FUNCTIONAL ANALYSIS OF THE PROMOTER OF THE 3-PHOSPHOGLYCERATE KINASE GENE OF *ASPERGILLUS NIDULANS*.

By Stephen John Streatfield

BSc (Hons.) Biochemistry,
University of York.

Thesis submitted for the degree of Ph.D. in the Department of Genetics, University of Leicester.

May, 1990.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

To my Mother and Father.

ABSTRACT

FUNCTIONAL ANALYSIS OF THE PROMOTER OF THE 3-PHOSPHOGLYCERATE GENE OF *ASPERGILLUS NIDULANS*.

By Stephen John Streatfield

Sequence elements contributing to high level expression of the 3-phosphoglycerate kinase (*pgk*) gene of *Aspergillus nidulans* have been investigated by the construction of gene fusion vectors using *lacZ* (β-galactosidase) of *Escherichia coli* as an assay system, and including the catabolic dehydroquinase (*qutE*) gene as a selective marker for the transformation of *A.nidulans*. The importance of targeting these constructs to a specific gene locus has been demonstrated by the analysis of a large number of transformed strains which reveals that the expression of the *lacZ* reporter gene is dependent upon the site of integration of the vector into the genome, and that when targeted to the *qutE* locus, is directly proportional to its copy number.

The analysis of transformed strains in which single copies of *lacZ* fusions have been targeted to the *qutE* locus has identified three constitutive positively acting sequence elements in the *pgk* gene. Firstly, sequence located between -161 and -120bp (possibly spanning -120) relative to the transcript start site is essential for expression, thus demonstrating that the putative core promoter, including potential TATA and CCAAT boxes, and a pyrimidine rich region, is not alone sufficient to enable expression. A comparison of this sequence to known promoter elements has revealed the presence of an octamer AAGCAAAT (-131 to -124), with a seven out of eight base pair match to the consensus octamer sequence ATGCAAAT characterized as being essential for the expression of several higher eukaryotic genes.

Whilst the octamer containing sequence contributes almost a residual 40% of the maximum expression recorded, a second region encompassing codons 14 to 183 and including the two introns of *pgk* has been shown to account for over 30% of the total activity. However, the full effect of the internal sequence may be greater, since the large β-galactosidase fusion proteins studied here have been shown to be unstable.

Thirdly, an extensive region extending from 0.5 to over 3Kb upstream of the start site has been shown to be required for full expression, accounting for almost 30% of the total recorded activity.

Furthermore, sequence located between -638 and -488bp, and including an eight bp consensus element TGAGGTGT common to the four cloned *A.nidulans* glycolytic genes, has been shown to modulate expression: increasing activity about 1.5-fold on gluconeogenic compared to glycolytic carbon sources.

In a separate exercise, the native *pgk* gene has been replaced with disrupted sequence to generate a pgk^- mutant strain of *A.nidulans*. This replacement has been achieved in a diploid strain, from which both pgk^+ and pgk^- segregants have been isolated, and the analysis of these has allowed the *pgk* locus to be assigned to chromosome VIII. The pgk^- mutant strain that has been isolated is deficient in 3-phosphoglycerate kinase activity, requires both acetate and glycerol for growth, conidiates poorly, and is poisoned by even moderate concentrations of hexoses.

Acknowledgements

I am grateful to all members of the Genetics Department at the University of Leicester, and in particular to those in the Laboratory of Dr. Clive Roberts, for help and advice throughout the course of my work. I would especially like to thank Dr. Raj Beri for guidence on several aspects of molecular biology and on the transformation of Aspergillus, Isobel Hopkins for technical suggestions on RNA work, and Dr. Clive Roberts and Dr. Susan Grant for advice on classical genetics methods.

The gene disruption presented in Chapter 3 was done in collaboration with an undergraduate project student, Rachel Mead. Also, the investigation of the expression of the *pgk-lacZ* fusion on different carbon sources described in Chapter 6, and the analysis of gene fusions including extensive *pgk* coding sequence described in Chapter 7, was in collaboration with Sian Toews.

The pAN923 vector series was provided by Dr. Cees van den Hondel of Wageningen Agricultural University, The Netherlands, and the synthetic oligonucleotide used in sequencing was made by John Keyte in the Department of Biochemistry, Leicester.

I am immensely grateful to Dr. Clive Roberts for supervising my work, in particular for his enthusiasm in discussing experimental data and for his suggestions on the writing of this Thesis. I also wish to thank Vidya Rajan for helping to type the manuscript and to prepare certain figures.

The work presented in this Thesis was supported by an MRC research studentship.

There are many people outside of the laboratory who have made my stay in Leicester very enjoyable, and I would particularly like to acknowledge Rob (and his bicycle), Fanis, Justin and Vidya, and training partners in the University and City Athletic Clubs. Finally, many thanks to my parents for their continual love and support, and to Vidya for inspiration in the darkest days of the write up.

Abbreviations

ATP	adenosine triphosphate
bp	base pair
BSA	bovine serum albumin
Ci	Curie
cpm	counts per minute
DNA	deoxyribonucleic acid
dNTP	2'-deoxy(N) 5'-triphosphate N= adenosine(A), cytidine(C), guanosine(G), thymidine(T)
EDTA	ethylenediamine tetra-acetic acid
EGTA	ethyleneglycol-bis(β-aminoethyl ether) tetra-acetic acid
IPTG	isopropyl-β-D-galactopyranoside
kb	kilo base
kD	kilo Daltons
keV	kilo electron Volt
kV	kilo Volt
mRNA	messenger ribonucleic acid
mw	molecular weight
nt	nucleotide
ONPG	o-nitrophenyl-β-D-galactopyranoside
paba	para amino benzoic acid
PEG	polyethylene glycol
pyro	pyrodoxin
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
SDS	sodium dodecyl sulphate
snRNA	small nuclear ribonucleic acid
SSC	saline sodium citrate
TEMED	N,N,N',N'-tetra-methylethylenediamine
tRNA	transfer ribonucleic acid
v/v	volume for volume
w/v	weight for volume
X-gal	5-bromo-4-chloro-3-indoyl β-D-galactopyranoside
μF	micro Farad

Contents

Chapter 4.

Construction of fusions of the *Aspergillus nidulans* 3-phosphoglycerate kinase gene promoter (5'*pgk*) to a "reporter" gene.

Chapter 5.

Analysis of transformed strains of *Aspergillus nidulans*.

Chapter 1.

Introduction.

1.1 Elements in the control of transcription in eukaryotic cells.

The regulation of gene expression has the potential to occur at several different levels in eukaryotic cells, from transcriptional initiation and termination through post-transcriptional modification of the message to its translation. The major control of eukaryotic gene expression is generally thought to be at the level of transcription. Certainly, since this represents the initial stage in the expression of a gene, tight control of transcription prevents any unnecessary synthesis of transcripts and is therefore of importance to the economy of the cell.

Prior to the initiation of transcription chromatin structure must be modified so that transcriptional control sequences are accessible to regulatory proteins. This process is little understood, but may involve specific DNA sequences binding to the nucleoskeleton so altering DNA conformation to reveal promoter sequences (Gasser and Laemmli, 1986; Cockerill and Garrard, 1986). However, the major thrust of investigations on transcriptional control have centred on the association of RNA polymerase with specific promoter sequences. Such interactions are mediated by the binding of many different transcription factors.

RNA polymerase is the central apparatus of the transcription process. In prokaryotes the single RNA polymerase consists of a core enzyme (M.W. of 395kD) of four polypeptides ($\alpha\alpha\beta\beta'$) and one of a selection of sigma factors (M.W. of 44 to 92kD) which controls the binding of the complex to specific promoters on the DNA, but is released once RNA synthesis has begun (Chamberlin, 1976).

A number of conserved sequences have been identified

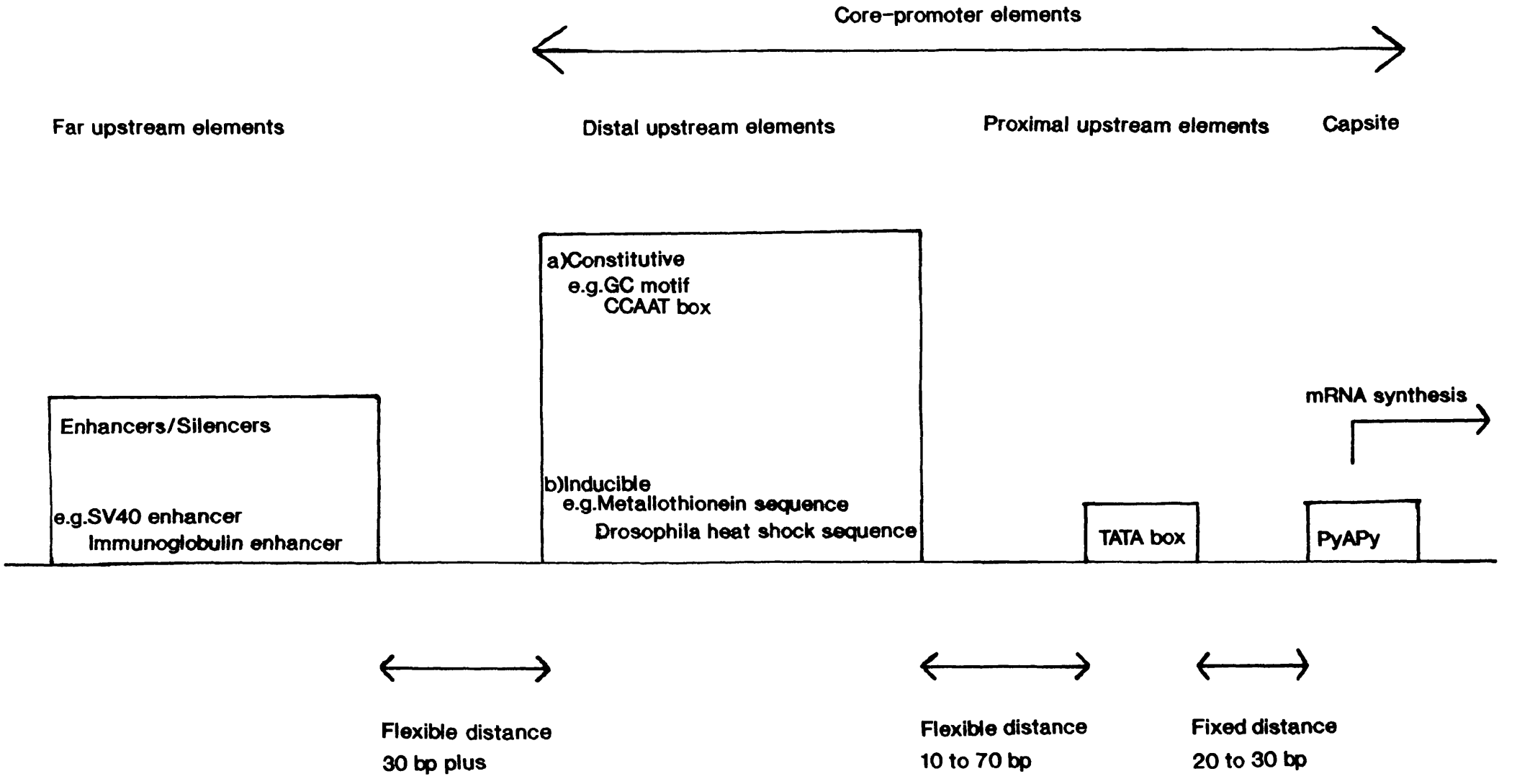
which are required upstream of the transcription initiation site to enable RNA polymerase to bind correctly. The Pribnow box, located at position -10 relative to transcription start, and the -35 box are the best characterized constitutive elements. In addition, positively acting elements located next to the constitutive sequences can stimulate transcription through binding activator proteins that can then interact with RNA polymerase to stabilize its binding. Conversely, negative acting elements may overlap the constitutive sequences, and the binding of repressor proteins to these sequence elements then inhibits RNA polymerase. Prokaryotic transcriptional control is reviewed by Ptashne (1986a).

In eukaryotes there are three types of RNA polymerase, each with a molecular weight in excess of 500kD and consisting of several subunits. RNA polymerase I is responsible for the synthesis of 18S, 28S and 5.8S rRNAs, accounting for between 50 and 70% of the total cellular RNA (Sollner-Webb and Tower, 1986). Whereas RNA polymerase I is located in the nucleolus, both RNA polymerase II and III are located in the nucleoplasm. RNA polymerase III transcribes 5SrRNA, tRNA and some snRNA genes, to give approximately 10% of the total cellular RNA (Ciliberto *et al.*, 1983). The remaining 20 to 40% of cellular RNA is transcribed from protein encoding genes by RNA polymerase II (La Thangue and Rigby, 1988).

A survey of the promoters of many RNA polymerase II transcribed genes demonstrates the presence of four distinct classes of elements (Figure 1.1), which I will now briefly review.

There is no extensive sequence conservation at transcription start sites. However, there is a bias towards A for the initiating base, with pyrimidine residues either side of this giving the capsite (reviewed by Chambon *et al.*, 1983). Although the rate of message synthesis usually decreases when this site is deleted, transcription is still observed from a new position, indicating a degree of

Figure 1.1
Conserved features of RNA polymerase II promoters.


Core-promoter elements
Far upstream elements
Distal upstream elements
Proximal upstream elements
Capsite
a)Constitutive
e.g.GC motif
CCAAT box
b)Inducible
e.g.Metallothionein sequence
Drosophila heat shock sequence
Enhancers/Silencers
e.g.SV40 enhancer
Immunoglobulin enhancer
TATA box
PyAPy
mRNA synthesis
Flexible distance
30 bp plus
Flexible distance
10 to 70 bp
Fixed distance
20 to 30 bp

flexibility in start point selection.

A more highly conserved element is the TATA or Hogness-Goldberg box reveiwed by Chambon *et al.*, (1983). It is a 7bp sequence positioned approximately 20 to 30bp upstream of the transcriptional start and tends to be surrounded by GC rich sequences. The distance between the TATA box and the start site is critical; deletions between the two having the effect of shifting the initiation of transcription an equivalent number of bases downstream. Furthermore, analysis of the TATA box by site directed mutagenesis has revealed that alterations within it not only lead to reduced message levels, but also to heterogeneity of start sites. Thus, the principal role of this element appears to be in positioning the point of transcription initiation. The transcriptional factor TFIID recognizes the TATA box, and analysis of the gene encoding it reveals homology to a region of sigma factor implicated in the recognition of the TATA like Pribnow box in prokaryotic promoters (Horikoshi *et al.*, 1989).

Many RNA polymerase II promoters also have conserved sequences located between 40 and 90bp 5′ of the transcription start site. These distal upstream elements can be categorized into two types, depending upon whether they are present within the promoters of constitutive or inducible genes. The first are required to maintain a basal level of expression and are common to several different genes. Examples include the GC motif present in the early promoter of SV40 (Hansen and Sharp, 1983), which has been shown to bind the transcription factor Sp1 (Dynan and Tjian, 1983), and the CCAAT box located upstream of the rabbit β-globin gene (Dierks *et al.*, 1983), which has been shown to bind different factors when located upstream of different genes (Jones *et al.*, 1985; Graves *et al.*, 1986). Deletion of these sequences significantly reduces the level of transcription. The second class of distal upstream elements are exemplified by sequences in the promoters of the metallothionein (Karin *et al.*, 1984) and Drosophila heat-shock (Pelham, 1982) genes. Deletion of these

sequences has demonstrated that they mediate the effect of physiological conditions on transcription. Distal upstream elements located 5′ to both constitutive and inducible genes are often present in multiple copies and in combination with one another (Lee *et al.*, 1987).

All of the elements discussed above are considered to constitute the core promoter. In addition to these, enhancers (and silencers) are often located even further upstream of the transcription initiation site. They are reviewed by Hatzopoulos *et al.*, (1988). Enhancer sequence elements show a wide variation in size and in position 5′ to the transcription start site. They have also been characterized at the 3′ end of genes beyond the polyadenylation site (Choi and Engel, 1986; Lusky *et al.*, 1983), within introns (Gillies *et al.*, 1983; Picard and Schaffner, 1984), and even within a coding region (Shout *et al.*, 1985). Experiments altering the position and orientation of enhancers have demonstrated that they can exert their effects up to 10Kb 5′ to the core promoter, and are also active when placed 3′ of a gene. They can act in either orientation, and even in association with heterologous genes. The analysis of heterologous gene fusions has shown that hormonal activation of transcription, and thus tissue specific expression, often depends upon specific enhancers (Voss *et al.*, 1986).

Several models have been proposed to explain the mode of action of enhancers. The location of nuclear scaffold attachment regions close to enhancers of three Drosophila genes (Gasser and Laemmli, 1986) and the κ-immunoglobulin gene (Cockerill and Garrard, 1986) has implicated enhancers in directing genes selected for expression to subcellular compartments enriched in the transcriptional machinery. Alternatively, the mapping of DNaseI hypersensitive sites to a number of elements including the μ-immunoglobulin enhancer (Mills *et al.*, 1983) indicates that enhancers may induce alterations in chromatin structure, so facilitating the association of RNA polymerase II with the gene. A related model proposes that enhancers induce conformational

changes in the DNA which may then stabilize transcriptional complexes. This view is supported by the mapping of topoisomerase II sites in the SV40 enhancer (Yang *et al.*, 1985). However, experiments in which the enhancer was topologically separated from the core promoter on a plasmid and yet still stimulated transcription, provide evidence against this hypothesis (Plon and Wang, 1986).

Each of the models discussed above is specific to particular enhancer sequences. However, a more general feature of enhancers, and indeed also of proximal and distal upstream elements, is their affinity for protein factors essential for efficient transcription. Evidence supporting this view is provided by competition assays with enhancer sequences (Mercola *et al.*, 1985), by the effect of mutations within enhancer sequence motifs on factor binding (Greenberg *et al.*, 1987) and, most directly, by the correlation of enhancer activity with the binding of specific factors (Seguin and Hamer, 1987; Bohman *et al.*, 1987). The hypothesis that enhancers may act to promote transcription by a similar mechanism to core promoter sequences is reinforced by the observation that an octamer motif originally characterized as a distal upstream element (Harvey *et al.*, 1982) has also been identified in enhancers (Nomiyama *et al.*, 1987; Parslow *et al.*, 1987). Furthermore, the 21bp repeats that are part of the SV40 core promoter can, like enhancers, work in either orientation (Everret *et al.*,1983; Gidoni *et al.*, 1985). However, the ability to act over large distances is unique to enhancers. The mechanism of action at a distance may rely on transcription factors binding to enhancers and then migrating to the site of initiation by sliding along the DNA molecule, or more probably, on the looping out of intervening sequences bringing the enhancer bound factor into the vicinity of the rest of the transcriptional complex (Ptashne, 1986b). The SV40 and cytomegalovirus enhancers have been shown to stimulate transcription *in vitro*, even when non-covalently attached to the β-globin promoter, an effect that is inconsistent with the sliding of transcription factors along the DNA molecule (Muller *et al.*, 1989). Furthermore, when

two identical promoters were placed nearby in a divergent orientation with the SV40 enhancer a great distance to one side, the promoters were shown to be equally stimulated, whereas if the sliding model were to operate, a polar effect upon the initiation of transcription would have been expected (Heuchel *et al.*, 1989).

The cooperative binding of transcription factors has been demonstrated in SV40. Insertions of nucleotides corresponding to half or whole turns of the DNA helix between sequence elements in the early gene promoter indicate that different factors must bind on the same face of the DNA molecule to exert their effect (Takahashi *et al.*, 1986). Thus, as with interactions between promoter sequence elements and RNA polymerase in prokaryotes, the critical feature of cis-acting elements that constitute eukaryotic promoters appears to be their ability to bind protein factors which can then interact with RNA polymerase II to produce the mature transcriptional complex.

1.2 Aspergillus nidulans as a model for eukaryotic gene expression.

Aspergillus nidulans is an ascomycete fungus noted for its rapid growth rate, colonial habit and uninucleate conidiospores. The organism has both haploid and diploid growth phases, and although it is homothallic, sexual crosses can be done by the use of genetic markers. These features, together with parasexual processes for mitotic recombination, make *A.nidulans* an attractive organism in genetics, and a powerful system for genetic analysis has been developed by Pontecorvo *et al.*, (1953). A genetic map of the eight chromosomes is given in Clutterbuck (1982). The application of classical genetics and biochemistry in this lower eukaryote has allowed the regulation of generally unlinked enzyme structural genes under common metabolic control to be investigated. Such studies have revealed complex interactions between many regulatory genes affecting responses to a variety of metabolic signals (Kinghorn and Pateman, 1977; Cove, 1979; McCullough *et al.*, 1977; Arst and Bailey, 1977).

More recently the development of a DNA mediated transformation procedure (Ballance *et al.*, 1983; Tilburn *et al.*, 1983; Yelton *et al.*, 1984) and the construction of cosmid cloning vectors (Yelton *et al.*, 1985; Turner and Ballance, 1986) have facilitated the isolation of specific genes by the repair of metabolic lesions in equivalent mutant strains of *A.nidulans*. Before these techniques were available, a few *A.nidulans* genes were isolated by the complementation of mutants of *Escherichia coli*, for example the aromatic amino acid biosynthetic (*aromA*) gene (Charles *et al.*, 1986), or of *Saccharomyces cerevisiae*, for example the alcohol dehydrogenase III (*alcC*) gene (McKnight *et al.*, 1985). More commonly, *A.nidulans* genes have been isolated using heterologous DNA probes for hybridization, as for the phosphoglycerate kinase (*pgk*) gene (Clements and Roberts, 1985).

The determination of the nucleotide sequences of genes

cloned from filamentous fungi has revealed a number of common features: core promoter elements of defined sequence and position, upstream activator sequences, transcript processing signals, and introns with consensus junction sites and internal sequence elements. These aspects of filamentous fungal genes are reviewed by Gurr *et al.* (1987), and the investigation of transcriptional control is discussed in detail in Section 1.4.

The combination of classical genetic studies and molecular analysis of cloned genes has allowed complex regulatory circuits in *A.nidulans* to be investigated in a similar fashion to those of prokaryotes and yeast, and these are reviewed by Davis and Hynes, 1989. In particular, regulatory genes have been cloned and analysed which encode factors that control the expression of the gene clusters for quinic acid utilization (Beri *et al.*, 1987) and proline utilization (Hull *et al.*, 1989). Other regulatory genes have also been analysed which encode factors that mediate more general aspects of gene expression, such as nitrogen metabolite repression (Caddick *et al.*, 1986).

The development of gene replacement strategies in *A.nidulans* (Miller *et al.*, 1985) provides a means by which gene function can be investigated, and the construction of *lacZ* fusion vectors (van Gorcom *et al.*, 1986) has facilitated the analysis of cis-acting promoter sequences (Hamer and Timberlake, 1987). Furthermore, *A.nidulans* undergoes differentiation during development, and the identification and molecular analysis of genes expressed during conidiospore formation are described by Mirabito *et al.*, (1989). Thus, *A.nidulans* combines a well characterized genetic system and proven techniques for gene manipulation including DNA mediated transformation, such as are also available in *E.coli* and *S.cerevisiae*, with some degree of the developmental differentiation associated with higher eukaryotes. The ease with which *A.nidulans* can be handled and investigated by both classical and molecular genetics, makes it a good choice for the study of eukaryotic gene expression.

1.3 Transformation of yeast and filamentous fungi.

The availability of an effective DNA mediated transformation procedure is important to facilitate the analysis of many aspects of gene function. For example the identity and integrity of cloned genes may be confirmed when well characterised mutant recipient strains are available for transformation. The development of plasmid and cosmid vectors has permitted the isolation of fungal genes by the complementation of mutant strains (Rine and Carlson, 1985; Yelton *et al.*, 1985; Atkins and Lambowitz, 1985; Vollmer and Yanofsky, 1986). When a gene has been cloned by a less direct route, such as hybridization to heterologous sequence, and no corresponding mutant strains are known, the function of the gene may be identified phenotypically by disruption of the resident gene to generate a null mutant (Shortle *et al.*, 1982; Rothstein, 1983; Paietta and Marzluf, 1985; Miller *et al.*, 1985). The related strategy of gene replacement permits specific genes to be exchanged for sequences with defined *in vitro* modifications (Struhl, 1983; Frederick *et al.*, 1989). This method is particularly useful for investigating the action of regulatory proteins, in apportioning functions to specific domains within large multifunctional proteins, and in analyzing sequences upstream of genes to define promoter elements. The introduction of multiple copies of cis-acting regulatory elements and of regulatory genes have allowed the *in vivo* binding sites for proteins controlling the initiation of transcription to be identified (reviewed by Davis and Hynes, 1989).

Fungal transformation procedures were pioneered in *Saccharomyces cerevisiae* by Hinnen *et al.*, (1978) and are reviewed by Boguslawski (1985). The standard transformation protocol depends upon the preparation of spheroplasts by incubation of cells harvested from an exponential growing culture with a commercial glucanase preparation to remove the cell wall. The spheroplasts are osmotically stabilized using sorbitol and are stimulated to take up DNA by including calcium ions and polyethylene glycol (PEG) in the

transformation mixture (Hinnen *et al.*, 1978). The use of spheroplasts obviates the difficulty of attempting to introduce DNA through the cell wall, and the calcium ions and PEG are believed to make the cell membrane permeable to large nucleic acid molecules. Whole cells can be regenerated from protoplasts by including the osmotic stabilizer sorbitol in the growth medium. Selection is enforced by regenerating whole cells on agar medium permitting only the growth of the desired transformants. This may be by selection for the independence of a particular nutrient, or by selection for a dominant character such as hygromycin B resistance, $HygB^r$ (Kaster *et al.*, 1984).

Transformation systems in filamentous fungi are reviewed by Turner and Ballance (1985) and Fincham (1989). In the case of *Neurospora crassa* spheroplasts are generally prepared from germinating conidia using a commercial enzyme preparation, Novozym 234, and are stabilized with sorbitol (Vollmer and Yanofsky, 1986). With *Aspergillus nidulans* both germinating conidia and fresh growing mycelium are commonly used to prepare spheroplasts using Novozym 234, and these are stabilized by the presence of sodium or potassium ions (Ballance *et al.*, 1983; Yelton *et al.*, 1984). As with *S.cerevisiae*, the uptake of DNA by spheroplasts is promoted by including calcium ions and PEG in the transformation reaction. Dimethyl sulphoxide is also used in the case of *N.crassa*. Transformants may be selected by the complementation of an auxotrophic mutant, or by the use of dominant selectable markers: for example the mutant β-tubulin gene (Ben^r) of *N.crassa* which provides benomyl resistance (Orbach *et al.*, 1986), or the oligomycin resistance ATP synthase subunit 9 gene (*oliC*) of *A.nidulans* (Ward *et al.*, 1986). When a transforming gene cannot be easily selected, the identification of colonies that have incorporated the gene can be aided by cotransforming with a second convenient selectable marker (Werners *et al.*, 1987). However, since cotransformation relies on the uptake of more than one DNA molecule into a competent cell, the resultant transformants are generally complex and not accessible for

Figure 1.2
Diagrammatic representation of possible events integrating vector sequence into the genome.

a)TypeI: a single crossover at the homologous host locus.

