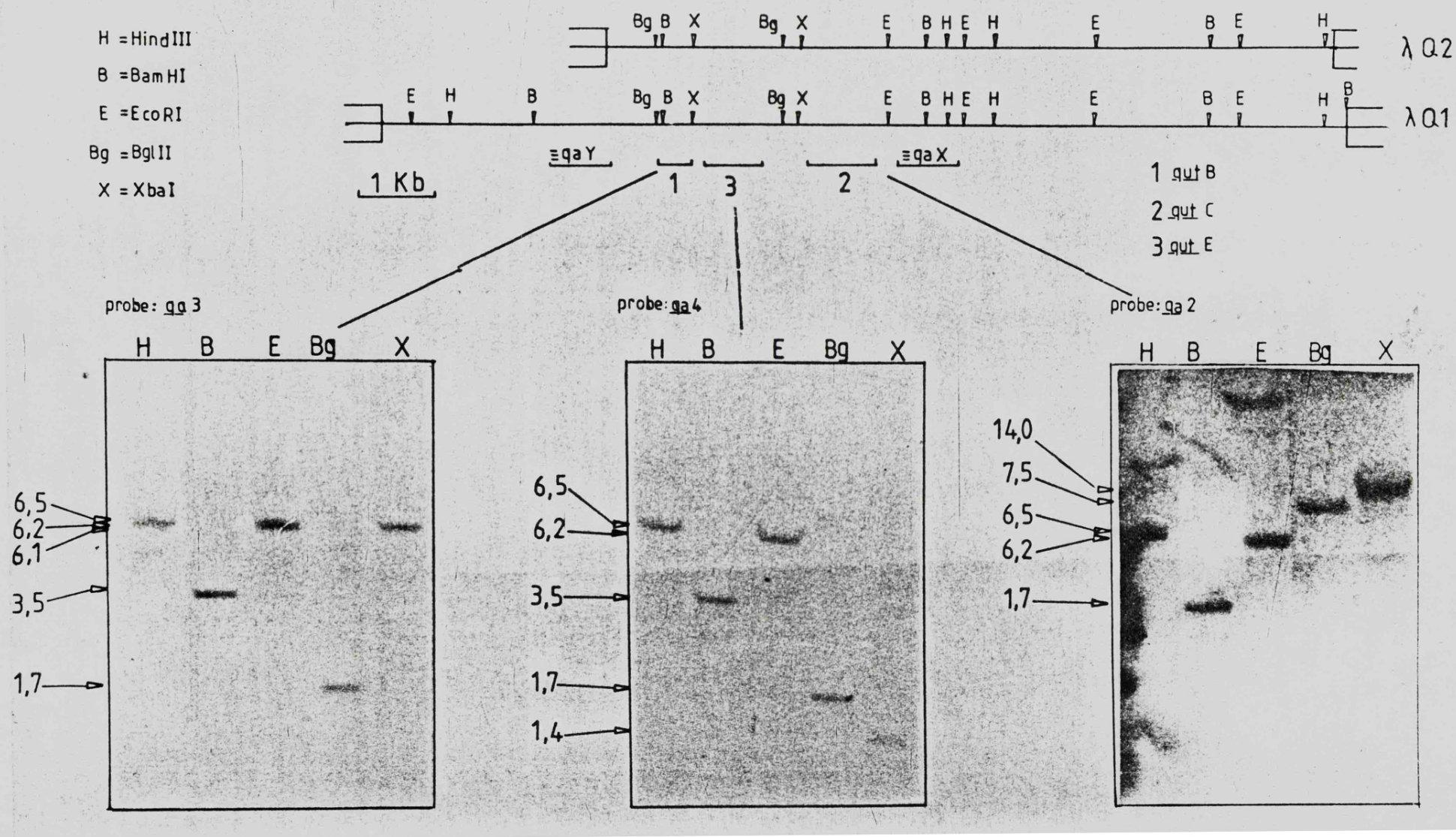
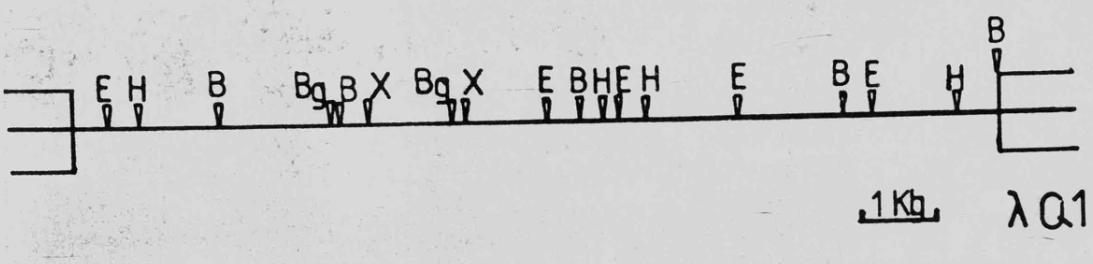

H = HindIII
B = BamHI
E = EcoRI
Bg = BglII
X = XbaI
λ Q2
λ Q1
1 Kb
≡qaY
≡qaX
1
3
2
1 qut B
2 qut C
3 qut E
probe: qa 3
probe: qa 4
probe: qa 2
H B E Bg X
6,5
6,2
6,1
3,5
1,7
1,4
14,0
7,5

FIGURE 8 Identification of sequence organization in *A. nidulans* genomic DNA and the cloned fragment λQ1

The five tracks on each autoradiograph represent *A. nidulans* genomic DNA digested with the restriction enzymes indicated, the products separated by electrophoresis and transferred to nitrocellulose filters. The filters were then probed with α-^{32}P dCTP labelled λQ1. Bands giving fainter signals are indicated by the arrows.

FIGURE 8


E H B Bg B X Bg X E B H E H E B E H B
1 Kb
λ Q1

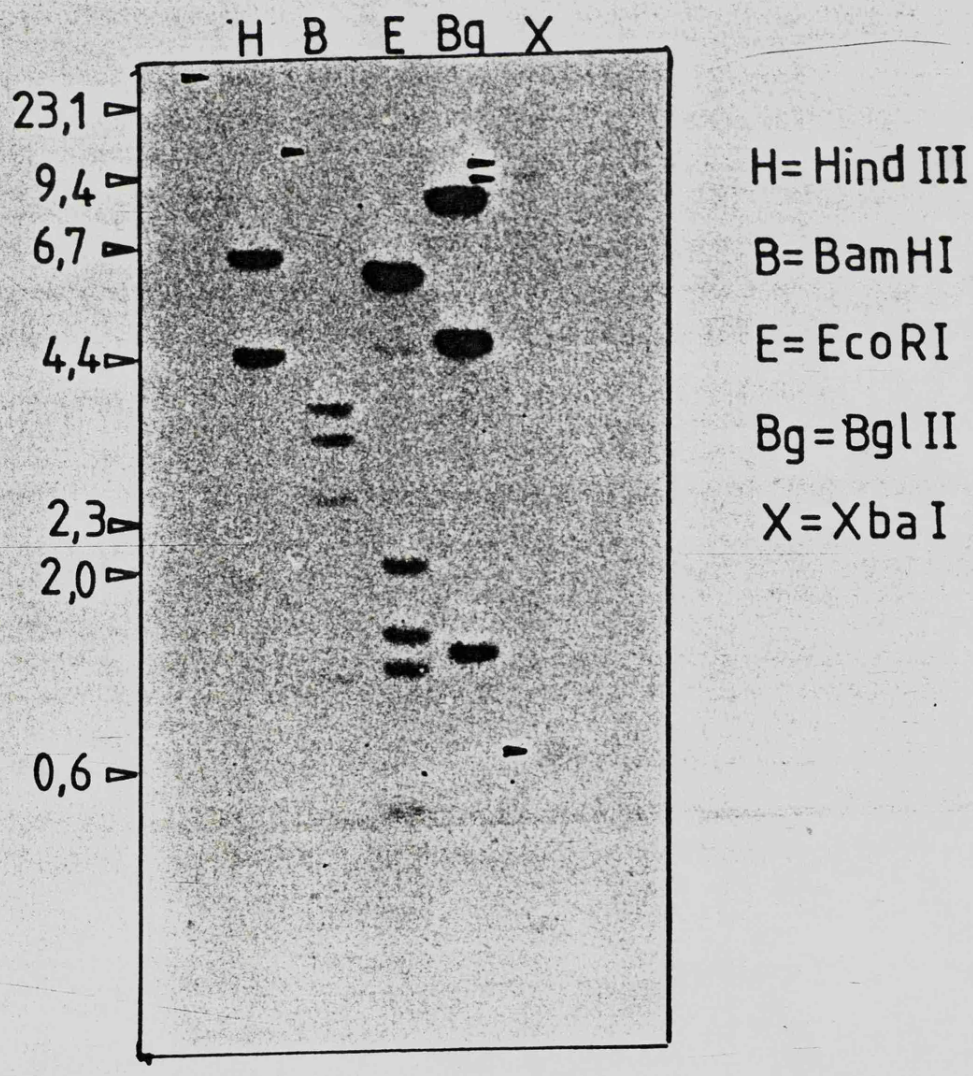

H B E Bg X
23,1
9,4
6,7
4,4
2,3
2,0
0,6
H= Hind III
B= BamHI
E= EcoRI
Bg= Bgl II
X= Xba I

FIGURE 9 Identification of the location of three *QUT* enzyme structural genes in the cloned λQ1 fragment.

DNA was digested with each of the five restriction enzymes listed and in all pair wise combination (listed below), and the products separated by electrophoresis. The upper panel shows a photograph of the gel and the various fragments obtained.

The restriction products were transfered to nitrocellulose filters which were hybridized with probes of the equivalent genes from *N. crassa* (see Figures 5 and 7b) and the resulting autoradiograph are shown in the lower panels. The positions of standard size markers are indicated by by the figure and arrows.

Tracks refer to λQ1 DNA digested as follows:

1.*Hind*; 2.*Hind*III-*Bam*HI; 3.*Hind*III-*Eco*RI; 4.*Hind*III-*Bgl*II; 5.*Hind*III-*Xba*I; 6.*Bam*HI; 7.*Bam*HI-*Eco*RI; 8.*Bam*HI-*Bgl*II; 9.*Bam*HI-*Xba*I; 10.*Eco*RI; 11.*Eco*RI-*Bgl*II; 12.*Eco*RI-*Xba*I; 13.*Bgl*II; 14.*Bgl*II-*Xba*I; 15.*Xba*I.

FIGURE 9

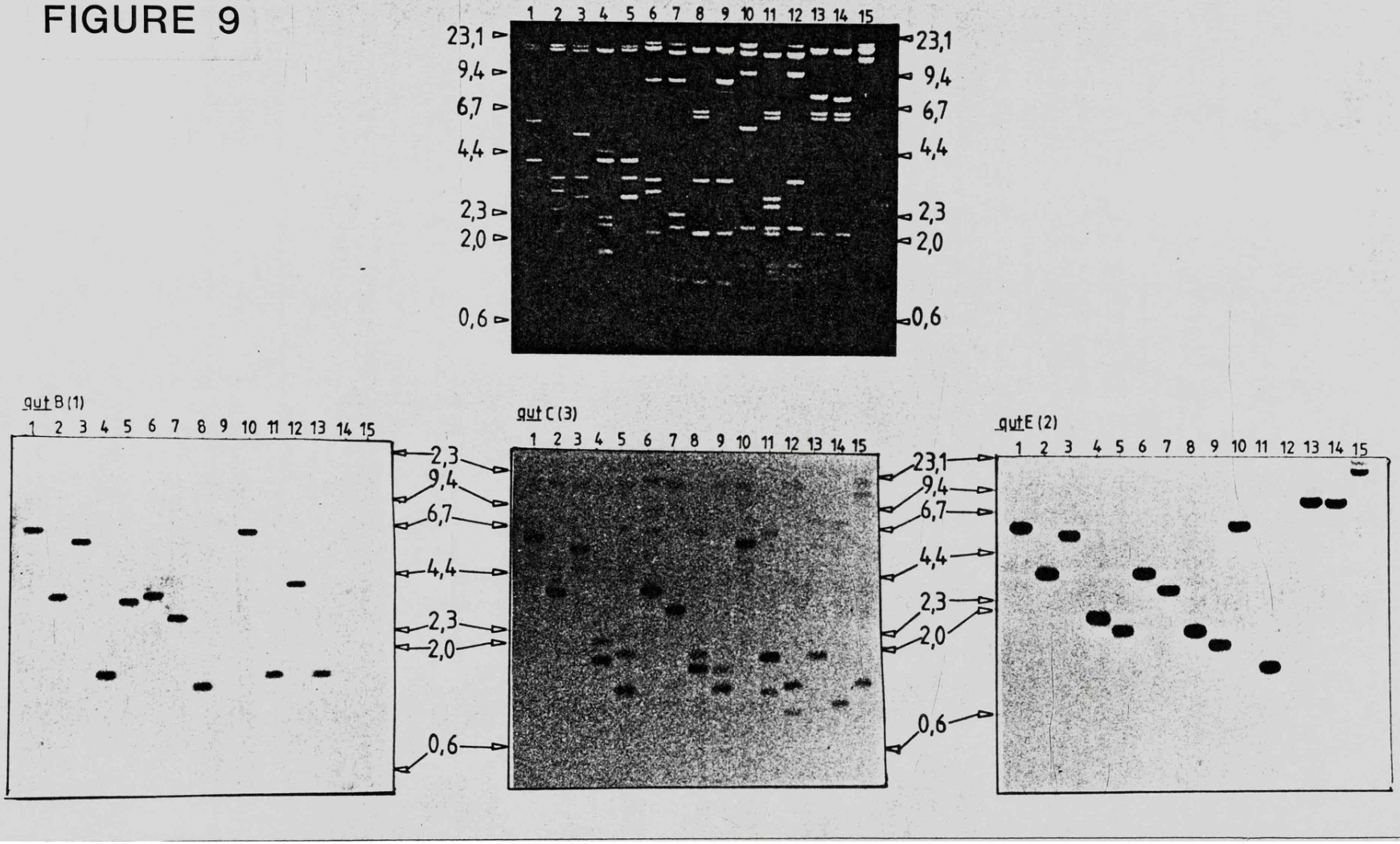
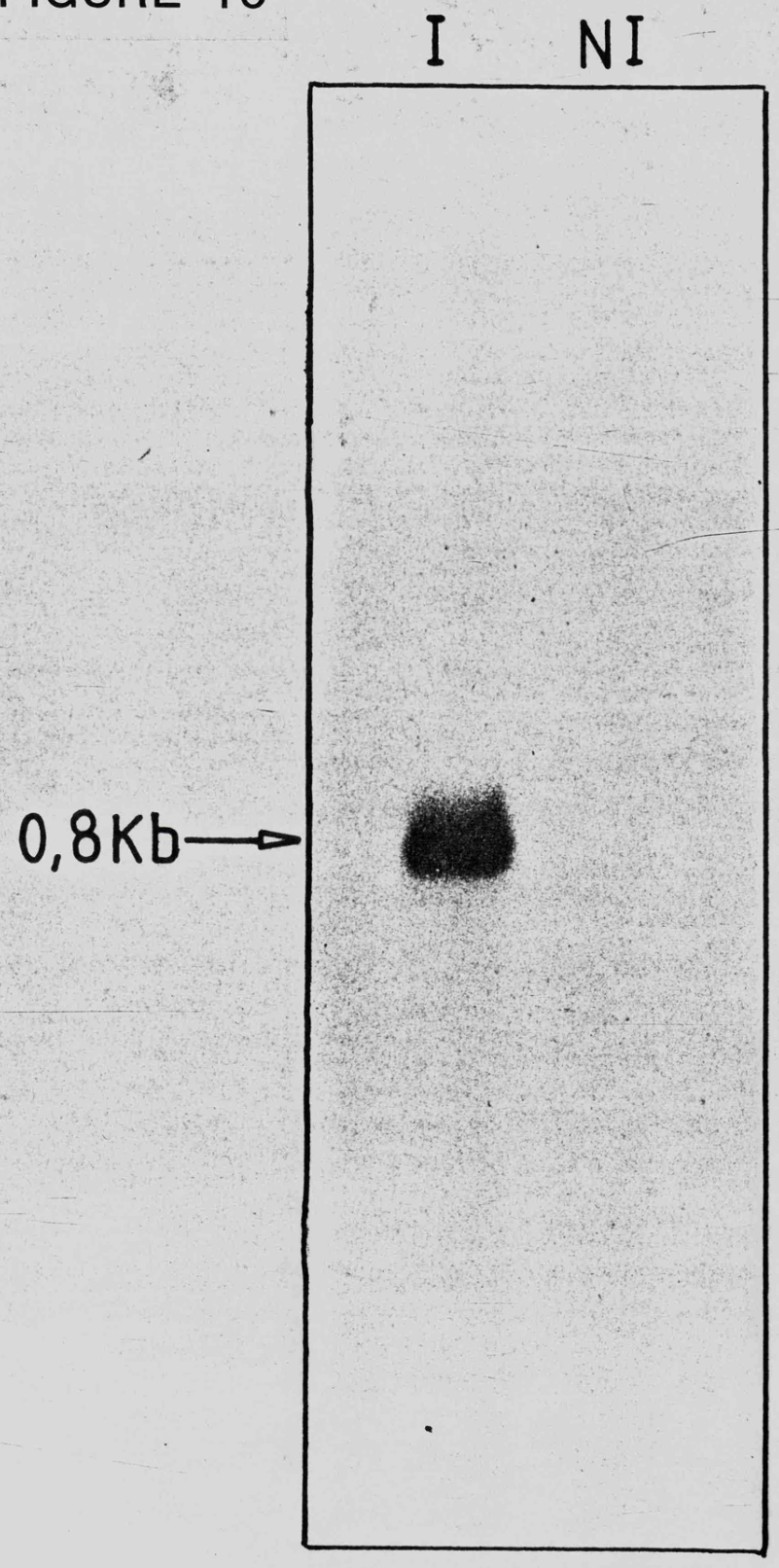

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
23,1
9,4
6,7
4,4
2,3
2,0
0,6
gut B (1)
gut C (3)
gut E (2)

FIGURE 10 Identification of dehydroquinase (*qut*E) mRNA in *A. nidulans*.