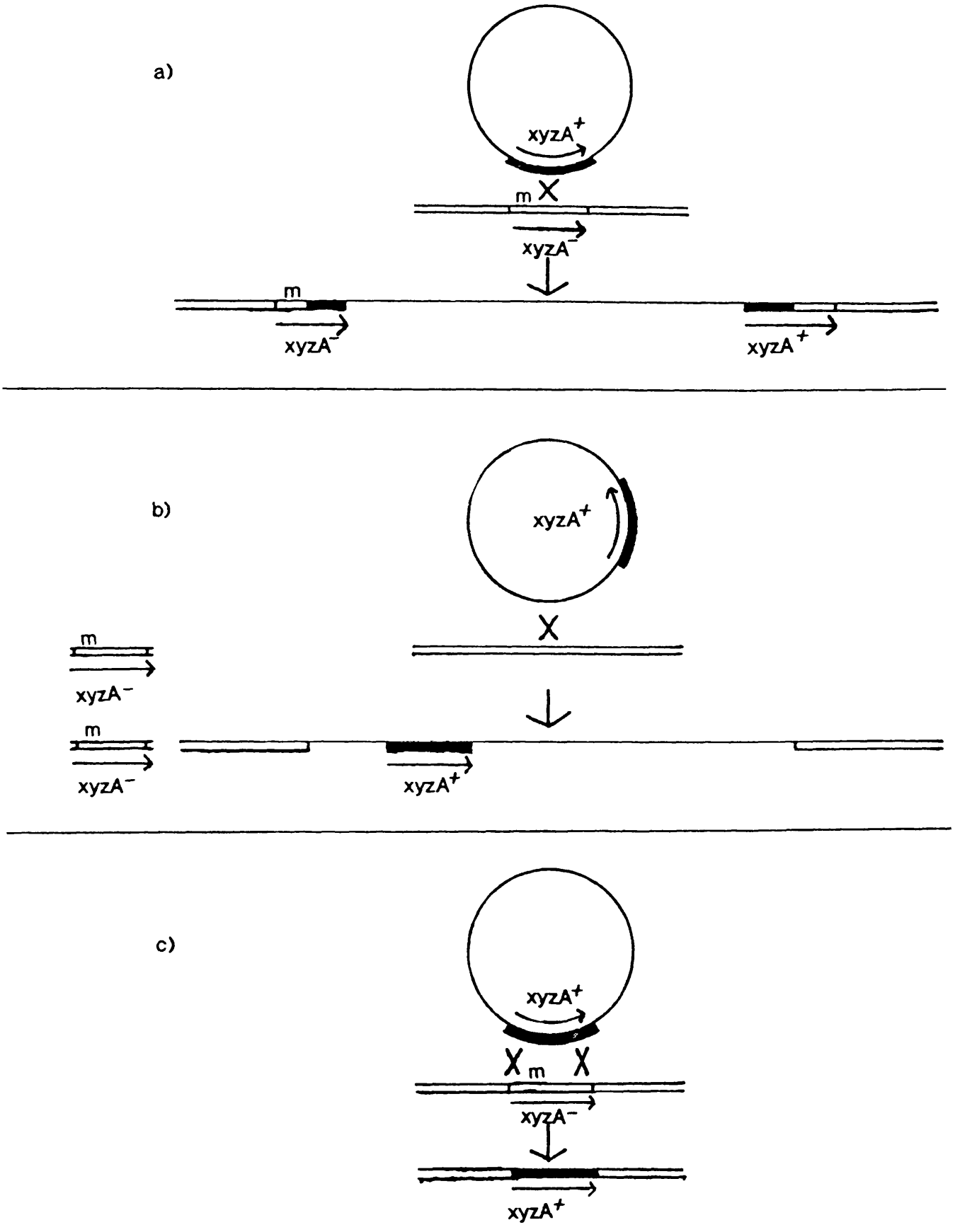
b)TypeII: a single crossover at a non-homologous host locus.

c)TypeIII: a double crossover at the homologous host locus.

The classification of integration events is as described by Hinnen *et al.* (1978).

Key: In this Figure and in Figure 1.3,
sequence of bacterial origin is represented as
sequence of fungal origin located on the vector as
and genomic fungal sequence as

A wild type fungal gene is depicted *xyzA*$^{+}$ and a mutant version of this gene *xyzA*$^{-}$, with m representing the site of mutation.


a)
xyzA+
m
xyzA−
m
xyzA−
xyzA+
b)
xyzA+
m
xyzA−
m
xyzA−
xyzA+
c)
xyzA+
m
xyzA−
xyzA+

further molecular analysis.

With *N.crassa*, the transformant colonies that are initially selected are likely to be heterokaryons because of the multinucleate nature of conidiospores and nuclear mixing within mycelial compartments. The resolution of these heterokaryons may require several rounds of reisolation from single conidiospores in order to ensure the segregation of nuclear components. In *A.nidulans*, although the transformed spheroplasts are likely to be multinucleate, the segregational nature of the organism ensures that heterokaryons are much less likely to be formed. If isolated, heterokaryons can be easily resolved due to the uninucleate nature of the conidiospore.

The stable heritable maintenance of transforming DNA in mitotic growth requires the selectable marker either to be on a replicating vector molecule, or to be incorporated into the genome of recipient cells. The first class of plasmids developed for the transformation of *S.cerevisiae* were yeast integrating plasmids (YIp). These are incapable of replicating and thus rely upon the integration of yeast DNA present in an *Escherichia coli* based plasmid into the yeast genome. This is most commonly achieved by a single recombination event between homologous yeast sequences on the vector and on a recipient chromosome, and results in the incorporation of the entire plasmid molecule at the relevant host locus (a type I event). Less frequently yeast sequences on the vector may replace the corresponding host sequence (a type III event), and integration may also occur at an unrelated locus (a type II event), most probably because of chance homology between vector and genomic sequences such as Ty or δ (Kingsman *et al.*, 1981). The types of vector integration events categorized above were originally identified by Hinnen *et al.* (1978) and are outlined in Figure 1.2. Although the loss of integrated sequence by intrachromosomal recombination is possible at low frequency (Ilgen *et al.*, 1979), multiple copies of a vector can be arranged in tandem at one site by the integration of several plasmids at a particular locus

(Orr-Weaver and Szostak, 1983).

Further advances have been made in yeast to generate episomal plasmids (YEp) and replicating plasmids (YRp). Yeast episomal plasmids replicate autonomously within the cell, a property they owe to the presence of a replication origin and stability sequences derived from the native 2μm plasmid of yeast (Broach *et al.*, 1979). Yeast replicating plasmids contain an autonomous replication sequence (*ARS*), originally identified by Stinchcomb *et al.* (1979), of which there are several types dispersed throughout the genome (Chan and Tye, 1980). A lack of stability and a bias in the segregation of YRp towards the mother cell during budding makes YRp less heritable than YEp (Murray and Szostak, 1983a).

In order to stabilize yeast replicating plasmids, centromeric sequences were introduced to produce yeast centromeric plasmids (YCp) (Clarke and Carbon, 1980; Fitzgerald-Hayes *et al.*, 1982). The presence of centromeric sequences increases plasmid stability to approach that of chromosomes and reduces the plasmid copy number from approximately sixty to one. The sequences are believed to interact with proteins involved in spindle formation, so allowing regular distribution during mitosis and meiosis. The stability of transforming DNA sequences has been further increased by including cloned yeast telomeric DNA sequences on the vectors to create yeast linear plasmids (YLp) (Murray and Szostak, 1983b) also referred to as yeast artificial chromosome (YAC) vectors. When the plasmids are very large (greater than 50Kb), they are mitotically stable and behave correctly during meiosis.

The development of yeast episomal plasmids based on the native 2μm plasmids of *S.cerevisiae*, promoted an extensive search for similar autonomously replicating molecules in filamentous fungi. Although no nuclear plasmids have been identified, mitochondrial plasmids have been discovered. The inclusion of mitochondrial plasmid DNA from *Neurospora intermedia* in an *E.coli* based construct with a *N.crassa*

selectable marker, led to an increased transformation frequency (Stohl and Lambowitz, 1983). However, a deletion derivative of this vector, lacking almost the entire mitochondrial plasmid sequence, had a comparable transformation frequency, indicating that modification of replication origins in sequence of *E.coli* or *N.crassa* origin had stimulated transformation (Stohl *et al.*, 1984).

In the absence of stable, autonomously replicating plasmids for filamentous fungi, the integration of vector sequences into the genome in the fashion of yeast integrating plasmids has provided a route by which filamentous fungal sequences can be transformed and selected in recipient organisms. Whatever the selective marker used, over 90% of the colonies initially observed remain very small and are incapable of further growth when subcultured (Buxton and Radford, 1984; Turner and Ballance, 1985). Such abortive transformants may result from transient expression without stable incorporation of transforming DNA. Mitotically stable transformants in which vector DNA has been incorporated into the genome have been categorized according to the types of DNA integration events described for yeast above. However, transformed strains show a considerable degree of meiotic instability (Tilburn *et al.*, 1983; Yelton *et al.*, 1984). The linkage of a gene on transforming DNA to the homologous chromosome locus has been investigated by crossing transformed strains to the wild type and analysing the progeny with respect to the selectable marker used in the transformation experiment. The transforming acetamidase (*amdS*$^+$) gene of *A.nidulans* has been mapped to several different chromosomes (Wernars *et al.*, 1986), and the transforming NADP dependent glutamate dehydrogenase (*am*$^+$) gene of *N.crassa* has been shown to give approximately equal numbers of integrants at the resident gene locus and elsewhere (Kinsey and Rambosek, 1984).

A more informative analysis of the fate of transforming DNA is obtained by restriction enzyme digestion of genomic DNA from transformants, and hybridization with defined probes following agarose gel electrophoresis to size

separate fragments. This kind of analysis has demonstrated that transformation of *N.crassa* tends to result in predominantly ectopic (type II) events, although replacement (type III) events are also common (Case *et al.*, 1979; Kim and Marzluf, 1988), whereas additive (type I) events are rarely observed (Selker *et al.*, 1987). In *A.nidulans* integration events resulting from homologous recombination at the host locus appear to be far more common than in *N.crassa*, with type I and III transformants predominating over type II transformants (Tilburn *et al.*, 1983; Yelton *et al.*, 1984; DeGraff *et al.*, 1988). Similar results have been obtained with *Penicillium chrysogenum* (Bull *et al.*, 1988). As in *S.cerevisiae* (Orr-Weaver and Szostak, 1983), tandemly repeated arrays of vector sequences have been identified at the homologous host locus in *A.nidulans* (Yelton *et al.*, 1984). Multiple copies of vector sequences have also been observed at non-homologous loci, possibly through an initial type II integration event followed by homologous recombination between the integrated sequence and further plasmid molecules (Wernars *et al.*, 1985).

Linearization of vector molecules at a site within the selective marker, so generating potentially recombinogenic homologous ends, has been shown to increase the frequency of type I transformation, or targetting in *S.cerevisiae* up to a thousand-fold (Orr-Weaver *et al.*, 1981). Such dramatic effects have not been observed in filamentous fungi, although as discussed in Chapter 5, Section 3, the site of linearization does appear to govern the degree of targeting to the homologous locus. Furthermore, the choice of recipient strain appears to affect the proportion of homologous recombination events (Kim and Marzluf, 1988).

An *A.nidulans* sequence designated *ansl* was identified by its ability to act as an *ARS* in *S.cerevisiae*. When this sequence was included on a vector carrying the *pyr-4* (orotidine decarboxylase) gene of *N.crassa*, the transformation frequency was increased 50 to 100-fold when the vector was introduced into the corresponding mutant strain (*pyrG*$^{-}$) of *A.nidulans* (Ballance and Turner, 1985).

The location of vector sequences in high molecular weight DNA prepared from transformed strains, indicated that the plasmid had integrated into chromosomal sequence. Increased vector integration may reflect the presence of multiple copies of *ansl* in the genome, so facilitating homologous recombination and an increased proportion of non-homologous (type II) integration events with respect to the *A.nidulans* selectable marker on the vector.

The development of efficient transformation systems has greatly facilitated the cloning of fungal genes. Yeast episomal plasmids based on *S.cerevisiae* 2μm plasmid sequences and also carrying an *E.coli* plasmid origin are ideal shuttle vectors in which a library of yeast DNA sequences can be made and transformed into a particular mutant strain. Plasmid DNA is prepared directly from transformed strains that contain sequence complementing the mutant, and is recovered in *E.coli*. The use of yeast integrating plasmids to clone genes is technically more difficult, but does allow the gene locus to be identified in the genome by genetic analysis of transformed strains (Boguslawski, 1985).

The construction of a cosmid vector has provided an efficient means by which *A.nidulans* genes can be cloned by complementation (Yelton *et al.*, 1985). The cosmid includes the trifunctional tryptophan biosynthetic gene (*trpC*), and initial selection is for tryptophan independent transformants in a strain that is *trpC*$^{-}$, and also mutant for the gene to be cloned. Subsequent selection is for complementation of the mutant phenotype. Location of the complementing gene within genomic sequences cloned into the cosmid is rapidly achieved by transforming DNA fragments derived from this cosmid into the *A.nidulans* mutant (Timberlake *et al.*, 1985). Subdividing a library of cosmid clones into pools, and transforming each pool into a particular mutant strain, and then sequentially repeating this process on sub-pools has allowed the cloning of genes which could not easily be isolated by direct selection (Atkins and Lambowitz, 1985; Vollmer and Yanofsky, 1986).

The disruption of a cloned gene is a commonly applied method of determining its function, and gene disruption and replacement strategies are outlined in Figure 1.3. Gene disruption has been achieved in *S.cerevisiae* by excising sequences from the gene to be disrupted and then transforming this into the organism on a vector carrying a selective marker. The type I integration of vector sequence at the locus corresponding to the defective gene generates two copies of the gene, each of which is non-functional. If the gene is essential for cell viability transformants can be obtained by transforming a diploid strain (Shortle *et al.*, 1982). By inserting the selectable marker into the cloned gene whose function is to be determined, the native gene can be disrupted at the host locus by gene replacement. This technique was first applied in *S.cerevisiae* (Rothstein, 1983), and has since been extended to *N.crassa* (Paietta and Marzluf, 1985), and *A.nidulans* (Miller *et al.*, 1985).

The replacement of chromosomal sequence with an *in vitro* modified cloned gene can be achieved through a two-step procedure, in which a vector carrying both a modified version of the gene and a selectable marker is transformed into the organism under study. A type I integration event at the host locus corresponding to the mutated gene, followed by deletion of the intervening marker sequence used for the initial selection of the transformants, may generate a strain in which the native gene has been replaced with the modified version. This two-step gene replacement strategy was developed in *S.cerevisiae* (Scherer and Davis, 1979), but has also been used in *A.nidulans* (Miller *et al.*, 1985). The extensive applications of gene disruption and gene replacement strategies are discussed further by Boguslawski (1985) and Fincham (1989).

Figure 1.3
The replacement of chromosomal sequence with a disrupted or an *in vitro* modified cloned gene.

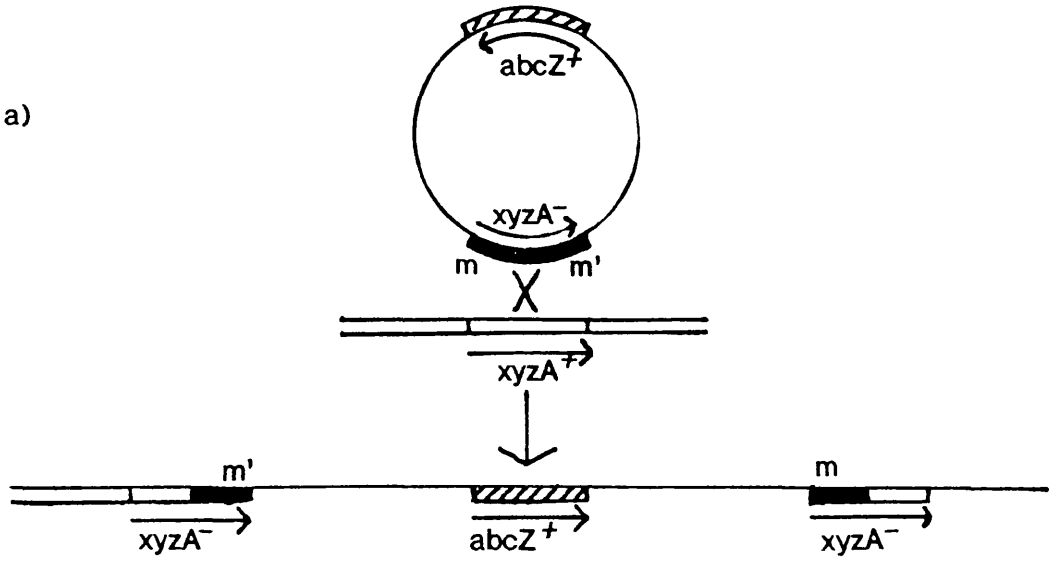
a) Gene disruption by the direct (one-step) method of Shortle *et al.* (1982). A single crossover at the homologous host locus, with *abcZ*$^{+}$ as the selective marker. The defective gene *xyzA*$^{-}$ is missing both N-terminal and C-terminal codons.

b) Gene disruption by the direct (one-step) method of Rothstein (1983). A double crossover at the homologous host locus, with *abcZ*$^{+}$ as the selective marker. The defective gene *xyzA*$^{-}$ is interrupted within coding sequence by the gene *abcZ*$^{+}$ (*xyzA*$^{-}$*::abcZ*$^{+}$).

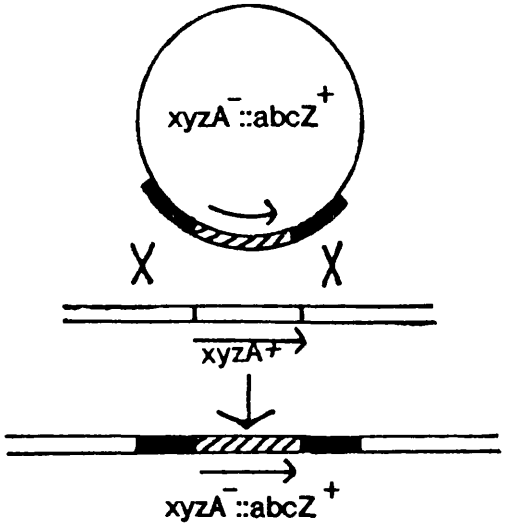
c) Gene replacement by the indirect (two-step) method of Scherer and Davis (1979). A single crossover at the homologous host locus, with *abcz*$^{+}$ as the selective marker which is subsequently lost through recombination.

Sequences are represented as described for Figure 1.2. Wild type fungal genes are depicted *abcZ*$^{+}$ and *xyzA*$^{+}$, and mutant versions of the latter *xyzA*$^{-}$ with m and m′ representing the sites of mutation.

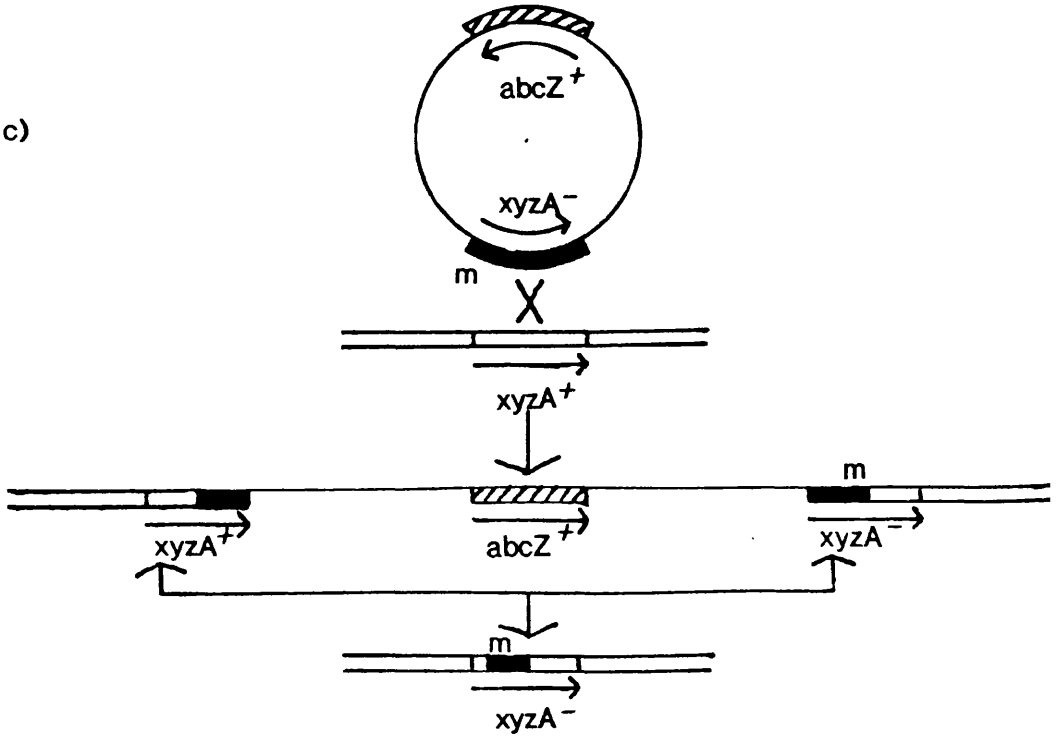
a)


abcZ+
xyzA−
m
m'
xyzA+
m'
m
xyzA−
abcZ+
xyzA−

b)


xyzA−::abcZ+
xyzA+
xyzA−::abcZ+

c)


abcZ+
xyzA−
m
xyzA+
m
xyzA+
abcZ+
xyzA−
m
xyzA−

1.4 Control of transcription in filamentous fungi.

The two filamentous fungi most intensively studied are the Ascomycetes *Neurospora crassa* and *Aspergillus nidulans*, although in recent years interest has also turned to fungi of industrial importance such as *Aspergillus niger* and *Penicillium chrysogenum*. *A.nidulans* and *N.crassa* have genome sizes of approximately 2.8×10^7 and 3.7×10^7bp respectively (Timberlake, 1978; Krumlauf and Marzluf, 1979), and their nuclear DNA is organised in a similar fashion to that in higher eukaryotes, with linear chromosomes associated with histones and chromatin being ordered into nucleosomes undergoing structural changes throughout the cell cycle (Noll, 1976). Furthermore, the DNA methylation pattern of filamentous fungi resembles that of higher eukaryotes (Russel *et al.*, 1985). However, the genomes of filamentous fungi lack extensive repeated DNA sequences common in higher eukaryotes (Timberlake, 1978, Krumlauf and Marzluf, 1980), with the exception of rRNA genes which account for about 5% of the total sequence (Free *et al.*, 1979).

Genetic analysis has revealed that the majority of genes encoding enzymes that act in successive reactions in a metabolic pathway are scattered around the genome, but there are examples of clusters of tightly linked genes with related metabolic function: most notably the genes for quinic acid utilization in *N.crassa* (Giles *et al.*, 1985) and in *A.nidulans* (Da Silva *et al.*, 1986; Hawkins *et al.*, 1988), and for nitrate utilization (Cove 1979), proline catabolism (Arst and Scazzocchio, 1985) and ethanol utilization (Pateman *et al.*, 1985) in *A.nidulans*. Furthermore, genes associated with the sporulation of *A.nidulans* are tightly linked in a gene cluster (Timberlake and Barnard, 1981). The molecular analysis of these clusters has revealed that component genes are transcribed to generate individual transcripts as found in higher eukaryotes. However, there are examples of a single gene encoding a multifunctional product. The first gene to be thus characterized was the *HIS4* gene of *Saccharomyces cerevisiae* shown to have three

functions on a single polypeptide (Bigelis *et al.*, 1977). Similar situations include the *arom* genes of *N.crassa* (Giles *et al.*, 1967) and *A.nidulans* (Charles *et al.*, 1986) involved in the aromatic amino acid pathway, and the *his-3* gene of *N.crassa* (Legerton and Yanofsky, 1985) involved in histidine biosynthesis. Comparison of the derived protein sequence of the *A.nidulans aromA* gene with that of the corresponding proteins of five genes of *Escherichia coli*, indicates the fusion of five separate units to give a single functional cistron (Hawkins, 1987).

The major thrust of work on the control of gene expression in fungi has centered on transcription, and specifically on the role of cis-acting consensus sequences. The bulk of data has been collected from *S.cerevisiae*. However, well over fifty genes have also been cloned and sequenced from filamentous fungi, predominantly from *N.crassa* and *A.nidulans*. Several of the features of eukaryotic promoters discussed in Section 1 are present upstream of fungal genes, along with other novel sequence elements (reviewed by Gurr *et al.*, 1987). However, there are few examples of the demonstration of the function of these consensus elements to directly implicate them in controlling transcription.

Many fungal genes exibit multiple transcription start sites, with as many as twenty sites for the *S.cerevisiae* iso-1-cytochrome C (*CYC1*) gene (McNeil and Smith, 1985). Examination of the sequences surrounding the start sites of fungal genes shows that the capsite is not rigidly conserved, (Brown and Lithgow, 1987; Gurr *et al.*, 1987), although there is a tendency towards the sequence PyAAG in the case of efficiently expressed yeast genes (Dobson *et al.*, 1982).

The TATA box, implicated in positioning the site of transcription initiation, is present upstream of most *S.cerevisiae* genes, although its precise location is often more variable than in higher eukaryotes, often lying over 100nt 5′ to the start site (Dobson *et al.*, 1982). Such

variation questions the role of this element in defining the transcription initiation point in *S.cerevisiae*, a doubt substantiated by experiments on the histidine biosynthetic (*HIS3*) gene (Chen and Struhl, 1985). Deletions in the region between the TATA box and the start site did not affect the site of initiation or the level of message produced. Furthermore, fusions of the TATA element of one gene to the coding region of another gave transcripts that initiated at the normal site. On the other hand constructs in which the TATA box was left intact, but sequences around the start site were modified, did not initiate correctly. In contrast, four TATA elements identified upstream of the iso-1-cytochrome C (*CYC1*) gene each independently define different subsets of transcripts (McNeil and Smith, 1986).

Some filamentous fungal genes have an appropriately placed TATA like sequence at approximately the consensus eukaryotic position, but others have an AT rich region located anywhere from 130nt upstream of the start site to the start site itself (Gurr *et al.*, 1987). The functional role of TATA like elements is uncertain, for the deletion of such a sequence from the 5′ region of the *A.nidulans* tryptophan biosynthetic (*trpC*) gene does not affect the expression of a *lacZ* reporter gene fused downstream (Hamer and Timberlake, 1987). The distal CCAAT element located approximately 80nt upstream of many higher eukaryotic genes is rare in *S.cerevisiae* genes, and although such a potential element has been identified upstream of some filamentous fungal genes, its position varies considerably (Gurr *et al.*, 1987). Thus, the function of sequence elements apparently like TATA or CCAAT 5′ of fungal genes have not been proved, and in many cases may simply reflect chance homology. However, a further sequence element unique to fungal genes, the CT motif, has been demonstrated to be involved in directing the initiation of transcription. This pyrimidine rich sequence immediately precedes or even includes the transcription start site and is most common 5′ to strongly expressed genes of *S.cerevisiae* and filamentous fungi, particularly those lacking canonical TATA and CCAAT sequences (Dobson *et al.*, 1982; Brown and Lithgow, 1987;

Gurr *et al.*, 1987). Functional analysis of several CT rich regions upstream of the multiple start sites of the *CYC1* gene of *S.cerevisiae* has demonstrated that this motif may be important in locating the transcription initiation sites (McNeil and Smith, 1985).

Sequence elements that mediate regulatory responses have been identified 5′ to many *S.cerevisiae* genes. These elements are termed upstream activator sequences (UAS) or upstream repressor sequences (URS) according to their effect on gene expression, and resemble the enhancers characteristic of higher eukaryotic genes, except that they are unable to act when they are positioned 3′ to a gene (Struhl, 1984; Guarente and Hoar, 1985). One of the best characterized UAS elements of *S.cerevisiae* is that which lies in the bidirectional promoter of the galactose utilization (*GAL1* and *GAL10*) genes. The positive regulatory protein encoded by the *GAL4* gene binds to this element, an association inhibited in the presence of glucose (Giniger *et al.*, 1985). A sequence element located upstream of *S.cerevisiae* a-specific genes provides an example of a URS. The product of the *MAT* α-2 gene represses the transcription of these genes by binding at this site (Johnston and Herskowitz, 1985). A further type of regulatory element, a downstream activator sequence (DAS), has been located within the coding region of the *S.cerevisiae* 3-phosphoglycerate kinase (*PGK*) gene (Mellor *et al.*, 1987), and is discussed in detail in Chapter 7.

Comparison of sequences upstream of coordinately regulated genes, together with an analysis of the expression of these genes under the control of native or modified promoters, has revealed the presence of UAS elements in filamentous fungi. Whether these sequence elements more closely resemble the UAS elements described in *S.cerevisiae*, or eukaryotic enhancers is not yet clear. The best characterized examples of such elements in filamentous fungi are in the quinic acid utilization genes of *N.crassa* (Giles *et al.*, 1985) and *A.nidulans* (Hawkins *et al.*, 1988), the alcohol utilization genes of *A.nidulans* (Gwynne *et al.*,

1987), and the acetamidase gene of *A.nidulans* (Hynes and Davis, 1986). Sequence analysis of the cloned genes encoding the positively acting regulatory proteins *qa-1F* and *qutA* for quinic acid utilization (Giles *et al.*, 1985; Beri *et al.*,1987), *prnA* for proline utilization (Scazzocchio, cited in Davis and Hynes, 1989), *amdR* for the induction of lactam and acetamidase catabolism (Andrianopoulos and Hynes, cited in Davis and Hynes, 1989) and *creA* for carbon catabolite repression (Dowzer and Kelly, 1989) have identified so called "zinc finger" sequences in the proteins indicative of DNA binding activity. In the case of *amdR*, DNA binding activity has been confirmed by gel retardation studies (van Heeswyck and Hynes, cited in Davis and Hynes, 1989). In addition to the activator proteins listed above, the *qa-1S* gene of *N.crassa* (Huiet and Giles, 1986) and *qutR* gene of *A.nidulans* (Grant *et al.*, 1988) appear to encode proteins involved in the repression of the enzyme structural genes of quinic acid catabolism.

Multiple copies of sequences present on transforming plasmids can be integrated into the genome of *A.nidulans*. This phenomenon , together with that of cotransformation of a selectable marker, has allowed multiple copies of the potential binding sequences of the *amdS* gene to be introduced into *A.nidulans*. Specific sequence elements have been shown to be capable of titrating the activator proteins *amdR*, *facB* and *amdA*, thus defining their sites of binding (reviewed by Davis and Hynes, 1989). Introduction of multiple copies of *amdR* or *facB*, so increasing the levels of regulatory proteins to match those of the binding sequences, have been shown to reverse this titration effect (Andrianopoulos and Hynes, 1988; Davis and Hynes, 1989). The *A.nidulans qutE* gene has been selectively overexpressed by introducing equal multiple copies of both *qutE* and the activator gene *qutA* (Beri *et al.*, 1990). However, the presence of multiple copies of either *qutE* or *qutA* alone is insufficient to increase *qutE* expression, indicating a direct correlation between the synthesis of activator protein and the number of control sequences upstream of *qut* cluster genes.