Northern blot analysis of total RNA isolated from induced (I) and non-induced (NI) mycelium of the wild-type strain R153 of *A. nidulans*. The DNA probe was the 1.1 Kb *Bgl*II-*Eco*RI fragment containing the *QUT*E gene. The size was determined in comparision with standard markers.

FIGURE 10



CHAPTER 6. LOCATION OF THE TWO CONTROL GENES IN GENOMIC DNA AND ISOLATION OF *QUT*D IN THE RECOMBINANT LAMBDA PHAGE λQ1

The successful cloning and physical mapping of the *A. nidulans QUT* structural genes was facilitated by the DNA sequence homology with the equivalent *QA* genes of *N. crassa* and suggested that the same strategy might be followed for the isolation and identification of the regulatory genes, *QUT*A and *QUT*D. These two genes exhibit characteristics in common with the *QA*1^S and *QA*1^F genes of *N. crassa*, suggesting the possibility that they have common modes of action and therefore are the equivalent genes in the two organisms. If this is true, then the two pairs of genes may also be expected to possess DNA sequence homology and thus the identification of *QUT*A and *QUT*D may be possible by DNA hybridization, using the *N. crassa* genes as probes.

1. Hybridization of *N. crassa QA*1^S and *QA*1^F sequences to *Aspergillus* genomic DNA.

To test whether the *N. crassa* DNA sequences from the two control genes hybridized *Aspergillus* DNA, genomic DNA was digested with the same five restriction enzymes previously used (Figure 7) and the products transferred to nitrocellulose filters after separation by gel electrophoresis. The *Neurospora* DNA probes originated from plasmids pMSK366 (Huiet, 1984) and pMSK375 (Schweizer *et al.*, 1981), which carry *N. crassa QA*1^S and *QA*1^F genes

respectively (Figure 5). In the initial hybridization experiment with the *QA1*^{S} gene, the entire insert from pMSK366 (4.9 Kb in length) was used, consisting almost entirely of unique *QA1*^{S} sequences (Figure 11, left hand panel). The plasmid pMSK375 contains a 5.4 Kb *Bam*H1 insert, which in addition to the *QA1*^{F} gene also includes a *N. crassa* tRNA^{leu} (Huiet *et al.*, 1984). In *Neurospora*, fragments containing this tRNA gene, hybridize to several sequences in the genome giving a smear of bands in the genomic Southern blots (Layne Huiet, personnal communication). It was therefore necessary to generate a DNA probe free of this tRNA gene. This was achieved by further excising the *N. crassa* DNA insert of pMSK375 to isolate the 2.1 *Bam*HI-*Xba*I fragment (Figure 11, right hand panel).

The results of hybridization, done at the standard conditions of stringency (see Chapter 2), to *Aspergillus* genomic DNA, indicate that both of the *N. crassa* genes exhibit homology to a number of *Aspergillus* genomic DNA sequences. However, in the case of *QA1*^{S}, when the 4.9 Kb *Eco*RI fragment from the plasmid pMSK366 is further restricted, and the 2.0 Kb *Xho*I-*Sal*I fragment is purified and used as an hybridization probe against an identical set of filters, a set of single bands was observed to hybridize in genomic DNA (Figure 11, centre panel).

2. Testing λQ1 for sequence homology with the *QA1*^{S} and *QA1*^{F} genes

Experiments done to test whether the two *Neurospora* control genes hybridized to the cloned DNA from

Aspergillus, showed that λQ1 contained sequences homologous to *QA1*^{F} but not *QA1*^{S}.

In the case of the *QA1*^{S} gene, the size of the bands appearing in the genomic Southern blot probed with the 2.0 kb *Xho*I-*Sal*I fragment from pMSK366, which is contained within the *QA1*^{S} gene, are:

*Hind*III	2.7 and 1.2 Kb
*Bam*HI	14.0 Kb
*Eco*RI	8.6 Kb
*Bgl*II	7.5 Kb
*Xba*I	14.0 Kb

None of these band sizes have any equivalents in the cloned λQ1 DNA restriction map, implying that the *Aspergillus* sequence identified by the 2.0 Kb *Xho*I-*Sal*I *N.crassa* probe was not contained within the cloned *A. nidulans* DNA. To confirm this result, a filter with bound λQ1 DNA which had been cleaved with the five restriction enzymes listed above, singly and in all pairwise combinations, was tested with the same 2.0 kb *N. crassa* probe. No hybridization was detected even after prolonged exposure of the X-ray film to the filters, in contrast to the strong hybridization signals generated in the genomic DNA hybridizations. To exclude the possibility that some of the bands appearing in the autoradiograph in which the whole of the 4.9 *Eco*RI insert of the plasmid pMSK366 was used as a probe, could correspond to λQ1 sequences, the same experiment was repeated using this larger fragment as hybridization probe. The result of this experiment was still clearly negative, demonstrating that there is no

homology between the DNA sequences of the *N. crassa* *QA1*^{S} gene and the λQ1 clone of *A. nidulans*.

However, in the case of the *QA1*^{F} gene, the 2.1 Kb *Bam*HI-*Xba*I fragment from pMSK375, which is wholly within the coding region of the *QA1*^{F} gene), strongly hybridizes to λQ1 DNA. It identifies a region immediately to the right of the sequence that, in preliminary experiments, appears to show homology to the *N. crassa* *QAY* gene, thus mapping the *QUT*D gene within the 1.7 *Eco*RI fragment (Figure 12).

3. Strategies for cloning the *QUT*A gene

Genetic analysis of the pleiotropic non-inducible mutant *qut*A4 (Hawkins *et al.*, 1982; Chapter 3), have shown that the *QUT*A gene is closely linked to *QUT*D and the three enzyme structural genes *QUT*B, *QUT*E and *QUT*C. Moreover, the *QA1*^{S} gene of *N. crassa* is postulated to be equivalent to the *QUT*A gene of *A. nidulans*. If this is correct, then the two genes should show sequence homology. The observation that a small 2.0 Kb DNA fragment from within the *QA1*^{S} gene of *N. crassa* strongly hybridizes to a unique fragment of *A. nidulans* genomic DNA sUggests a strategy to isolate this fragment and thus the *QUT*A gene. Since the gene is not present in the genomic clone λQ1 it must of necessity lie, either to the "left" or to the "right" of the λQ1 region at a distance defined by the close linkage of *QUT*A to the *QUT* locus. One should thus be able to recover the *QUT*A gene from a genomic DNA library by "chromosome walking" using fragments from the left and right ends of λQ1 as probes to extend the cloned regions in both

directions.

The terminal fragments used as hybridization probes for this exercise were the 2.4 Kb *Bam*HI fragment (left hand) and the 1.1 Kb *Hind*III-*Bam*HI fragment (right hand) from λQ1. These fragments were isolated and subcloned into pBR322 to generate plasmids pAJS1 and pAJS2 respectively (Figure 12). DNA prepared from these two plasmids was digested with the appropriate endonucleases to produce substantial quantities of the two DNA probes, which were used to screen the same λ-phage genomic library from which λQ1 and λQ2 recombinant phage were isolated. In three attempts, on each occasion screening 10 *A. nidulans* genome equivalents, no plaques were recovered that gave consistent hybridization signals in subsequent cycles of purification.

An alternative route was then attempted using the *N. crassa* 2.0 Kb *Xho*I-*Sal*I DNA fragment from pMSK366, which identifies a unique band in *A. nidulans* genomic DNA upon hybridization, as a probe to screen the library. Five putative positive phage were initially isolated, but again no consistent hybridization signals were observed in subsequent cycles of purification, indicating that these phage did not contain DNA homologous to the *QA*1^{S} gene.

These probes are currently being used in Newcastle to attempt to isolate the *QUT*A gene from newly constructed libraries of *A. nidulans* genomic DNA in phage-λ.

4. Discussion

The identification of a region within the cloned λQ1 DNA sequence of *A. nidulans* that strongly hybridizes to the positivly acting regulatory gene of *N. crassa*, *QA*1^{F},

is postulated to correspond to the *QUT*D gene locus. Confirmation of this proposal, will eventually be obtained from transformation experiments testing complementation of a *qut*D$^-$ functional lesion, by expression of the putative cloned *QUT*D gene carried in this fragment. Additional evidence, would also be supplied by analysis of the sequence of this putative gene compared with that of the *N. crassa* *QA*1^F gene.

The DNA sequence identified in Southern blots of *A. nidulans* genomic DNA probed with the *QA*1^S gene of *N. crassa* are not contained within the cloned λQ1 DNA. Hence, the minimal distance between the putative *QUT*D gene, and *QUT*A, which are genetically closely linked, is at least 4.0 Kb, which is the distance between *QUT*D and the nearest end (at the right hand) of the cloned fragment in λQ1. This observation stands in contrast to the otherwise compact organization of the other *QUT* genes (Chapter 5 and Figure 7). In *N. crassa* the physical distance between the two regulatory genes is only 0.3 Kb.

Comparision of the organization of the quinic acid gene clusters in the genomes of *A. nidulans* and *N. crassa* provides both striking similarities and differences. The structural genes, including the two genes *QAX* and *QAY* whose function have not been characterized, and their hybridization equivalents in *Aspergillus*, appear to occur in the same order: ≡*QAX*, *QUTE*, *QUTC*, *QUT*B and ≡*QAY* in *Aspergillus*; *QAX*, *QA*2, *QA*4, *QA*3 and *QAY* in *Neurospora*. The *QUT*D and *QUT*A genes, on the other hand, do not appear to follow the same relative order (Figures 5 and 7). This difference in the organization is that in *Neurospora* the two control genes *QA*1^S, *QA*1^F lie very close to *QAY* and in

that order (Figure 5), whereas in *Aspergillus QUTD* (≡*QA1*^{F}) appears at the other distal end in relation to the structural genes, and the *QUTA* gene at a much greater distance (at least 4 Kb) than in *N. crassa*. This suggests that major alterations in the regulation of expression of the structural genes may have occurred in the evolution of the two filamentous fungi from their primordial ancestor.

Several observations now seem to point in this direction. For example, attempts to isolate constitutive mutants in the *QUTA* gene have been unsuccessful (Chapter 4), in striking contrast to the situation in *N. crassa* where *qa1*$^{S-}$ mutants spontaneously revert to constitutivity (Valone *et al.*, 1971; L. Huiet, personnal communication). This may be interpretated that the *Neurospora QA1*^{S} protein possesse a domain, which when disrupted by mutations in the corresponding genetic region(s), readly confers constitutivity to the mutant strain. In fact, Giles *et al.* (1985) report that in the *qa1*^{C} mutants sequenced so far, the mutations conferring constitutivity all fall within the amino (NH_2-) terminal region of the repressor protein. Giles and co-workers further discussed the idea that by analogy to mutations in certain prokaryotic repressor genes, for example the *lacI* gene (Miller & Reznikoff, 1978), and to mutants in *gal*80 in *Saccharomyces cerevisiae* (Nogi & Fukasawa, 1984), the two types of mutations detected in the *QA1*^{S} gene, could define two domains in the protein: one the inducer binding domain, at the C-terminal, and identified by the *qa1*$^{S-}$ mutations, and the second for the repressor target domain, in the NH_2-terminal, and identified by the *qa1*^{C} mutations.

The lack of success in the isolation of constitutive

mutants in *A. nidulans*, has been discussed (Chapter 4) and may indicate that the *QUTA* protein plays a more complex rôle than that of the *QA1*^{S} protein in *Neurospora*, where the inactivation of the repressor protein does not affect the ability of the mutant strain achieve enzyme induction and thus utilize quinic acid as carbon source.

Both the *QA1*^{S} and *QA1*^{F} genes of *N. crassa* show homology to several sequences within the *Aspergillus* genome, giving signals corresponding to strong hybridizations in genomic Southern blot analysis. However, when the *QA1*^{S} DNA probe is reduced to just its 3' end which is thought to code for the inducer-binding domain, only one sequence in *Aspergillus* is recognized. An attractive interpretation is that the 5' end, which may code for the DNA binding domain (Giles *et al.*, 1985), has significant DNA sequence homology to several regions of the *A. nidulans* genome, and is thus identifying various genes coding for DNA binding proteins in *Aspergillus*. Reports on the structure and organization of fungal DNA binding proteins are scarce at this time. However, Laughon & Gestland (1984) reported that the NH_2- terminal region of the activator protein encoded by the *GAL4* gene, contains an amino acid sequence which resembles the two-helix DNA-binding domain found in some prokaryotic repressor and activator proteins. The DNA probes from the *QA1*^{S} and *QA1*^{F} genes that gave multiple bands in the *Aspergillus* genomic Southern blot hybridizations were both from the NH_2- terminal ends of the coding sequence of these two *Neurospora* genes, and suggests that they may identify sequences coding potential DNA binding domains that are common to proteins coded by different regulatory

genes.

FIGURE 11 Identification on *A. nidulans* genomic DNA of sequences showing homology to the *N. crassa* control genes, *QA1*^{S} and *QA1*^{F}.

The physical organization of the *N. crassa* regulatory region is shown in the top panel and the extent of the DNA probes used indicated.

A. nidulans genomic DNA was digested with the five restriction enzymes used previously (Figure 7), the products separated by electrophoresis, transferred to nitrocellulose filters and tested by hybridization with the probes described above.

DNA fragments identified in the resulting autoradiograph are shown together with estimated sizes determined against known standard markers.

FIGURE 11

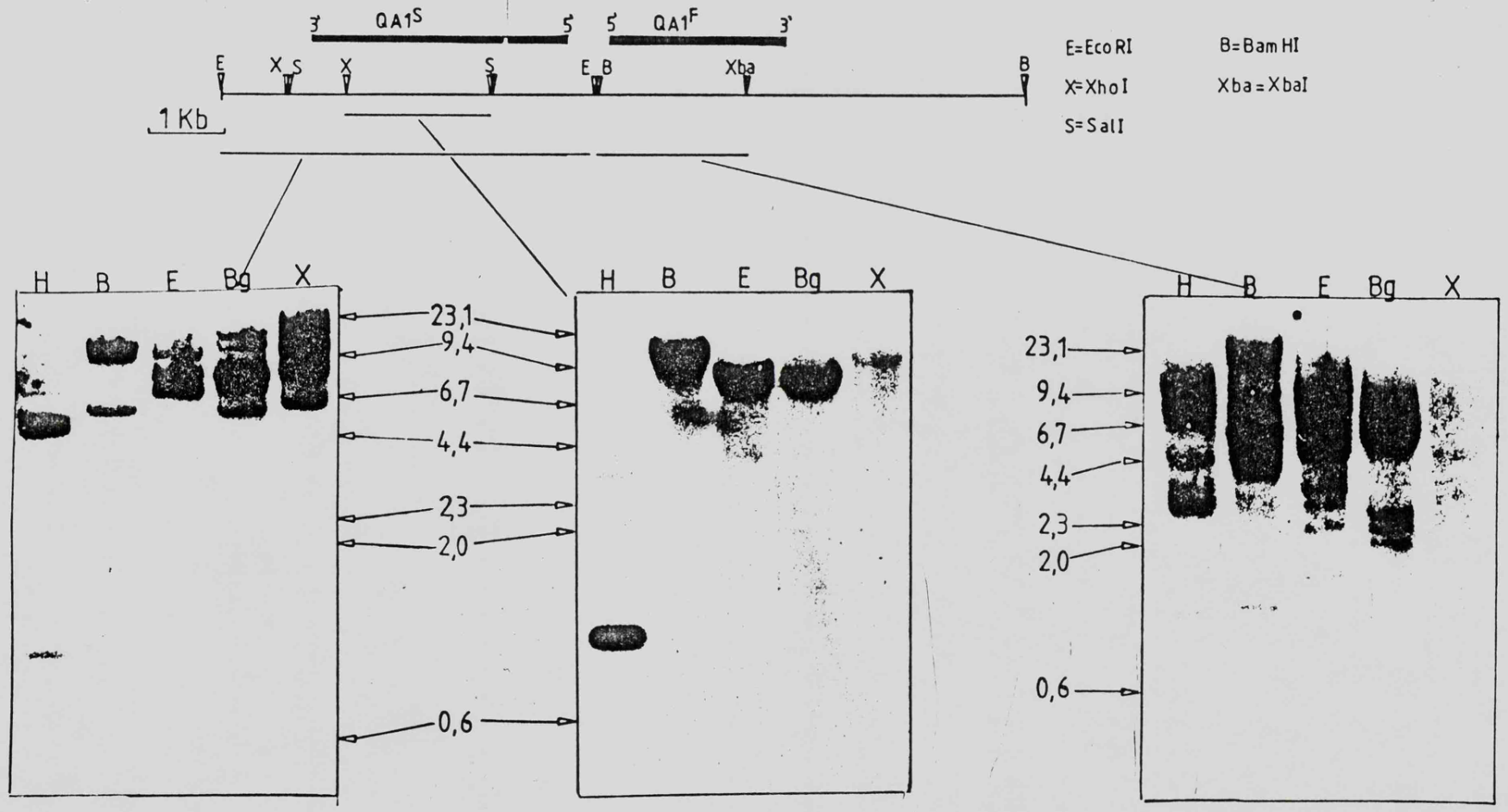
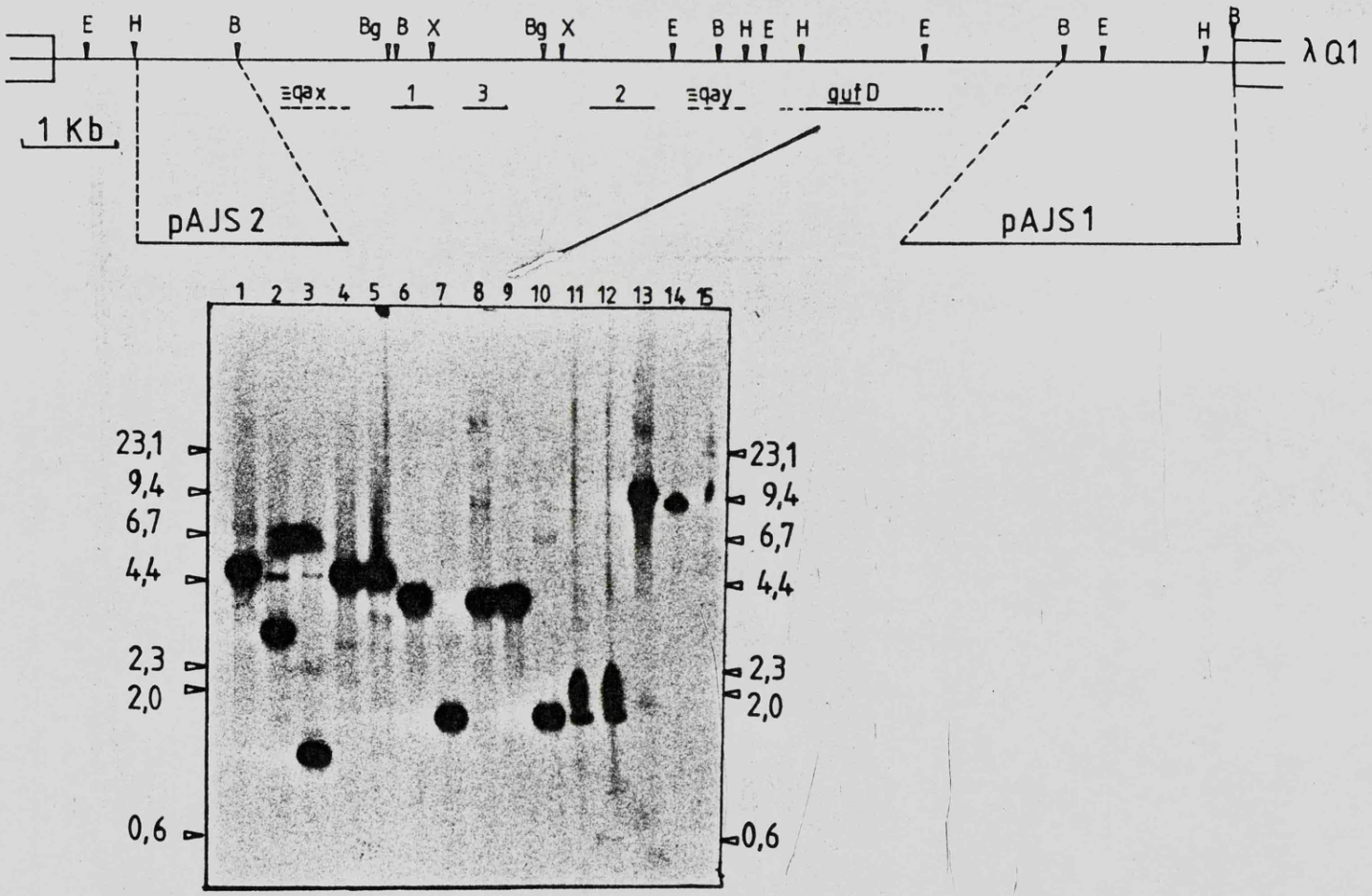

3' QA1S 5'
5' QA1F 3'
E X S X S E B Xba B
1 Kb
E=Eco RI
B=Bam HI
X=Xho I
Xba=Xba I
S=Sal I
H B E Bg X
23,1
9,4
6,7
4,4
2,3
2,0
0,6
H B E Bg X
H B E Bg X
23,1
9,4
6,7
4,4
2,3
2,0
0,6

FIGURE 12 Identification of the *QUTD* gene in λQ1.

The upper panel indicates the restriction map of λQ1.
The λQ1 DNA was digested with the five restriction enzymes indicated singly and in all pairwise combinations, separated by electrophoresis, transfered to nitrocellulose filters and hybridized with $^{\alpha-32}$P-dCTP labelled 2.1 Kb *Bam*HI-*Xba*I fragment from within the *QA1*^{F} gene of *N. crassa* (Figure 11). The positions of standard size markers are indicated by the figure with arrows.
Order of digests is: 1.*Hind*III; 2.*Hind*III-*Bam*HI; 3.*Hind*III-*Eco*RI; 4.*Hind*III-*Bgl*II; 5.*Hind*III-*Xba*I; 6.*Bam*HI; 7.*Bam*HI-*Eco*RI; 8.*Bam*HI-*Bgl*II; 9.*Bam*HI-*Xba*I; 10.*Eco*RI; 11.*Eco*RI-*Bgl*II; 12.*Eco*RI-*Xba*I; 13.*Bgl*II; 14.*Bgl*II-*Xba*I; 15.*Xba*I.
The two end fragments from the cloned DNA to be used for the chromosome "walk" are also indicated (pAJS1 and pAJS2).

FIGURE 12

H = Hind III
B = Bam HI
E = EcoRI
Bg = BglII
X = XbaI



CHAPTER 7. SEQUENCING OF THE *QUT* ENZYMES STRUCTURAL GENES: ANALYSIS OF THE CATABOLIC DEHYDROQUINASE GENE *QUTE*

The model proposed for the regulation of the expression of the *QUT* genes coding for the enzymes catabolizing quinic acid to PCA, postulates that the two regulatory proteins exert their activity upon the DNA, by interacting with a region 5'-upstream of the start of transcription of each of the genes. An analysis of the DNA sequences of these promoter regions, might be expected to reveal features in common not only amongst the *QUT* structural genes but also with those of other fungal genes under similar physiological controls (Chapter 1). Such common sequence motifs in the promoter regions would give an indication of those regions recognized by the regulatory proteins, and also of the interaction with RNA polymerase II to initiate transcription.

Furthermore, there is some evidence that the regulation of expression of the *QUT* genes by the product of the *QUT*A gene in *A. nidulans* and the *QA* genes by the *QA*1^{S} protein in *N. crassa* may operate differently (Chapters 4 and 6). If this view is correct, one may expect to find extensive DNA sequence homology in the coding sequences between the *QUT* and *QA* structural genes, since they code for equivalent enzymes, but less so in the promoter regions, as a reflection of the different modes of regulation.

The plasmid pEH1, contains a 6.5 Kb fragment of *A. nidulans* DNA from λQ1 which includes the three enzyme structural genes (Chapter 5). Inspection of the

restriction map of this insert reveals a wealth of strategically well placed endonuclease restriction sites, enabling the isolation of small, but well defined fragments in the range 1.1 to 0.5 Kb for subcloning into M13 phage sequencing vectors (Figure 13). Thus, directional sequencing was chosen as the quickest method for determining the DNA sequence of the three genes.

1. Directional sequencing

Each sequencing reaction permitted some 300 to 400 bases to be read from each end of the restriction fragments by using sequencing gels subjected to electrophoresis for 3 or 6 hours. To extend these sequences, the same DNA templates were used in conjunction with an internal specific primer, consisting of a synthetic 17-mer oligonucleotide designed to anneal to a sequence proximal to the point where reading of the initial sequencing gels had ended ("primer extension"). It was expected that this technique applied to small sized fragments would yield sequence data which would reveal overlap of complementary sequences read from opposite ends of each strand of the fragment.

Twelve 17-mer synthetic oligonucleotide primers were used. However, only six enabled primer extensions of the previously determined sequence. A common observation in the unsuccessful sequencing reactions, was the excessive number of bands appearing in the sequencing tracks. This might be due to annealing of the synthetic primer at several sites, causing the initiation of DNA synthesis at several different positions on the DNA template, and thus

the production of a series of staggered overlapping oligonucleotides. Charles *et al.* (in press) experienced similar difficulties in determining the DNA sequence of the *arom* gene complex of *A. nidulans*. Since this strategy has been successfully used for sequencing procaryotic genes, one possible explanation is that related sequences occur in the *Aspergillus* cloned fragments which are sufficiently homologous to some of the synthetic primers to permit their annealing to the DNA templates at more than one site.

Fragment E (≈1.1 Kb *Bgl*II-*Eco*RI, Figure 13), which hybridizes to the *QA*2 gene (≡*QUT*E) of *N. crassa*, was sequenced from the *Eco*RI site to a total lengh of 492 bases (which includes the use of an internal synthetic primer). The data reveal DNA sequence homology to the 3' end of the *Neurospora QA*2 gene (Giles *et al.*, 1985), thus indicating that the *QUT*E gene is transcribed left to right, that is in the direction from the *Bgl*II to the *Eco*RI sites (Figure 13).

2. Random fragment sequencing

The difficulties encountered with the use of internal synthetic primers required a change of strategy and DNA sequence was determined from randomly generated fragments. For this purpose, the plasmid pEH1 was treated with ultrasonic vibration to introduce multiple breaks at random in the circular molecule. The DNA fragments were separated by electrophoresis and those in the size range 800 to 1000 bp isolated. These fragments were incubated with the Klenow fragment of DNA polymerase plus all four

dNTPs, to repair single stranded ends, and ligated into the *Sma*I site of the phage vector M13mp19. The ligation products were then used to transform strain JM101 of *E. coli*. A total of 357 colourless "plaques" on X-GAL indicator plates (revealing disruption of the function of the β-Galactosidase gene by insertion of the foreign fragments) were recovered (Chapter 2). These M13 phage recombinants were then screened to identify those carrying DNA sequences containing the *A. nidulans QUT* genes, using two restriction fragments as probes: the 3.5 and 1.5 Kb *Bam*HI fragment from plasmid pEH1 (Figure 13). To identify those phage carrying DNA sequences coding for the *QUT*E gene, the same 357 phages were screened by hybridization with the 1.1 Kb *Bgl*II-*Eco*RI fragment (Figure 13). The results of these tests revealed that 101 phages contained DNA sequences from the two *Bam*HI restriction fragments, of which 31 specifically hybridized to the 1.1 Kb *Bgl*II-*Eco*RI fragment. The remaining 70 recombinant phages, which contain fragments from the *QUT*B and *QUT*C genes, are currently being sequenced by Dr. A. Hawkins in the Department of Genetics, University of Newcastle.

DNA was prepared from the 31 phages, containing fragments of the *QUT*E gene and used as templates in sequencing reactions. Of these templates, 8 did not give readable sequence. In 11 it was not possible to align the sequence with those previously obtained from fragment E (see above), and these phage are thought to contain *Aspergillus* sequences of fragment E with the *QUT*E gene sequences at the opposite end of the universal primer site. The remaining 12 templates yielded sequences overlapping with one another and some with the fragment E

sequence, to give a total lengh of 2186 of unique sequence (Figure 13). It is noticeable that several of these 12 fragments are clustered around the *Eco*RI site (Figure 13). A possible explanation is that this region is especially sensitive to breakage by ultra-sonic vibration, therefore generating fragments that are represented more frequently than expected from strictly random shearing of the plasmid as discussed by Deninger (1983).

3. Analysis of the *QUTE* gene DNA sequence.

a) Analysis of the coding sequence

Inspection of Figure 13, shows that both strands were only sequenced for regions containing the *QUTE* gene. Confirmation of the sequence determined over these single stranded regions was obtained by Dr. A. Hawkins in Newcastle, using specific internal primers in sequencing reactions extending the sequence determined from clone 43 (Figure 13). The entire sequence of both strands of fragment E has thus been determined.

The 2186 base sequence of DNA determined from the *A. nidulans* contains an open reading frame of 462 bases (Figure 14), which include the sequence that had been previously identified as having homology with the 3' end of the *QA2* gene (see above). The calculated molecular weight of a protein of the derived amino acid sequence is approximately 18,000. This is in good agreement with the observed MW for the deydroquinase subunit (18-20,000; Hawkins *et al.*, 1982 and Chapter 1).

Comparative analysis of the *Aspergillus* and

Neurospora sequence data using a dot-matrix computer programme (White *et al.*, 1984) (Figure 15) shows that there are regions of considerable homology with both the 5' and the 3' end of the *Neurospora QA2* gene (Giles *et al.*, 1985). However, the dot-matrix also suggests that some evolutionary rearrangements have occurred between these two genes, in that the *QA2* gene appears to contain an additional DNA sequence at approximately the centre of its reading frame. This is illustrated in Figure 16 which shows alignment of the *QUTE* and the *QA2* gene DNA sequences in regions which exhibit homology. Hence, the *Neurospora QA2* gene contains a 75 bp sequence, coding for 25 amino acids, which is absent from the *QUTE* gene of *Aspergillus*. This observation indicates that this sequence was either deleted from the *QUTE* gene or that the *QA2* gene has gained this insertion. The latter hypothesis is preferred, since it is easier to envisage that an insertion would be less likely to disrupt the sequence encoding for an active enzyme. At the 5' end of the two genes, the *QA2* gene contains six additional nucleotides, corresponding to two amino acids, adjacent to the methionine (ATG) start codon. The *QUTE* gene is in turn 21 bp longer at its 3' end than the equivalent gene from *N. crassa*; that is an extra 7 amino acids. In total the coding region of the *QUTE* gene contains 459 bases (153 amino acids), 60 bases (20 amino acids) less than the *N. crassa QA2* gene. A summary of the sequence rearrangements between the polypeptides coded by the two genes is schematically presented in Figure 17.

Considering the regions in the coding sequence which show homology (Figure 16), the 5' region has 60% DNA homology and the 3' region 65%. In the derived

polypeptides, conservation is 63% and 65% respectively (Figure 14, the underlined residues). An interesting feature in the 5' coding region of the two genes is that there is a block of 13 nucleotides which show total homology, but which are out of frame (Figure 16). This sequence is located between positions 73 and 85 in the *QUT*E gene, and positions 72 and 84 in the *QA*2 gene. The introduction of a cytosine base at position 72 in the *QUT*E gene sequence induces the frameshift, but the initial reading frame is restored by the inclusion of a thymine at position 85 in the *QA*2 gene sequence.