There is no requirement for vector sequence to be introduced into filamentous fungal genomes at the homologous gene locus, and this property of non-homologous integration has allowed genes to be transferred between different filamentous fungal species. Inter-species studies have been undertaken with several different genes and species, with for example the nitrate reductase gene of *A.nidulans* having been transformed into seven other species of filamentous fungi (Daboussi *et al.*, 1989). Even genes containing introns, such as *amdS* of *A.nidulans*, and those encoding regulatory proteins, such as *nit-2* of *N.crassa* which can activate *areA*-controlled genes *A.nidulans, have been transformed across filamentous fungal species, even though the sequences of nit-2* and *areA* have diverged considerably (Davis and Hynes, 1987). This heterologous expression may reflect a relaxation of the precise amino acid sequence in activator proteins, the overall protein conformation and acidity being the governing factors in nucleic acid binding, as proposed for *S.cerevisiae* (Giniger and Ptashne, 1987; Sigler, 1988). More generally, when analysing the expression of structural genes, comparisons of promoters to identify conserved sequences may be combined with studies of heterologous expression to determine whether these sequences are consensus elements required for the efficient expression of transcription.

In summary, the *in vitro* manipulation of cloned DNA sequences, combined with the reintroduction of sequences into the host organism, has facilitated the detailed molecular analysis of several filamentous fungal genes to complement previous genetic studies. Although conserved sequences have been identified upstream of many genes, and considerable progress has been made in elucidating complex regulatory circuits, a more rigorous functional analysis of promoter sequences is required before a specific role can be assigned to the several consensus elements described above.

1.5 The 3-phosphoglycerate kinase *pgk* gene of *Aspergillus nidulans*.

The glycolytic conversion of glucose to pyruvate with the concomitant production of ATP is an essential pathway in both aerobic and anaerobic organisms. Under aerobic conditions pyruvate is metabolized by the citric acid cycle and the electron transport chain, whilst under anaerobic conditions pyruvate is fermented: either to ethanol, as in yeast, or to lactate, as in muscle.

The enzyme 3-phosphoglycerate kinase (*PGK*) catalyses the high energy phosphoryl transfer reaction:

1,3-diphosphoglycerate + ADP $\rightleftharpoons$ 3-phosphoglycerate + ATP

in which the phosphoryl group from the acyl phosphate of 1,3-diphosphoglycerate is transferred to ADP. Magnesium is required as a cofactor in this reaction. The synthesis of glucose from non-carbohydrate precursors by the gluconeogenic pathway also requires *PGK* to catalyze the reverse reaction to generate 1,3-diphosphoglycerate.

The enzyme is a monomer with a molecular mass of approximately 45kDa (Scopes, 1973) and the complete amino acid sequence has been determined for *PGK* isolated from horse muscle (Banks *et al.*, 1979), human (Huang *et al.*, 1980) and yeast (Watson *et al.*, 1982). These data reveal that the primary structure of the enzyme is highly conserved, the yeast enzyme being 65% identical to the horse and human enzymes. Furthermore, X-ray crystallographic analysis reveals a very similar structure for the horse muscle and yeast proteins (Banks *et al.*, 1979; Watson *et al.*, 1982). The protein has two distinct similarly sized domains, one corresponding to C-terminal residues, and one to N-terminal residues plus a short α-helical stretch of C-terminal residues. The nucleotide binds to the C-domain and the phosphoglycerate substrate to the N-domain (Banks *et al.*, 1979; Watson *et al.*, 1982), and it is postulated that the two domains swing together by a hinge-bending mechanism

to bring the substrates into close association.

Mutant strains of *Escherichia coli* (Irani and Maitra, 1974), *Saccharomyces cerevisiae* (Lam and Marmur, 1977) and *Pseudomonas aeruginosa* (Banerjee *et al.*, 1987) deficient in *PGK* activity have been isolated, and the defective phenotype mapped to a single locus in each organism. In the mouse two enzyme deficient phenotypes, one of these testis specific, have been mapped (Nielson and Chapman, 1977; Eicher *et al*; 1978). The phenotypes of the *E.coli*, *S.cerevisiae* and *P.aeruginosa pgk* mutants with respect to growth on different carbon sources are discussed in Chapter 3.

A cDNA clone encoding human *PGK* has been isolated and sequenced (Michelson *et al.*, 1983), as has the gene encoding yeast *PGK* (Hitzeman *et al.*, 1982). The yeast gene, which has no introns, is expressed at a very high level, producing 1-5% of total cellular message and protein (Holland and Holland, 1978), and sequence elements (two TATA boxes and a CCAAT box) have been located upstream of the gene which were predicted to be responsible for this strong expression (Dobson *et al.*, 1982; Hitzeman *et al.*, 1982). However, gene fusions of 5'*pgk* sequence to the coding sequences of other genes resulted in a 20-fold drop in expression relative to *pgk* (Tuite *et al.*, 1982; Mellor *et al.*, 1983), and subsequent experiments indicate the presence of a downstream activator sequence located within the coding sequence of *pgk* (Mellor *et al.*, 1987). Evidence for this positively acting sequence is discussed in Chapter 7, as is the characterization of an upstream activator sequence 5' to the yeast *pgk* gene (Stanway *et al.*, 1987).

More recently, the *pgk* genes of the filamentous fungi *Aspergillus nidulans* (Clements and Roberts, 1986), *Penicillium chrysogenum* (van Solingen *et al.*, 1988) and *Trichoderma reesei* (Vanhanen *et al.*, 1989) have been cloned using the corresponding yeast gene as a hybridization probe. Analysis of the sequence of the *A.nidulans* gene revealed two 57bp introns that interrupt N-terminal codons (Clements and Roberts, 1986) and have intron/exon junctions conforming

closely to the eukaryotic consensus sequence (Mount, 1982). Similarly located introns are present in the *P.chrysogenum* and *T.reesei* genes. The introns of the *A.nidulans pgk* gene both contain the sequence AGCCT__AAAC in their 5' region, a feature that indicates that they may share a common origin. An internal sequence corresponding to the filamentous fungal consensus PyGCTAACN (Gurr *et al.*, 1987) and required for lariat formation is also present in both the introns of *A.nidulans pgk*. Comparison of the derived amino acid sequence in *A.nidulans* with that of the known structure of the yeast protein (Watson *et al.*, 1982), suggests that the two introns probably fall between sequences encoding separate secondary structural elements within the N-domain of the enzyme.

The sequence of the *A.nidulans* gene reveals a 421 amino acid protein with 81% homology to the *P.chrysogenum* enzyme, 77% to the *T.reesei* enzyme, 68% to the *S.cerevisiae* enzyme and 64% to the mammalian enzymes. There is a clear preference in the codon usage of the *A.nidulans pgk* gene, with a bias against A and for C at the third base of each codon. This correlates to the preferred codon usage observed with other highly expressed filamentous fungal genes (Gurr *et al.*, 1987).

As in the case of *S.cerevisiae* (Dobson *et al.*, 1982; Hitzeman *et al.*, 1982), a TATA like sequence and a CCAAT box have been identified upstream of the *A.nidulans pgk* gene (Clements and Roberts, 1986). However, in contrast to the situation in yeast, these elements are located in the consensus positions for eukaryotic promoters of -32 to -38bp and -78 to -81bp relative to the transcript start site respectively. Also, a pyrimidine rich sequence is located encompassing and immediately downstream of the putative TATA box of the *A.nidulans* gene. Such an element is common to several filamentous fungal genes, and the more highly expressed genes of *S.cerevisiae* (Gurr *et al.*, 1987; Dobson *et al.*, 1982). Finally, the *A.nidulans pgk* gene has an 8bp direct repeat CTGGTCTC located between nucleotides -22 to -29 and -50 to -57. However, this feature is not common to

other fungal gene promoters.

Preliminary studies have indicated that *PGK* mRNA is constitutively synthesized at a high level (Clements, 1986). Mapping the *A.nidulans PGK* message reveals a single transcription initiation site 32bp upstream of the translation start codon, and three major polyadenylation regions, each of which would generate a message of about 1500nt (Clements and Roberts, 1986). The most abundant polyadenylation site is located about 115nt downstream of the translation stop codon. A nucleotide sequence identical to the consensus higher eukaryotic poly(A) addition signal AAUAAA (Proudfoot and Whitelaw, 1988) is present 16nt before this polyadenylation site within a potential hairpin-loop structure. One of the other sites is also preceded by the sequence AAUAAA and a potential hairpin structure but no loop.

The DNA sequence surrounding the translation initiation site of *pgk* is CAACAATGTC. As for the eukaryotic consensus sequence (Kozak, 1981) an A is located in the -3 position.

The major aim of my work was to investigate promoter sequences thought to be responsible for the high level of transcription of *pgk* in *A.nidulans*.

1.6 Objectives of research.

The cloning, sequencing and transcript mapping of the single *pgk* gene of *Aspergillus nidulans* has allowed sequence elements to be identified upstream of the gene with homology to eukaryotic transcriptional control sequences (Clements and Roberts, 1986). The principal aim of my work is to analyse the function of these putative promoter sequences by constructing fusions to the *lacZ* gene of *Escherichia coli* to provide a convenient reporter function. The pAN923 vector series (van Gorcom et al., 1986) provides starting material for the construction of a gene fusion which will include all 5'*pgk* elements postulated to be important for stimulating transcription upstream of *lacZ*. The correct initiation of translation in Aspergillus will be ensured by including sequence around the translation initiation site of *pgk*. Subsequently, exonuclease digests starting from the 5' extent of *pgk* sequence on the construct will generate a series of promoter deletions with a diminishing extent of 5' sequence.

The inclusion of the catabolic dehydroquinase gene (*qutE*) on the constructs will allow transformed strains to be selected by growth on quinic acid when the vectors are introduced into a *qutE*$^-$ mutant strain of *A.nidulans*. Furthermore, the presence of the *qutE* gene on the vectors should result in a proportion of the transformants having vector sequence targeted to the *qutE* locus. The analysis of genomic DNA prepared from transformed strains will identify the modes of vector integration, and β-galactosidase assays will allow an examination of the effects of copy number and site of vector integration upon expression.

Transformants with a single copy of a vector molecule integrated at the *qutE* locus will be identified, and the expression of the *pgk-lacZ* fusions monitored by conducting β-galactosidase assays on these strains. By assaying single copy targeted integrants, levels of expression will be related directly to the extent of 5'*pgk* sequence present upstream of *lacZ* without concern over the effects upon

expression of copy number or the site of vector integration. Thus, the 5′ extent of the *pgk* promoter should be determined, and sequences essential for expression identified.

A downstream activator sequence has been identified within the coding region of the *Saccharomyces cerevisiae pgk* gene (Mellor *et al.*, 1987). In order to investigate whether such an element is located within the coding sequence of the *A.nidulans* gene, further vectors will be constructed which will include extensive *pgk* coding sequence upstream of *lacZ*, and β-galactosidase expression will be monitored in transformed strains as described above.

No *pgk*$^-$ mutant strain of *A.nidulans* has been isolated, and consequently neither the position of the *pgk* gene on the *A.nidulans* genetic map, nor the precise phenotype of such a mutant is known. It is my aim to disrupt the *pgk* gene in a diploid strain of *A.nidulans*. Analysis of haploid segregants from a diploid strain in which one of the two *pgk* genes has been inactivated should allow the *pgk* gene to be located to a particular linkage group, and might also enable the recovery of a *pgk*$^-$ mutant strain.

Chapter 2.

Materials and Methods.

2.1 Genetic materials and methods associated with *Aspergillus nidulans*.

2.1.1 *A.nidulans* strains.

All *A.nidulans* strains used were originally derived from the Glasgow stock of Pontecorvo *et al.* (1953). The genotype and origins of the strains used are listed below. Strains in the R21 and R153 genetic background are expected to be translocation free (Fantes and Roberts, 1973).

Strain	Genotype	Reference
WA53	*pyrG89; pyroA4; qutE208*	1
R21*bgaA4*	*pabaA1, yA; bgaA4*	*2*
G191	*pyrG89, pabaA1; fwA; uaY9*	3
R153	*wA3; pyroA4*	4
R153*acuF305*	*wA3; pyroA4; acuF305*	*4*
R153acuD306	*wA3; pyroA4; acuD306*	4
R153*acuN356*	*wA3; pyroA4; acuN356*	4
R153*acuH329*	*wA3; pyroA4; acuH329*	4
R153*acuJ302*	*acuJ302; wA3; pyroA4*	4
R153*acuL304*	*wA3; pyroA4; acuL304*	4

References: 1) Da Silva *et al.*, 1986
2) Fantes and Roberts, 1973
3) Ballance and Turner, 1985
4) Armitt *et al.*, 1976

A *pyrG* strain requires uracil, a *pabaA* strain para-amino benzoic acid and a *pyroA* strain pyrodoxine. A *qutE* strain does not utilize quinic acid for growth, and an *acu* strain is acetate non-utilizing. A *bgaA* strain has a lesion in β-galactosidase. Strains *yA*, *wA* and *fwA* strains have yellow, white and fawn conidiospores respectively. The wild type is green.

2.1.2 Media.

The defined minimal medium (MM) used throughout was based on that of Pontecorvo *et al.* (1953), and was prepared as a 10x solution of stock salts (see below) to which sterile 1M $MgSO_4$ was added to give a final concentration of 10mM after autoclaving (Armitt *et al.*, 1976).

Stock salts solution.

Substance	Final amount per litre
$NaNO_3$	6.00g
KCl	0.52g
KH_2PO_4	1.52g
Trace elements solution	1.0ml

The solution was brought to pH6.5 with NaOH.

Trace elements solution (Armitt *et al*, 1976).

Substance	Amount per litre
$FeSO_4.7H_2O$	1.00g
$ZnSO_4.7H_2O$	8.80g
$CuSO_4.5H_2O$	0.40g
$MnSO_4.4H_2O$	0.15g
$Na_2B_4O_7.10H_2O$	0.10g
$(NH_4)_6Mo_7O_{24}$	0.05g

Minimal agar (MA) medium for plates was solidified with 1.5% (w/v) agar, whilst that for liquid culture contained the wetting agent Tween 80 at 10^{-5} (v/v).

The complete medium used was Malt Extract Agar (MEA):

Substance	Amount per litre
Malt extract	20.0g
Bacto-peptone	1.0g
Agar	15.0g

Nutritional supplements were added to both defined and complete medium as required; para-amino benzoic acid to a final concentration of $1.0mgl^{-1}$, pyridoxin.HCl to $0.5mgl^{-1}$ and uracil to $1.1gl^{-1}$.

All media were sterilized by autoclaving at 15lbs pressure for 15 minutes and carbon sources then added aseptically from sterile stock solutions as given below:

Carbon source	Stock solution	Final concentration
Glucose	1M	0.02M
Glycerol	2M	0.05M
Acetate	2M	0.1M
Sucrose	1M	0.05M
Fructose	1M	0.05M
Quinic acid	20%(w/v)	1% (w/v)

Quinic acid was prepared as a 20% (w/v) stock solution and brought to pH6.5 with NaOH. Stock solutions of 20% (w/v) quinic acid, 1M glucose, 2M glycerol, 2M acetate (pH6.5 with NaOH), 1M sucrose and 1M fructose were prepared in distilled water and sterilized by autoclaving at 15lbs pressure for 10 minutes.

2.1.3 Growth and storage of strains.

All cultures were grown at $37^{o}C$ unless otherwise stated. Liquid cultures were grown on a rotary shaker operating at 250 revolutions per minute in an Ehrlenmeyer flask with four vertical indentations to aerate and suspend the culture. The culture volume did not exceed one fifth of the volume of the flask used. Stock cultures of strains were maintained on MEA slants unless otherwise stated. Long term storage of strains was on silica gels (Roberts, 1969).

2.1.4 Preparation of conidiospore suspension.

Dilute suspensions of conidiospores were spread on the surface of MEA plates and incubated for two days at 37°C. Conidia were harvested by flooding the plates with 15ml sterile Tween-Saline solution (0.8% (w/v) NaCl; 0.025% (v/v) Tween-80) and drawing a sterile wire across the surface of the plate. The suspensions were collected into sterile disposable universal tubes and violently agitated on a Vortex mixer to disrupt conglomerates of conidia. The suspensions were washed by two serial centrifugations in phosphate buffer (20mM KH_2PO_4; 50mM Na_2HPO_4; 50mM NaCl; 0.4mM $MgSO_4.7H_2O$, final pH 7.2) at 3300rpm, and were finally resuspended in 10ml of the same buffer. The concentration of conidial suspensions were estimated by making serial dilutions and counting the conidia present using a haemocytometer examined under a light microscope.

2.1.5 Genetic methods.

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 MA 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 1ml of phosphate buffer (described above), and a sample of the resultant ascospore suspensions streaked on MEA plates to identify hybrid cleistothecia by the presence of recombinant strains for conidial colour markers. 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 incubating overnight. Small blocks of agar containing mycelium were then transferred to thick MA 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 MA by repeated transfer of agar blocks containing mycelium taken from the edge of the growing colonies.

Heterozygous diploid strains were selected by plating suspensions of conidiospores harvested from heterokaryons in MA (Roper, 1952), and dilute suspensions of conidiospores were replated for single colony isolation on the same medium.

Diploid strains were haploidized by stab inoculating conidial suspensions onto MEA plates supplemented with the nutritional requirements of the component haploid strains from which the diploid was constructed, and containing benomyl ($0.7\mu gml^{-1}$) which produces greatly enhanced mitotic segregation (Hastie, 1970). Haploid segregants were purified by streaking conidiospores on MEA plates and then tested on appropriate media for the segregation of the genetic markers in the diploid.

2.1.6 Screening for β-galactosidase production.

Two colony staining techniques were developed to detect β-galactosidase production by A.*nidulans* strains.

In the first method conidia were stab inoculated onto MA glucose plates supplemented with the necessary nutrients and brought to a pH of 6.0 by 0.1M citric acid/ 0.2M Na_2HPO_4. The media also contained the chromogenic substrate X-gal at a final concentration of $40\mu gml^{-1}$. The plates were incubated at 37^oC for 1 day, after which clean sterile 9cm filter paper discs were applied to the surface of the media and the plates were incubated at 37^oC for a further day. Subsequently, the filters were stripped from the surface of the plates, so removing conidia which would otherwise hinder

observation, and the plates (or the face of the filters that had been in contact with the medium) were examined. The intensity of blue staining was assumed to indicate the degree of β-galactosidase production by *A.nidulans* strains.

In the second method (Fantes and Roberts, 1973) conidia were stab inoculated onto MA glucose plates supplemented with the necessary nutrients and incubated at 37^0C for 2 days, with clean sterile 9cm filter paper discs being applied after the first day as above. Once the filters had been removed a 0.5M solution of sodium phosphate (pH8.3) containing 5% (v/v) of a methanolic solution of $8mgml^{-1}$ 6-bromo-β-naphthyl-β-D-galactopyranoside was poured onto each plate, and the plates left at 37^0C for 2 to 4 hours. Excess substrate solution was then poured off, 4mls of a 0.5% (w/v) aqueous solution of Fast Blue B salt added, and the plates left at room temperature for 10 minutes. The degree of β-galactosidase activity is indicated by the intensity of purple staining around the colonies.

2.1.7 Preparation of *A.nidulans* cell-free extracts for enzyme assays.

Liquid minimal media was inoculated with a suspension of conidiospores to yield 10^6 spores per ml, and 400ml cultures were typically grown in 2l indented flasks at 37^oC on rotary shakers for 15 to 24 hrs depending upon the carbon source. Mycelium was harvested by filtration on Whatman No.1 filters, and washed first with water, and then with the relevant extraction buffer. These are 0.1M triethanolamine pH 7.5, 1mM DTT, 1mM EDTA for phosphoglycerate kinase assays, and 50mM sodium phosphate pH 7.6, 1mM $MgCl_2$ for β-galactosidase assays. The thin pad was then immersed in liquid nitrogen in a pre-cooled mortar and ground to a fine powder with a cold pestle. The frozen powdered mycelium was resuspended in 1ml extraction buffer, allowed to thaw and extracted by gently shaking the slurry in a polypropylene tube on ice for about 30 minutes, after which time cell debris was removed by centrifugation for 10 minutes in an Eppendorf microfuge at 4^oC. The supernatant was decanted,

stored on ice, diluted x10 in extraction buffer, and the enzyme assayed within a maximun of 5 hours after extraction.

2.1.8 Phosphoglycerate kinase assays.

The activity of 3-phosphoglycerate kinase was assayed spectrophotometrically by the decrease in optical density at 366nm in a system coupled with D-glyceraldehyde-3-phosphate dehydrogenase (GAPD) using the procedure described by Kulbe and Bojanovski (1982). Reagents were added to a quartz cuvette in the following order:

Stock solution	Amount added(μl)	Final concentration
100mM Triethanolamine pH7.5	925	92.5mM
150mM $MgSO_4$	10	1.5mM
65mM EDTA	10	0.65mM
125mM ATP	10	1.25mM
27mM NADH	10	0.27mM
1Uμl^{-1} GAPD in water	5	5 Units
A.nidulans extract (dil 1/10)	5-20	-
Distilled water to 1ml		

The reaction mix was monitored for about 30 seconds to obtain a base rate before the addition of 20μl of 175mM 3-phospho-D-glycerate to the sample cuvette to give a final concentration of 3.5mM. The reactions were carried out at room temperature and it was determined that the rate of reaction was proportional to the amount of cell free extract added.

The concentration of soluble protein was assayed in each cell-free extract by the method of Bradford (1976). The presence of EDTA in the system would interfer with the Lowry method. A calibration curve was constructed using bovine serum albumin standards each time protein estimations were performed.

One unit of enzyme activity is defined as the amount of the enzyme required for the formation of 1μmole of D-glyceraldehyde-3-phosphate and NAD^+ per minute. This is calculated using the molar extinction coefficient of NADH at 366nm which is $3.4 \times 10^3 M^{-1} cm^{-1}$. Specific activity is expressed as units per mg protein.

2.1.9 β-Galactosidase enzyme assays.

The activity of β-galactosidase was assayed spectrophotometrically by the increase in optical density at 420nm with o-nitrophenyl-β-D-galactopyranoside (ONPG) as substrate, after the method of Wallenfels (1962). The assays were performed at a constant temperature of 25°C. Reagents were added to a quartz cuvette in the following order:

Stock solution	Amount added(μl)	Final concentration
100mM sodium phosphate pH7.6	500	50mM
10mM $MgCl_2$	100	1mM
A.nidulans extract (dil 1/10)	25-100	-
Distilled water to 0.9ml		

The reaction mix was monitored for about 30 seconds to obtain a base rate before the addition of 100μl of 10mM ONPG to give a final concentration of 1mM. It was determined that the reaction rate recorded was proportional to the amount of cell free extract added.

The concentration of soluble protein was assayed in each cell free extract by the method of Lowry *et al.* (1951), and a calbration curve was constructed using bovine serum albumin standards each time protein estimations were performed.

For the purposes of this study, where relative activities were required, it was convenient to define units of β-galactosidase acivity in the manner of Fantes (1972).

One unit of β-galactosidase activity produced a change in optical density at 420nm of 0.001 per minute. The specific activity is defined as the acitivity per mg of protein assayed. That is, specific activity = $\Delta OD_{420} x 10^3 min^{-1} mg^{-1}$. One unit of activity is equivalent to a rate of hydrolysis of ONPG of $0.188 nmol min^{-1}$.

2.1.10 DNA mediated transformation of *A.nidulans*.

General aspects of *A.nidulans* transformation are discussed in Chapter 1, section 3, and the transformation procedure used throughout was adapted from that of Ballance *et al.* (1983).

Cellophane discs the size of a 9cm petri dish base and sterilzed by autoclaving, were placed on the surface of 20 appropriately supplemented MEA glucose plates, and approximately 10^6 conidiospores were evenly dispersed over each plate using a sterile glass spreader. Cultures were incubated for 15 hours at 37°C, after which the cellophane discs bearing young mycelium were divided equally between 4 clean petri dishes each containing a 15ml solution of the cell wall dissolving enzyme complex Novozym 234 ($5 mg ml^{-1}$) in an osmotic stabilizer, 0.6M KCl, and placed at 30°C on a slow rotary shaker for about 1 hour. The cellophane discs were then removed from the resulting protoplast suspension and rinsed in 15ml of 0.6M KCl in a petri dish to remove any adhering protoplasts. The suspensions of protoplasts were passed through a nylon filter (Gallenkamp GMX-500-V) and a sintered glass filter (porosity 1) to remove cell debris. The protoplasts were pelleted by centrifugation at room temperature at 1500xg for 5 minutes and washed three times, twice with 0.6M KCl and once with 0.6M KCl, 50mM $CaCl_2$. The protoplasts were resuspended in 500μl of 0.6M KCl, 50mM $CaCl_2$, and the protoplast concentration was estimated by making serial dilutions and counting the protoplasts present using a haemocytometer examined under a light microscope. The volume of the protoplast suspension was then adjusted to give $1-5x10^7$ protoplasts per ml. Not more than 20μg of DNA in no more than 40μl of TE buffer pH8, was added to 200μl

aliquots of the protoplast suspension in 20ml sterile disposable universal tubes, followed by the addition of 50μl of 25% PEG6000, 50mM $CaCl_2$, 10mM Tris.HCl, pH7.5. After 20 minutes on ice a further 2ml of the PEG solution was added and the mixture left at room temperature for 5 minutes. Following the addition of 4ml 0.6M KCl, 50mM $CaCl_2$, the total transformation mixture was added to 100ml of apropriately supplemented molten minimal medium containing 0.6M KCl and 2% (w/v) agar and held at 48^{o}C. This molten agar was poured as a top layer onto similarly supplemented minimal medium plates containing 0.6M KCl and 1.5% (w/v) agar. When selecting for uracil independent transformants, plates were incubated at 37^{o}C for 2 to 3 days, and when selecting for transformants growing on quinic acid, plates were incubated at 37^{o}C for 3 to 4 days.

The efficiency of protoplast regeneration was assessed by adding 100μl aliquots of 10^{-2}, 10^{-3} and 10^{-4} dilutions of the final transformation mixture (including the PEG solution and the KCl/$CaCl_2$ solution) to molten complete medium containing 0.6M KCl and 2% (w/v) agar, pouring as a top layer onto similar complete medium plates, and comparing the number of colonies obtained after 48 hours incubation at 37^{o}C with the previous visual estimation of the number of protoplasts present.

All transformants obtained were replated by spreading dilute suspensions of conidiospores for single colony isolation on selective medium. This procedure was generally repeated at least once, and the strains isolated stored by transferring conidiospores from a single colony to selective media slants to guard against the loss of transforming DNA integrated into the genome.

2.2 Genetic materials and methods associated with Escherichia coli.

2.2.1 *E.coli* strains.

The *E.coli* strain DH5 (*supE44, hsdR17, recA1, endA1, gyrA96, thi-1, relA1*; Hanahan, 1983) was used throughout as a host for the construction and maintenance of recombinant plasmid vectors. Strain JM83 (*ara, Δlac-pro, strA, thi, Φ80dlacZΔM15*; Vieira and Messing, 1982) was used to test for the expression of *lacZ* fusions in *E.coli*, and strain JM101 (*supE, thi, Δlac-proAB, [F' traD36, proAB*$^{+}$*, lacI*^{q}*lacZΔM15]* Messing, 1983) for propogation of recombinant M13 phage for DNA sequencing.

2.2.2 Media.

The growth media used are given below showing the final concentration of components.

Luria Broth.

Tryptone	1.0% (w/v)
Yeast extract	0.5% (w/v)
NaCl	0.5% (w/v)

M9 Minimal Medium.

Na_2HPO_4	0.6% (w/v)
KH_2PO_4	0.3% (w/v)
NH_4Cl	0.1% (w/v)
Glucose (autoclaved separately)	0.4% (w/v)
$MgSO_4$ (autoclaved separately)	10mM
$CaCl_2$ (autoclaved separately)	1mM
As required:	
Thiamine (filter sterilized)	$10\mu gml^{-1}$
Amino acids	$40\mu gml^{-1}$

ψb medium (for DNA mediated transformation).

Bacto yeast extract	0.05% (w/v)
Bacto tryptone	0.02% (w/v)
$MgSO_4.7H_2O$	0.02% (w/v)

pH7.6 with KOH.

Agar plates were prepared by solidifying liquid media with 1.5% (w/v) Difco Bacto agar, and soft agar was prepared by solidifying liquid media with 0.6% (w/v) Difco Bacto agar. The antibiotics ampicillin and chloramphenicol were added to media as required to final concentrations of $25\mu gml^{-1}$. Ampicillin stock solutions were filter sterilized, and chloramphenicol solutions were prepared in ethanol.