The distribution of codon usage in the *QUT*E gene (Figure 18) does not show preferential codon usage, in contrast to the *QA* genes of *N. crassa*, which show bias against those codons containing an adenine (A) in third position (Rutledge, 1984). However, the small size of the derived polypeptide (183 residues), makes it difficult to draw at any firm conclusion on this point. Codon usage is normally thought to play a rôle in modulating the efficiency of translation of the mRNAs at the ribosomes. In the yeast *S. cerevisiae* codon bias is found to be more extreme in genes that are constitutively expressed at high levels than in those expressed at low levels (Bennetzen & Hall, 1982). In some highly expressed genes 96% of the amino acid residues are encoded by only a select 25 of the 61 possible coding triplets, and these preferred codons correspond to the anti-codons of the most common isoacceptor tRNA species in the cell.

b) Analysis of the 5' and 3' flanking sequences

The analysis of the *QUTE* gene non-coding flanking sequences reveals some interesting features (Figure 14). The sequence found around the start codon (first methionine) in the translated sequence shows agreement with the sequence proposed by Kozac (1981) for the initiation of translation by eukaryotic (mainly mammalian) ribosomes and the consensus sequence from several filamentous fungi genes (Figure 19).

QUTE:	A C A A T G G
Filamentous fungi:	A C A A T G $^{T}_{G}$
Kozac sequence:	$^{A}_{G}$ N N A T G G

A TATA sequence is found at position 232, 84 bp upstream from the start of translation (Figure 14). A second sequence TAAATAA, is found at position 260, 68bp upstream from the start of translation. However, this second sequence appears downstream of the two CAAG sequences found in this promoter region. This is in opposition to the situation in yeast, where this TATA sequence is found upstream of the CAAG sequences (Sentenac & Hall, 1982), and is apparently required for the proper positioning of RNA Polymerase II (McKnight & Kingsbury, 1982). On balance the first TATA sequence (232), therefore appears the most likely to have significance in the *QUTE* promoter. In higher eukaryotes, a "TATA" box is typically found about 30 bp upstream of the site of initiation of transcription, but in the yeast *S. cerevisiae* the sequence is found at variable distances from the transcription

start sites (Sentenac & Hall, 1982).

A number of yeast genes that encode abundant mRNAs, seem to have their transcription start sites at or near the sequence CAAG and a pyrimidine rich block is usually found upstream of the CAAG sequence. This sequence seems to be involved in specifying trancriptional initiation or capping of mRNAs (McKnight & Kingsbury, 1982). A GAAG sequence is found at the start of transcription of the glucoamylase gene of *Aspergillus niger* (Boel *et al.*, 1984) and the *TRPC* gene of *A. nidulans* (Mullaney *et al.*, 1985) but not in any of the *N. crassa QA* genes (Alton *et al.*, 1982; Rutledge, 1984).

Two such CAAG sequences appear in the promoter region of the *QUTE* gene at positions 239 and 252. However, since the initiation site(s) is yet to be determined, speculation on the importance of these sequences has to be treated with caution. Both sequences are very near the putative "TATA" box, which seems to devalue their potential importance. However, the sequence just upstream from the first CAAG sequence, starting at position 199, is 73% T/C rich (30 of 41 nucleotides).

More importantly however, are two sequences at around positions 168 and 190. The sequence centered at nucleotide 168 contains the "CAAT" box (Breathnach & Chambon, 1981). This sequence has been recognized in many higher eukaryotes has being important for the binding of the RNA Polymerase II to the DNA, and is usually located 70 to 80 bp upstream from the start of transcription. The *Aspergillus QUTE* gene strikingly shows almost complete sequence homology to the canonical sequence postulated for mammalian genes (Benoist *et al.*, 1980):

QUTE:	G G G C A A T C T
Mammalian genes:	G G Py C A A T C T

Downstream from this "CAAT" box, there is a sequence centered at position 190, which shows a perfect dyad of symmetry. Such sequences, are generally recognized in both eukaryotes and prokaryotes, as candidate sequences for the binding of proteins to DNA. The possible rôle of this sequence in the regulation of expression of the *QUTE* gene is discussed below.

The 3' non-coding flanking sequences that are transcribed by RNA Polymerase II in eukaryotes, have been less well characterized than the 5' non-coding regions. A sequence commonly found near the termination of transcription of yeast genes TATGT <...> TTT (Zaret & Sherman, 1982), is also found in the 3' non-coding region of the *QUTE* gene (Figure 14). Another sequence, AATAAA followed by several T/G rich blocks of sequence, is typically found near the sites of polyadenylation in eukaryotic genes (Proudfoot & Brownlee, 1976; Birnstiel *et al.*, 1985) finds a near consensus sequence around position 945 in the *QUTE* gene. However, until the end of the *QUTE* gene mRNA transcripts are determined, these interpretations are purely hypothetical.

4. Discussion

The DNA sequence of the *QUTE* gene coding for the catabolic dehydroquinase from *A. nidulans* has been determined and analysed. DNA sequence homology with the equivalent gene from *N. crassa*, *QA2*, has been found as

expected from the DNA hybridization experiments described in Chapter 5. However, the two genes exhibit evolutionary rearrangements in their internal structure. Hence, the *QA2* gene possesses 75 extra nucleotides at approximately the middle of the gene, whereas the *Aspergillus QUTE* gene is 21 bp longer at the 3' end. The two regions that show good DNA sequence homology probably include the sequences encoding the amino acids in the catalytic sites of the dehydroquinase enzyme, and which therefore have been conserved. However, the discovery that the two genes code for structurally different polypeptides, may provide an explanation for why the *N. crassa QA2* gene does not complement the *A. nidulans qut*E208 deficiency (Hayley Whittington, unpublished results), which however is efficiently transformed by the *QUTE* gene (Chapter 5). The enzyme is multimeric in both organisms (Hawkins *et al.*, 1982; Chapter 1) consisting of an aggregate of 10-12 monomers of approximately 20,000 molecular weight. The lack of complementation of the *qut*E208 lesion by the *Neurospora* gene may be due to inactivation of a functional polymeric protein, due to competition from defective, and structurally heterologous, *Aspergillus* polypeptides. This negative result is in contrast with other genes in filamentous fungi, in which functionally homologous genes are sufficiently similar in structure to allow heterologous expression. For example, the *N. crassa PYR4* gene complements an *A. nidulans pyr*G mutation (Balance *et al.*, 1983) and conversely the *A. nidulans ARGB* gene complements a *N. crassa arg*-12 mutation (Berse *et al.*, 1983).

The *N. crassa QA2* gene 5' non-coding flanking

sequence, contains a sequence, 11 bp upstream of the initiation of translation codon (ATG) which shows homology to the bacterial canonical sequence for the 3' end of the 16S ribosomal RNA (Alton *et al.*, 1982). This sequence is absent from the *QUTE* promoter region, and this may explain why by contrast, the *Neurospora QA2* gene is more highly expressed in *E. coli* (Chapter 5), possibly due to a proper alignment of the *QA2* transcript on the *E. coli* ribosomes.

The sequence CCACC, which is suggested (Sargan *et al.*, 1982) to function as ribosome binding site for eukaryotes, is absent from the *QUTE* gene, in common with the situation found with the *QA* genes (Rutledge, 1984).

The *QUTE* promoter region however, is more typically eukaryotic than in the *QA* genes. The *Neurospora* genes lack most of the consensus sequences usually described as playing important regulatory rôles in orientating transcription, namely the "TATA" and "CAAT" boxes (Alton *et al.*, 1982; Rutledge, 1984). By contrast, other genes from filamentous ascomycetes (*N. crassa* NADP-specific ADP/ATP carrier protein: Arends & Sebald, 1984; *A. niger* glucoamylase: Boel *et al.*, 1984; *A. nidulans SPO* C1-C gene: Mullaney *et al.*, 1985) have one or both of these sequences. None of the promoter regions of these genes has been subjected to functional analysis, but since these promoter motifs are also found in yeast and in the higher eukaryotes, one assumes that they also play a vital rôle in the filamentous fungi.

The most important feature of the analysis of the 5' control region of the *QUTE* gene for the theme of this thesis however, was the discovery of a dyad of symmetry, approximately 120 nucleotides 5' to the start of

translation. Such sequences are usually interpreted as possible sites in the DNA that are recognized by proteins binding to DNA.

Studies of the sensitivity of the *N. crassa QA* genes to digestion *in vitro* by DNase I (Giles *et al.*, 1985), reveal that under inducing conditions, hypersensitive sites are found in the promoter regions of these genes, and that these sites fall within or very near to dyads of symmetry which also exhibit considerable sequence homology to one another. Furthermore, if the same studies are carried out in $qa1^F$ mutants (activator deficient), these sites are no longer sensitive to DNase I nuclease activity. Giles and his co-workers conclude that these sequences were recognized by the activator protein, coded by the $QA1^F$ gene, which caused a relaxation of the chromatin, normally allowing RNA polymerase to attach to the coding DNA strand, and thus conferring sensitivity to attack by DNase I.

Giniger *et al.* (1985) studying the rôle of the *GAL4* activator protein in regulating expression of the *GAL1* and *GAL10* structural genes, showed that this regulatory protein binds to four sites in the upstream sequence to activate transcription of the two genes. The sequence of these sites each display 2-fold rotational symmetry. However when a single near-consensus synthetic oligonucleotide was installed in front of the *CYC1* gene, it brought the *CYC1* gene under galactose-dependent transcription.

The dyad of symmetry found 120 bp upstream from the translation start codon (ATG), shows considerable homology not only to the *N. crassa QA* genes consensus sequence, but

also to the *GAL* gene consensus sequence in *S. cerevisiae* (Figure 20). Surprisingly, the degree of homology is the same betwen the *QUT*E sequence and the *N. crassa* sequence, and with the *GAL* sequences. Furthermore, the conserved nucleotides appear to be the same in the three organisms. The significance of these evolutionary conserved nucleotides is difficult to assess. Could it be that DNA binding activator proteins recognize similar DNA sequences? The observation that the *QA*1^{F} gene of *N. crassa* hybridizes to several different sequences in genomic DNA from *Aspergillus* may provide an argument in this direction. If this suggestion is correct, then a given activator protein would require additional interaction to provide specificity for particular genes would require the interaction of an additional co-activator. In the *QUT* system of *A. nidulans*, this funtion could possibly be played by the inducer quinic acid interacting with the *QUT*D activator protein or the *QUT*A protein, which would be responsible for guiding the activator protein, coded by the *QUT*D gene, to the promoter regions of the *QUT* structural genes.

Inspection of the 5' non-coding region of the *QUT*E gene has only revealed one such dyad of symmetry. In *N. crassa* the *QA* genes can contain one or more such sequences (Giles *et al.*, 1985). The yeast *GAL*1 and *GAL*10 genes are closely linked, and are transcribed divergently sharing a common 5' non-coding regulatory sequence. Four very tightly linked sequences showing dyads of symmetry are found in this region (Giniger *et al.*, 1985). When each of these sequences is placed singly upstream of the *GAL*1 gene, levels of transcription are reduced several fold.

However, if a consensus sequence oligonucleotide is substituted, transcription regains the wild-type level. It was thus concluded that the presence of these separate sequences was required as a mechanism for of regulation, to permit titration of transcription when levels of activator protein become depleted. In the light of these reports, it is possible that the other *QUT* structural genes may also contain multiple putative sites for binding of the *QUT*D activator protein. This will be assessed by determining the DNA sequence of their 5' control regions, a project which was started during the work for this thesis, and is now being completed by Dr. A. Hawkins in Newcastle.

Some 22 bp upstream from the centre of the rotational symmetry of the dyad, one finds a "CAAT" box exhibiting in its flanking nucleotides remarkably good homology with the canonical sequence for higher eukaryotes (Benoist *et al.*, 1980). Such sequences are commonly recognized as being involved in the attachment of RNA polymerase II to the DNA. If this is the rôle of this sequence in the upstream control region of the *QUT*E gene, then one observes that it is strategically located, just prior to the putative binding site for the *QUT*D activator protein. Hence, in the absence of the activator protein, the dyad would form a "hairpin loop" thus preventing the RNA Polymerase II from proceeding along the DNA to start transcription. However in the presence of the activator, the tertiary structure of the DNA chromatin at this locus would be relaxed, allowing the movement of the RNA polymerase.

Interestingly, the 5' control regions of the *QUT*E and *QA*2 genes do not show evolutionary conservation. Whether

this divergence indicates major differences in the regulation of transcription, specific to either each of the microorganisms in general or the quinic acid utilization cluster in particular, is not known.

FIGURE 13 **Strategy for sequencing the *QUTE* gene of *A. nidulans*.**

The restriction map of pEH1, which contains the 6.5 Kb *Hind*III fragment of λQ1, is shown. Below this map, (a) shows the fragments subcloned into phage M13 for directional sequencing and (b) shows the length of sequence determined in fragment E and adjacent fragment B and Eco. The sites of the specific internal primers used are indicated (*). The extent and direction of DNA sequence determined from the randomly generated fragments of fragment E are shown (c) and the numbers identify the particular sub-clones used.

Determination of the second strand was done by the use of internal specific primers in the sequencing reactions using clone 43 as DNA template (d).

Restriction enzymes are:

H : *Hind*III

B : *Bam*HI

E : *Eco*RI

Bg : *Bgl*II

X : *Xba*I

FIGURE 13

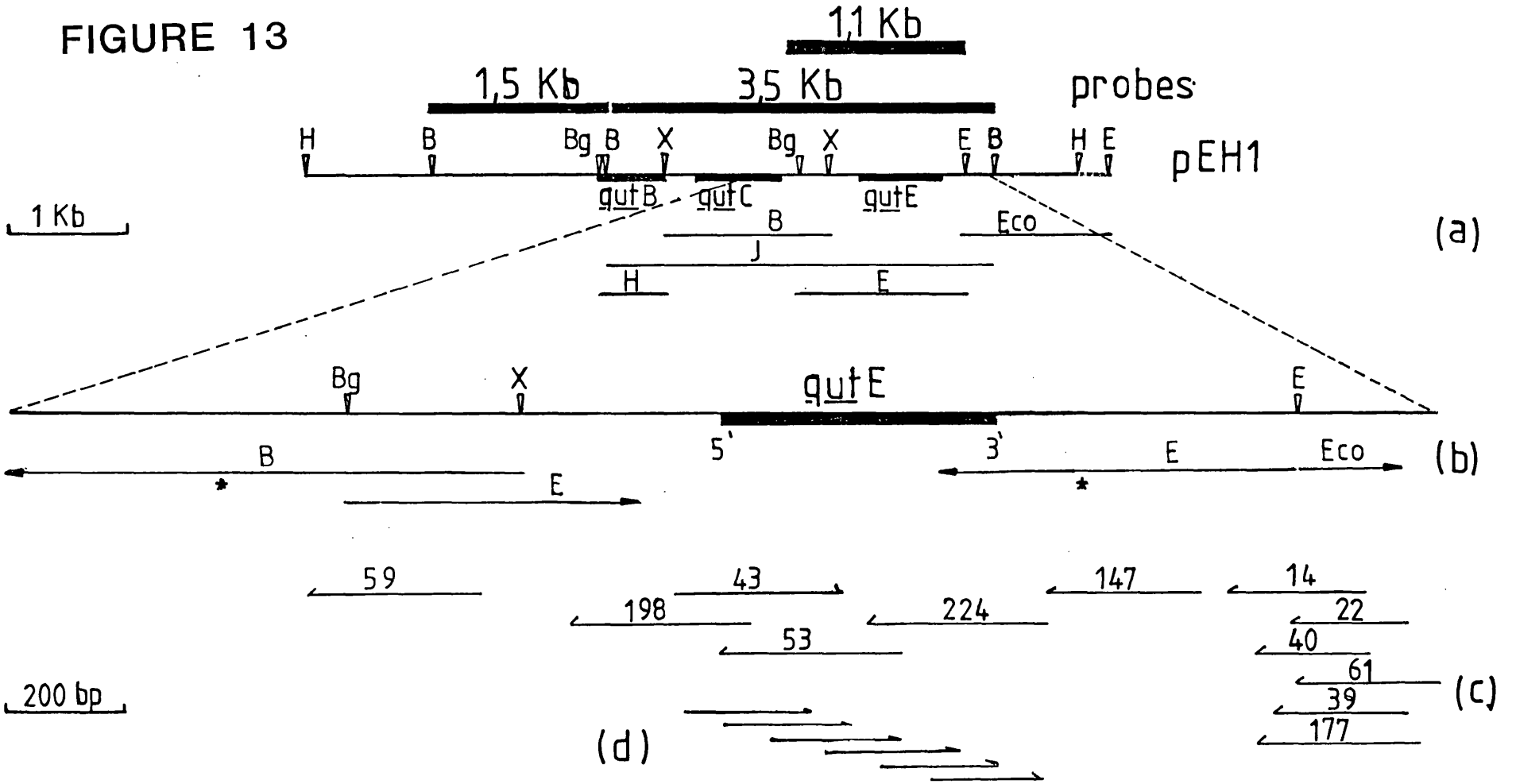

1,1 Kb
1,5 Kb
3,5 Kb
probes
H
B
Bg
B
X
Bg
X
E
B
H
E
pEH1
gutB
gutC
gutE
1 Kb
B
Eco
(a)
J
H
E
Bg
X
gutE
E
5'
3'
B
E
E
Eco
(b)
59
43
147
14
198
224
22
53
40
61
200 bp
39
(c)
177
(d)

FIGURE 14 DNA sequence of the *QUTE* gene of *A. nidulans* including the 5' and 3' non-coding, flanking sequences.

The amino acid residues indicated by underlining are identical with those in *N. crassa*. Nucleotides shown in the box (around position 520) indicate position where the additional 75 nucleotides of the *QA2* gene are to be found.

FIGURE 14

```
        10         20         30         40         50         60
aaaacggtttataccatagcgcttgagaaggcgcgacaaggaattgactggtcctaatgc

        70         80         90        100        110        120
ttctagaaggttcttagaagaaacagtacctggagatcaattgatgtctttcaattttga

       130        140        150        160        170        180
ccgccccatactgcaaccttcatcttatgccgttttgactgggggcaatctctgttgcagc

       190        200        210        220        230        240
cggcagagcgttctgccaccactttcccttatccgtttccccgcaatttatactccaaag

       250        260        270        280        290        300
tccagagacaaggactaaataacaatctacccccccaaaaccaatcggccgtaaaccttca

       310        320        330        340        350        360
ttgctactctgcacaATGGAAAAATCAATCCTCCTCATCAATGGCCCGAACCTCAACCTC
                 MetGluLysSerIleLeuLeuIleAsnGlyProAsnLeuAsnLeu

       370        380        390        400        410        420
TTGGGCACCCGCGAACCGCACATCTACGGCTCAACCACGCTCTCCGACGTCGAAGAATCC
LeuGlyThrArgGluProHisIleTyrGlySerThrThrLeuSerAspValGluGluSer

       430        440        450        460        470        480
TCCAAAGGTCACGCAGCTTCGCTCGGCGCCTCCCTGCAAACATTTCAGTCAAACCATGAA
SerLysGlyHisAlaAlaSerLeuGlyAlaSerLeuGlnThrPheGlnSerAsnHisGlu

       490        500        510        520        530        540
GGTGCCATTGTTGACCGAATCCATGCCGCTCGCGGGAACACCGACGCGATCATCATCAAC
GlyAlaIleValGluArgIleHisAlaAlaArgGlyAsnThrAspAlaIleIleIleAsn

       550        560        570        580        590        600
CCCGGTGCATATACACACACCTCTGTCGCTATTCGGGATGCCCTGTTAGGAGTTGAGATT
ProGlyAlaTyrThrHisThrSerValAlaIleArgAspAlaLeuLeuGlyValGluIle

       610        620        630        640        650        660
CCCTTCATCGAACTCCATGTGAGCAATGTCCATGCACGAGAGCCGTTCAGACATCACAGT
ProPheIleGluLeuHisValSerAsnValHisAlaArgGluProPheArgHisHisSer

       670        680        690        700        710        720
TACTTTAGCGACAAAGCGTCAGGCATCATTGTGGGACTGGGTGTGTATGGGTACAAGGTC
TyrPheSerAspLysAlaSerGlyIleIleValGlyLeuGlyValTyrGlyTyrLysVal

       730        740        750        760        770        780
GCAGTGGAGCATGTTGCGTTGAACTTCAAGCCTCTGGAGAAAAAGGCTGCTCTTTAAatt
AlaValGluHisValAlaLeuAsnPheLysProLeuGluLysLysAlaAlaLeu***

       790        800        810        820        830        840
cttcttggattatgatactcctgcgacgtatcatgtaatttatgtcaaacaaagtttgct

       850        860        870        880        890        900
ccatataaaatcatctggggttgagtgattctgtcgctatctgcgactcacgcggtagtg

       910        920        930        940        950        960
tattccccggcgacgttaatgtttgttgtgtcggaaggagaaatgaatcagtctggctgg

       970        980        990       1000       1010       1020
tgtaccgatacagctcgggattgcggcttccggaacgaatcccagattctgcatgaattc

      1030       1040       1050       1060       1070       1080
gggctggttgtccatcctagtgctagttagttttataacttgcgtgcagaagagaggttg

      1090       1100
ttacctatctgtcccatcaagc
```

FIGURE 15 **DNA sequence comparision between the coding sequences of the *QUTE* gene of *A. nidulans* and the *QA2* gene of *N. crassa*.**

Computer program was set to record a minimal match of 9 homologous nucleotides in a window of 12 (White *et al.*, 1984).

FIGURE 15

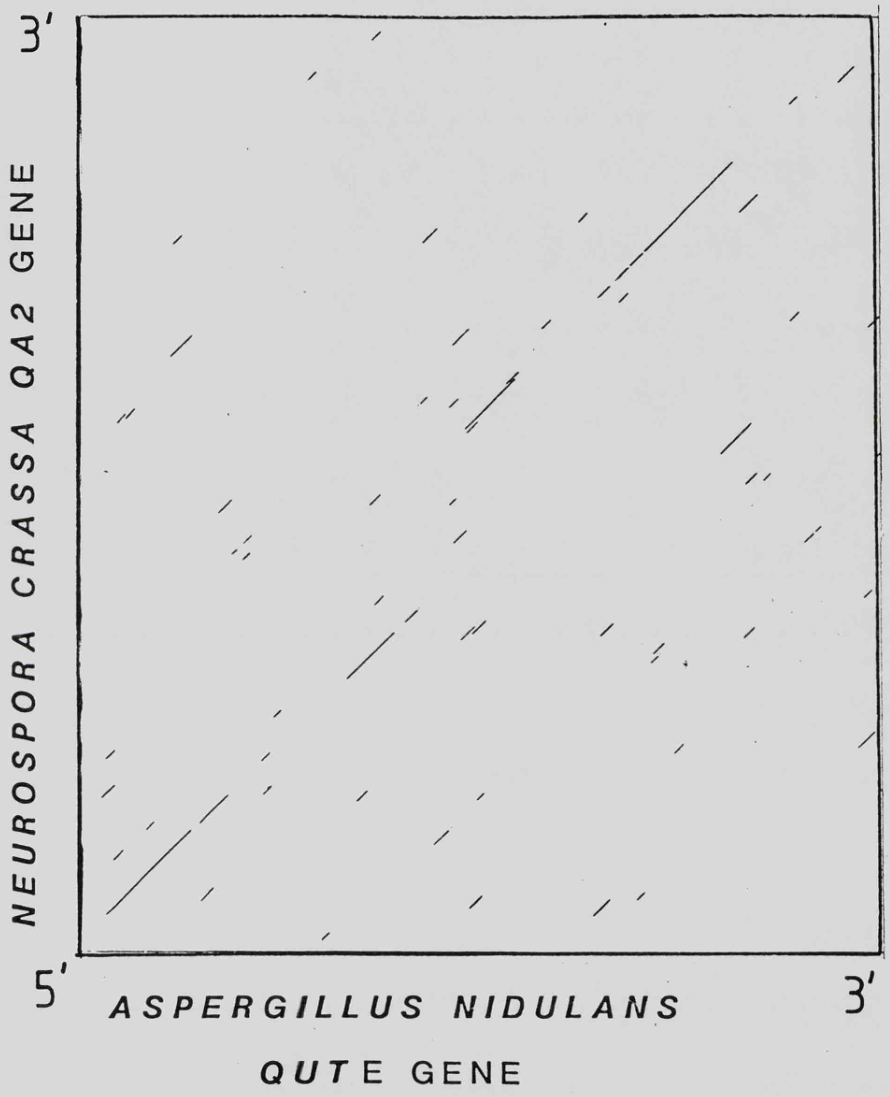

3'
NEUROSPORA CRASSA QA2 GENE
5'
3'
ASPERGILLUS NIDULANS
QUTE GENE

FIGURE 16 Alignment of the coding sequences of the *QUTE* gene of *A. nidulans* and the *QA2* gene of *N. crassa*.

The sequences are aligned in a way to maximize the homology. Stars (*) indicate identical bases and dashed lines (---), regions of the sequence non-existing from the two equivalent genes

FIGURE 16

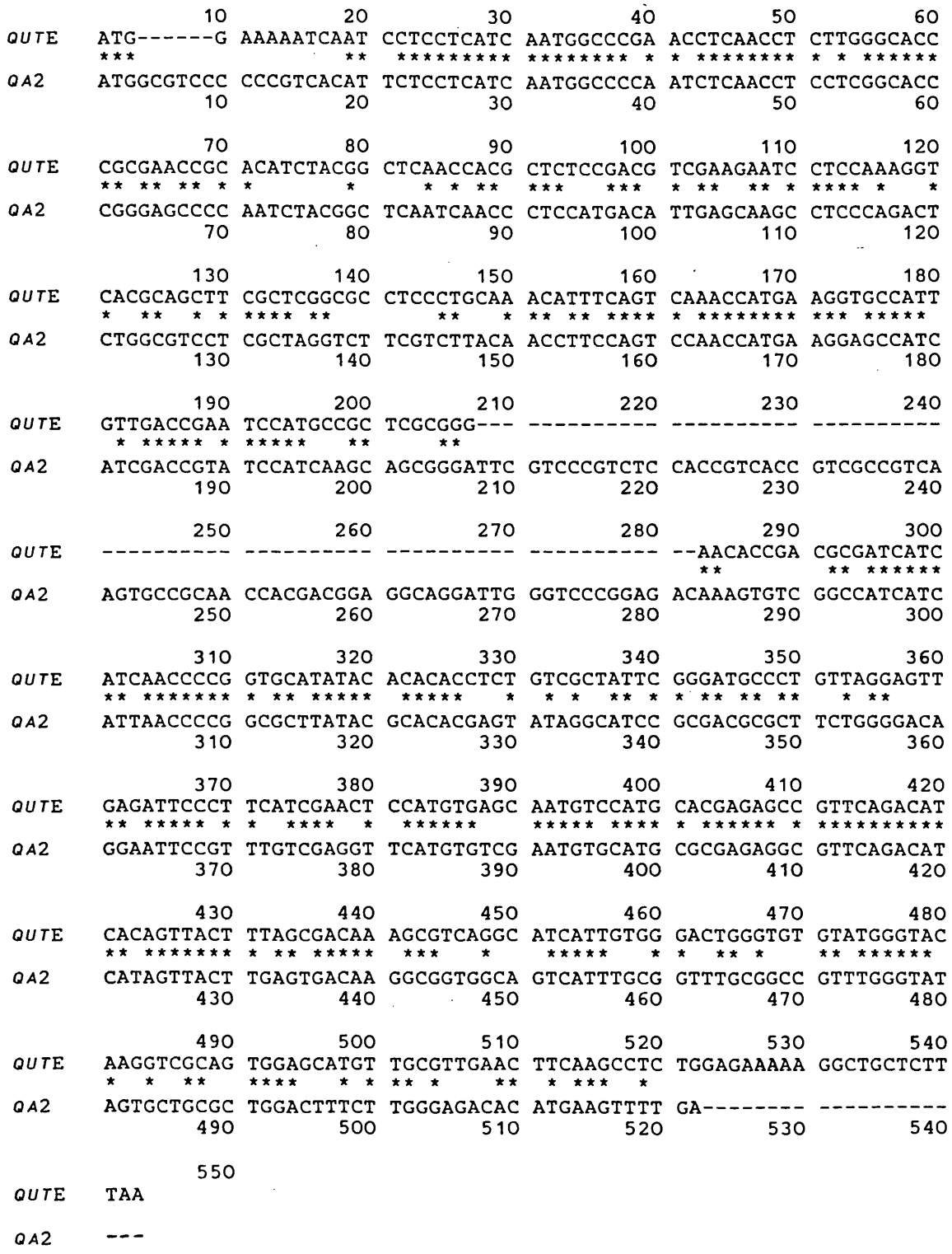
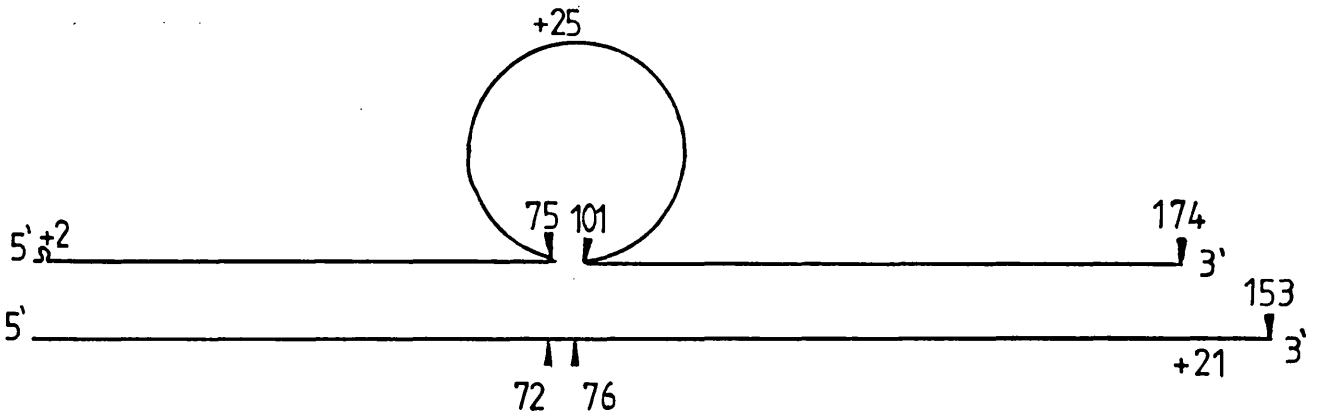

10 20 30 40 50 60
QUTE ATG------G AAAAATCAAT CCTCCTCATC AATGGCCCGA ACCTCAACCT CTTGGGCACC
QA2 ATGGCGTCCC CCCGTCACAT TCTCCTCATC AATGGCCCCA ATCTCAACCT CCTCGGCACC
10 20 30 40 50 60
70 80 90 100 110 120
QUTE CGCGAACCGC ACATCTACGG CTCAACCACG CTCTCCGACG TCGAAGAATC CTCCAAAGGT
QA2 CGGGAGCCCC AATCTACGGC TCAATCAACC CTCCATGACA TTGAGCAAGC CTCCCAGACT
70 80 90 100 110 120
130 140 150 160 170 180
QUTE CACGCAGCTT CGCTCGGCGC CTCCCTGCAA ACATTTCAGT CAAACCATGA AGGTGCCATT
QA2 CTGGCGTCCT CGCTAGGTCT TCGTCTTACA ACCTTCCAGT CCAACCATGA AGGAGCCATC
130 140 150 160 170 180
190 200 210 220 230 240
QUTE GTTGACCGAA TCCATGCCGC TCGCGGG--- ---------- ---------- ----------
QA2 ATCGACCGTA TCCATCAAGC AGCGGGATTC GTCCCGTCTC CACCGTCACC GTCGCCGTCA
190 200 210 220 230 240
250 260 270 280 290 300
QUTE ---------- ---------- ---------- ---------- --AACACCGA CGCGATCATC
QA2 AGTGCCGCAA CCACGACGGA GGCAGGATTG GGTCCCGGAG ACAAAGTGTC GGCCATCATC
250 260 270 280 290 300
310 320 330 340 350 360
QUTE ATCAACCCCG GTGCATATAC ACACACCTCT GTCGCTATTC GGGATGCCCT GTTAGGAGTT
QA2 ATTAACCCCG GCGCTTATAC GCACACGAGT ATAGGCATCC GCGACGCGCT TCTGGGGACA
310 320 330 340 350 360
370 380 390 400 410 420
QUTE GAGATTCCCT TCATCGAACT CCATGTGAGC AATGTCCATG CACGAGAGCC GTTCAGACAT
QA2 GGAATTCCGT TTGTCGAGGT TCATGTGTCG AATGTGCATG CGCGAGAGGC GTTCAGACAT
370 380 390 400 410 420
430 440 450 460 470 480
QUTE CACAGTTACT TTAGCGACAA AGCGTCAGGC ATCATTGTGG GACTGGGTGT GTATGGGTAC
QA2 CATAGTTACT TGAGTGACAA GGCGGTGGCA GTCATTTGCG GTTTGCGGCC GTTTGGGTAT
430 440 450 460 470 480
490 500 510 520 530 540
QUTE AAGGTCGCAG TGGAGCATGT TGCGTTGAAC TTCAAGCCTC TGGAGAAAAA GGCTGCTCTT
QA2 AGTGCTGCGC TGGACTTTCT TGGGAGACAC ATGAAGTTTT GA-------- ----------
490 500 510 520 530 540
550
QUTE TAA
QA2 ---

FIGURE 17 **Schematic representation of the amino acid sequence organization of the *QUTE* gene of *A. nidulans* and the *QA2* gene of *N. crassa*.**

Arrows indicate extremities of homologous sequence and numbers amino acid residues.