2.2.3 Growth and storage of strains.

Plasmid bearing strains were grown in and maintained on media containing the appropriate antibiotic and were incubated at 37°C. Cultures were maintained on Luria agar plates at 4°C when in current use, and frozen at -20°C in Luria broth containing 20% (v/v) glycerol for long term storage.

2.2.4 Plasmids and phage.

The origin of plasmids used for the construction of an *A.nidulans pgk* gene disruption vector, and *pgk-lacZ* fusion vectors are given below.

Plasmid	Reference
pPGK2	Clements and Roberts, 1985
pDJB1	Ballance *et al.*, 1983
pAL3.3-B3, -B4	Hawkins *et al.*, 1985
pAN923-41A, -42A, -43A	van Gorcom *et al.*, 1986

The phage λ clone λPGK3 (Clements and Roberts, 1985) was also used for the construction of a *pgk-lacZ* fusion vector. DNA fragments were subcloned into the single stranded phage M13mp18 (Messing, 1983) so that their seqeunce could be determined.

2.2.5 DNA mediated transformation of *E.coli*.

During the construction of recombinant plasmid vectors, competent cells were prepared from the *E.coli* strain DH5 and were transformed according to the method of Hanahan (1983). In order to prepare competent cells, a single isolated colony of DH5 was used to inoculate 10ml of ψb medium (see above) and incubated at 37°C for 3 hours with vigourous shaking. 1ml of this culture was used to inoculate 100ml of ψb medium, which was incubated at 37°C with vigourous shaking until the culture reached an optical density at 550nm of about 0.5. The culture was chilled on ice for 5 minutes, dispensed into 50ml sterile polypropylene tubes and the cells pelleted at 3.5K, 15 minutes at 4°C. The cells were then resuspended in two-fifths of the original volume of TfbI solution (30mM potassium acetate, 100mM $RbCl_2$, 10mM $CaCl_2$, 50mM $MnCl_2$, 15% glycerol (v/v), pH5.8 with 0.2M acetic acid), left on ice for 5 minutes, and the cells again recovered by centrifugation as above. The cells were gently resuspended in one-fifteenth the original volume of TfbII solution (10mM MOPS, 75mM $CaCl_2$, 10mM $RbCl_2$, 15% glycerol (v/v), pH6.5 with KOH), left on ice for 15 minutes, aliquoted into 100μl samples and frozen in liquid nitrogen and stored at -70°C.

Prior to transforming the cells, they were thawed and then left on ice for 10 minutes. Up to 50ng of DNA in a 40μl volume was added to the cells, and the mixture left on ice for 30 minutes. The cells were then heat shocked at 42°C for 90 seconds and returned to ice for 90 seconds. Four volumes of ψb medium was added, and the cells incubated at 37°C for 50 minutes, after which they were spread plated onto Luria agar plates with appropriate antibiotics and incubated at 37°C overnight.

When testing for the expression of *pgk-lacZ* fusions in *E.coli*, the strain JM83 was transformed by the method of electroporation (Dower *et al.*, 1988). 1ml of an overnight culture of JM83 was used to inoculate 300ml Luria broth which was incubated at 37°C with vigourous shaking until the culture reached an optical density of about 0.6 at 600nm. The culture was chilled on ice and four serial washes conducted in decreasing volumes of ice-cold sterile distilled water until the cells were finally resuspended in a volume of 500μl. 1ng of DNA in a volume of 5μl water was added to 40μl of cells in an electroporation cuvette and a pulse (1.5kV, 25μF) delivered, immediately after which 1ml of the regenerating medium, SOC (2% tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 20mM $MgCl_2$, 20mM glucose), was added to the cells. The cells were then incubated at 37°C for 60 minutes with vigourous shaking, after which they were spread plated onto Luria agar with appropriate antibiotics and 40μgml^{-1} of the chromogenic substrate X-gal, and incubated at 37°C overnight.

2.3 Recovery of nucleic acids.

Proteins were extracted from nucleic acid solutions by the addition of an equal volume of phenol reagent (100g phenol, 100ml chloroform, 4ml isoamyl alcohol, 0.1g 8-hydroxyquinoline) saturated with 10mM Tris.HCl, pH7.5, agitation on a Vortex mixer for 1 minute and brief centrifugation. The upper (aqueous) phase was collected taking care not to disturb the interphase.

Nucleic acid solutions were concentrated after the addition of 0.1 volumes of 3M sodium acetate pH5.2, by precipitation with 2 volumes of cold ethanol. The solution was mixed and chilled in a methanol/dry ice bath for 10 minutes or at $-20^{o}C$ overnight. The precipitated nucleic acid was collected by centrifugation at 12000xg for 15 minutes at $4^{o}C$, rinsed in cold 70% (v/v) ethanol, dried under a vacuum and resuspended in the appropriate volume of sterile distilled water or TE buffer (10mM Tris.HCl, pH 8.0, 1mM EDTA).

2.4 DNA preparation.

2.4.1 Small scale preparation of plasmid DNA from *E.coli*.

The method used followed a modification of the procedure of Birnboim and Doly (1979). Overnight cultures of the plasmid carrying strain of *E.coli* were grown in Luria broth containing the appropriate antibiotic. A 1.5ml sample in an Eppendorf tube was centrifuged for 1 minute to pellet the cells, 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, 4mgml^{-1} freshly added lysozyme) and held at room temperature for 5 minutes, after which 200μl of a freshly prepared solution of 0.2M NaOH, 1% SDS (w/v) was added. The solutions were mixed by inverting the tube gently several times and transferred to ice for 5 minutes. 150μl of 3M potassium acetate (pH4.8 with glacial acetic acid) was then added and mixed. The tube was placed on ice for 10 minutes, and then centrifuged for 5 minutes in a Eppendorf centrifuge in a 4^{o}C cold room. The supernatant was transferred to a new Eppendorf tube, phenol-chloroform extracted and the DNA precipitated with ethanol at room temperature for 2 minutes. The sample was resuspended in 50μl TE buffer containing 20μgml^{-1} of DNAase-free RNAase.

2.4.2 Large scale preparation of plasmid DNA from *E.coli*.

The method used for the large scale preparation of plasmid DNA was essentially a scaled up version of the method of Birboim and Doly (1979), with the additional steps of chloramphenicol induced amplification of plasmid copy number, and purification by caesium chloride gradient centrifugation.

Initially, 25ml of Luria broth containing the appropriate antibiotic was inoculated with 0.1ml of an overnight culture of an *E.coli* strain carrying the required plasmid, and was incubated at 37^{o}C with vigourous shaking until the culture reached an optical density of about 0.6 at

600nm. The whole culture was then used to inoculate 500ml Luria broth (with appropriate antibiotic), and this was then incubated at 37°C with vigourous shaking until the culture reached an optical density of 0.4 at 600nm, after which chloramphenicol was added to a final concentration of 170μgml^{-1}, and the culture incubated for another 12 to 16 hours.

The cells were pelleted by centrifugation at 4000xg for 5 minutes at 4°C, washed in 100ml of an ice-cold solution of STE (0.1M NaCl, 10mM Tris-HCl pH8, 1mM EDTA), resuspended in 8ml of ice-cold lysis buffer (50mM glucose, 10mM EDTA, 25mM Tris-HCl pH8.0, 4mgml^{-1} of freshly added lysozyme) and held on ice for 5 minutes. To this solution 16ml of a freshly prepared solution of 0.2M NaOH, 1% SDS (w/v) was added, the solution gently mixed and held on ice for 10 minutes. Then 12ml of ice-cold potassium acetate (pH4.8 with glacial acetic acid) was added, mixed and left on ice for 10 minutes to precipitate the chromosomal DNA and proteins. This precipitate was removed by centrifugation at 15000xg for 20 minutes at 4°C, and remaining nucleic acids were then precipitated by the addition of a 0.6 volume of propan-2-ol, mixing and standing at room temperature for 15 minutes. The precipitate was collected by centrifugation at 12000xg for 30 minutes at room temperature, and the pellet was rinsed with cold 70% ethanol, dried under vacuum and finally resuspended in 8ml TE buffer.

Eight grams of AR grade caesium chloride was dissolved in the above nucleic acid solution, and 0.64ml of a 10mgml^{-1} solution of ethidium bromide added to give a concentration of 600μgml^{-1}. The solution was brought to a refractive index of 1.386 by the addition of further CsCl. The solution was transferred to a Beckman polyallomer "Quickseal" tube, which was heat sealed and centrifuged at 20°C at 55000rpm overnight in a Beckman Type-65 rotor. The plasmid band was collected, 3 volumes of TE added, and the DNA precipitated with 0.1 volumes 4M sodium acetate pH6 and 0.6 volumes propan-2-ol. The plasmid DNA was then resuspended in 500μl of TE buffer, phenol-chloroform

extracted and the DNA precipitated with ethanol and resuspended in TE buffer.

The concentration and quality of plasmid preparations was determined by optical density at 260nm and 280nm. An O.D. of 1 at 260nm corresponds to approximately 50μg ml^{-1} DNA, and a 260nm to 280nm ratio of 1.8 indicates a pure preparation of DNA (Maniatis *et al.*, 1982).

2.4.3 Small scale preparation of chromosomal DNA from *A.nidulans.*

Chromosomal DNA was prepared from a dense suspension of conidiospores by a method adapted by Raj Beri from Morris (1978). Conidiospores were harvested from the surface of a single MEA plate as described above (Section 2.1.4). The conidiospores were washed by two serial centrifugations in cold sterile spermidine lysis buffer (5mM spermidine, 100mM KCl, 10mM EDTA, 10mM Tris-HCl pH 7.5, 250mM sucrose), and were finally resuspended in 0.8ml of the same buffer. The conidiospores were then vigorously agitated on a Vortex mixer for about 2 minutes with 3g acid washed sterile glass beads (0.45mm) in a glass McCartney bottle, when the formation of empty conidia could be seen under a light microscope. A further 0.8ml of spermidine lysis buffer was then added to the preparation of "ghost cells" and the supernatants dispensed into two Eppendorf tubes, and each incubated at 60^{o}C for 20 minutes with 67μl of a 10% solution of SDS (w/v). Next, 222μl of 4M sodium acetate pH6 was added to each tube, mixed and held on ice for 30 minutes, prior to centrifugation at 12000xg in a microfuge at room temperature for 5 minutes and recovery of the supernatants. These supernatants were then incubated with proteinase K at a final concentration of 200μgml^{-1} and DNAase-free RNAase at a final concentration of 100μgml^{-1} at 37^{o}C for 1 hour. The DNA was then precipitated by the addition of 0.1 volumes of 4M sodium acetate pH6 and 0.6 volumes of propan-2-ol, and collected by centrifugation at 12000xg in an Eppendorf microfuge for 5 minutes. The pellets were rinsed with 70% ethanol, dried under vacuum and pooled by resuspending in

400μl of TE buffer. The solution was then extracted with phenol-chloroform and the DNA precipitated as before with sodium acetate and propan-2-ol, washed with 70% ethanol and resuspended in 50μl TE buffer.

DNA preparations were stored at -20^{o}C and typically yielded 20 to 50μg of high molecular weight genomic DNA.

2.5 Analysis of nucleic acids by agarose gel electrophoresis.

Horizontal agarose gels were prepared and run in electrophoresis buffer (40mM Tris base, 5mM sodium acetate, 2mM EDTA; pH8 with glacial acetic acid) containing ethidium bromide at 0.5μgml^{-1}. The concentration of the agarose used was in the range 0.4 to 2.0% (w/v) depending upon the size of the DNA fragments to be separated.

Molecular weight marker DNA standards were provided by *HindIII* digestion of λ DNA. DNA samples were mixed with 0.1 volumes of loading buffer (25% Ficoll, 0.001% Orange C, 200mM EDTA) and electrophoresis continued until the orange dye had travelled to almost the end of the gel.

DNA binding ethidium bromide in agarose gels was visualized using a short wavelength ultra-violet transilluminator, and photographed using a polaroid MP-3 camera and a Kodak Tmax 100 film (10.2 x 12.7cm).

2.6 DNA manipulation techniques.

2.6.1 Digestion of DNA with restriction enzymes.

DNA samples were routinely incubated with a 2 to 5 fold excess of restriction enzymes (manufacturer's definitions) to ensure total digestion of DNA. Typical reaction mixtures (10μl to 20μl total volume) consisted of 0.2 to 1μg DNA, the appropriate restriction enzyme buffer (manufacturer's conditions), and were incubated at 37^{o}C for a minimum of 1 hour. When digesting *A.nidulans* genomic DNA, a second aliquot of restriction enzyme was subsequently added, and the reaction incubated at 37^{o}C for at least a further hour.

2.6.2 Digestion of DNA with Bal31 exonuclease.

Bal31 exonuclease treatment of a plasmid was achieved by linearizing the plasmid DNA with an appropriate restriction endonuclease, and the DNA reisolated by phenol-chloroform extraction, ethanol precipitation and resuspension of the DNA in 50μl of 500μgml^{-1} bovine serum albumin solution. The total reaction volume was made up to 100μl by adding 50μl of 2x reaction buffer (24mM $CaCl_2$, 24mM $MgCl_2$, 0.4M NaCl, 40mM Tris-HCl pH8, 2mM EDTA) and incubated at 30^{o}C for 3 mintues , after which 0.5 Units of Bal31 was added and samples withdrawn from 30^{o}C incubation at one minute intervals. The reactions were terminated by adding EGTA to a final concentration of 20mM and placing the tube on ice.

2.6.3 Recovery of DNA fragments from agarose gels.

The method used is based on that described by Maniatis *et al.* (1982). Electrophoresis of the agarose gel was in the dark, and the position of the desired fragments visualized with a portable long wave length ultra violet lamp. Following electrophoresis, the gel was sliced laterally alongside the fragment to be recovered and also longitudinally from the ends of the first cut at the lower molecular weight end of the gel. A piece of dialysis tubing

that had been cut 5mm wider than the lane or trough containing the desired fragment and boiled for 2 minutes in TE buffer was inserted into the lateral cut. The DNA fragment was electrophoresed onto the membrane at 100V, and without switching off the current, the membrane was quickly transferred into an Eppendorf tube. One corner of the membrane was trapped by the cap of the tube and the DNA collected by centrifugation for 1 minute in an Eppendorf microfuge. The membrane was washed with TE buffer to collect further DNA and spun again, after which the DNA was ethanol precipitated and resuspended in the required volume of TE buffer.

2.6.4 Fill-in of recessed ends of DNA fragments.

Filling in of 3′ recessed ends of DNA fragments derived by restriction endonuclease digestion was achieved by incubating the recovered fragment in 18μl fill-in buffer (10mM Tris-HCl pH 7.5, 10mM $MgCl_2$, 4mM spermidine, 5mM DTT) with 2μl of Chase mix (10mM Tris-HCl pH8.0, 0.1mM EDTA, 0.25mM dATP, 0.25mM dGTP, 0.25mM dCTP, 0.25mM dTTP) and with 1 Unit of the Klenow fragment of DNA polymerase I at 37^{o}C for 30 minutes. The reaction mixture was then extracted with phenol-chloroform, and the DNA precipitated with ethanol and resuspended in the required volume of TE buffer.

2.6.5 Ligation of DNA fragments.

Double stranded restriction endonuclease fragments recovered from agarose gels were ligated to similarly recovered plasmid DNA molecules in a 10 to 20μl reaction volume. When DNA fragments had been generated by digestion of plasmids with two different restriction endonucleases the ratio of insert to vector DNA molecules was equimolar. However, a five fold excess of insert was usually used when the DNA fragment had been generated by a single restriction endonuclease. This alleviated the problem of recircularization of the vector molecule. The total concentration of vector plus insert DNA was usually 1 to 20ngμl^{-1}, and this was incubated in ligation buffer

(manufacturer's conditions) overnight at 15°C after the addition of 1 Unit of T4 DNA ligase.

Calf intestinal alkaline phosphatase was also used in some cases to reduce the degree of recircularization of the vector molecule. The endonuclease digested plasmid DNA molecule recovered from an agarose gel was incubated in dephosphorylation buffer (10mM $ZnCl_2$, 10mM $MgCl_2$, 100mM Tris-HCl pH8) with the appropriate amount of calf intestinal phosphatase (manufacturer's conditions) for 30 minutes at 37°C. Subsequently SDS and EDTA were added to final concentrations of 0.5% and 5mM respectively and proteinase K to a final concentration of $100\mu g ml^{-1}$. The mixture was incubated for a further 30 minutes at 56°C in order to digest the phosphatase. Prior to conducting ligation reactions, DNA was recovered from the reaction mixture by phenol-chloroform extraction and ethanol precipitation, and was resuspended in the appropriate volume of TE buffer.

2.6.6 Phosphorylation of DNA linker molecules.

DNA linker molecules were phosphorylated prior to ligation to a DNA fragment by incubating 1 to $2\mu g$ of linker DNA in linker-kinase buffer (66mM Tris-HCl pH7.6, 1mM ATP, 1mM spermidine, 10mM $MgCl_2$, 15mM DTT, 0.2mg ml^{-1} BSA) with 2 Units of T4 DNA kinase at 37°C for 1 hour.

2.7 Analysis of DNA by transfer onto nylon membranes and hybridization using radioactive probes.

2.7.1 Southern blot filter hybridization.

The method used to transfer DNA from an agarose gel to a membrane for hybridization is based on that of Southern (1975). Following agarose gel electrophoresis, non-essential areas of the gel were removed and the 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, 1.5M NaCl to neutralize the gel. A plate was laid across a tray filled with 10x SSC (1.5M Na citrate, 0.15M NaCl), such that the plate was 2 to 3cm above the surface of the liquid. A strip of Whatman 3MM paper was placed over the plate and in contact with the 10x SSC in the tray so as to act as a wick for the uptake of the salt solution. The gel was inverted and laid on the wick, taking care to avoid trapping air bubbles between the gel and the 3MM paper. First a Hybond-N nylon filter and then 2 pieces of 3MM paper cut to size slightly larger than the gel, were soaked in 2x SSC and laid on top of the gel, again removing any air bubbles between the gel and the nylon filter. A stack of paper towels was placed above the 3MM paper, covered with a glass plate, and compressed with a 500g weight. Transfer of DNA was continued for at least 6 hours, replacing the stack of paper towels at intervals. The positions of the loading slots in the gel were marked on the filter, and the filter was then washed in 6x SSC, blotted dry on 3MM paper, wrapped in Saran wrap and placed on a short wavelength ultra-violet transilluminator for 30 seconds to bind the single stranded DNA to the filter.

2.7.2 Preparation of genomic DNA for the estimation of fragment copy number by "dot-blot" analysis.

An equal volume of small scale preparations of genomic DNA in TE buffer from each *A.nidulans* strain was added to an equivalent amount of 2x denaturation solution (2M NaCl, 0.2M

NaOH, 20mM EDTA). A piece of Hybond-N nylon filter was placed on impervious paper and an aliquot of each denatured DNA preparation pipetted onto the filter to form dots of approximately equal size. After the dots had dried at room temperature the filter was washed in 6x SSC, blotted dry, wrapped in Saran wrap, and the DNA bound to the filter by exposure to short wavelength ultra violet as for Southern blots above.

2.7.3 Oligolabelling of DNA fragments to be used as probes.

The method used to radioactively label DNA fragments was based on that of Feinberg and Vogelstein (1984). The DNA fragment to be used as probe was denatured by heating at 100°C for 5 minutes directly before use. The following reagents were then added to a screw capped Eppendorf tube in this order:

Distilled water	(to give a total volume of 25μl)
OLB buffer	5μl
BSA (10mgml^{-1})	1μl
DNA fragment	1 to 50ng in up to 16μl
α-^{32}P-dCTP (10mCiml^{-1})	2.5μl
Klenow fragment	2 Units (0.5μl)

The reaction mixture was either incubated at 37°C for 30 minutes or at room temperature for at least 5 hours, after which the reaction was stopped by the addition of 100μl of stop buffer (20mM NaCl, 20mM Tris-HCl pH 7.5, 2mM EDTA, 0.25% SDS (w/v), 1μM dCTP).

OLB is made up from three solutions:

2 volumes solution A:	1.25M Tris-HCl pH8, 125mM $MgCl_2$, 0.18% (v/v) β-mercaptoethanol, 0.5mM dATP, 0.5mM dGTP, 0.5mM dTTP
5 volumes solution B:	2M HEPES (pH6.6 with NaOH)
3 volumes solution C:	Hexadeoxynucleotides suspended in 3mM Tris-HCl pH7.7, 0.2mM EDTA at 90 OD units per ml

2.7.4 DNA hybridization of Southern blot and dot blot filter membranes.

The filters were cut into convenient sized strips and all reactions carried out in hybridization boxes made of perspex and fitted with a rubber seal. The filters were prehybridized by incubation with gentle agitation at 65°C for at least 1 hour in 25ml FHM solution (15mM Na_2H/NaH_2PO_4 pH7, 0.27M NaCl, 1.5mM EDTA, 0.5% (w/v) Marvel milk, 1% (w/v) SDS, 6% (w/v) PEG 6000). The radioactively labelled probe solution was heated at 100°C for 5 minutes to denature the DNA and then added directly to the prehybridization mix. Hybridizations were generally conducted with gentle agitation at 65°C overnight.

Unbound labelled DNA was removed from the filters by washing twice in 2x SSC, 0.5% (w/v) SDS for 5 minutes at 65°C, and then four times in 0.1x SSC, 0.5% (w/v) SDS for 30 minutes at 65°C. If detectable levels of ^{32}P were still present in the final wash solution, further washes at the latter, higher stringency were conducted. The filters were then dried at room temperature, mounted onto 3MM paper and covered with Saran wrap.

2.7.5 Autoradiography.

The filters were placed in a cassette against FUJI RX-100 X-ray film backed by an intensifying screen, and the film was exposed at -80°C for as long as necessary to obtain a clear signal.

2.7.6 Stripping DNA probe from the filter membrane.

When it was necessary to hybridize DNA on a filter membrane with a second DNA probe, the first probe was removed by incubation for 30 minutes at 45°C in 0.4M NaOH, followed by a similar incubation in 0.1x SSC, 0.1% (w/v) SDS, 0.2M Tris-HCl pH7.5. Removal of radioactive isotope was confirmed by autoradiography.

2.7.7 Quantification of radioactively labelled probe DNA hybridized to DNA on filter membranes by radioisotope scintillation counting.

The binding of DNA probe to genomic DNA present on a filter was quantified by excising specific regions of the filter, previously identified by autoradiography, and recording α-particle decay using a TRI-CARB 2000 CA liquid scintillation analyser. The portions of the filter were placed in scintillation vials and kept beneath scintillation fluid, and counts were recorded over the range 5 to 1700 keV for 10 minutes.

2.8 Preparation of *A.nidulans* RNA.

The method used for RNA extraction was based on that of Chomczynski and Sacchi (1987).

In work with RNA all solutions and plastic disposable tubes and tips where possible were treated with 0.1% diethylpyrocarbonate for 12 hours before autoclaving. All glassware and ceramics were baked for 4 hours at 250°C before use.

A.nidulans total cellular RNA was prepared from young (15 hour) growing mycelium, which was harvested and immediately washed in ice cold water, rapidly frozen in liquid nitrogen and powdered in a previously heat treated pestle and mortar under liquid nitrogen.

For each 0.1g (wet weight) mycelium, the powdered fragments were resuspended in 1ml of extraction buffer (4M guanidinium thiocyanate, 25mM sodium citrate pH7, 0.5% (w/v) sarkosyl, 0.1M β-mercaptoethanol), and 0.1 volumes of 2M Na acetate pH4 was added. Immediately, 1ml phenol and 0.2ml chloroform were added, the mixture shaken vigorously and held on ice for 15 minutes. The aqueous phase was recovered by centrifugation in polypropylene tubes for 10 minutes at 10000rpm, and RNA precipitated by adding 0.6 volumes of isopropanol, holding at -20°C for 1 hour and repeating the spin. The pellet so obtained was then resuspended in 0.3ml extraction buffer, precipitated again with isopropanol, ethanol washed and dried, and resuspended in water to give a $1mgml^{-1}$ solution, which was stored at -20°C.

For long term storage, 2 volumes of ethanol and 0.1 volumes of 3M sodium acetate pH5.2 were added to the RNA solution, which was stored at -20°C. The RNA precipitate was recovered by centrifugation and was resuspended in water as required.

2.9 Analysis of RNA by transfer on to nitrocellulose membranes and hybridization using radioactive probes.

2.9.1 Northern blot filter hybridization.

To 4.8μl of 1mgml^{-1} RNA the following reagents were added in the order given: 10μl deionized formamide, 2μl of 10x MOPS buffer (0.2M MOPS, 0.05M Na acetate, 0.01M EDTA, pH7) and 3.2μl formaldehyde. The sample was incubated at 65^{o}C for 5 minutes, immediately cooled on ice, and 0.1 volumes loading buffer (25% Ficoll, 0.001% Orange C, 200mM EDTA) and 1μl of 1mgml^{-1} ethidium bromide were added. The samples were then loaded onto a 1% agarose gel made by dissolving the agarose in MOPS buffer containing 17.9% formaldehyde. Electrophoresis was at 100V for about 3 hours with circulating buffer, after which the RNA was immediately transferred onto nitrocellulose in precisely the same way as for DNA onto Hybond-N (Section 2.7.1.), except that the transfer solution was 20x SSC. Following transfer, the RNA was bound to the filter by baking the filter between sheets of 3MM paper at 80^{o}C for 2 hours.

To provide a molecular weight marker for the gel, 10μg of phage λ DNA digested with the restriction endonuclease *HindIII* was denatured by incubating for 5 minutes at 65^{o}C in 0.15M NaOH, 0.01M EDTA pH8. Loading buffer and ethidium bromide were added to the single stranded DNA as for the RNA samples, and the marker was loaded in a lane set well to one side of the gel. Following electrophoresis the portion of the gel with marker was excised and washed in four changes of water and two changes of 0.1M ammonium acetate. The single stranded DNA was then stained with ethidium bromide by gently shaking the gel fragment for 1 hour in a solution of 0.5μgml^{-1} ethidium bromide in 0.1M ammonium acetate /0.01M β-mercaptoethanol, and then partially destained in the same buffer lacking ethidium bromide. The marker lane was visualized under short wavelength ultra-violet illumination and photographed.

2.9.2 Preparation of RNA for quantification by "dot-blot" analysis.

Serial dilutions of RNA preparations were made in water (diethylpyrocarbonate treated), and 5μl aliquots spotted onto nitrocellulose paper which had been pre-washed in 20x SSC. The RNA spots were allowed to dry at room temperature before being fixed to the filter by baking at 80^{o}C for 2 hours.

2.9.3 Hybridization of RNA Northern blot and RNA dot blot membranes.

Prior to hybridizing RNA to single stranded DNA probe prepared as described in Section 2.7.3., the filters were incubated with gentle agitation at 42^{o}C for about 6 hours in 20mls of a solution of 50% formamide, 5x SSC, 50mM Na phosphate pH6.5, 0.02% BSA, 0.02% Ficoll, 0.02% PVP, 250μgml^{-1} salmon sperm DNA (freshly added). Radioactively labelled single stranded probe DNA (prepared as described in Section 2.7.3) was then added to the solution along with 5mls of 75% (w/v) PEG 6000, and RNA on the filters hybridized at 42^{o}C with gentle agitation overnight.

Unbound labelled DNA was removed from the filters by washing twice in 2x SSC, 0.5% (w/v) SDS for 5 minutes at 50^{o}C, and then four times in 0.1x SSC, 0.1% (w/v) SDS for 30 minutes at 50^{o}C. The filters were then dried at room temperature, mounted onto 3MM paper, covered with Saran wrap, and exposed to X-ray film as decribed for the autoradiography for Southern blot filters in Section 2.7.5.

When it was necessary to hybridize RNA on a filter membrane with a second DNA probe, the first probe was removed by incubation for 1 to 2 hours at 65^{o}C in 0.1X SSC, 0.1% (w/v) SDS. Removal of the first radioactively labelled probe was confirmed by autoradiography.

2.9.4 Densitometric analysis of autoradiographs of RNA filters.

Profiles were obtained of autoradiographic images of RNA filters by scanning the developed X-ray films with an LKB Ultrascan XL laser densitometer. In order to give a linear response to radioactive decay, the X-ray film was made sensitive by low level exposure to light prior to autoradiography.

2.10 The sequencing of DNA subcloned into M13 phage.

2.10.1 Preparation of replicative form M13 vector DNA.

Phage M13mp18 (Messing, 1983) was picked as a fresh plaque and used to inoculate 20mls of Luria broth seeded with one drop of an overnight culture of the *E.coli* strain JM101. The cells were grown with vigourous agitation at 37°C for 7 to 8 hours, pelleted by centrifugation, and the supernatant recovered. The supernatant was then used to infect a 2 litre culture of JM101 grown in Luria broth to an optical density of 0.5 at 550nm, and grown overnight at 37°C. The cells were recovered by centrifugation and replicative form DNA isolated as described in Section 2.4.2.