FIGURE 17



QA2

QUTE

FIGURE 18 Codon usage in the *QUTE* gene of *A. nidulans*.

FIGURE 18

Amino acid	Codon	Frequency	Amino acid	Codon	Frequency
PHE	UUU	2/5 (40%)	ALA	GCU	5/16 (31%)
	UUC	3/5 (60%)		GCC	4/16 (25%)
LEU	UUA	1/15 (7%)		GCA	4/16 (25%)
	UUG	2/15 (13%)		GCG	3/16 (19%)
	CUU	1/15 (7%)	HIS	CAU	6/10 (60%)
	CUC	7/15 (47%)		CAC	4/10 (40%)
	CUA	- (0%)			
	CUG	4/15 (27%)	GLN	CAA	1/2 (50%)
				CAG	1/2 (50%)
ILE	AUU	4/13 (31%)			
	AUC	9/13 (69%)	ASN	AAU	2/8 (25%)
	AUA	- (0%)		AAC	6/8 (75%)
MET	AUG	1/1 (100%)	LYS	AAA	4/7 (57%)
				AAG	3/7 (43%)
VAL	GUU	3/9 (33%)			
	GUC	2/9 (22%)	ASP	GAU	1/4 (25%)
	GUA	- (0%)		GAC	3/4 (75%)
	GUG	4/9 (45%)			
			GLU	GAA	6/11 (55%)
TYR	UAU	2/5 (40%)		GAG	5/11 (45%)
	UAC	3/5 (60%)			
			CYS	UGU	- (0%)
SER	UCU	1/13 (8%)		UGC	- (0%)
	UCC	4/13 (31%)			
	UCA	4/13 (31%)	TRP	UGG	- (0%)
	UCG	1/13 (8%)			
	AGU	1/13 (8%)	ARG	CGU	- (0%)
	AGC	2/13 (14%)		CGC	2/6 (33%)
				CGA	2/6 (33%)
PRO	CCU	1/6 (17%)		CGG	1/6 (17%)
	CCC	2/6 (33%)		AGA	1/6 (17%)
	CCA	- (0%)		AGG	- (0%)
	CCG	3/6 (50%)			
			GLY	GGU	4/13 (31%)
THR	ACU	- (0%)		GGC	5/13 (38%)
	ACC	4/7 (57%)		GGA	2/13 (15%)
	ACA	2/7 (29%)		GGG	2/13 (16%)
	ACG	1/7 (14%)			

FIGURE 19 Analysis of the flanking sequences surrounding the first methionine (ATG) in the coding sequences of a selection of genes from filamentous fungi.

FIGURE 19

A. nidulans	*QUTE*:	CACA ATG GAA
	ADH:	AAGG ATG TCT
	PGK:	AACA ATG TCT
	TRPC:	GATC ATG GCA
A. awamori	glucoamylase:	AGCA ATG TCG
N. crassa	*QA2*:	CACA ATG GCG
	QA3:	CACC ATG TCG
	QA4:	CGCC ATG CCG
	ATP/ADP:	CACA ATG GCG
	proteolipid:	AAAA ATG GCC
	AM:	CAAA ATG TCT
	histone H3:	CACA ATG GCC
	" H4:	CAAA ATG ACT
	TRP1:	CACA ATG TCG

CONSENSUS SEQUENCE : CACA ATG $\begin{matrix}T\\G\end{matrix}$CG

KOZAC SEQUENCE : NANN ATG GNN

REFERENCES:
PGK: John Clements, PhD Thesis, Leicester University (in preparation)
ADH: McKnight *et al.*, 1985
TRPC: Mullaney *et al.*, 1985
Glucoamylase: Boel *et al.*, 1984
QA2: Alton *et al.*, 1982
QA3: Altom *et al.*, 1982
QA4: Rutledge, 1984
ATP/ADP: Arends & Sebald, 1984
Proteolipid: Viebrock *et al.*, 1982
AM: Kinnaird & Fincham, 1983
Histone H3: Wouldt *et al.*, 1983
Histone H4: Wouldt *et al.*, 1983
TRP1: Schechtman & Yanofsky, 1983

Kozac sequence :Kozac, 1980

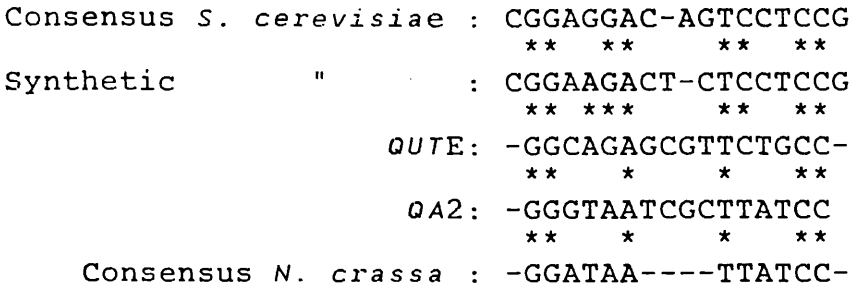
FIGURE 20 **Nucleotide conservation (*) of sequences showing a dyad of symmetry in the *GAL* system of *S. cerevisiae*, the *QUTE* gene of *A. nidulans* and the *QA* gene cluster of *N. crassa*.**

References:

QA : Giles *et al.*, 1985

GAL : Giniger *et al.*, 1985

FIGURE 20


Consensus S. cerevisiae : CGGAGGAC-AGTCCTCCG
** ** ** **
Synthetic " : CGGAAGACT-CTCCTCCG
** *** ** **
QUTE: -GGCAGAGCGTTCTGCC-
** * * **
QA2: -GGGTAATCGCTTATCC
** * * **
Consensus N. crassa : -GGATAA----TTATCC-

CHAPTER 8 . DISCUSSION

The work presented in this thesis has been concerned with the analysis of the regulation of expression of the *QUT* gene cluster, which mediates the first three steps in the catabolism of quinic acid in the filamentous fungus *Aspergillus nidulans*.

Analysis of quinic acid non-utilizing (*qut*) mutants has confirmed that all three enzyme structural genes are tightly linked. A mutant strain (*qut*E208) deficient for catabolic dehydroquinase has been isolated, and subsequent recombination analysis, established that the *QUT*E gene is also part of the cluster. Furthermore, the gene cluster has now been mapped about 12.5 map units, equidistant between the *ORN*B and *FWA* genes, confirming its location on chromosome VIII.

Analysis of three pleiotropic, non-inducible *qut* mutants, has shown on the basis of growth tests of heterozygous diploid strains and enzyme assays that at least two genes, called *QUT*A and *QUT*D, regulate expression of the structural genes. One mutant, *qut*D8 is fully recessive to wild-type and does not exhibit gene dosage effects upon enzyme induction when combined in a number of different heterozygous diploid strains containing the wild-type *QUT*D allele. It is therefore deduced that the product coded by the *QUT*D gene is not rate-limiting in enzyme induction and is abundantly synthesized. It is concluded that *qut*D8 is a loss-mutation affecting a function essential for enzyme induction, and that the *QUT*D

product is an activator protein necessary for expression of the *QUT* enzyme structural genes.

In contrast, the second class of non-inducible mutants, *qut*A identify a gene coding for a regulatory protein that has a more complex function in enzyme induction. One mutant allele, *qut*A4 is partially dominant to wild-type and a second mutant *qut*214, most probably also a *qut*A mutation, is almost fully dominant. The *qut*A4 mutation exhibits gene dosage effects observed as reduced enzyme induction in heterozygous diploid strains and is epistatic to the expression of the *QUT*D wild-type allele. One possibility is that the *qut*A4 mutation codes for an altered but functionally active protein insensitive to quinic acid and thus acting as an irreversible superrepressor. In this case, the *QUT*A gene can be interpreted to code for a negatively acting repressor protein, which in the absence of the inducer quinic acid binds to the 5' control region of the *QUT*D gene and possibly also to the promoter regions of the three enzyme structural genes, to prevent their expression. It is predicted that the *QUT*A gene is transcribed constitutively.

An alternative interpretation is that *qut*A4 and *qut*214 are deficient in a function that is also necessary for enzyme induction. In this case, the *QUT*A gene can be interpreted to code for a complex regulatory protein which upon binding quinic acid undergoes conformational change such that it becomes a co-activator necessary to the function of the *QUT*D activator protein. The interaction between the *QUT*A protein-quinic acid complex and the *QUT*D activator protein is assumed to occur on the promoter

regions of the *QUT* enzyme structural genes. Thus, the *qut*A4 and *qut*214 mutant phenotype, could be interpreted as identifying mutant alleles deficient in binding quinic acid or binding to the *QUT*D activator protein or both functions. Hence, the gene dosage effects observed with the *qut*A4 heterozygous diploids may be due to the *QUT*A gene product being rate-limiting for induction of the three structural gene enzymes. In this situation, the *QUT*D activator protein, which is not rate-limiting, provides the coarse control of enzyme synthesis, whereas the *QUT*A protein determines the fine control modulating expression of the *QUT* enzyme structural genes.

Several observations now seem to indicate that the *QUT*A gene may in fact function as a co-activator for the function of the *QUT*D protein. The *qa*1^S mutant alleles of *N. crassa* (thought to code for superrepressor proteins that are insensitive to quinic acid) readily yield revertants selected by growth on quinic acid many of which are constitutive for enzyme formation. These constitutive strains prove to be double mutants, containing a new mutation in addition to the original *qa*1^S mutation (Valone *et al.*, 1971; Partridge *et al.*, 1972). Genetic analysis, confirmed by DNA sequencing, of *qa*1^C constitutive mutants of *N. crassa*, have revealed that in contrast to the original *qa*1^S mutations, the new mutations conferring this phenotype map in a discrete region within the 5' coding region of the *QA*1^S gene thus inactivating the NH_2-terminal end, and presumed DNA binding domain of the *QA*1^S repressor protein (Giles *et al.*, 1985). These observations are consistent with the *QA*1^S gene coding for a repressor protein. In contrast, although both the *qut*A4 and *qut*214

strains yield revertants for growth on quinic acid, neither has generated a constitutive strain amongst ≈ 500 revertants tested (Dr S. Grant, personal communication). This observation suggests that the *QUT*A gene product is indispensable and thus required for the expression of the three *QUT* enzyme structural genes.

Hybridization of the two *N. crassa* control genes, *QA*1^S (using the 4.9 Kb *Eco*RI fragment which contains the whole gene) and *QA*1^F (using the 2.1 Kb *Bam*HI-*Xba*I fragment from within the 5' coding sequence), to *A. nidulans* genomic DNA identifies in both cases homology with several different regions of the *Aspergillus* genome. It is possible that these signals reflect sequences coding for DNA binding proteins. In the case of the *N. crassa* activator gene, *QA*1^F, the sequence conservation found with the dyads of symmetry in the *QUT*E gene in *Aspergillus* (Chapter 7), the *QA* genes in *Neurospora* (Giles *et al.*, 1985) and the *GAL* genes in *Saccharomyces* (Giniger *et al.*, 1985), supports this idea. In the case of the *qut* system of *A. nidulans*, the *QUT*D activator protein would thus require the *QUT*A co-activator to confer specificity for the promoter regions of the *QUT* structural genes. This could be the function of the *QUT*A protein which, either directs the *QUT*D activator protein to specific *QUT* enzyme structural gene promoters or stabilizes their interaction at these promoters

The different interpretations of the model are amenable to test by analysis of mRNA synthesis in haploid and heterozygous diploid strains. If the *QUT*A product functions as a repressor protein, the prediction is that the *QUT*A gene would be transcribed constitutively and the

*QUT*D gene inducible. In the case of the *QA* system of *Neurospora*, transcription of the $QA1^F$ (activator) gene is not detected in $qa1^S$ (repressor) mutants (Patel *et al.*, 1981). However, in the galactose utilization *(GAL)* system of *S. cerevisiae* regulation by the *GAL*80 repressor protein of the *GAL*4 activator protein is at the post-translational level (Matsumoto *et al.*, 1981; Laughon & Gestland, 1982). If on the other hand, the *QUT*A product operates solely as a co-activator, then one may expect the *QUT*D gene to be constitutively expressed in addition to the *QUT*A gene.

The discovery that in *Aspergillus* at least two genes regulate the expression of the *QUT* structural genes, falls within the general pattern of genetic regulation found in other fungal systems. In most of the systems studied, several genes appear to be required for the efficient control of expression, where a network of regulatory events is observed (Chapter 1).

Analysis of the 5' and 3' non-coding regions flanking the *QUT*E gene reveals the presence of the majority of the sequence motifs considered to play an important rôle in directing RNA polymerase II onto the DNA and in modulating the rate of mRNA synthesis in eukaryotes in general (Figure 21; and Chapter 7). Thus, one finds at approximately 150 bp upstream of the start of translation codon (ATG), a CAAT box (Benoist *et al.*, 1980), which is thought to be important for the binding of the RNA polymerase to the DNA. A sequence showing a near perfect dyad of symmetry is found 11 bp downstream of this CAAT box; this sequence exhibits considerable nucleotide

conservation with sequences postulated to be important for binding of certain activator proteins in fungi. A TATA box is found some 82 bp upstream of the first ATG codon; this sequence is suggested to be important in regulating the rate of RNA synthesis by RNA polymerase II in yeast and higher eukaryotes (McKnight & Kingsbury, 1982). Good homology of the nucleotides flanking the first methionine in the coding sequence is found with the proposed consensus sequence for eukaryotes (Kozac, 1981); this sequence is said to confer recognition by the ribosomes of which ATG is to initiate translation on mRNA. As for the 3' non-coding, flanking region, one finds the sequence TAGT<...>TTT, which is commonly found near the termination of transcription of yeast genes (Zater & Sherman, 1982) and 16 bp in the 3' direction the sequence AAATGAA followed by T/G rich blocks, showing near perfect agreement with the consensus sequence typically found near the sites of polyadenylation in eukaryotic genes (Proudfoot & Brownlee, 1976; Birnstiel *et al.*, 1985)). Hence, the *QUT* system of *A. nidulans* appears to be a good model for the analysis of the mechanisms operating in the control of transcription in eukaryotes, namely protein interaction on the DNA and the control signals in the promoter sequences.

Several goals may be identified for the future development of the project. It is necessary to clone the *QUT*A gene in addition to the *QUT*D gene. In the case of the *QUT*D gene, it is necessary to subclone the λQ1 DNA sequence which hybridizes with the $QA1^F$ gene of *N. crassa* (Figure 12) and test for transformation of *qut*D mutant

strains of *Aspergillus* to prove functional expression and verify the cloning of the entire gene. Identification of the *QUT*A gene requires the use of the terminal end fragments from the cloned DNA as probes to isolate sequences flanking the cloned genes, and verification of the *QUT*A gene will again require transformation of *qut*A mutants of *Aspergillus*, to test for the functional expression of the putative genes. As the *qut*A mutants are expected to be partially dominant over the wild-type in heterozygous diploids, efficient expression of the wild-type function in the transformed strains can be somewhat reduced. However, transformation experiments resulting in multiple insertions of the wild-type gene should over come this problem by over production of the *QUT*A protein.

The long term goal is the use of the proteins encoded by the *QUT*A and *QUT*D regulatory genes, for *in vitro* studies of how and where the proteins interact with one another and with the DNA. Such experiments should elucidate the importance of sequences such as the dyad of symmetry found in the 5' control region of the *QUT*E gene (Figure 21), in the regulation of the expression of the genes controlling the catabolism of quinic acid in *A. nidulans*.

FIGURE 21 Physical organization of the quinic acid utilization (*qut*) gene cluster of *A. nidulans*.

Broken blocks indicate the boundaries of genes which are not yet determined precisely.