2.10.2 Ligation of DNA fragments into M13 and transformation of JM101.

Ligations were conducted as described in Section 2.6.5, after which the DNA was ethanol precipitated, using 20μg of tRNA as carrier to aid the precipitation, and resuspended in 10μl of water. The products of the ligation reactions were transformed into the *E.coli* strain JM101 by the method of electroporation, basically as described in Section 2.2.5 for the transformation of plasmid DNA into strain JM83. However, in this case following the addition of the regenerating medium SOC to the electroporated cells, 300μl of fresh cells were added along with 30μl of 25mgml^{-1} IPTG, 30μl of 25mgml^{-1} X-gal and 4.5ml of molten soft luria agar. The mixture was then poured onto luria agar plates and incubated at 37°C overnight. White and blue plaques developed overnight corresponding to recombinant and non-recombinant M13 transformants respectively.

2.10.3 Preparation of recombinant single-stranded M13 DNA sequencing templates.

Phage particles from white plaques were transferred with sterile toothpicks into 250μl aliquots of phage buffer (6mM Tris-HCl pH7.2, 10mM $MgSO_4$, 0.005% (w/v) gelatin), 50μl

of which were then used infect 2ml volumes of a JM101 culture (one drop of an overnight culture into 20mls of Luria broth). The cultures were incubated at 37°C with vigourous shaking for 6 to 8 hours, after which the cells were spun down by centrifugation in an Eppendorf microfuge for 5 minutes, and 1.2ml of the supernatant containing M13 phage was recovered. Phage particles were precipitated by the addition of 0.25 volumes of a solution of 20% (w/v) PEG 6000, 2.5M NaCl and holding at room temperature for 30 minutes, and phage was pelleted by centrifugation in an Eppendorf microfuge for 10 minutes. The supernatant was removed by aspiration and the phage pellet resuspended in 200μl of water. DNA was extracted from the virus particle, first by phenol and then by chloroform treatment, and then the DNA was recovered by ethanol precipitation and resuspended in 30μl of water.

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

Sequencing of M13 recombinant clones was based on the dideoxyribonucleotide chain termination method of Sanger *et al.* (1977), as modified by Biggin *et al.* (1983) and further adapted by A.Hawkins (Univ. of Newcastle-upon-Tyne). Typically, 5μl of sequencing template (Section 2.10.3) was annealed to the 17mer sequencing primer by incubating with 5μl of primer mix (1ngμl^{-1} primer in 20mM Tris-HCl pH8.5, 10mM $MgCl_2$) at 65°C for 1 hour. Either commercial universal primer was used, or a specific primer synthesized locally by the method of Matthes *et al.* (1984) with modifications to the wash cycle as described by Sproat and Gait (1985). Locally synthesized primers were ethanol precipitated twice before use.

For each template four tubes were prepared containing 2μl of the annealed mixture plus 2μl of either a G, A, T or C nucleotide mix shown below.

Nucleotide mixtures for sequencing reactions (μl volumes).

	G	A	T	C
0.5mM dGTP	25	500	500	500
0.5mM dTTP	500	500	25	500
0.5mM dCTP	500	500	500	25
10mM ddGTP	16	0	0	0
10mM ddATP	0	1	0	0
10mM ddTTP	0	0	50	0
10mM ddCTP	0	0	0	8
TE buffer	1000	1000	500	1000

Next, 2μl of freshly prepared Klenow mix (17μl water, 1μl Klenow polymerase (5Units μl^{-1}, 2μl α-^{35}SdATP (10Ciml^{-1}) was added to each tube. The solutions were mixed gently and then incubated at 37^{o}C for 20 minutes. This was followed by the addition of 2μl of sequence chase mixture (0.5mM each of dGTP, dATP, dTTP and dCTP) to each reaction and a further 20 minute incubation at 37^{o}C.

The reactions were stopped by the addition of 4μl of formamide dye mix (stock solution containing 10ml deionized formamide, 10mg xylene cyanol FF, 10mg bromophenol blue, 0.2ml 0.5M EDTA pH8), and the reactions were placed in a boiling water bath for 5 minutes in order to denature the DNA molecules before loading them onto a sequencing gel.

2.10.5 Sequencing gels.

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

	0.5x mix	2.5x mix
Acrylamide	2.05g	0.46g
Bis acrylamide	0.11g	0.02g
Urea	18.0 g	4.0 g
Sucrose	-	0.4 g
10x TBE	1.8ml	2.0ml
Water to a final volume	36.0ml	8.0ml

10x TBE: 1M Tris-borate pH8.3, 20mM EDTA

Preparation and running of 36x20cm 6% polyacrylamide 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 under vacuum before the addition of:

	0.5x mix	2.5x mix
25% Ammonium persulphate	72μl	16μl
TEMED	72μl	16μl

The gradient gels were poured as follows: 4mls of the 0.5x mix was drawn up in a glass 10ml pipette followed by 6mls of the 2.5x mix. The mixture was poured down the centre of prepared sequencing gel plates carefully, avoiding air bubbles, and was topped up with 0.5x mix down each side of the gel mould. The glass plates were then laid horizontally, a "shark's tooth" comb inserted with the flat face down the gel, and the gel firmly clamped with "bulldog" clips. Once the gel was set, the comb was taken out, the slot washed out with 0.5x TBE, and the comb reinserted with the teeth piercing the gel to a distance of about 1mm.

The samples were loaded, each in a separate slot, and electrophoresed at 1300V with 0.5x TBE running buffer in the upper reservoir and 1x TBE running buffer in the lower reservoir, until the leading dye was about to run off the gel. The glass plates were disassembled and gels were then fixed in 10% acetic acid for 15 minutes, transferred onto Whatman 3MM paper, dried using an ATTO gel drier, and autoradiographed at room temperature for as long as necessary to obtain a clear signal.

2.11 Computing.

The analysis of DNA sequences, in particular to identify restriction endonuclease sites and consensus elements, was conducted using programs written by or for the University of Wisconsin Computer Group, and run on the University of Leicester DEC VAXcluster.

This thesis was compiled using the WORD-11 word processing package which was run on the DEC VAXcluster. In addition, the program packages Corel Draw (Corel Systems Corporation) and Harvard Graphics, both run on an IBM personal computer, were used to produce certain figures.

2.12 Source of materials.

Reagents and chemicals were of analytical grade and most were supplied by Fisons plc and Sigma Chemical Co. Ltd., U.K. Chemicals obtained from other sources are listed below:

Chemical	Source
Radionucleotides	Amersham, England
Acrylamide	BDH Ltd., England
TEMED	Bio-Rad, Richmond, USA
C.I. alkaline phosphatase	Boehringer Mannheim GmbH
HGT agarose	FMC BioProducts, USA
Restriction enzymes T4 DNA Ligase Deoxyribonucleoside triphosphates T4 DNA kinase	GIBCO BRL Ltd., Scotland
IPTG X-gal	Novo Biochem
NovoZym 234	Novo Biolab., Denmark
Klenow fragment	Pharmacia Ltd., England

2.13 Containment and safety.

All experiments conducted in the work described in this thesis were done with reference to the Genetic Manipulation Advisory Group's guidelines on safety and containment conditions for such work. All procedures fall within the category of good microbial practice.

Chapter 3.

Disruption of the 3-phosphoglycerate kinase gene in *Aspergillus nidulans*.

3.1 Introduction.

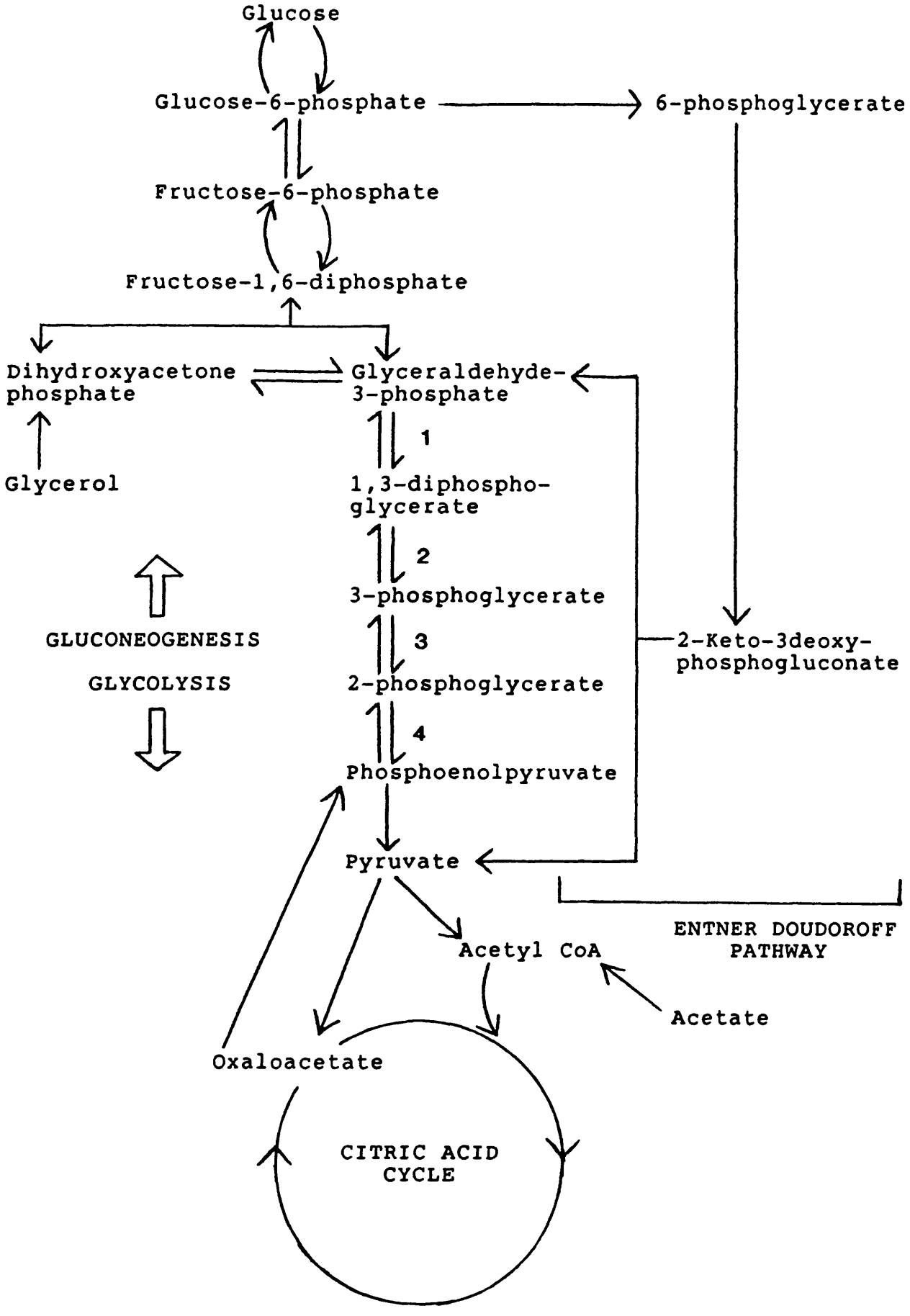
Mutants deficient in 3-phosphoglycerate kinase have been isolated and the defective phenotype mapped to a single locus in *Escherichia coli* (Irani and Maitra, 1974), *Saccharomyces cerevisiae* (Lam and Marmur, 1977) and *Pseudomonas aeruginosa* (Banerjee *et al.*, 1987). In the mouse an enzyme deficient phenotype has been mapped to a locus on the long arm of the X-chromosome (Nielson and Chapman, 1977), and in addition a testis specific enzyme has been mapped to an autosomal site (Eicher *et al.*, 1978).

The *E.coli* and *S.cerevisiae pgk* mutants will grow on a respiratory substrate such as succinate or ethanol if supplemented with a small amount of sugar, or a precursor of sugar such as glycerol. However, they are unable to grow on any single glycolytic carbon source such as glucose, or an oxidative substrate such as pyruvate, or tricarboxylic acid cycle intermediates. The inhibition of growth by normal concentrations of glucose or other sugars is a characteristic of certain glycolytic mutants, and is thought to be due to the accumulation of toxic levels of glycolytic intermediates.

Surprisingly, although the mutant of *P.aeruginosa* is also unable to grow on gluconeogenic precursors such as acetate, it will grow on glycolytic precursors. This mutant was found to use the Entner-Doudoroff pathway which bypasses the glycolytic intermediates (Figure 3.1). Thus, a strain deficient in the *PGK* enzyme could still form the citric acid cycle precursor pyruvate from glucose. However, since this alternative pathway is unidirectional, and the gluconeogenic pathway is interrupted, the mutant is unable to grow on gluconeogenic precursors. Since the *E.coli* and *S.cerevisiae pgk* mutants will not grow on glycolytic precursors, it can

Figure 3.1
The relationship of the Entner-Doudoroff pathway to glycolysis.

The acetate non-utilizing mutants of *A.nidulans* (*acuN*, *acuH*, *acuJ and acuL*) may correspond to deficiencies in the enzymes glyceraldehyde-3-phosphate dehydrogenase (1), 3-phosphoglycerate kinase (2), phosphoglyceromutase (3) and pyruvate kinase (4) at the points indicated.


Glucose
Glucose-6-phosphate
6-phosphoglycerate
Fructose-6-phosphate
Fructose-1,6-diphosphate
Dihydroxyacetone phosphate
Glyceraldehyde-3-phosphate
Glycerol
1
1,3-diphospho-glycerate
2
3-phosphoglycerate
GLUCONEOGENESIS
3
2-Keto-3deoxy-phosphogluconate
GLYCOLYSIS
2-phosphoglycerate
4
Phosphoenolpyruvate
Pyruvate
ENTNER DOUDOROFF PATHWAY
Acetyl CoA
Acetate
Oxaloacetate
CITRIC ACID CYCLE

be concluded that the Entner-Doudoroff pathway is not functional in either of these organisms. It has been reported that *Aspergillus niger* can utilize the Entner-Doudoroff pathway (Elzaing *et al.*, 1973), but it is not known whether the pathway functions in *Aspergillus nidulans*.

Analysis of a large number of acetate non-utilizing (*acu*$^-$) mutants of *A.nidulans* has identified 13 gene loci (Armitt *et al.*, 1976). No enzyme deficiency has been discovered for 4 of these genes and if *A.nidulans* could utilize the Entner-Doudoroff pathway, they may represent lesions in certain enzymes essential for glycolysis and gluconeogenesis. If this is the case the mutants would be able to grow on a glycolytic, but not a gluconeogenic carbon source, and would each lack activity for a specific glycolytic enzyme that is bypassed by the Entner-Doudoroff pathway. No *pgk*$^-$ mutant has been identified in *A.nidulans*, so to determine if any of the acetate non-utilizing mutants corresponded to a *PGK* deficiency, they were examined for growth on either glucose or acetate, and for *PGK* activity (Section 2). Should none of the above strains prove to be a *pgk*$^-$ mutant, a gene disruption strategy was devised to generate the mutant. This would allow the phenotype of a *pgk*$^-$ mutant to be characterized and also enable the *pgk* gene to be mapped in *A.nidulans*.

In *S.cerevisiae* cloned DNA sequences can be introduced into chromosomes by homologous recombination (Hinnen *et al.*, 1978), and this property has all owed the development of one-step (Rothstein, 1983) and two-step (Scherer and Davis, 1979) procedures for the precise replacement of chromosomal sequences with sequences which have been modified *in vitro*. A one-step gene disruption strategy has also been deployed in *Neurospora crassa* (Paietta and Marzluf, 1985), even though vector integration at non-homologous sites is much more common in *N.crassa* than in *S.cerevisiae*. Both the direct (one-step) and indirect (two-step) procedures have been used to disrupt genes in *A.nidulans* (Miller *et al.*, 1985). The disruption of fungal genes using the phenomenon

of integrative transformation is discussed in more detail in the Introduction.

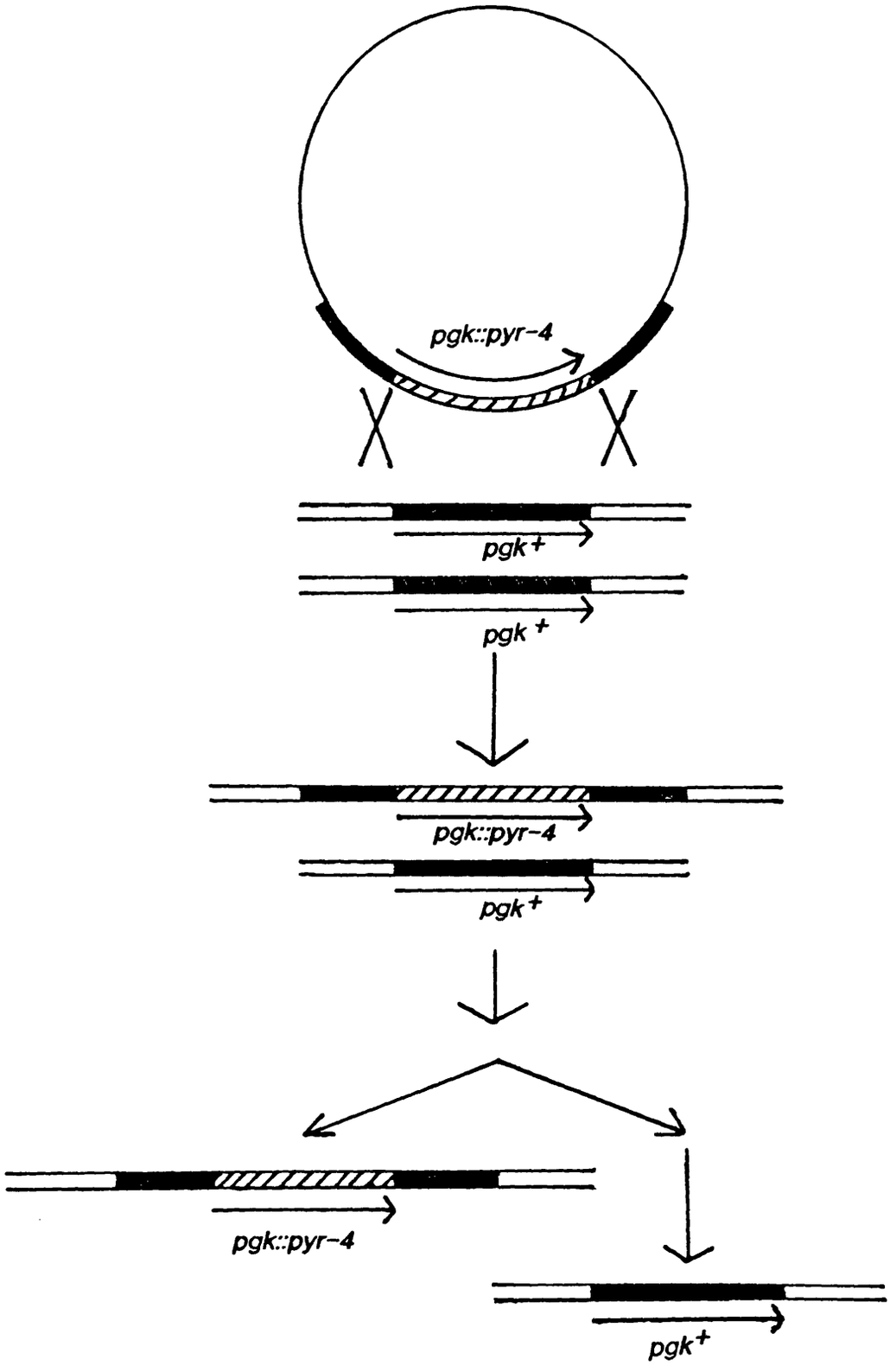
The strategy that was designed to disrupt the resident *pgk*$^+$ gene of *A.nidulans*, known to be present at a single locus (Clements and Roberts, 1985), is outlined in Figure 3.2. A homozygous *pyrG*$^-$*/pyrG*$^-$ diploid strain of Aspergillus was to be transformed with a construct in which the cloned *A.nidulans pgk* gene was interrupted by the *pyr-4* (orotidine decarboxylase) gene of *N.crassa* (Newbury *et al.*, 1986), which complements the *pyrG*$^-$ enzyme lesion. When transformed on a vector the *pyr-4* gene does not target DNA to the *A.nidulans pyrG* gene locus (Turner and Ballance, 1985), there being very little homology between the genes (Oakley *et al.*, 1987). Thus, it should be possible to select *pyr-4*$^+$ transformants for uracil independence in which one of the two resident *pgk* genes has been replaced by the interrupted sequence.

The precise composition of a growth medium for a *pgk*$^-$ mutant strain of *A.nidulans* is not known, and thus the regeneration medium to select *pgk::pyr-4* transformants can not be defined. The use of a diploid strain in which the disrupted *pgk* mutant gene, reasonably assumed to be recessive, can be rescued by the presence of the wild type gene overcomes this difficulty. Transformants would be regenerated on glucose medium lacking uracil, to select for incorporation of *pyr-4* sequence. Subsequently, the *pgk*$^-$ mutant strain might be identified in haploid segregants from the transformed diploid strain using benomyl to increase haploidization by disruption of the mitotic spindle (Hastie, 1970). In order to isolate a haploid strain in which the *pgk* gene had been disrupted (*pgk::pyr-4*), several different combinations of carbon source would be used to supplement media lacking uracil and thus selecting the *pyr-4*$^+$ marker. Furthermore, by including genetic markers on a number of the chromosomes of the original diploid strain, the *pgk* gene may be located to a particular linkage group should any of the markers segregate with *pgk::pyr-4* upon haploidizing.

Figure 3.2
Disruption of the 3-phosphoglycerate kinase gene in a diploid strain of A.*nidulans*.

The A.*nidulans pgk* gene on the vector is interrupted by the *N.crassa pyr-4* gene. A double homologous recombination event between *pgk* sequence present on the vector and one of the native *pgk* loci in a homozygous $pyrG^{-}/pyrG^{-}$ diploid strain would generate a uracil independent transformant with one native and one disrupted *pgk* locus. Subsequent haploidization and chromosome segregation would give both uracil independent (*pgk::pyr-4*) and uracil dependent (pgk^{+}) haploid strains.

Key: Sequence of bacterial origin is represented as ———
the *N.crassa pyr-4* gene as 
the A.*nidulans pgk* gene as 
and other A.*nidulans* sequence as ═══


pgk::pyr-4
pgk+
pgk+
pgk::pyr-4
pgk+
pgk::pyr-4
pgk+

3.2 Assay for 3-phosphoglycerate kinase activities in the unidentified acetate non-utilizing mutants.

The four acetate non-utilizing (*acu*$^-$) mutants of *Aspergillus nidulans*, for which no enzyme deficiency has been assigned, were tested for growth on acetate or glucose, and cell free extract prepared from glucose grown mycelium was assayed for 3-phosphoglycerate kinase activity. Other *acu*$^-$ mutants lacking phosphoenolpyruvate carboxykinase (*acuF*) or isocitrate lyase (*acuD*) were included in the analysis as was the wild type (*acu*$^+$) strain R153.

As expected , none of the *acu*$^-$ mutants would grow on 0.1M acetate, and they all grew normally on 0.02M glucose. The wild type strain grew on both carbon sources. When assayed for 3-phosphoglycerate kinase activity, following growth of mycelium in liquid culture with glucose as carbon source, all of the *acu*$^-$ mutants were found to have a level of enzyme activity comparable to the wild type strain (Table 3.1). Thus, none of the uncharacterized *acu*$^-$ mutants correspond to a *pgk*$^-$ mutant.

Table 3.1

The 3-phosphoglycerate kinase activities recorded in certain *acu*$^-$ mutant strains.

The strains were grown in liquid medium supplemented with 0.02M glucose. In this and in subsequent experiments, cell free extracts were prepared from mycelium by grinding under liquid nitrogen, and were assayed for 3-phosphoglycerate kinase activity in a system coupled with glyceraldehyde-3-phosphate dehydrogenase (Kulbe and Bojanovski, 1982). Protein concentrations in the extracts were estimated by the method of Bradford (1976).

One unit of *PGK* specific activity was defined as that required for the formation of 1μmole of glyceraldehyde-3-phosphate and NAD^- per minute per mg protein.

Strain	Enzyme lesion	*PGK* specific activity (μmolmin^{-1}mg^{-1})
R153 *acu*$^{+}$	None	1.10
acuD306	Isocitrate lyase	1.40
acuF305	Phosphoenolpyruvate carboxykinase	1.20
acuH329	Not known	1.15
acuJ302	Not known	1.25
acuL304	Not known	1.00
acuN356	Not known	1.15

3.3 Construction of a phosphoglycerate kinase gene disruption vector and transformation into a diploid strain of *Aspergillus nidulans*.

In order to construct a phosphoglycerate kinase gene disruption vector which could be selected in *Aspergillus nidulans*, the *pyr-4* gene of *Neurospora crassa* encoding orotidine decarboxylase (Newbury *et al.*, 1986) was introduced into the coding sequence of the *pgk* gene of *A.nidulans* (Clements and Roberts, 1986). The route of vector construction is outlined in Figure 3.3.

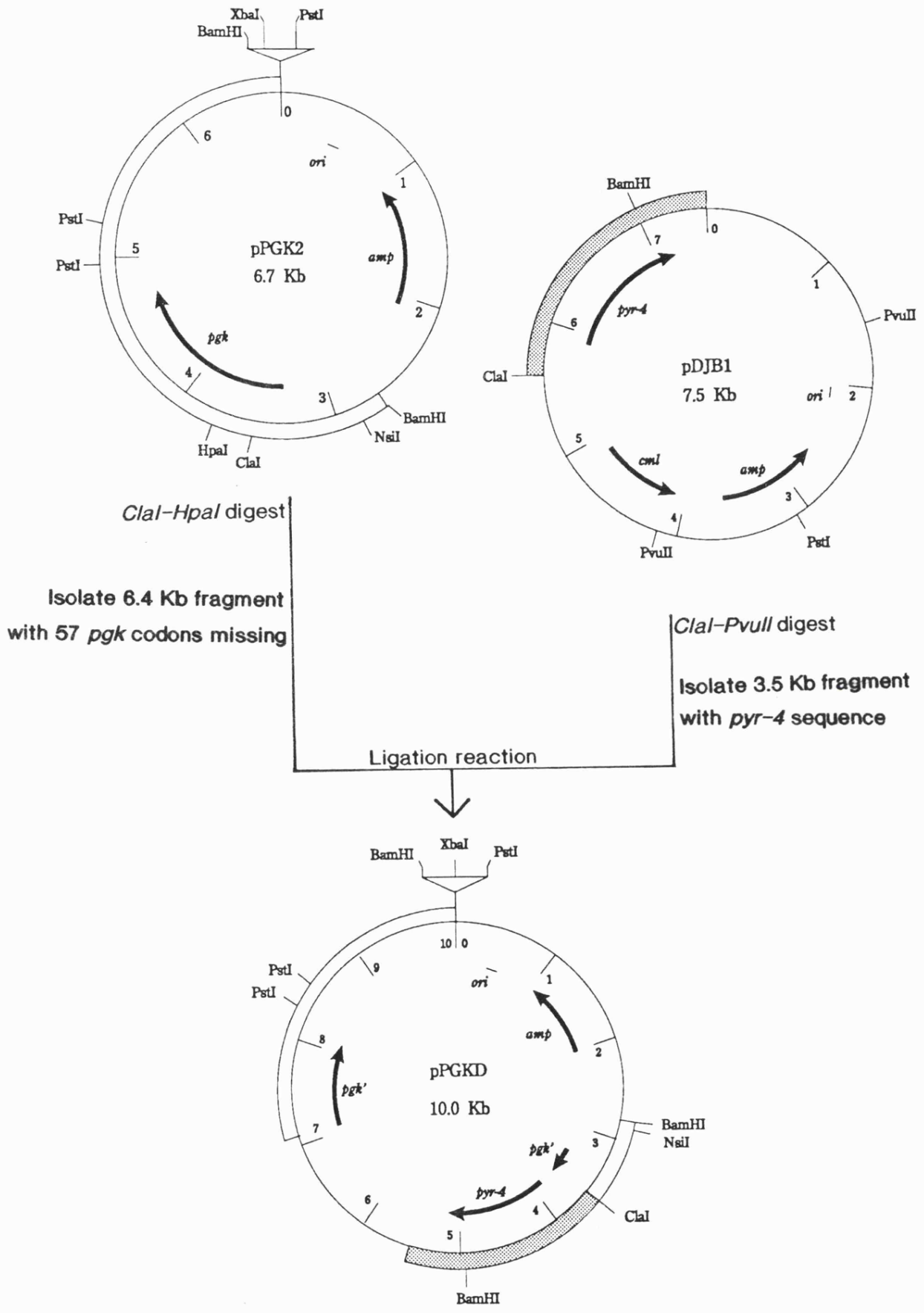
The plasmid pDJB1 (Ballance *et al.*, 1983) was digested with the restriction enzymes *ClaI* and *PvuII*, and the plasmid pPGK2 (Clements and Roberts, 1985) with *ClaI* and *HpaI*. The fragments that were generated by these digests were separated by agarose gel electrophoresis and the 3.5Kb fragment of pDJB1 (including *pyr-4*) together with the 6.4Kb fragment of pPGK2 (missing 171bp of *pgk* coding sequence) were recovered from the gel and incubated in a ligation reaction. The reaction mixture was then introduced into the *Escherichia coli* strain DH5, and transformants were selected for resistance to ampicillin. Plasmid DNA was isolated from several bacterial colonies and screened by restriction enzyme digest analysis for recombinant constructs. In this way, a vector of the desired type was isolated, with the *pyr-4* gene of *N.crassa* and 1.7Kb of bacterial plasmid sequence having replaced codons 54 to 110 of *pgk*. This construct was designated pPGKD.