Restriction enzymes used were as follows:

H : *Hind*III

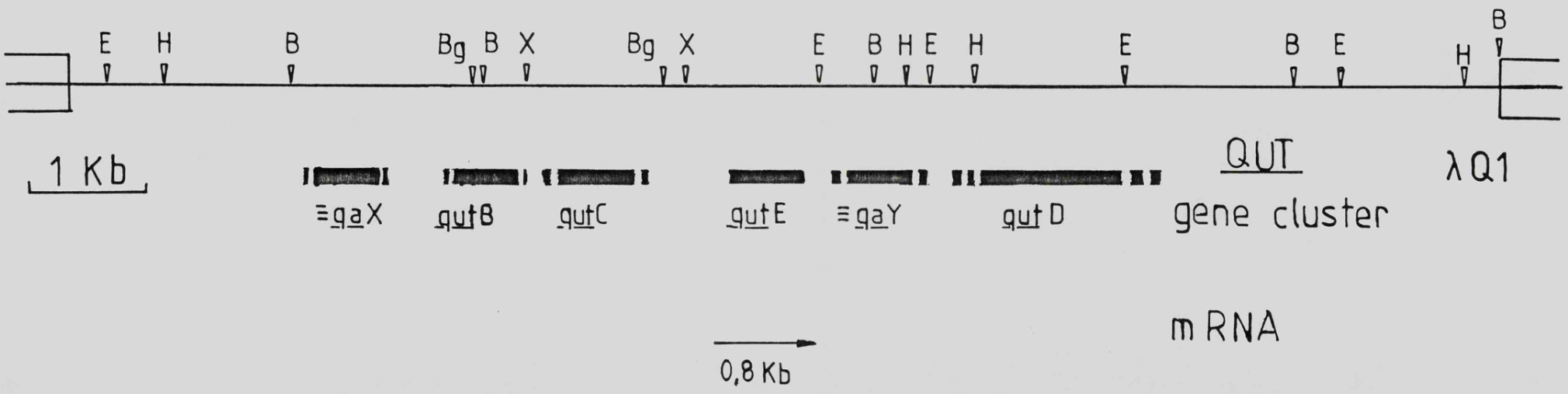
B : *Bam*HI

E : *Eco*RI

Bl : *Bgl*II

X : *Xba*I

FIGURE 21



5' control region of QUT E

GGCAATCT<11>GGCAGAGCGTTCTGCC<32>TATA<78>CACA 5' ATGGAA...
met glu

3' control region of QUTE

...CTT 3' TAA<118>TAGT<22>TTT<16>AAATGAA
Leu ***

REFERENCES

AAIJ, C. & BORST, P. (1972): "The gel electrophoresis of DNA". Biochem. Biophys. Acta, 269 : 192-200.

ALTON, N. K.; BUXTON, F.; PATEL, V.; GILES, N. H.; & VAPNEK, D. (1982): "5'-untranslated sequences of two structural genmes in the *qa* gene cluster of *Neurospora crassa*". Proc. Natl. Acad. Sci.USA, 79 : 1955-1959.

APIRION, D (1965): "The two-way selection of mutants and revertants in respect of acetate utilization and resistance to fluoroacetate in *Aspergillus nidulans*". Genetical Research, 6 : 317-329.

APPLEYARD, R. K. (1954): "Segregation of new lysogenic types during growth of a doubly lysogenic strain derived *from Escherichi coli* K12". Genetics, 39 : 440-452.

ARENDS, H. & SEBALD, W (1984): "Nucleotide sequence of the cloned mRNA and gene of the ADP/ATP carrier from *Neurospora crassa*". The EMBO Journal, 3 : 377-382.

ARMITT, S.; McCULLOUGH, W.; & ROBERTS, C. F.(1976): "Analysis of acetate non-utilizing (*acu*) mutants in *Aspergillus nidulans*". J. Gen. Microbiol., 92 : 263-282.

ARST, N. H. (1981): "Aspects of the control of gene expression in fungi". Sym. Soc. Gen. Microbiol., 31 : 131-160.

ARST, H. N. & 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". Mol. Gen. Genet. 163 : 17-22.

BALLANCE, D. J.; BUXTON, F. P. & TURNER, G. (1983): "Transformation of *Aspergillus nidulans* by the orotidine-5'-phosphate decarboxylase gene of *Neurospora crassa*". Biochem. Biophys. Res. Commun., 112 : 284-289.

BAILEY, C. & ARST, H. N. (1975): "Carbon catabolite repression in *Aspergillus nidulans*". Eur. J. Biochem., 51: 573-577.

BENNETZEN, J. L. & HALL, B. D. (1982): "Codon selection in yeast". J. Biol. Chem., 257 : 3026 : 3031.

BENOIST, C.; O'HARE, K.; BREATHNACH, R. & CHAMBON, P. (1980): "The ovalbumin gene-sequence of putative controlo regions". Nucleic Acids Res., 8 : 127-142.

BENTON, W. D. & DAVIS, R. W. (1977): "Screening λ recombinant clones by hybridization to single plaques in situ". Science, 196 : 180-182.

BERGEN, L. G.; UPSHALL, A. & MORRIS, N. R. (1984): "S-phase, G_2 and nuclear division mutants of *Aspergillus nidulans*". J. Bacteriol., 159 : 114-119.

BERLYN, M.B. & GILES, N. H. (1972): "Studies of aromatic biosynthetic and catabolic enzymes in *Ustilago maydis* and mutants in *Ustilago violacea*". Genet. Res., 19 : 261-270.

BERSE, B.; DMOCHOWSKA, A.; SKRZYPEK, M.; WEGLENSKI, P.; BATES, M. A. & WEISS, R. L. (1983): " Cloning and characterization of the ornathine carbamoyl-tranferase gene from *Aspergillus nidulans*". Gene, 25 : 109-117.

BIGELIS. R.; KEESEY, J. & FINK, G. R. (1977): "The *his*4 fungal gene cluster is not polycistronic". In Wilcox, G.; Abelson, J. & Fox, C. F. (eds). Molecular approaches to eucaryotic genetic systems. ICN-UCLA Symposia on Molecular and Cellular Biology. Academic Press, New York, 8 : 179-187

BIRBOIM, H. C. & DOLY, J. (1979): "A rapid alkaline extraction procedure for screening recombinant plasmid DNA". Nucleic Acids Res., 7 : 1513-1523.

BIRNSTIEL, M. L.; MEIRAD, B. & STRAB, K. (1985): "Transcription termination and 3' processing: the end is in sight!". Cell, 41 : 349-359

BOEL, E.; HANSEN, I.; HJORT, I.; HØEGH, I. & FIIL, N. P. (1984): "Two different types of intervening sequences in the glucoamylase gene from *Aspergillus niger*". The EMBO Journal, 7 : 1581-1585.

BOLIVAR, F.; RODRIGUEZ, R. L.; BETLACH, M. C. & BOYER, H. W. (1977a): "Construction and characterization of new cloning vehicles. I. Ampicillin-resistant derivatives of the plasmid pMB9". Gene, 2 : 75-93.

BOLIVAR, F.; RODRIGUEZ, R. L.; GREENE, P. J.; BETLACH, M. C.; HEYNEKER, H. L. & BOYER, H. W. (1977b): "Construction and characterization of new cloning vehicles. II. A multipurpose cloning system". Gene, 2 : 95-113

BRAMMAR, W. J. (1982) in Williamson, R. (ed). "Vectors based on bacteriphage lambda". Genetic engineering, vol. 3. Academic Press, London; 53-82.

BREATHNACH, R. & CHAMBON, P. (1981): "Organization and expression of eukaryotioc split genes coding for proteins". Ann. Rev. Biochem., 50 : 349-383.

CASE, M. E. & GILES, N. H. (1975): "Genetic evidence of the organization and action of the *qa*-1 gene product: a protein regulating the induction of three enzymes in quinatre metabolism in *Neurospora crassa*". Proc. Natl. Acad. Sci. USA, 72 : 553-557.

CASE, M. E. & GILES, N. H. (1976): "Gene order in the *qa* gene cluster of *Neurospora crassa*". Molec. Gen. Genetics, 147 : 83-89.

CHALLEF, R. S. (1974): "The inducible quinate-shikimate catabolic pathway in *Neurospora crassa*: genetic organization". J. Gen. Microbiol., 81 : 337-355.

CHAMPE, S.; KURTZ, M. B.; YAGER, L. N.; BUTRICK, N. J. & AXELROD, D. F. (1981): "Spore formation in *Aspergillus nidulans*: competence ans other developmental processes". In Turian, G. & Hohl, H. (eds.). The fungal spore: morphogenic controls. Acxademic Press, New York, 255.

CHARLES, I. G.; KEYTE, J. W.; BRAMMAR, W. J. & HAWKINS, A. R. (in press): "Nucleotide sequence encoding the biosynthetic dehydroquinase function of the penta-functional *AROM* locus of *Aspergillus nidulans*".

CIRIACY, M. (1979): "Isolation and characterization of further cis- and trans- acting regulatory elements involved in the synthesis of glucose-repressible alcohol dehydrogenase (ADHII) in *Saccharomyces cerevisiae*". Molec. Gen. Genetics.

CLARKE, L. & CARBON, J. (1978): "Selection of specific clones from colony banks by supression or complementation tests". Methods in Enzymol., 68 : 396-408.

CLEMENTS, J. M. & ROBERTS, C. F. (1985): "Molecular cloning of the 3-phosphoglycerate kinase (*PGK*) gene from *Aspergillus nidulans*". Current Genetics, 9 : 293-298.

CLUTTERBUCK,A. J. (1977): "The genetics of conidiation of *Aspergillus nidulans*". In Smith, J. E. & Pateman, J. A. (eds.): Genetics and physiology of *Aspergillus*. Academic Press, London, 305.

CLUTTERBUCK, A. J. & ROPER, J. A. (1966): "A direct determination of nuclear distribution in heterokaryons of *Aspergillus nidulans*. Genetical Res., 7 : 185-194.

CLUTTERBUCK, A. J. & COVE, D. J. (1976): "The genetic loci of *Aspergillus nidulans*". In Laskin, A. I. & Lechevalier, H. (eds.): CRC handbook of microbiology, vol. VI. Ohio, CRC Press, 665-676.

COVE, D. J. (1977): "The genetics of Aspergillus nidulans". In Smith, J. E. & Pateman, J. A. (eds.): Genetics and physiology of *Aspergillus*. Academic Press, London

DENINGER, P. L. (1983): "Random subcloning of sonicated DNA: application to shotgun DNA sequence analysis". Anal. Biochem., 129 : 216-223.

DENIS, C. L. & YOUNG, E. T. (1983): "Isolation and characterization of the positive regulatory gene *ARD*1 from *Saccharomyces cerevisiae*". Mol. Cell. Biol., 3 : 360-370.

deLEEUW, A. (1968): "Gene-enzyme relationships in aromatic biosynthesis in yeast". Phd thesis, Yale University.

DOUGLAS, H. C. & HAWTHORNE, D. C. (1964): "Enzymatic expression and genetic linkage of genes controlling galactose utilization in *Saccharomyces*". Genetics, 49 : 837-844.

DOUGLAS, H. C. & HAWTHORNE, D. C. (1966): "Regulation of genes controlling synthesis of the galactose pathway genes in yeast". Genetics, 54 : 911-916.

DRETZEN, G.; BELLARD, M.; SASSONE-CORSI, P. & CHAMBON, P. (1981): "A reliable method for the recovery of DNA fragments from agarose and acrylamide gels". Anal. Biochem., 112 : 295-298.

ENGLESBERG, E.; IRR, J.; POWER, J. & JEE, N. (1965): "Positive control of enzyme synthesis by gene C in the L-arabinose system". J. Bacteriol., 90 : 946-957

GILES, N. H (1978): "The organization, function and evolution of gene clusters in eukaryotes". The american naturalist, 112 : 641-658.

GILES, N. H.; CASE, M. E.; PARTRIDGE, C. W. & AHMED, S. I. (1967a): "A gene cluster in *Neurospora crassa* coding for an aggregate of five aromatic synthetic enzymes". Proc. Natl. Acad. Sci. USA, 58 : 1453-1460.

GILES, N. H.; PARTRIDGE, C. W. H.; AHMED, S. I. & CASE, M. E. (1967b): "The occurrence of two dehydroquinases in *Neurospora crassa*, one constitutive and one inducible". Proc. Natl. Acad. Sci. USA, 58 : 1930-1937.

GILES, N. H.; CASE, M. E.; BAUM, J.; GEEVER, R.;HUIET, L.; PATEL, V. & TYLER, B. (1985): "Genetic regulatory mechanisms in the quinic acid (*QA*) gene cluster of *Neurospora crassa*". UCLA symposium on molecular genetics of filamentous fungi.

GINIGER, E.; VARNUM, S. M. & PTASHNE, M. (1985): "Specific DNA binding of *GAL4*, a positive regulatory protein of yeast". Cell, 40 : 767-774.

HASHIMOTO, H.; KIKUCHI, Y.; NOGI, Y. & FUKASAWA, T. (1983): "Regulation of expression of the galactose gene cluster in *Saccharomyces cerevisiae*. Isolation and characterization of the regulatory gene *GAL4*". Mol. Gen. Genet., 191 : 31-38.

HASTIE, A. C. (1970): "Benelate-induced unstability of *Aspergillus*". Nature (London), 226 : 771.

HAWKINS, A. R.; GILES, N. H. & KINGHORN, J. R. (1982): "Genetical and biochemical aspects of quinate breakdown in the filamentous fungi *Aspergillus nidulans*". Biochem. Genet., 20 : 271-286.

HIUET, L. (1984): "Molecular analysis of the *Neurospora qa*-1 regulatory region indicates that two interacting genes control *qa* gene expression". Proc. Natl. Acad. Sci. USA, 81 : 1174-1178.

HOHN, B. (1979): "*In vitro* packaging of λ and cosmid DNA". Methods Enzymol., 68 : 299-309.

HOHN, B. & MURRAY, K. (1977): "Packaging recombinant DNA molecules into bacteriophage *in vitro*". Proc. Natl. Acad. Sci. USA, 74 : 3259-3263.

HOPPER, J. E. & ROWE, L. B. (1978): "Molecular expression and regulation of the galactose pathway genes in *Saccharomyces cerevisiae*". J. Biol. Chem., 253 : 7566-7569.

HYNES, M. J.; CORRICK, C. M. & KING, J. A. (1983): "Isolation of genomic clones containing the *and*S gene of *Aspergillus nidulans* and their their use in the analysis of structural and regulatory mutations". Molec. Cell. Biol., 8 : 1430-1439.

JACOB, F. & MONOD, J. (1961): "Genetic regulatory mechanisms in the synthesis of proteins". J. Mol. Biol., 3 : 318-356.

JEFFREYS, A. J.; WILSON, V.; WOOD, D.; SIMONS, J. P.; KAY, R. M. & WILLIAMS, J. G. (1980): "Linkage of adult α- and β- globin genes in *X. laevis* and gene duplication by tetraploidization. Cell, 21 : 555-564.

JONES, E. W. & FINK, G. R. (1983): "Regulation of aminoacid and nucleotide biosynthesis in yeast". Strathern, J. N.; Jones, E. W. & Broach, J. R. (eds.): The molecular biology of the yeast *Saccharomyces*, metabolism and gene expression. Cold Spring Harbor Laboratory, CSH, New York, 181-299.

JOHNSTON, S. A. & HOPPER, J. E. (1982): "Isolation of the yeast reguilatory gene *GAL4*, and its dosage effects on the structural genes of the galactose system". Proc. Natl. Acasd. Sci. USA, 79 : 6971-6975.

KAFER, E. (1965): "The origins of translocations in *Aspergillus nidulans*". Genetics, 52 : 217-232.

KARN, J.; BRENNER, S.; BARNET, L. & CESARINI, G. (1980): "Novel bacteriophage λ cloning vector". Proc. Natl. Acad. Sci. USA, 77 : 5172-5176.

KINGHORN, J. R. & PATEMAN, J. A. (1977): "Nitrogen metabolism". In Smith, J. E. & Pateman, J. A. (eds.): Genetics and physiology of *Aspergillus*. Academic Press, London, 147-202.

KINGHORN, J. R. & HAWKINS, A. R. (1982): "Cloning and expression in *Escherichia coli* K12 of the biosynthetic dehydroquinase function of the *arom* cluster gene from the eukaryote, *Aspergillus nidulans*". Mol. Gen. Genet., 186 : 145-152.

KINNAIRD, J. H. & FINCHAM, J. R. S. (1983): "The complete nucleotide sequence of the *Neurospora crassa am* (NADP-specific glutamate dehydrogenase) gene". Gene, 26 : 253-260.

KOZAC, M. (1981): "Possible rôle of flanking sequences in recognition of the AUG initiator codon by eukaryotic ribosomes". Nucleic Acids Res., 9 : 5233-5252.

KUSHNER, S. R. (1978): "An improved method for transformation of *Escherichia coli* with colE1 derived plasmids". In Boyer, H. W. & Nicosia, S. (eds.): Genetic engineering. North Holland, Biochemical Press, Amsterdam, 17-23.

KUSHNER, S. R.; HAUTALA, J. A.; JACOBSON, J. W.; GILES, N. H. & VAPNEK, D. (1977): "Expression of the structural gene for catabolic dehydroquinase of *Neurospora crassa* in *Escherichia coli* K12". Genetic Interaction and Gene Expression. Brookhaven symposia in biology, No29.

LAUGHON, A. & GESTLAND, R. F. (1982): "Isolation and preliminary characterization of the *GAL4* gene, a positive regulator of transcription in yeast". Proc. Natl. Acad. Sci. USA, 79 : 6827-6831.

LAUGHON, A. & GESTLAND, R. F. (1984): "Primary structure of the *Saccharomyces cerevisiae GAL4* gene". Mol. Cell. Biol., 4 : 260-267.

LOEN, W. A. M. & BRAMMAR, W. J. B. (1980): "A bacteriophage lambda vector for cloning large DNA fragments made with several restriction enzymes". Gene, 10 : 249-259.

LOWRY, O. H.; ROSEBROUGH, N. J.; FARR, A. L. & RANDAL, R. J. (1951): "Protein measurement with the folin phenol reagent". J. Biol. Chem., 193 265-275.

LUCCHINI G.; HINNEBUSCH, A. G.; CHEN, C. & FINK, G. R. (1984): "Positive regulatory interactiuons of the *HIS4* gene of *Saccharomyces cerevisiae*". Mol. Cell. Biol., 4 : 1326-1333.

LUMSDEN, J. & COGGINS, J. R. (1977): "The sub-unit structure of the *arom* multienzyme complex of *Neurospora crassa*". Biochem. J., 161: 599-607

MANIATIS, T.; FRITSCH, E. F. & SAMBROOK, J. (1982): "Molecular cloning, a laboratory manual". Cold Spring Harbor Laboratory.

MATHES, HY. W. D.; ZENKE, W. M.; GUNDSTROM, T.; STRAUSS, A.; WINLZERITH, M. & CHAMBON, P. (1984): "Simultaneous rapid chemical synthesis of over one hundred oligonucleotides on a microscale". The EMBO Journal, 3 : 801-805.

MATSUMOTO, K.; TOH-E, A. & OSHIMA, Y. (1978): "Genetic control of galactokinase synthesis in *Saccharomyces cerevisiae*: evidence for constitutive expression of the positive regulatory gene *GAL4*". J. Bacteriol., 134 : 446-457.

McCULLOUGH, W.; PAYTON, M. A. & ROBERTS, C. F. (1977): "Carbon metabolism in *Aspergillus nidulans*". In Smith, J. E. & Pateman, J. A. (eds.): Genetics and physiology of *Aspergillus nidulans*. University Press, London.

McKNIGHT, G. L.; KATO, H.; UPSHALL, A.; PARKER, M. D.; SAARI, G. & O'HARA, P. J. (1985): "Identification and molecular analysis of a third *Aspergillus nidulans* alcohol dehydrogenase gene". The EMBO J., 4 : 2099-1985.

MESSING, J.; CREA, R. & SEEBURG, P. H. (1981): "A system for shotgun DNA sequencing". Nucleic Acids Res., 9 : 309-321.

MESSING, J. VIEIRA, J. (1982): "A new pair of M13 vectors for selecting either DNA strand of double-digested restriction fragments".
Gene, 19 : 269-276.

METZENBERG, R. L. & CHIA, W. (1979): "Genetic control of phosphorus assimilation in *Neurospora crassa*. Dose dependent dominance and recessiveness in constitutive mutants". Genetics, 93 : 625-643.

MILLER, J. H. & REZNIKOFF (1978): "The operon". New York, Cold Spring Harbor laboratory.

MORRIS, N. (1976): "Mitotic mutants of *Aspergillus nidulans*". Genet. Res., 26 : 237-254.

MORRIS, N.; LAI, L. & OAKLEY, C. E. (1979): "Identification of a gene for F-Tubulin in *Aspergillus nidulans*". Cell, 16 : 437-442.

MORRIS, N.; KIRSCH, D. & OAKLEY, B. R. (1982): "Molecular and genetic methods for studying mitosis and spindle proteins in *Aspergillus niduilans*". Methods in Enzymol., 25 : 107-130.

McKNIGHT, S. L. & KINGSBURY, R. (1982): "Transcriptional control signals of a eukaryotic protein-coding gene". Science, 217 : 316-324.

MULLANEY, E. J.; HAMER, J. E.; ROBERTI, K. A.; YELTON, M. M. & TIMBERLAKE, W. E. (1985): "Primary structure of the *trp*C gene from *Aspergillus nidulans*". Mol. Gen. Genet., 199 : 37-45.

NOGI, Y. & FUKASAWA, T. (1984): "Nucleotide sequence of the yeast regulatory gene *GAL*80". Nucleic Acids Res., 12 : 9287-9298.

NOGI, Y.; SHIMADA, H.; MATZUZAKI, Y.; HASHIMOTO, H. & FUKASAWA, T. (1984): "Regulation of expression of the galactose gene cluster in *Saccharomyces cerevisiae*. II The isolation and dosage effect of the regulatory gene *GAL*80". Mol. Ge. Genet., 195 : 29-34.

OAKLEY, B. R. & MORRIS, N. R. (1980): "Nuclear movement is F-tubulin dependent in *Aspergillus nidulans*". Cell, 19 : 255-262.

ORSTON, L. N. (1971): "Regulation of catabolic pathways in *Pseudomonas*". Bacteriol. Rev., 35 : 87-116.

PALL, M. L. (1981): "Adenosine 3',5'-phosphatase in fungi". Microbiological Reviews, 45 : 462-480.

PARTRIDGE, C. W.; CASE, M. E. & GILES, N. H. (1972): "Direct induction in wild-type *Neurospora crassa* of mutants (*qa*-1^C) constitutive for the catabolism of quinate and shikimate". Genetics, 72 : 411-417.

PATEL, V. B.; SCHWEIZER, M.; DYKSTRA, C. C.; KUSHNER, S. R. & GILES, N. H. (1981): "Genetic organization and transcriptional regulation in the *qa* gene cluster of *Neurospora crassa*". Proc. Natl. Acad. Sci. USA, 9 : 5783-5787.

PENN, M. D.; THIREAS, G. & GREER, H. (1984): "Temporal analiysis of general control of amino acids biosynthesis in *Saccharomyces cerevisiae*: rôle of positive regulatory genes in initiation and maintenance of mRNA derepression". Mol. Cell. Biol., 4 : 520-528.

PATEMAN, J. A.; DOY, C. H.; OLSEN, J. E.; NORRIS, V.; CREASER, E. H. & HYNES, M. (1983): "Regulation of alcohol dehydrogenase (*ADH*) and aldehyde dehydrogenase (*aldDH*) in *Aspergillus nidulans*". Proc. R. Soc. London, B217 : 243-264.

PONTECORVO, G.; ROPER, J. A.; HEMMONS, J. M.; MacDONALD, K. D. & BUFTON, A. W. J. (1953): "The genetics of *Aspergillus nidulans*". Adv. Genet., 5 : 141-238.

PONTECORVO, G. & KAFER, E. (1958): "Genetic analysis based on mitotic recombination". Adv. Genet., 9 : 7.

POST-BEITTENMILLER, M. A.; HAMILTON, R. W. & HOPPER, J. E. (1984): "Regulation of basal and induced levels of the *MEL*1 transcript in *Saccharomyces cerevisiae*". Mol. Cell. Biol., 4 : 1238-1245.

PROUDFOOT, N. J. & BROWNLEE, G.G. (1976): "3' non-coding region sequence in eukaryotic messenger RNA". Nature (London), 263 : 211-214.

RENZ, T. & KURZ, C. (1984): "A colorimetric method for DNA hybridization". Nucleic Acids Res., 12 : 3435-3444.

ROBERTS, C. F. (1963): "The genetic analysis of carbohydrate utilization in *Aspergillus nidulans*". J. Gen. Microbiol., 31 : 45-48.

ROBERTS, C. F. (1969): "Silica gel stock mcultures of *Aspergillus nidulans*". *Aspergillus* News Letters, 10 : 29.

ROPER, J. A. (1952): "Production of heterozygous diploids in filamentous fungi". Experientia, Basel, 8 : 14-15.

RUTLEDGE, B. J. (1984): "Molecular characterization of the *qa*-4 gene of *Neurospora crassa*". Gene, 32 : 275-287.

SANGER, F.; NICKLEN, S. & GOULSON (1977): "DNA sequencing with chain-termination inhibitors". Proc. Natl. Acads. Sci. USA, 74 : 5463-5467.

SARGAN, D. R.; GREGORY, S. P. & BUTTERWORTH, P. H. W. (1982): "A possible novel interaction between the 3'-end of 18S ribosomal RNA and the 5'-leader sequence of many eukaryotic messenger RNAs". FEBS Lett., 147 : 133-136

SCALENGHE, F.; TURCO, E.; EDSTROM, J. E.; PIROTTA, V. & MELLI, M. (1981): "Microdissecytion and cloning of DNA from a specific region of *Drosophila melanogaster* polytene chromosomes". Chromosoma, 82 : 205-216.

SCHECHTMAN, M. G. & YANOFSKY, C. (1983): "Structure of the trifunctional *Trp*-1 gene from *Neurospora crassa* and its aberrant expression in *Escherichia coli*". J. Mol. Appl. Genet., 2 : 83-99.

SCHWEIZER, M.; CASE, M. E.; DYKSTRA, C. C.; GILES, N. H. & KUSHNER, S. R. (1981): "Idfentification and characterization of recombinant plasmids carrying the complete *qa* gene cluster from *Neurospora crassa* including the *qa*-1$^+$ regulatory gene". Proc. Natl. Acad. Sci. USA, 78: 5086-5090.

SHINE, J. & DALGRANO, L. (1975): "Determinant of cistron specificity in bacterial ribosomes". Nature, 254 : 34-36.

SENTENAC, A. & HALL, B. (1982): "Yeast nuclear polymerases and their rôle in transcription". In Strathern, J. N.; Jones, E. & Broach, J. R. (eds.): The molecular biology of the yeast *Saccharomyces*: metabolism and gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 561-606.

SMITH, J. E. & PATEMAN, J. A. (1977) (eds): "Genetics and physiology of *Aspergillus*". Academic Press, London.

STADEM (1980): "A new computer method for the storage and manipulation of DNA gel reading data". Nucleic Acid Res., 8 : 3673-3694.

StJOHN, T. P. & McDONNEL, R. W. (1979): "Isolation of galactose-inducible DNA sequences from *Saccharomyces cerevisiae* by differential plaque filter hybridyzation". Cell, 16 : 443-452.

StJOHN, T. P. & DAVIS, R. W. (1981): "The organization and transcription of the galactose gene cluster of *Saccharomyces*". J. Mol. Biol., 152 : 285-315.

TIMBERLAKE, W. E.; GWYNE, D. I.; HAMER, J. E.; MILLER, B. L.; MILLER, K. Y.; MULLANEY, E. M.; deSOUZA, E. R.; YELTO, M. M. & ZIMMERMAN, C. R. (1983): "Gene regulation during conidiation in *Aspergillus nidulans*". In Goldberg, R. B. (ed.): Plant molecular biology. Alan R, Liss, New York, 179

THOMAS, P. S. (1980): "Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose". Proc. Natl. Acad. Sci. USA, 77 5201-5205.

TOH-E, A.; INOUYE, S. & OSHIMA, Y. (1981): "Structure and function of the *PHO*82-*pho*4 locus controlling the synthesis of repressible acid phosphatase of *Saccharomyces cerevisiae*". J. Bacteriol., 145 : 221-232.

VALONE, J. A.; CASE, M. E. & GILES, N. H. (1971): "Constitutive mutants in a regulatory gene exerting positive control of quinic acid catabolism in *Neurospora crassa*". Proc. Natl. Acad. Sci. USA, 68 : 1555-1559.

VAPNEK, D.; HAUTALA, J. A.; JACOBSON, J. W.; GILES, N. H. & KUSHNER, S. R. (1977): "Exprtession in *Escherichia coli* K12 of the structural gene for catabolic dehydroquinase of *Neurospora crassa*". Proc. Natl. Acad. Sci. USA, 74 : 3508-3512.

VIEBROCK, A.; RERZ, A. & SEBALD (1982): "The imported preprotein of the proteolipid subunit of the mitochondrial ATP synthase from *Neurospora crassa*. Molecular cloning and sequencing of the mRNA". The EMBO Journal, 1 : 565-571.

WAHL, G. M.; STERN, M. & STARK, G. R. (1979): "Efficient transfer of large DNA fragments from agarose gels to diazobenzloxymethal-paper and rapid hybridization by using dextran sulfate". Proc. Natl. Acad. Sci. USA, 76 : 3683-3687.

WHITE, C. T.; HARDIES, S. C.; HUTCHINSON III; C. A. & EDGELL; M. H. (1984): "The diagonal-transverse homology search algorithm for locating similarities between two sequences". Nucleic Acid Res., 12 : 751-766.

WOODWARD, V. W.; deZEEUW, J. R. & SRB, A. M. (1954): "The separation and isolation of particular biochemical mutants of *Neurospora* by differential germination of conidia followed by filtration and selective plating". Proc. Natl. Acad. Sci. USA, 40 : 192-200.

WOUDT, L. P.; PASTINK, A.; VEMPERS-VEENSTRA, A. E.; JANSEN, A. E. M.; MAGER, W. H. & PLANTA, R. J. (1983): "The genes coding for histone H3 and H4 in *Neurospora crassa* are unique and contain intervening sequences". Nucleic Acids Res., 11 : 5347-5360.

YAMAMOTO, K. R.; ALBERTS, B. M.; BENZINGER, R.; LAWHORNE, L. & TREIBER, G. (1970): "Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification". Virology, 40 : 734-744.

ZARET, K. S. & SHERMAN, F. (1982): "DNA sequences required for efficient transcription termination in yeast". Cell, 28 : 563-573.

GENETIC AND MOLECULAR ANALYSIS OF THE QUINIC ACID UTILIZATION *(QUT)* GENE CLUSTER IN *ASPERGILLUS NIDULANS*

by

ANTONIO JORGE FRANCISCO DA SILVA

ABSTRACT

The first three steps in the utilization of quinic acid in *Aspergillus nidulans* are catalyzed by highly inducible enzymes encoded by a gene cluster containing three enzyme structural genes and two trans-acting regulatory genes. The gene cluster has been mapped meiotically between the *ORNB* and *FWA* genes in linkage group VIII.

The regulatory gene *QUTD* codes for an activator protein, whose function seems to depend upon interaction with the product of the *QUTA* gene. The *QUTA* protein thus has a complex rôle, acting both in repression of gene expression in the absence of the inducer quinic acid and in mediating the rôle of the *QUTD* activator protein, upon induction.

To test models for the regulation of expression of the *QUT* enzyme structural genes, the gene cluster has been physically isolated from an *A. nidulans* genomic DNA library in an λ phage replacement vector, using DNA fragments from the equivalent genes in *Neurospora crassa* as heterologous hybridization probes. An unique 13.2 Kb DNA sequence from *A. nidulans* was thus isolated and the three structural genes mapped within a 3.5 Kb *Bam*HI fragment. An adjacent region of the cloned DNA shows hybridization to the *N. crassa* activator gene, $QA1^F$, thus possibly identifying the *QUTD* gene. No hybridization could be detected with the *N. crassa* repressor gene, $QA1^S$, which was initially suggested to be equivalent to the *QUTA* gene of *A. nidulans*. Evidence against the hypothesis that the *QUTA* gene encodes a protein with only properties of a repressor is discussed.

The *QUTE* gene which codes for dehydroquinase has been sequenced. Major rearrangements in the coding sequence are found in comparison with the equivalent gene of *Neurospora*. Analysis of the promoter region of *QUTE* reveals that it contains a CAAT sequence exhibiting perfect homology with the consensus sequence for eukaryotes, and in close proximity 5' to a sequence showing a dyad of symmetry. The dyad exhibits homology to sequences in other fungal genes identified as binding activator proteins. The importance of these two promoter elements in the control of the expression of the *QUTE* gene is discussed.