To provide a recipient strain for the disrupted *pgk* gene, it was necessary to construct a suitable heterozygous diploid strain of *A.nidulans*. One haploid component was QG716, originally designed as a recipient strain for *pgk-lacZ* fusions (Chapter 4). The genotype of this strain is *pyrG89, pabaA1, yA; bgaA4; qutE208* and the route by which it was made is described in Chapter 4, Section 4. The second haploid strain used to construct a diploid was isolated from a cross between the strains G191 (*pyrG89, pabaA1; fwA*) and R153 (*wA3; pyroA4*). The required strain

Figure 3.3
Construction of the vector pPGKD.

The vector pPGK2 consists of a 4.0Kb fragment spanning the *pgk* gene of *A.nidulans* subcloned into the plasmid pUC13 (Clements and Roberts, 1985), and the vector pDJB1 consists of a 1.85Kb fragment spanning the *pyr-4* gene of *N.crassa* subcloned into the tet gene of the plasmid pBR325 (Ballance *et al.*, 1983).
The restriction enzymes *HpaI* and *PvuII* produce blunt ends.

Key: Sequence of bacterial origin is represented as
A.nidulans sequence as
and *N.crassa* sequence as


XbaI
PstI
BamHI
0
1
2
3
4
5
6
ori
amp
pgk
pPGK2
6.7 Kb
PstI
PstI
HpaI
ClaI
BamHI
NsiI
BamHI
PvuII
ClaI
pyr-4
cml
amp
ori
pDJB1
7.5 Kb
PvuII
PstI
ClaI-HpaI digest
Isolate 6.4 Kb fragment
with 57 pgk codons missing
ClaI-PvuII digest
Isolate 3.5 Kb fragment
with pyr-4 sequence
Ligation reaction
BamHI
XbaI
PstI
10
9
8
7
6
5
4
3
2
1
0
ori
amp
pgk'
pgk'
pyr-4
pPGKD
10.0 Kb
PstI
PstI
BamHI
NsiI
ClaI
BamHI

had the genotype *pyrG89; wA3; pyroA4; fwA* and was designated R153-SJS1.

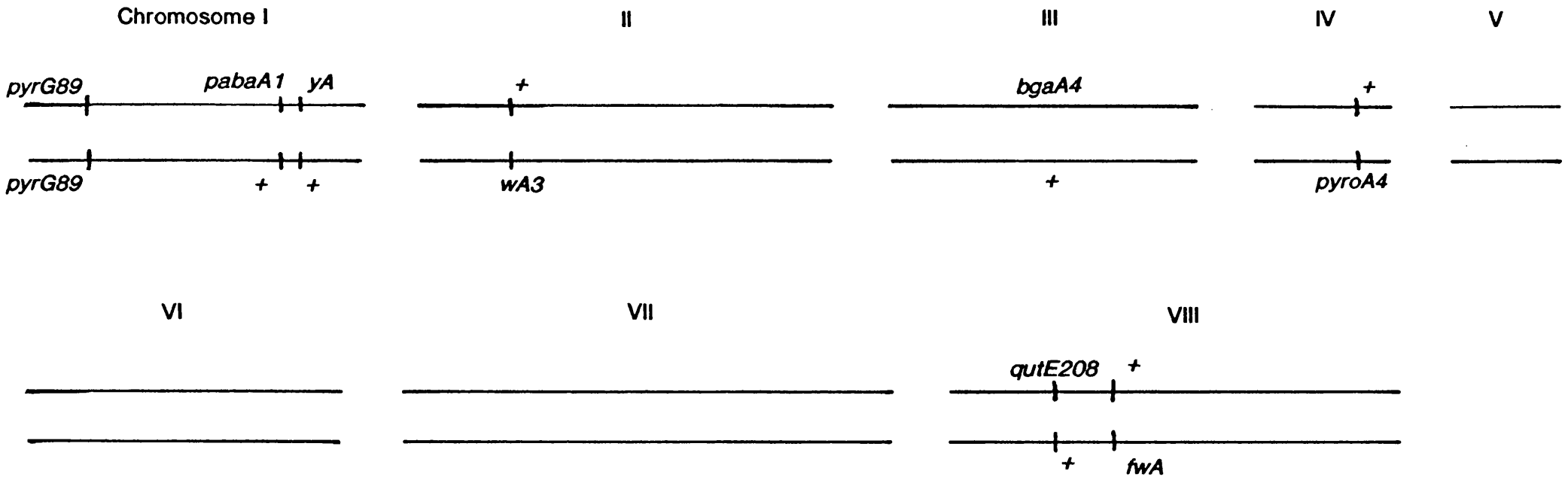
Heterokaryons were established between the two haploid strains QG716 and R153-SJS1 on minimal agar plates supplemented with uracil, and diploids were isolated by the conventional procedure. Four green strains were isolated that were putative diploids, and these were tested for growth with and without uracil and were grown in the presence of the mitotic spindle inhibitor benomyl. All four strains were uracil dependent and were induced to sector by benomyl. It was concluded that these strains were diploids of the required type, and one of them, designated Z10, was chosen as a recipient strain for the disrupted *pgk* gene. Genetic markers are present on 5 of the 8 *A.nidulans* linkage groups, and the genotype of the strain is shown in Figure 3.4.

The construct pPGKD, carrying the disrupted *pgk* gene, was transformed into the diploid strain Z10 by selection for uracil independence. In experiments with *pgk-lacZ* fusion vectors it had been shown that linear vectors transformed into *A.nidulans* at a much higher frequency than circular molecules (Chapter 5). Therefore, both circular and linear vector molecules were used in each transformation experiment to disrupt the *pgk* gene. The vector was linearized at a unique *XbaI* restriction enzyme site located in what was originally the polylinker sequence of pUC13, into which *pgk* sequence had been subcloned to give pPGK2 (Clements and Roberts, 1985). Complete digestion of vector molecules was checked by examining an aliquot by agarose gel electrophoresis, and the remaining digest reaction mixture was used directly in the transformation experiments.

The transformation procedure was adapted from that of Ballance *et al.* (1983), and transformed strains that had incorporated the *pyr-4* gene were selected by uracil independent growth upon glucose media. The transformation procedure was attempted on three occassions. In one case no transformants were isolated, and in a second only two were

Figure 3.4
Genetic map of the diploid strain Z10.

The relative lengths of the chromosomes are deduced from linkage data (Clutterbuck, 1982).


Chromosome I
II
III
IV
V
pyrG89
pabaA1
yA
+
bgaA4
+
pyrG89
+
+
wA3
+
pyroA4
VI
VII
VIII
qutE208
+
+
fwA

recovered; one from the experiment with circular vector molecules and one with linear molecules. In a final experiment, 3 more transformed strains were isolated with circular vector and 31 with linear vector. Thus, as with experiments using *pgk-lacZ* fusion vectors, transforming with linearized molecules appears to significantly increase the transformation frequency. In the final experiment the transformation frequency was increased from 0.054 to 0.55 colonies/μg DNA/10^7 spheroplasts: a factor of 10.

Transformed strains were serially replated four times on selective medium. However, several of these strains appeared to be unstable, failing to give uniform green colonies when replated. In all, 3 strains transformed with the circular vector and 14 transformed with the linear vector were taken forward for further analysis as mitotically stable transformants. Since these transformants had all been selected and grew normally upon glucose media, it is evident that only one, or perhaps neither, of the resident *pgk* genes in the diploid strain had been disrupted.

3.4 Analysis of haploid segregants from the transformed diploid strains.

Since an attempt to isolate *pgk*$^-$ (*pgk::pyr-4*) haploid segregants was likely to prove difficult, it was decided to first seek to recover *pgk*$^+$ segregants. This was achieved by inoculating the diploid transformants onto complex medium with glucose as carbon source and supplemented with uracil in the presence of benomyl. It is expected that under these conditions only *pgk*$^+$ (*pyr-4*$^-$) haploid sectors should be recovered. If the disrupted *pgk* sequence from the vector had replaced the native wild type gene on one of a pair of chromosomes in the diploid, then haploid segregants carrying *pyr-4+* would also harbour the disrupted *pgk* gene, and thus they would not be expected to grow on glucose. However, haploid segregants lacking *pyr-4+* would have the resident wild type *pgk* gene and thus would be expected to grow on glucose (Figure 3.2). Such a result would generate a characteristic segregation pattern should the *pgk* gene lie in one of the genetically marked chromosomes in the diploid strain.

Amongst the transformed diploid strains, 3 of those that are likely to have incorporated the circular vector carrying the disrupted *pgk* gene, and 14 of those that should have incorporated the linear vector, were inocluated onto the medium described above. Haploid segregants isolated in this way were then screened on test media to classify them with respect to the genetic markers in the diploid strain. The presence of the *bgaA4* allele was determined by growth on a medium containing X-gal, as described in Chapter 4, Section 5. The data are presented in Table 3.2.

First, the diploid strains transformed with linear vector are considered. Two (PGKD:L20 and L29) of the 14 examined gave rise to only uracil dependent haploid segregants expected as a result of the disrupted *pgk* gene replacing the native gene. With L20 the genetic markers on linkage groups I, II, III, IV and VIII segregated independently of uracil requirement, implying vector

Table 3.2
Analysis of haploid segregants derived from transformed diploid strains.

The diploid strain Z10 was transformed with the *XbaI* linearized *pgk* gene disruption vector pPGKD to give the series of transformed strains PGKD:L3 to :L31, and with the circular vector to give the transformed strains PGKD:C2 to :C4. Haploid sectors were recovered by inoculating the diploid transformants onto complex medium with glucose and uracil in the presence of benomyl.
Uracil independence demonstrates the presence of the *N.crassa pyr-4* gene. The segregation of the *yA* marker was not scored for all segregants, because white segregants could have either the *yA* or the wild type allele, although it could be assumed that *yA* segregated with *pabaA*.

Diploid strain	Uracil dependence	Number of haploid sectors in each class with respect to the chromosomal markers											
		Chromosome I				II		III		IV		VIII	
		pabaA/+		*yA*/+		*wA*/+		*bgaA*/+		*pyroA*/+		*qutE*/+	
PGKD:L3	ura^- 7	3	4	3	1	3	4	5	2	3	4	7	0
	ura^+ 15	6	9	3	0	12	3	7	8	10	5	0	15
:L10	ura^- 6	0	6	0	1	5	1	3	3	1	5	4	2
	ura^+ 16	16	0	11	0	5	11	7	9	6	10	10	6
:L11	ura^- 0	0	0	0	0	0	0	0	0	0	0	0	0
	ura^+ 19	14	5	7	3	9	10	13	6	5	14	10	9
:L12	ura^- 6	5	1	3	0	3	3	4	2	5	1	6	0
	ura^+ 18	15	3	8	0	10	8	4	14	9	9	0	18
:L13	ura^- 11	5	6	4	2	5	6	6	5	8	3	11	0
	ura^+ 11	9	2	4	2	5	6	7	4	5	6	0	11
:L15	ura^- 0	0	0	0	0	0	0	0	0	0	0	0	0
	ura^+ 24	14	10	10	10	4	20	12	12	9	15	6	18
:L16	ura^- 0	0	0	0	0	0	0	0	0	0	0	0	0
	ura^+ 20	11	9	7	5	8	12	14	6	5	15	6	14
:L20	ura^- 22	15	7	10	1	11	11	10	12	8	14	16	6
	ura^+ 0	0	0	0	0	0	0	0	0	0	0	0	0
:L23	ura^- 15	6	9	2	5	8	7	7	8	7	8	3	12
	ura^+ 36	24	12	13	2	21	15	16	20	9	27	18	18
:L25	ura^- 8	4	4	2	1	5	3	4	4	4	4	0	8
	ura^+ 15	9	6	6	1	8	7	10	5	4	11	15	0
:L26	ura^- 0	0	0	0	0	0	0	0	0	0	0	0	0
	ura^+ 22	12	10	9	4	9	13	4	18	9	13	11	11
:L28	ura^- 14	6	8	3	0	11	3	9	5	5	9	0	14
	ura^+ 8	6	2	2	1	5	3	3	5	4	4	8	0
:L29	ura^- 22	11	11	9	1	12	10	11	11	8	14	0	22
	ura^+ 0	0	0	0	0	0	0	0	0	0	0	0	0
:L31	ura^- 9	6	3	6	3	0	9	5	4	3	6	3	6
	ura^+ 15	3	12	0	0	15	0	5	10	6	9	5	10
:C2	ura^- 24	14	10	12	7	5	19	12	12	13	11	11	13
	ura^+ 20	14	6	11	4	5	15	8	12	9	11	10	10
:C3	ura^- 9	8	1	5	0	4	5	2	7	5	4	3	6
	ura^+ 16	14	2	8	2	6	10	9	7	6	10	6	10
:C4	ura^- 8	7	1	5	0	3	5	4	4	1	7	2	6
	ura^+ 23	17	6	7	3	13	10	13	10	9	14	10	13

integration and *pgk* gene disruption on one of the unmarked chromosomes (V, VI or VII), whilst with L29 a genetic marker on linkage group VIII (*qutE*$^{+}$) segregated with uracil requirement, all haploid strains recovered being capable of growth on quinic acid. The other marker on chromosome VIII (*fwA*) was not tested for segregation, since only white, yellow and green sectors were analysed. This result is contradictory since in the two transformants described above *pgk* is implied to be located on different chromosomes, but it has been shown that a *pgk* gene probe only identifies one locus in *A.nidulans* (Clements and Roberts, 1985).

With a further 8 of the diploid transformants (PGKD:L3, L10, L12, L13, L23, L25, L28 and L31) both uracil dependent and uracil independent haploid strains were recovered, implying that integration of the *pyr-4* gene into the genome had not resulted in the replacement of one of the native *pgk* alleles with the disrupted *pgk* sequence from the vector. However, for 5 of these transformants (L3, L12, L13, L25 and L28) the genetic marker on linkage group VIII (*qutE*) again segregated with *pyr-4*. In contrast, in L10 the markers on group I (*pabaA* and *yA*) segregated with *pyr-4*, and in L31 the marker on group II (*wA*) segregated with *pyr-4*, but in L23 no marker segregated with *pyr-4*. Thus, the *pyr-4* sequence has integrated into chromosome VIII in 5 of the above cases, into chromosome I once, chromosome II once and into either chromosome V, VI or VII once.

With the four remaining transformants (PGKD:L11, L15, L16 and L26) only uracil independent (*pyr-4*$^{+}$) haploids were recovered, and these showed no bias with respect to other markers. The most probable explanation for this result is that *pyr-4* sequence has integrated into both of an unidentified pair of chromosomes.

Summing up the above data, *pyr-4* sequence has integrated into chromosome VIII in 6 cases, apparently disrupting the resident *pgk* gene in one of these. It has integrated into chromosome I once, chromosome II once and chromosomes V, VI or VII twice, again apparently disrupting

pgk in one of the latter cases. There are also four examples of integration into both chromosomes of an unspecified homologous pair. Therefore, there is a clear bias towards integration of *pyr-4* sequence into chromosome VIII, this being the case for 6 out of 10 of the transformants that can be characterized. Since *pyr-4* is known not to direct vector DNA to any particular site in the *A.nidulans* genome (Turner and Ballance, 1985), and in any event its homologue (*pyrG*) is located on chromosome I, it is most probable that *pgk* sequence in the vector is responsible for the targeting effect that is observed. However, the probable integration of vector sequence at the correct locus only appears to have replaced the resident *pgk* gene with a disrupted sequence in one case (PGKD:L29).

The anomalous result of an apparent *pgk* gene disruption accompanied by integration of *pyr-4* sequence on another chromosome (PGKD:L20) could be explained by the integration event resulting in the disruption of an essential gene other than *pgk* on either chromosome V, VI, or VII. Such a diploid transformant would generate only uracil dependent haploid strains in which none of the genetic markers used above would segregate with uracil dependence; a characteristic of transformant PGKD:L20.

Subsequent to the above analysis of transformed strains that had incorporated linear vector molecules, those that had been transformed with the circular plasmid were analysed. Three transformed strains that were thought to have incorporated circular vector (PGKD:C2, C3 and C4) were examined as above. In each case both uracil dependent and independent haploid strains were isolated. However, *pyr-4* did not segregate with any of the genetic markers, implying the integration of vector sequence into chromosomes V, VI or VII.

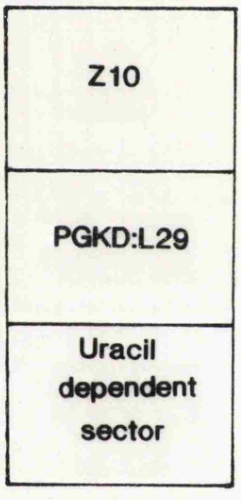
In order to confirm that *pyr-4* sequence had been incorporated into the transformed strain PGKD:L29, which is believed to be a diploid heterozygous at the *pgk* locus, genomic DNA was isolated from the transformant, a uracil

dependent haploid sector derived from it, and the original recipient diploid strain Z10. The DNA samples were denatured and spotted onto a Hybond-N filter, prior to hybridizing to a radioactively labelled probe. The DNA fragment used as probe was isolated from an agarose gel as a 1.1Kb *HindIII* fragment of the vector pDJB1, and it consisted entirely of *Neurospora crassa pyr-4* sequence. An autoradiographic image of the filter is presented in Figure 3.5. As expected, the *pyr-4* probe sequence only hybridized to DNA isolated from the transformed diploid strain, implying that only this strain had *pyr-4* sequence integrated into the genome. Since DNA isolated from the uracil dependent haploid strain did not hybridize to the probe, it can be concluded that *pyr-4* sequence on the vector molecule only integrated into one of an homologous pair of chromosomes in the transformed diploid strain from which it was derived. This reinforces the genetic evidence provided above, indicating that vector sequence has been incorporated into one of the two number VIII chromosomes.

Figure 3.5

The incorporation of *N.crassa pyr-4* sequence into the transformed strain PGKD:L29.

The autoradiograph shows the hybridization of an α^{32}P-dCTP labelled 1.1Kb *HindIII* DNA fragment consisting entirely of *pyr-4* sequence, with genomic DNA isolated from the recipient diploid Z10, the transformed diploid PGKD:L29, and a uracil dependent haploid sector derived from it.


Z10
PGKD:L29
Uracil dependent sector

3.5 Isolation and analysis of a *pgk* mutant strain of Aspergillus nidulans.

From the work described in Section 4, the transformed diploid strain PGKD:L29 is believed to be heterozygous with respect to the *pgk* locus (Figure 3.2). By haploidizing on medium supplemented with glucose and uracil, sectors were recovered which were uracil requiring (*pyr-4*$^-$) and (presumed) *pgk*$^+$. In order to isolate sectors of the other predicted class *pgk::pyr-4*, the diploid transformant was now haploidized on defined medium lacking uracil to select for the *pyr-4*$^+$ function, and containing both a glycolytic and a gluconeogenic carbon source (0.05M sucrose plus 0.1M acetate), and white and yellow sectors were observed.

When the white and yellow sectors formed on the above haploidization medium were replated by mass transfer of conidiospores onto media supplemented with various carbon sources, it was found that segregants fell into two groups. One group formed strongly conidiating yellow or white colonies on all of the carbon sources both with and without uracil. However, the other group similarly formed yellow or white colonies on media supplemented with uracil, but dull brown poorly conidiating colonies on media lacking uracil but supplemented with either fructose and acetate or glycerol and acetate. These data are summarized in Table 3.3. From these results it was deduced that the original sectors were probably heterokaryons. The dull brown colonies that grew on media supplemented with fructose and acetate or glycerol and acetate were thought to be *pgk*$^-$ mutant haploid strains. Their brown colouration does not correspond to the fawn marker, since they conidiated very poorly. These isolates were replated directly several times onto media in which the concentrations of glycerol and acetate were varied. The strains grew best on medium supplemented with 0.1M acetate and 0.05M glycerol, and growth was shown to be enhanced in the presence of peptone (0.1 to 1.0% w/v).

Subsequently, these strains were tested for growth on medium supplemented with 0.1M glycerol and 1% quinic acid.

Table 3.3

Growth of segregants isolated from the diploid transformed strain PGKD:L29 on a variety of carbon sources.

The original sectors were isolated on medium lacking uracil and containing both a glycolytic and a gluconeogenic carbon source (0.05M sucrose plus 0.1M acetate). They are thought to have been heterokaryons, since upon replating onto media supplemented with the carbon sources shown in the table the segregants were shown to be one of two classes.
The degree of colony growth was indicated by +, ++, +++, and no growth by -. Strongly conidiating colonies (+++) were yellow or white (Class I) and poorly conidiating colonies (++ or +) were dull brown (Class II).

Carbon source in agar media	Media supplemented with uracil	Class of segregant	
		I	II
Sucrose (0.05M) Acetate (0.1M)	Yes	+++	+++
Sucrose (0.05M)	No	+++	-
Acetate (0.1M)	No	+++	-
Sucrose (0.05M) Acetate (0.1M)	No	+++	-
Glucose (0.05M) Acetate (0.1M)	No	+++	-
Fructose (0.05M) Acetate (0.1M)	No	+++	+
Glycerol (0.05M) Acetate (0.1M)	No	+++	++

None of the strains grew under these conditions, indicating that they are *qutE*$^-$ mutants, and that the inability to grow on quinic acid has segregated with uracil independence. One of the strains (PGKD:L29:160) was isolated by replating for single colonies on 0.1M acetate and 0.05M glycerol with 1% peptone.

In order to determine whether the haploid strain isolated above is a *pgk*$^-$ mutant, cell free extract was prepared from mycelium grown in liquid culture supplemented with acetate and glycerol, and was assayed for 3-phosphoglycerate kinase activity in a system coupled with glyceraldehyde-3-phosphate dehydrogenase. The diploid transformed strain PGKD:L29, from which the putative mutant was derived, a *pgk*$^+$ (*pyr-4*$^-$) haploid strain derived from this transformant, and the original recipient diploid strain Z10, were also grown in the same medium and analysed in the same way. The data are given in Table 3.4. The recipient diploid strain, the transformed diploid strain and the *pgk*$^+$ (*pyr-4*$^-$) haploid strain all contained 3-phosphoglycerate kinase activity, although to varying degrees. The recipient diploid strain had approximately three times the activity of the diploid transformant and twice the activity of the *pgk*$^+$ (*pyr-4*$^-$) haploid strain, even though on the basis of chromosome complements, both the recipient strain and the *pgk*$^+$ (*pyr-4*$^-$) haploid strain might have been expected to have double the activity of the diploid transformant. Most strikingly, the uracil independent strain that was recovered on media supplemented with glycerol and acetate had only residual enzyme activity, thus confirming that it is a *pgk*$^-$ mutant.

Although a *pgk*$^-$ mutant had been recovered from a diploid strain transformed with the disrupted *pgk* gene, it was not known whether the disrupted sequence replaced the native gene on one of the chromosomes in the fashion predicted (Figure 3.2). In order to characterize the nature of the disruption event at the molecular level, genomic DNA was prepared from each of the strains that had been assayed for *PGK* enzyme activity. The DNA samples were digested with

Table 3.4

The 3-phosphoglycerate kinase activities in the diploid transformant PGK:L29 and in uracil dependent and uracil independent haploid segregants.

The strains were grown in liquid medium supplemented with 0.1M acetate, 0.05M glycerol, until the turbidity of the cultures appeared similar. Cell free extracts were prepared and assayed for *PGK* activity under standard conditions (Table 3.1).

Strain	Time of culture growth (hrs.)	*PGK* specific activity ($\mu molmin^{-1}mg^{-1}$)
Recipient diploid (Z10)	18	2.00
Transformed diploid (PGKD:L29)	21	0.71
Uracil dependent haploid	21	0.95
Uracil independent haploid	45	0.14

the restriction enzyme *BamHI*, and analysed by agarose gel electrophoresis and DNA hybridization. The genomic DNA digests were transferred from the gel onto a Hybond-N filter, and were incubated in a hybridization reaction with a radioactively labelled DNA probe at 65^{o}C for 15 hrs. A 1.5Kb *HindIII-PstI* restriction fragment consisting entirely of *pgk* sequence was used as probe. Subsequently the filters were washed at high stringency (four 30 minute washes in 0.1X SSC, 0.5% SDS), and an autoradiographic image was obtained by exposing the filter to X-ray film (Figure 3.6). The sizes of fragments that would be predicted to hybridize to the *pgk* probe if a classical gene replacement had occurred are given in the accompanying table.

Although the recipient diploid strain and the *pgk*$^{+}$ (*pyr-4*$^{-}$) haploid showed that the *pgk* probe only hybridizes to the native locus (as predicted), the diploid transformant and the *pgk::pyr-4* haploid derived from it both give complex hybridization patterns that do not correspond to those expected from a classical replacement event. The intensity of bands on the autoradiograph implies multiple integration events. The transformed diploid strain has clearly segregated to give one haploid strain with the wild type *pgk* allele, and another in which the *pgk* locus has been disrupted, although not by a straightforward gene replacement.

In Section 4 the uracil dependent phenotype (*pgk*$^{+}$, *pyr-4*$^{-}$ genotype) was shown to segregate with the ability to grow on quinic acid (*qutE*$^{+}$ genotype), and in this Section uracil independent strains were shown to be incapable of growth on quinic acid as a gluconeogenic carbon source in the presence of glycerol, implying that vector sequence had integrated into chromosome VIII carrying the *qutE* mutant gene. As a result of the DNA hybridization experiment described above, it is clear that the *pgk* locus has been disrupted through the integration of vector sequence into one chromosome in the diploid transformed strain. Taken together, these experiments clearly show that the *pgk* gene is located on chromosome VIII.

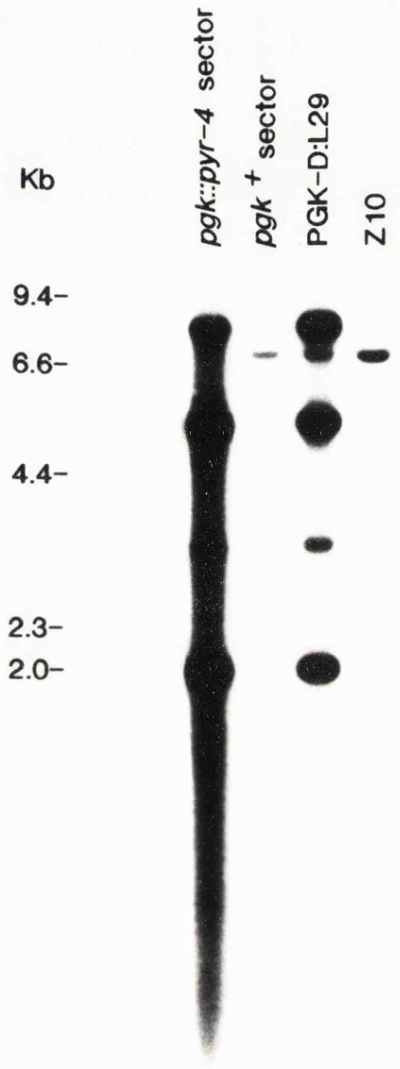
Figure 3.6

DNA analysis of the transformed strain PGKD:L29, to determine the type of vector integration event.

The autoradiograph shows the hybridization of an α^{32}P-dCTP labelled 1.5Kb *HindIII-PstI* DNA fragment consisting entirely of *pgk* sequence, with *BamHI* digests of genomic DNA prepared from the recipient diploid Z10, the transformed diploid PGKD:L29, and uracil dependent (*pgk*$^+$) and uracil independent (*pgk::pyr-4*) haploid sectors derived from it. Phage λ DNA digested with *HindIII* provided the molecular weight marker (not shown on the autoradiograph).

The table shows the molecular weights (Kb) of the fragments predicted for a classical gene replacement and of those calculated from the autoradiograph.

Strain	Fragment sizes (Kb)	
	Predicted	Calculated from autoradiograph
Recipient diploid (Z10)	6.7	7.1
Transformed diploid strain (PGKD:L29)	6.7, 5.1, 4.9	8.0, 7.1, 5.4, 3.3, 2.0
Uracil dependent (*pgk*$^+$) sector	6.7	7.1
Uracil independent (*pgk::pyr-4*) sector	5.1, 4.9	8.0, 5.4, 3.3, 2.0


Kb
9.4–
6.6–
4.4–
2.3–
2.0–
pgk::pyr-4 sector
pgk + sector
PGK-D:L29
Z10

3.6 Attempts to precisely replace the native pgk gene with the disrupted *pgk* sequence.

In the transformed diploid strain PGKD:L29 the native *pgk* gene on one chromosome has been disrupted through the integration of vector sequence. However, this has not occurred as a result of a classical gene replacement, and the manner in which vector sequence has integrated into the chromosome is unclear from genomic DNA analysis.

The transformed strain PGKD:L29 was isolated when the vector pPGKD, carrying disrupted *pgk* sequence was introduced into the diploid strain Z10, following linearization at a unique *XbaI* site in the vector pPGKD. This site lies in sequence of polylinker origin, though immediately adjacent to Aspergillus *pgk* 3' sequence. The proportion of type I integration and of gene replacement events have been shown to be significantly increased in *Saccharomyces cerevisiae* by linearizing the vector within sequence homologous to that to which it is to be targeted in the genome (Orr-Weaver *et al.*, 1981). A similar effect was also observed in *Aspergillus nidulans* when linear *pgk-lacZ* fusion vectors were cut within *qut* gene cluster sequence (Chapter 5). However, in each of these cases the vector was linearized within sequence homologous to the target sequence, whereas in the transformation experiment with pPGKD (Section 3) the vector was cut at an *XbaI* site that was originally located in the polylinker sequence of pUC13. In an attempt to generate a strain in which the native *pgk* gene had been replaced by disrupted sequence in a defined manner (a classical gene replacement), vector molecules that had been cut within *A.nidulans* sequence flanking *pgk* were introduced into the diploid strain Z10.

Initially a 5.5Kb *NsiI-PstI* restriction enzyme fragment spanning the entire disrupted *pgk* sequence was isolated from the vector pPGKD (Figure 3.3). The *NsiI* site is located approximately 0.5Kb upstream of the transcript start site of the *pgk* gene, and the *PstI* site is about 0.15Kb downstream of the translation stop codon. Three attempts were made to

introduce this fragment into the diploid *A.nidulans* strain Z10 by selecting for uracil independence. The first two were unsuccessful, but on the third occassion eight transformants were recovered. These were replated on selective medium to isolate single colonies from which genomic DNA was prepared and analysed as above (Section 5). Although certain fragments were conserved, none of the hybridization patterns corresponded to that which would be predicted for a classical gene replacement (data not shown).

In a second attempt to precisely replace the native *pgk* gene with the disrupted sequence, the vector pPGKD was linearized at the unique *NsiI* restriction enzyme site within *A.nidulans* sequence upstream of the disrupted gene, and was introduced into the diploid strain Z10. Two transformants were recovered and replated to isolate single colonies. Genomic DNA was prepared from both of these transformed strains and analysed as above (Section 5). The hybridization patterns indicated that in neither of these transformed strains had the disrupted *pgk* sequence replaced the native gene through a classical gene replacement on either chromosome of an homologous pair (data not shown).

In summary, introducing the disrupted *pgk* sequence into a diploid strain of *A.nidulans*, either on a fragment with free potentially recombinogenic ends spanning the sequence, or on a vector linearized within *A.nidulans* sequence upstream of the gene, did not lead to the replacement of the native *pgk* gene through a classical gene replacement.

3.7 Summary.

A mutant strain of *Aspergillus nidulans* deficient for the enzyme 3-phosphoglycerate kinase (*PGK*) has not previously been isolated, and consequently the *pgk* locus has not been mapped. Four acetate non-utilizing mutants of *A.nidulans*, for which no enzyme deficiency had yet been assigned, were tested for *PGK* activity. They were each comparable to the wild type strain, indicating that none of them are a pgk^- mutant strain.

In order to engineer a pgk^- mutant strain of *A.nidulans*, and thus map the gene locus to a specific chromosome, a gene disruption strategy was devised, the intention of which was to replace the native *pgk* gene with *pgk* sequence that had been interrupted with a selectable marker. A vector was made in which the *pyr-4* gene of *Neurospora crassa* was inserted into *pgk* coding sequence. This construct was transformed into a diploid strain of *A.nidulans* that was homozygous for a $pyrG^-$ allele, and had genetic markers to distinguish five of the eight chromosomes pairs.

By using benomyl to induce haploidization, diploid transformants were forced to segregate haploid strains on media containing uracil and glucose to select for the wild type pgk^+ allele. Analysis of the products revealed a clear bias in the segregation of genetic markers, with uracil requirement ($pyr\text{-}4^-$) segregating with the marker on chromosome VIII ($qutE^+$) in 6 out of the 10 cases that could be characterized. This indicates that the *pgk* locus is on chromosome VIII.

Subsequently, further haploid sectors were isolated from one particular transformed strain that had generated only uracil requiring segregants on glucose medium, indicating disruption of the native *pgk* locus on one chromosome of an homologous pair. However, this time sectors were isolated on media containing benomyl to induce haploidization, but lacking uracil and supplemented with a

variety of carbon sources to select for the *pgk::pyr-4* allele. In this way a strain was isolated that was uracil independent, but would only grow on media suplemented with both a glycolytic (fructose or glycerol) and a gluconeogenic (acetate) carbon source. The strain would not grow on glycerol plus quinic acid medium, indicating that it is a *qutE*$^-$ mutant and that the inability to grow on quinic acid had segregated with uracil independence. When this strain was tested for *PGK* activity, it was clearly shown to be a *pgk*$^-$ mutant, and restriction enzyme plus DNA hybridization analysis of genomic DNA confirmed that the *pgk* locus on one chromosome of an homologous pair had been disrupted. Since the uracil dependent phenotype had previously been shown to segregate with the genetic marker on chromosome VIII, the *pgk* locus could be confidently assigned to this chromosome.

Although the *pgk* locus had been disrupted, it was clear form the DNA hybridization pattern obtained that a straightforward replacement of native sequence with vector sequence had not occurred. In order to investigate the effect of transforming with molecules with free recombinogenic ends originating from *A.nidulans* sequence, the diploid strain was transformed with a fragment consisting only of the disrupted *pgk* sequence, and in a separate experiment with the whole vector linearized within 5'*pgk* sequence. Analysis of genomic DNA isolated from transformed strains indicated that in neither case was the resident *pgk* gene replaced with the disrupted sequence.

Chapter 4.

Construction of fusions of the *Aspergillus nidulans* 3-phosphoglycerate kinase gene promoter (5'*pgk*) to a "reporter" gene.

4.1 Introduction.

Sequence elements essential for the correct expression of a cloned gene can generally be identified by modifying the promoter *in vitro* and then introducing the modified clones into a suitable host cell, which ideally lacks the relevant gene. The synthesis of message and protein can then be followed and compared between differently modified clones. Alternatively, a normal host cell can be used, provided the modified clones are introduced at high copy number, and message and protein levels can then be compared to the untransformed strain. It is in this way that sequences important for the expression of the 3-phosphoglycerate kinase gene in *Saccharomyces cerevisiae* were identified (Ogden *et al.*, 1986). However, no autonomosly replicating vectors have been developed for filamentous fungi, and at the start of my investigation no *pgk*$^-$ mutant strain of *Aspergillus nidulans* was available.

In the absence of an appropriate mutant host cell the modifications must be applied to a construct in which the promoter of the gene of interest is fused to a structural gene which is not normally expressed in the recipient cell. The availability of a reliable assay for the heterologous "reporter" gene is an important aspect of such a strategy.

Promoter fusions to the β-galactosidase (*lacZ*) gene of *Escherichia coli* have provided a convenient method for studying gene expression in prokaryotes and yeast (reviewed by Casadaban *et al.*, 1983), in Drosophila (Lis *et al.*, 1983) and in mammalian cells (Nielsen *et al.*, 1983). Expression of β-galactosidase can be detected *in vivo* by using the chromogenic substrate X-gal, and by assaying β-galactosidase activity in cell free extracts, the regulatory effect of

promoter sequences fused to *lacZ* can be determined quantitively.

Such methods have been applied in Aspergillus, where an in phase translational fusion of *lacZ* into the protein coding region of the tryptophan biosynthetic gene (*trpC*) resulted in the production of a functional β-galactosidase fusion protein when the construct was transformed into *A.nidulans* (van Gorcom *et al.*, 1985). Production of fusion protein was abolished when *trpC* transcription and translation initiation signals were removed, indicating that β-galactosidase expression was dependent upon *A.nidulans* control sequences. A set of expression vectors have been constructed in which a unique restriction enzyme site has been positioned in front of the *lacZ* gene in each translational reading frame (van Gorcom *et al.*, 1986). The 3' end of the *trpC* gene was introduced downstream of *lacZ*, to allow for correct termination of transcription in *A.nidulans*. The *argB* gene was included elsewhere on the plasmids to provide a selectable marker for transformation of an *argB*$^-$ mutant strain. Using these constructs expression of β-galactosidase driven by the *A.nidulans trpC* and glyceraldehyde-3-phosphate dehydrogenase (*gpd*) gene promoters were compared (van Gorcom *et al.*, 1986), and promoter sequences required for expression of *trpC* were identified (Hamer and Timberlake, 1987).

As discussed in the Introduction, I proposed to analyse elements of potential importance in giving the high level of expression observed with *pgk* by constructing vectors with intact and modified *pgk* promoter sequences fused to *lacZ* and introducing these into *A.nidulans*.

4.2 Construction of *pgk-lacZ* gene fusions.

4.2.1 Requirements of fusion vectors.

Gene fusions to *lacZ* can be constructed with either transcription initiation control signals or both transcription and translation initiation control signals from another gene. Here I chose to make transcription-translation fusions to the *Aspergillus nidulans pgk* gene, with the aim of ensuring efficient initiation of translation of the fusion message in Aspergillus. Furthermore, by including a transcription termination sequence from an *A.nidulans* gene, the fusion message should terminate correctly in Aspergillus. The β-galactosidase protein produced by *lacZ* fusions can apparently tolerate any number of N-terminal codons from the foreign gene, and lack up to the first 27 codons from the *lacZ* structural gene and yet retain enzymatic activity (Casadaban *et al.*, 1983). However, such fusions must have the DNA coding sequences in the correct reading frame.

Ideally the fusion vectors should also contain three unique restriction enzyme sites to enable further modification and manipulation. One of these, positioned upstream of any putative control sequences in the *pgk* promoter would provide a start point for exonuclease digestion into the promoter, with the aim of determining its 5′ extent. By including two other restriction enzyme sites spanning the *lacZ* gene, provision would be made for the replacement of *lacZ* with other coding sequences of interest.

The fusion vectors had also to include a marker to select transformants in *A.nidulans*. As is discussed in the Introduction, this marker would potentially target vector sequence to the corresponding locus in the genome, so providing a constant chromosomal enviroment for the comparison of β-galactosidase expression from differently modified constructs. It was decided to use the catabolic dehydroquinase (*qutE*) gene of *A.nidulans* in order to select transformants by growth on quinic acid as carbon source.

It was also necessary to include the replication origin and an antibiotic resistance gene of a bacterial plasmid so that the vectors could be propagated and selected in *Escherichia coli*.

4.2.2 An attempt to construct fusion vectors by using DNA linker molecules.

Initially two strategies were followed in an attempt to construct a gene fusion with the features described above.

The first step, outlined in Figure 4.1, was common to both strategies and involved subcloning sequence upstream of the *pgk* gene into a vector carrying the *A.nidulans qutE* gene, the antibiotic resistance genes for ampicillin and chloramphenicol, and the replication elements of pBR322. This was successfully achieved by ligating the 0.7Kb *BamHI-HindIII* restriction enzyme digest fragment from the vector pPGK2 (Clements and Roberts, 1985) with the 7.45Kb *BglII-HindIII* fragment of pAL3.3-B3 (Hawkins *et al.*, 1985). The *BamHI-HindIII* fragment of pPGK2 consists of 638bp of sequence 5′ to the transcription start site, the 32bp leader sequence and the 7 N-terminal codons of *pgk*. This vector was designated p5′PGK-QUTE.

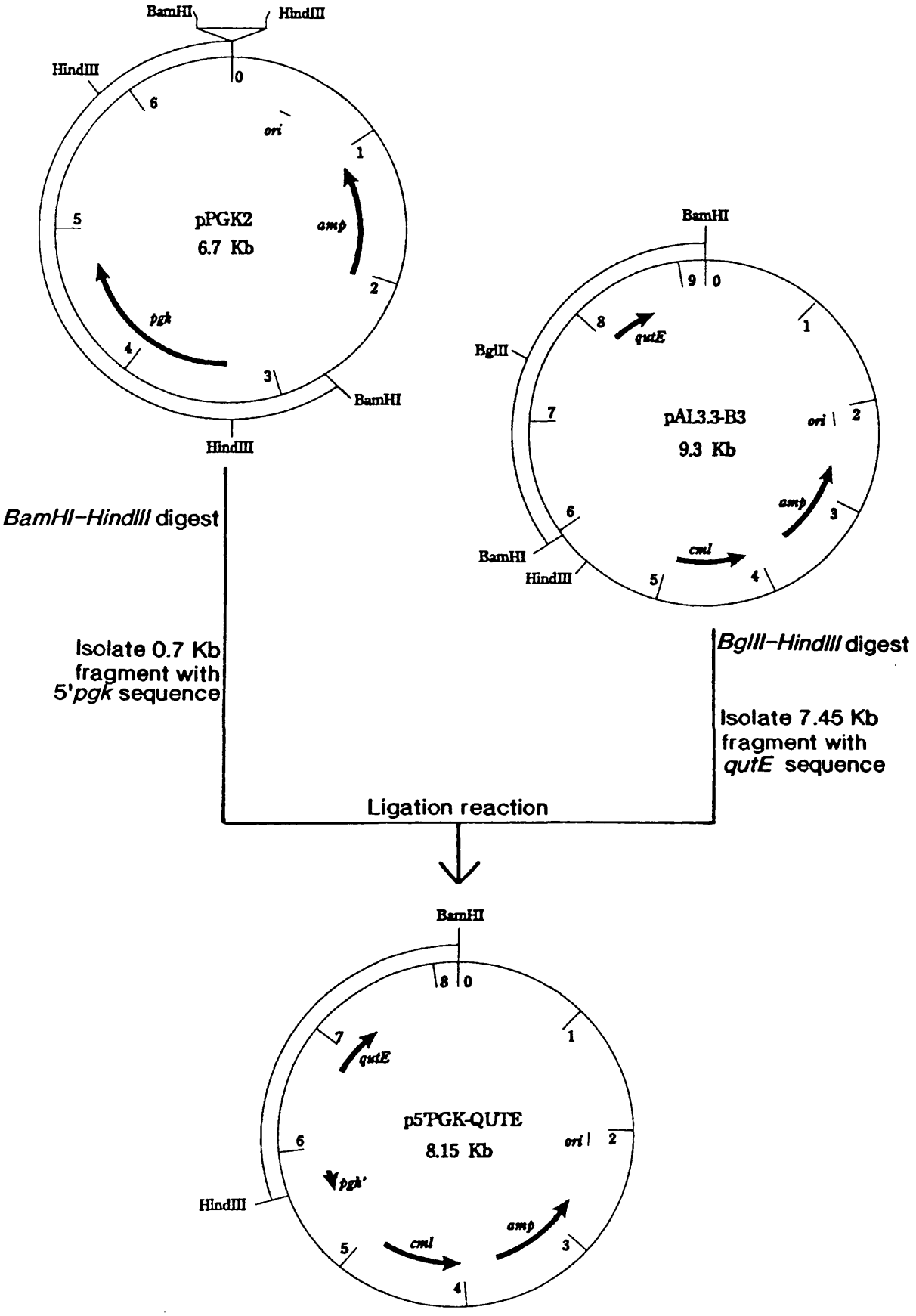
The second step of each strategy involved fusing the *lacZ* gene of *E.coli* to the *pgk* promoter on p5′PGK-QUTE. The *lacZ* gene, lacking its 7 N-terminal codons, is present on the pAN923 plasmid series, where it is fused 11bp downstream of its 3 translational stop codons to the 50 C-terminal codons and 0.55Kb of 3′ sequence of the *trpC* gene of *A.nidulans* (van Gorcom *et al.*, 1986). A *BamHI* site is present in each translational phase upstream of *lacZ* on the three vectors in the series (pAN923-41A, -42A and -43A), and an *XbaI* site lies downstream to the *trpC* 3′ sequence. By introducing this 3.9Kb *BamH-XbaI* fragment from each of the pAN923 vectors in the correct orientation into the unique *HindIII* site on p5′PGK-QUTE, vectors should be constructed with the *A.nidulans pgk* promoter and 7 N-terminal codons fused in each translational reading frame to *lacZ*.

Figure 4.1
Construction of the vector p5'PGK-QUTE.

The vector pPGK2 consists of a 4.0Kb fragment spanning the *pgk* gene of *A.nidulans* subcloned into the plasmid pUC13 (Clements and Roberts, 1985), and the vector pAL3.3-B3 consists of a 3.3Kb fragment spanning the *qutE* gene of *A.nidulans* subcloned into the plasmid pBR325 (Hawkins *et al.*, 1985).
The restriction enzymes *BamHI* and *BglII* produce complementary single stranded ends.

Key: In this figure, and in all subsequent figures in this Thesis:

sequence of bacterial origin is represented ———
A.nidulans sequence is represented ═══


BamHI
HindIII
HindIII
0
6
ori
1
5
pPGK2
6.7 Kb
amp
2
pgk
4
3
BamHI
HindIII
BamHI
9
0
8
qutE
1
BglII
7
pAL3.3-B3
9.3 Kb
ori
2
amp
3
6
cml
BamHI
HindIII
5
4
BamHI-HindIII digest
Isolate 0.7 Kb fragment with 5'pgk sequence
BglII-HindIII digest
Isolate 7.45 Kb fragment with qutE sequence
Ligation reaction
BamHI
8
0
7
qutE
1
p5'PGK-QUTE
8.15 Kb
6
ori
2
pgk'
HindIII
cml
amp
5
3
4

The most direct way to do this would be to use the Klenow fragment of DNA polymerase I of *E.coli* to fill in the single stranded ends of both *HindIII* linearized p5'PGK-QUTE and the *BamHI-XbaI lacZ* fragments from the pAN923 vectors. Ligation reactions between the blunt ended molecules should then generate the required constructs. However, such reactions may be very inefficient (Maniatis *et al.*, 1982), and therefore alternative routes were explored.

The first strategy that was attempted is outlined in Figure 4.2. The *BamHI-XbaI lacZ* fragment was incubated in a fill-in reaction with the Klenow fragment of DNA polymerase I to convert single stranded ends to blunt ends. The fragment was recovered from the reaction mixture by extraction with phenol and chloroform and by precipitation with ethanol, and was then incubated in a ligation reaction with a synthetic *HindIII* linker molecule, the blunt ends of which had been phosphorylated by T4 DNA kinase. The modified fragment was recovered by phenol/chloroform extraction and ethanol precipitation, digested with *HindIII*, recovered again, and incubated in a ligation reation with *HindIII*-linearized, alkaline phosphatase-treated p5'PGK-QUTE. However, in spite of varying the conditions of the two ligation reactions, no colonies were isolated carrying the desired vector when the putative constructs were transformed into *E.coli*. The Klenow fill-in reaction and the blunt end ligation of linker molecules to the *lacZ* fragment were difficult steps to monitor and are considered to be the two most likely points of failure in the procedure.

The second strategy that was attempted, outlined in Figure 4.3, used another synthetic DNA linker molecule with *XbaI* and *BamHI* single stranded ends and an internal *HindIII* site. This molecule was made by annealing two complementary oligonucleotides and phosphorylating the ends with T4 DNA kinase. Addition of linkers to the 3.9Kb *BamHI-XbaI lacZ* fragment and digesting with *HindIII* should generate a *lacZ* fragment with *HindIII* single stranded ends, as sought in the previous strategy. However, in this case a blunt end

Figure 4.2
Proposed route of construction of fusion vectors using blunt ended DNA linker molecules.

The vector series pAN923-41A, -42A, and -43A was constructed by van Gorcom *et al.* (1986). The synthetic DNA linker molecule was obtained from BRL. It includes an internal *HindIII* restriction enzyme site and is a 10mer with the sequence:

HindIII
↓
d5′-CCAAGCTTGG-3′
3′-GGTTCGAACC-5′d

Linkers were incubated with T4 DNA kinase to phosphorylate their 5′ ends prior to ligating with *lacZ* fragment.

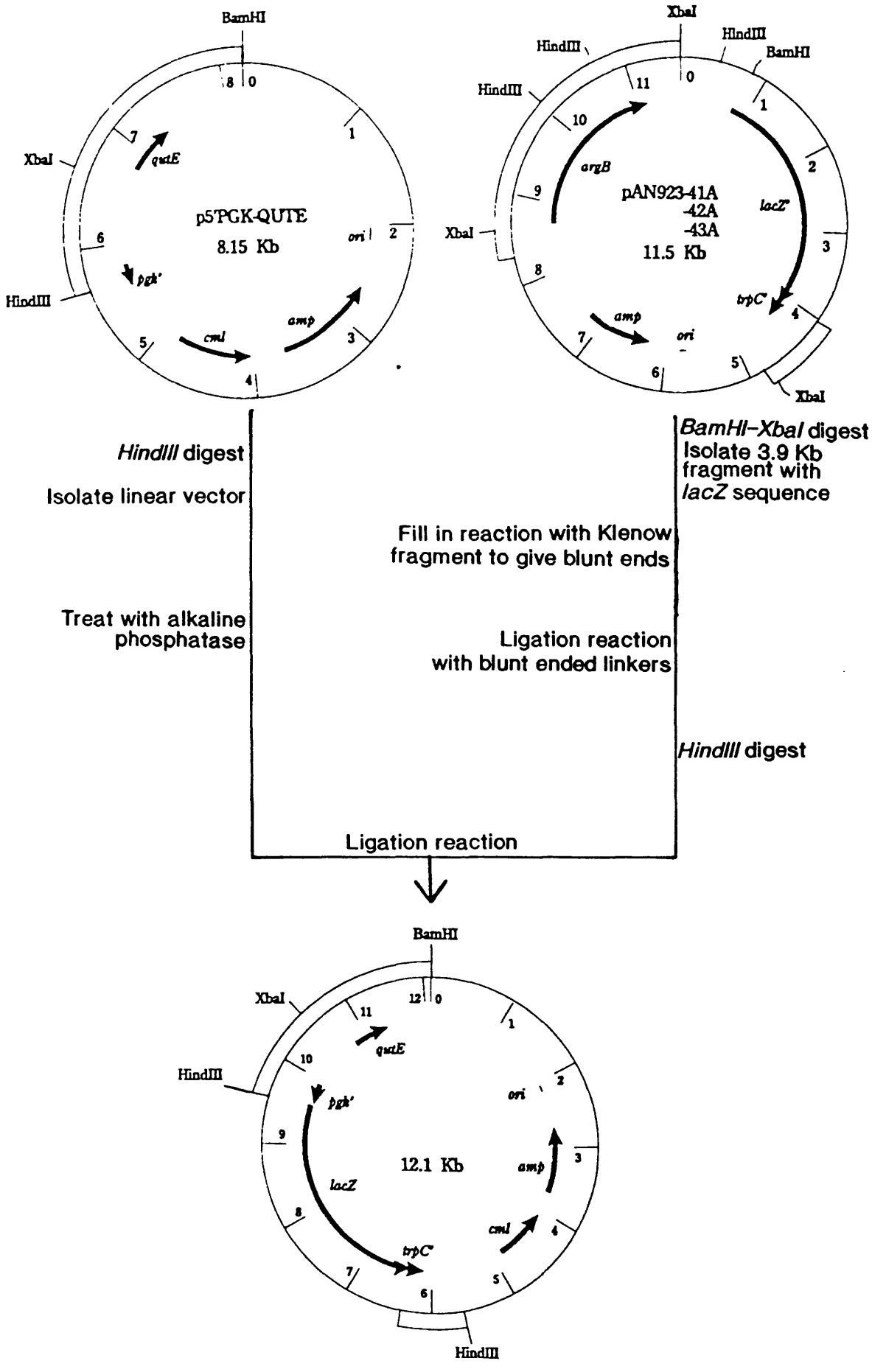
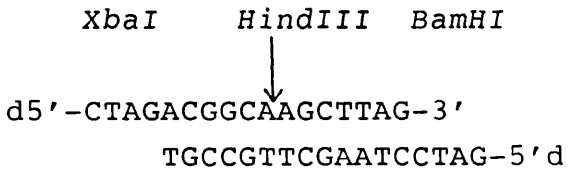
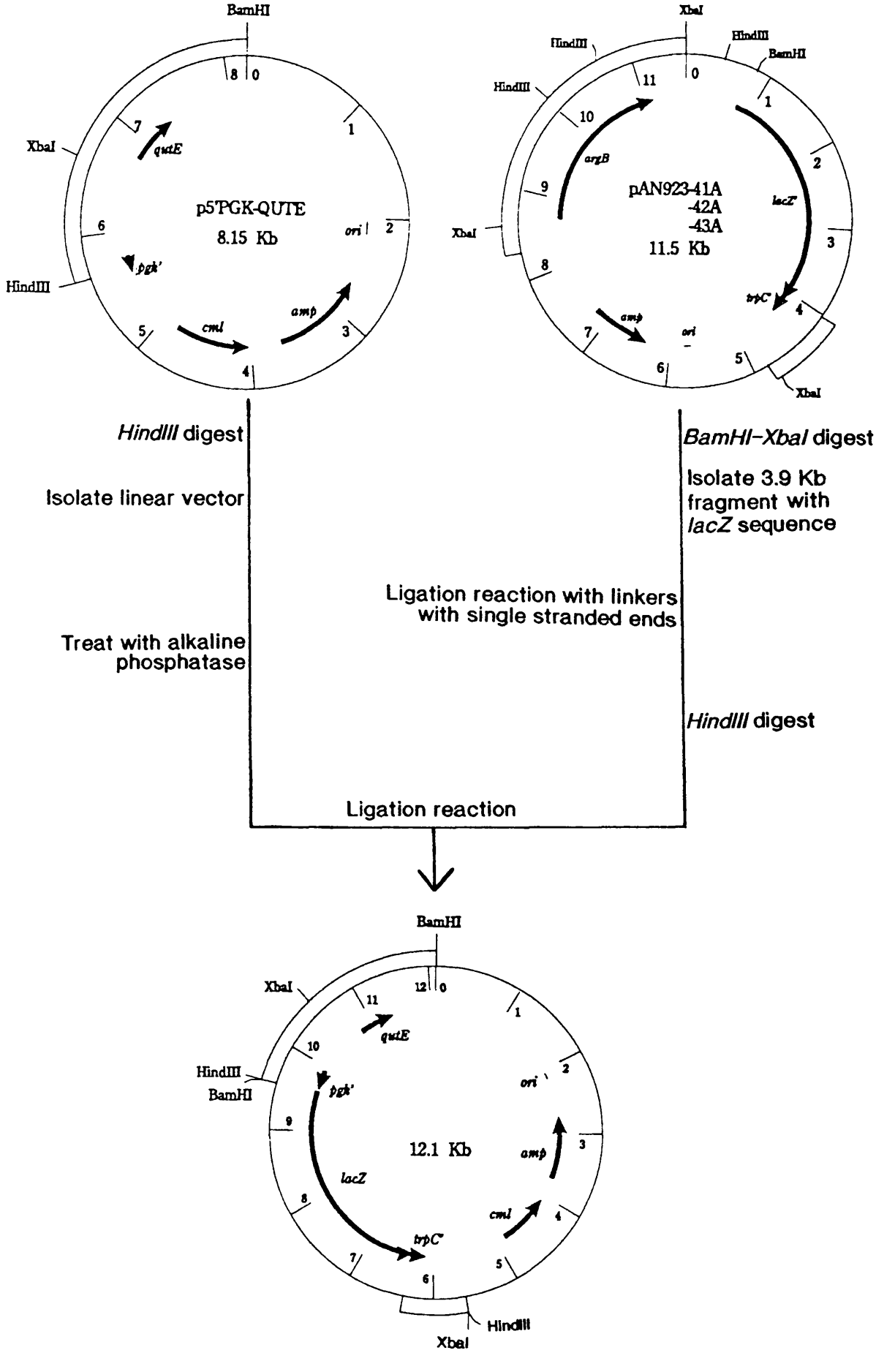

BamHI
XbaI
HindIII
p5'PGK-QUTE
8.15 Kb
qutE
pgk'
cml
amp
ori
pAN923-41A
-42A
-43A
11.5 Kb
argB
lacZ'
trpC'
HindIII digest
Isolate linear vector
Treat with alkaline phosphatase
BamHI-XbaI digest
Isolate 3.9 Kb fragment with lacZ sequence
Fill in reaction with Klenow fragment to give blunt ends
Ligation reaction with blunt ended linkers
HindIII digest
Ligation reaction
12.1 Kb
lacZ

Figure 4.3
Proposed route of construction of the fusion vectors using DNA linker molecules with cohesive single stranded ends.

The vector series pAN923-41A, -42A and -43A was constructed by van Gorcom *et al.* (1986). The DNA linker molecule was designed to have single stranded *XbaI* and *BamHI* ends and to include an internal *HindIII* site. It was made by annealing two oligonucleotides and phosphorylating the ends with T4 DNA kinase. The linker had the sequence:




BamHI
XbaI
HindIII
p5'PGK-QUTE
8.15 Kb
qutE
pgk'
cml
amp
ori
XbaI
HindIII
HindIII
BamHI
pAN923-41A
-42A
-43A
11.5 Kb
argB
lacZ'
trpC'
amp
ori
XbaI
HindIII digest
Isolate linear vector
Treat with alkaline phosphatase
BamHI-XbaI digest
Isolate 3.9 Kb fragment with lacZ sequence
Ligation reaction with linkers with single stranded ends
HindIII digest
Ligation reaction
BamHI
XbaI
HindIII
BamHI
12.1 Kb
qutE
pgk'
lacZ
trpC'
cml
amp
ori
HindIII
XbaI

ligation reaction was avoided. The fragment was then ligated into *HindIII* linearized, alkaline phosphatase treated p5'PGK-QUTE as before. The linkers were successfully attached to *lacZ* fragment, as demonstrated by the change in mobility of fragment following ligation (Figure 4.4). However, no transformants carrying the desired construct were isolated from *E.coli* transformed with the *lacZ* fragment/linear p5'PGK-QUTE ligation mixes, even when the extent of alkaline phosphatase treatment of the vector molecule was varied.

Since linker molecules were known to have ligated to the *lacZ* fragment, the most likely point of failure in this strategy is in the alkaline phosphatase treatment of vector molecules. Insufficient enzyme would result in preferential vector recircularization, whilst too much may lead to exonuclease digestion into the single stranded ends of the DNA molecule.

4.2.3 The construction of fusion vectors.

In light of the difficulties experienced in attempting to construct fusion vectors by the strategies described above, a new approach was devised which obviated the use of DNA linkers and the alkaline phosphatase treatment of vector molecules. This successful strategy is outlined in Figures 4.5 and 4.6.

In order to generate a *qutE* vector with a unique *BamHI* site the plasmid pAL3.3-B4 (Hawkins *et al.*, 1985) was digested with *EcoRI* and the larger fragment carrying the *qutE* gene of *A.nidulans* was re-ligated to give a 7.15Kb vector, which was designated pAL3.3-B4Δ. Then the 3.95Kb *BamHI-XhoI* fragments from each of the three pAN923 vectors (van Gorcom *et al.*, 1986) were introduced into a 6.85Kb *BamHI-SalI* fragment of pAL3.3-B4Δ, generating a series of vectors carrying the *qutE* gene of *A.nidulans* and the *lacZ* gene of *E.coli*. In these constructs a unique *BamHI* site lies upstream of the *lacZ* gene in each translational reading frame. I designated these vectors pLACZ-QUTE1, 2 and 3.

Figure 4.4
The ligation of linker molecules to *lacZ* fragment.

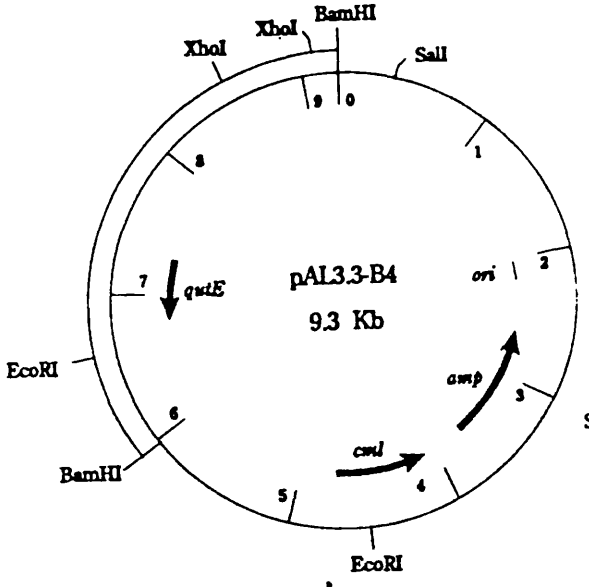
The lanes are:

1- 0.1g of the 3.9Kb *BamHI-XbaI* fragment of pAN923-43A
2- 0.1g of this fragment incubated in a ligation reaction
3- 0.6g of this fragment incubated in a ligation reaction with DNA linker molecules with *BamHI-XbaI* single stranded ends.

The DNA fragments were separated by agarose gel electrophoresis. The original fragment is faintly visible in lane 1, and the modified fragment is clearly visible with greater molecular weight in lane 3.

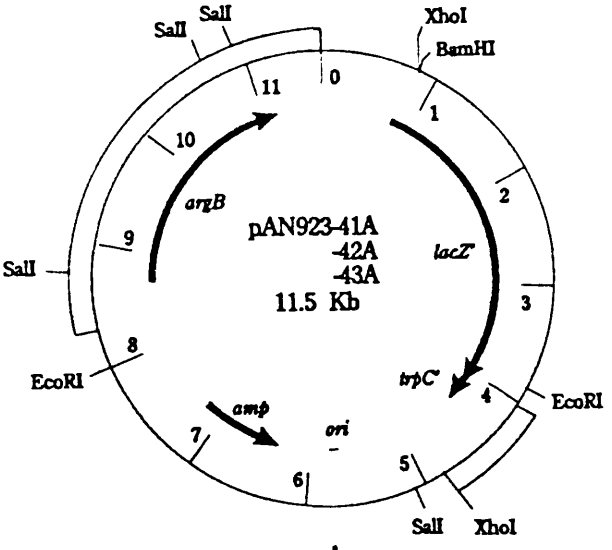
Figure 4.5
Construction of the vectors pLACZ-QUTE1, 2 and 3.

The vector pAL3.3-B4 consists of a 3.3Kb fragment spanning the *qutE* gene of *A.nidulans* subcloned into the plasmid pBR325 (Hawkins *et al.*, 1985), and the vector series pAN923-41A, -42A and -43A was constructed by van Gorcom *et al.* (1986).
The restriction enzymes *SalI* and *XhoI* produce complementary single stranded ends.



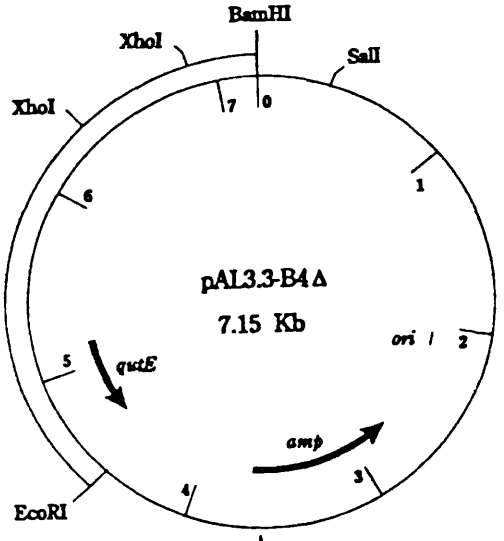
EcoRI digest
Isolate 7.15 Kb fragment
with *qutE* sequence

Self-ligation
reaction



BamHI-XhoI digest

Isolate 3.95 Kb fragment
with *lacZ* sequence



BamHI-SalI digest
Isolate 6.85 Kb fragment
with *qutE* sequence

Ligation reaction

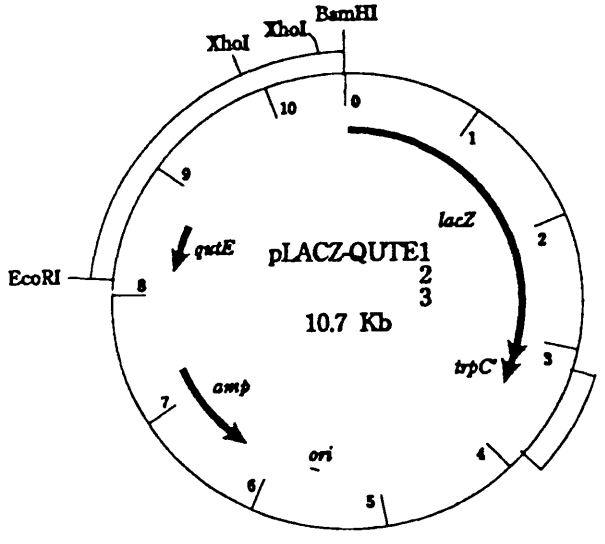
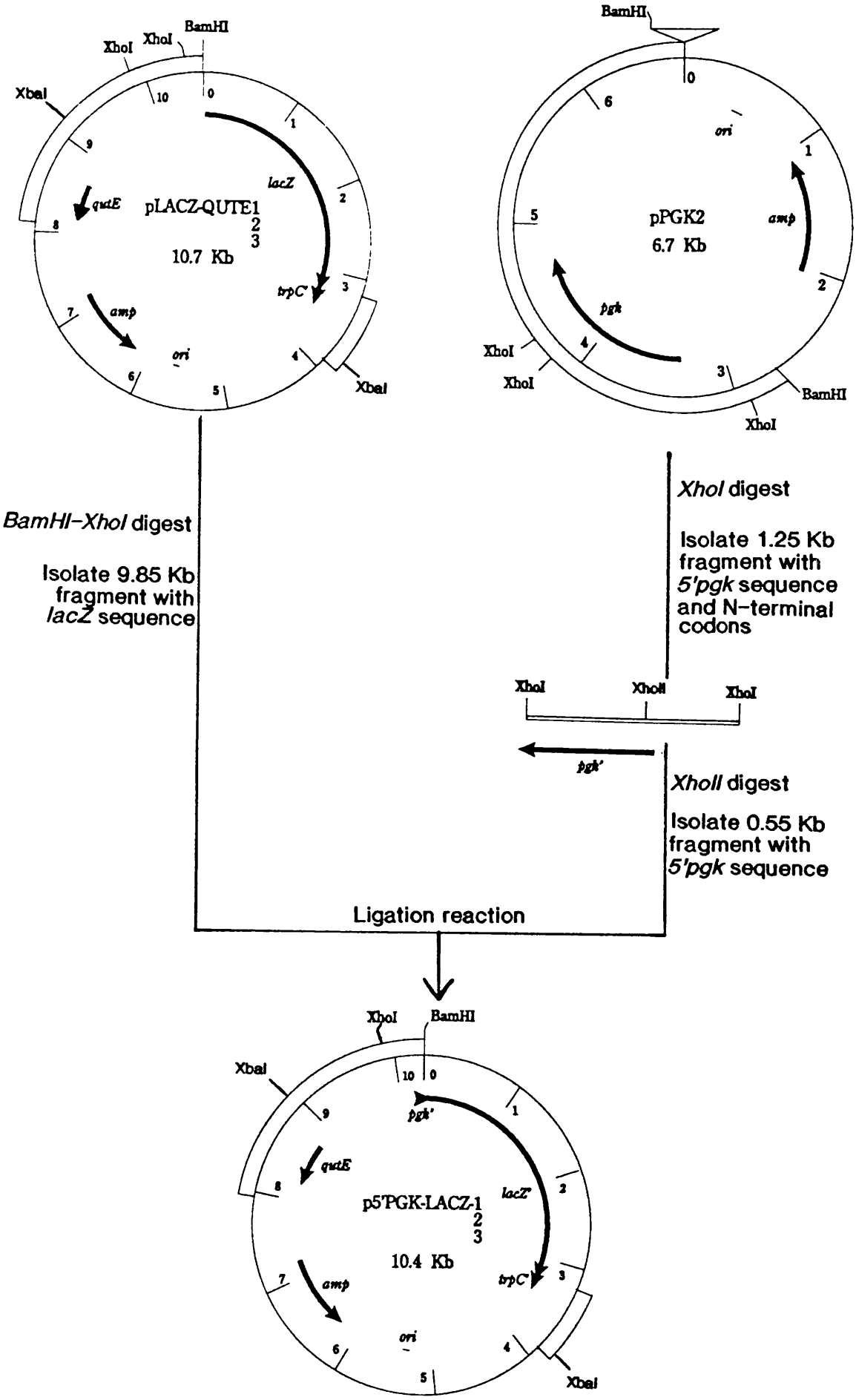


Figure 4.6
Construction of the vectors p5'PGK-LACZ1, 2 and 3.

The vector pPGK2 consists of a 4.0Kb fragment spanning the *pgk* gene of *A.nidulans* subcloned into the plasmid pUC13 (Clements and Roberts, 1985).
The 0.55Kb *XhoI-XhoII* fragment with 5'*pgk* sequence could not be isolated directly because of the presence of numerous other *XhoII* sites in the vector pPGK2 (not shown in the Figure). A *XhoI-XhoII* digest of pPGK2 generates additional fragments of approximately the size of the one to be isolated. Therefore, a 1.25Kb *XhoI* fragment was isolated spanning 5'*pgk* sequence and N-terminal codons, and then the required 0.55Kb fragment isolated by digesting this first fragment with *XhoII*.
The restriction enzymes *BamHI* and *XhoII* produce complementary single stranded ends, and in the case of this ligation reaction the *BamHI* site is reconstituted.


XhoI
XhoI
BamHI
XbaI
pLACZ-QUTE1
2
3
10.7 Kb
qutE
lacZ
trpC'
amp
ori
XbaI
BamHI
pPGK2
6.7 Kb
ori
amp
pgk
XhoI
XhoI
BamHI
XhoI
XhoI digest
Isolate 1.25 Kb fragment with 5'pgk sequence and N-terminal codons
XhoI
XhoII
XhoI
pgk'
XhoII digest
Isolate 0.55 Kb fragment with 5'pgk sequence
BamHI-XhoI digest
Isolate 9.85 Kb fragment with lacZ sequence
Ligation reaction
XhoI
BamHI
XbaI
pgk'
qutE
lacZ'
p5'PGK-LACZ-1
2
3
10.4 Kb
amp
trpC'
ori
XbaI

The 0.55Kb *XhoI-XhoII* fragment of the vector pPGK2, which consists of 487bp of sequence 5′ to the *pgk* transcription start site, the 32bp leader sequence and the 13 N-terminal codons of the structural gene (Clements and Roberts, 1985), was ligated into the 9.85Kb *BamHI-XhoI* fragment isolated from each of the pLACZ-QUTE plasmids (Figure 4.6). This vector series was designated p5′PGK-LACZ1, 2 and 3. The construct p5′PGK-LACZ2 is predicted to have *pgk* fused to *lacZ* in the correct reading frame.

In all the work described in this Section, newly made vectors were verified by comparing the fragments generated by restriction enzyme digests with the predicted fragment sizes. A photograph of an agarose gel following electrophoresis of restriction enzyme digests of the final constructs (p5′PGK-LACZ1, 2 and 3) is presented in Figure 4.7, together with the fragment sizes deduced from the gel and the predicted fragment sizes.

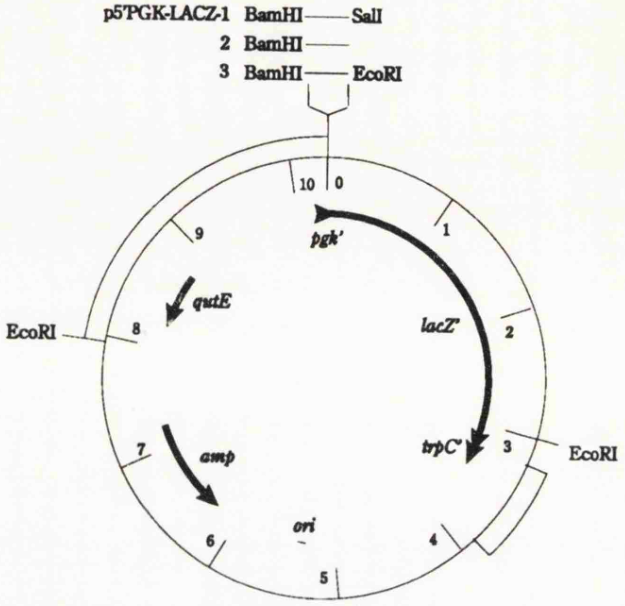
The final constructs have the *A.nidulans pgk* promoter, the untranslated leader sequence and 13 N-terminal codons fused in each translational reading frame to *lacZ*. The *qutE* gene is also present on the constructs to allow for selection in *A.nidulans*, and the ampicillin resistance gene and replication origin of pBR322 are present to allow for plasmid selection and propagation in *E.coli*. A unique *XhoI* site lies upstream of putative control sequences in the *pgk* promoter, and provides an entry point for exonuclease digestion into potential promoter sequences.

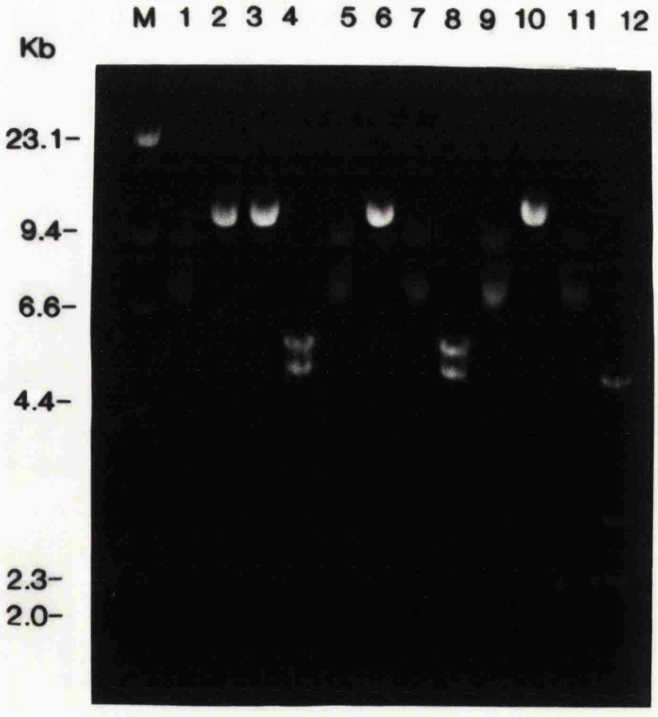
Figure 4.7
Analysis of the 5'*pgk-lacZ* fusion vectors by agarose gel electrophoresis of restriction enzyme digest products.

Each of the vectors had been purified by centrifugation to equilibrium in cesium chloride - ethidium bromide gradients. A sample of 0.5μg of vector DNA was incubated in each digest reaction, and all of this was loaded onto the gel.
The lanes are:
M -0.5μg of phage lambda DNA digested with *HindIII*.
1,2,3,4 -p5'PGK-LACZ1: undigested, *BamHI*, *SalI*, *EcoRI*.
5,6,7,8 -p5'PGK-LACZ2: undigested, *BamHI*, *SalI*, *EcoRI*.
9,10,11,12 -p5'PGK-LACZ3: undigested, *BamHI*, *SalI*, *EcoRI*.

The table shows the fragment sizes predicted from a diagram of the vectors, and those calculated from the restriction enzyme analysis.

Vector digest	Predicted fragment sizes (Kb)	Observed fragment sizes (Kb)
p5'PGK-LACZ1: *BamHI*	10.4	10.5
: *SalI*	10.4	10.5
: *EcoRI*	5.5, 4.9	5.8, 5.2
p5'PGK-LACZ2: *BamHI*	10.4	11.0
: *SalI*	No site	No site
: *EcoRI*	5.5, 4.9	5.7, 5.1
p5'PGK-LACZ3: *BamHI*	10.4	11.5
: *SalI*	No site	No site
: *EcoRI*	4.9, 3.0, 2.5	5.0, 3.0, 2.4


p5'PGK-LACZ-1 BamHI SalI
2 BamHI
3 BamHI EcoRI
10
0
1
2
3
4
5
6
7
8
9
pgk'
lacZ'
trpC'
ori
amp
qutE
EcoRI
EcoRI


M 1 2 3 4 5 6 7 8 9 10 11 12
Kb
23.1-
9.4-
6.6-
4.4-
2.3-
2.0-

4.3 Expression of gene fusions in *Escherichia coli*.

Prior to transforming the fusion constructs into *Aspergillus nidulans*, they were transformed into *Escherichia coli* to test if *lacZ* could be expressed when located downstream of a strong fungal promoter in a prokaryotic host.

The *E.coli* strain JM83, which is deleted for the *lac* operon (Viera and Messing, 1982), was transformed with the *pgk-lacZ* fusions p5'PGK-LACZ1, 2 and 3, and transformants were selected on plates containing ampicillin and X-gal. Incubation of the strain transformed with p5'PGK-LACZ2 gave faint blue colonies after 24 hours and dark blue colonies after 48 hours, but colonies with p5'PGK-LACZ1 or 3 remainded clear. Thus, only transformants carrying the vector with the *pgk* promoter fused to *lacZ* in the reading frame predicted as being in phase produced β-galactosidase. This observation does not necessarily prove that expression of *lacZ* in *E.coli* is being driven by the *A.nidulans pgk* promoter, since read-through from pBR325 sequence could generate *lacZ* expression. However, it does confirm the success of the gene fusion strategy.

To test if expression of *lacZ* depended on the presence of *pgk* promoter sequence upstream, the vectors pLACZ-QUTE1, 2 and 3 (which lack 5'*pgk* sequence upstream of *lacZ*) were transformed into JM83 and colonies selected as before. After 48 hours incubation all transformants remained clear.

The implication of these observations is that the expression of *lacZ* in *E.coli* transformants is being driven by the *A.nidulans pgk* promoter. However, to confirm that Aspergillus transcriptional control sequences are operating in *E.coli*, the transcript start site would have to be mapped; an exercise not undertaken here.

4.4 Construction of an *Aspergillus nidulans* host strain for transformation of the gene fusions.

In order to determine the activity of the β-galactosidase fusion protein in *Aspergillus nidulans*, a recipient strain was required lacking both catabolic dehydroquinase (*qutE*$^-$) and native β-galactosidase (*bgaA*$^-$). The fusion vectors could then be selected in Aspergillus by growth with quinic acid as carbon source, and the β-galactosidase activity of the fusion protein could be measured in the absence of background from the native enzyme.

The A.*nidulans qutE*$^-$ mutant strain WA53 (Da Silva *et al.*, 1986) was crossed with the *bgaA*$^-$ mutant strain R21*bgaA* (Fantes and Roberts, 1973) and *qutE*$^-$, *bgaA*$^-$ progeny were identified by screening for growth on quinic acid and for β-galactosidase production. One such strain, which was designated QG716, was chosen for subsequent transformation experiments. The genotype of this strain is *pyrG89, pabaA1, yA; bgaA4; qutE208*.

4.5 Design of a method to detect the production of β-galactosidase in *Aspergillus nidulans*.

Before introducing the *lacZ* fusion constructs into *Aspergillus nidulans*, a method had to be developed which would easily identify transformants with β-galactosidase activity. Two procedures for detecting enzyme activity in colonies grown on solid media were examined.

A histochemical technique, outlined in Chapter 2, for determining whether a colony grown on an agar medium is producing β-galactosidase had been developed in the laboratory and had been used to study the expression of the β-galactosidase gene (*bgaA*) in *A.nidulans* (Fantes and Roberts, 1973). Expression of the gene was shown to be repressed by glucose, glycerol and sucrose, and induced by lactose and galactose.

This histochemical technique was applied to distinguish between a wild-type and a *bgaA*$^-$ mutant strain grown on different carbon sources. Since the mutant strain does not grow at all well on lactose or galactose alone, it was necessary to provide an alternative carbon source such as glycerol to facilitate colony growth. The results are presented in Table 4.1 and verify those of Fantes and Roberts (1973) given above. Although the method enabled colonies producing β-galactosidase to be detected, it was often difficult to distinguish the degree of staining. Moreover, the procedure would be time consuming if used to to screen a large number of *A.nidulans* transformants.

The second method that was developed to detect β-galactosidase activity in *A.nidulans* used the chromogenic substrate X-gal and is described in Chapter 2. Suprisingly, when a wild-type strain was grown on Aspergillus minimal agar plates in the presence of X-gal, colonies stained blue irrespective of whether glucose or galactose was the carbon source. This is contrary to the observation of Fantes and Roberts (1973), that β-galactosidase is induced by galactose but is repressed by glucose when the wild-type is grown in

Table 4.1

The effect of carbon source upon the production of β-galactosidase in A.*nidulans*.

The wild type *A.nidulans* strain R21 and the mutant strain R21*bgaA* (Fantes and Roberts, 1973), were tested for β-galactosidase production on the stated carbon sources. The strains were incubated at 37°C for 3 days on Aspergillus minimal agar plates and were examined using the histochemical staining technique of Fantes and Roberts (1973), as described in Chapter 2.

The intensity of staining was indicated by +, ++, +++, and no colony staining by -. Poor colony growth is indicated by (*).

Carbon source for growth	Intensity of colony staining	
	Wild type	*bgaA*$^-$ mutant
Glycerol (0.04M)	-	-
Galactose (0.02M)	+++	- (*)
Galactose (0.02M) Glycerol (0.04M)	++	-
Galactose (0.02M) Fructose (0.02M)	+	-
Lactose (0.02M)	++	- (*)
Lactose (0.02M) Glycerol (0.04M)	+	-

liquid minimal medium. The *bgaA*$^-$ mutant strain did not stain blue when it was grown on either carbon source. However, when grown on the bacterial minimal agar medium M9 in the presence of X-gal, neither strain stained blue when glucose was the carbon source and only the wild-type strain stained blue with galactose. These observations verify the results of van Gorcom *et al.* (1985) concerning the staining of Aspergillus colonies when grown on M9 agar media in the presence of X-gal.

One of the principal differences between the Aspergillus and bacterial minimal agar media is the pH by which the media is shifted during the growth of an Aspergillus colony. After two days of growth the bacterial medium was shown to have shifted from a pH of 7.2 to 5.5, whereas the Aspergillus medium had shifted from a pH of 6.5 to 7.0. These changes are thought to reflect the nature of the nitrogen sources in the media. In the Aspergillus medium nitrate ions are replaced by hydroxide ions during growth, whilst in the bacterial medium, ammonium ions are replaced by hydrogen ions.

To test whether the pH difference was responsible for the discrepancy in staining on X-gal, the wild-type and *bgaA*$^-$ mutant strains were grown on Aspergillus minimal agar medium which had been buffered with citric acid/sodium hydroxide over a range of pH values. The results are presented in Table 4.2. At a pH in the range of 4.8 to 6.8, neither strain stained blue when grown on glucose, and only the wild-type stained blue on galactose. However, at a pH of 7.2 or 7.6, the wild-type stained blue on either carbon source. Thus, the pH of the medium does appear to affect the staining of Aspergillus colonies on X-gal. In all subsequent tests, strains were grown on Aspergillus media buffered at a pH of 6.0 with citric acid/ sodium hydroxide, and with glucose as carbon source.

Table 4.2

The effect of pH upon the production of β-galactosidase in A.nidulans.

The wild type *A.nidulans* strain R21 and the mutant strain R21*bgaA* (Fantes and Roberts, 1973), were tested for β-galactosidase production on Aspergillus minimal agar media over the stated range of pH values with glucose or galactose as carbon source. The media was buffered with 0.1M citric acid, 0.2M sodium hydrogen phosphate. The strains were incubated at 37^{o}C for 3 days and were examined by including the chromogenic substrate X-gal (40μgml^{-}1) in the growth media, as described in Chapter 2.

The *bgaA* strain grew poorly on galactose.

pH of media	Intensity of colony staining			
	Wild type		*bgaA*$^-$ mutant	
	Glucose (0.02M)	Galactose (0.02M)	Glucose (0.02M)	Galactose (0.02M)
4.8	-	+++	-	-
5.2	-	+++	-	-
5.6	-	+++	-	-
6.0	-	+++	-	-
6.4	-	+++	-	-
6.8	-	+++	-	-
7.2	+	++	-	-
7.6	+	+	-	-

4.6 Expression of gene fusions in *Aspergillus nidulans*.

The gene fusion vectors p5'PGK-LACZ1, 2 and 3 (described in Section 2), which have the promoter and N-terminal sequence of the *Aspergillus nidulans* 3-phosphoglycerate kinase gene fused in each translational reading frame to the *lacZ* structural gene of *Escherichia coli*, were transformed into the recipient *bgaA*$^-$, *qutE*$^-$ *A.nidulans* strain QG716. Several transformation experiments were done selecting for growth on quinic acid, and 8 strains transformed with p5'PGK-LACZ1, 30 with p5'PGK-LACZ2 and 8 with pPGK5'LACZ3 were isolated. Circular vector molecules were used for the p5'PGK-LACZ2 transformtion experiments, but for reasons described in Chapter 5, linear vector molecules were used for the p5'PGK-LACZ1 and 3 transformations.

The transformed strains were serially replated twice on selective medium to isolate single colonies, and were then screened for the production of β-galactosidase on Aspergillus minimal medium at pH6.0 with glucose as carbon source and containing $40\mu gml^{-1}$ X-gal. Of the 30 strains transformed with p5'PGK-LACZ2 12 stained blue, whereas all 16 of the p5'PGK-LACZ1 and p5'PGK-LACZ3 transformants were negative. Thus, as expected, only transformants with *pgk* fused to *lacZ* in the correct reading frame expressed β-galactosidase activity, although even then, only in 40% of the isolates. Some of the transformants grown on X-gal plates are shown in Figure 4.8.

In order to verify that *lacZ* sequence had been stably transformed into *A.nidulans*, genomic DNA was prepared from the original strain and from all transformed strains and was denatured and spotted onto a Hybond-N filter. The DNA on the filter was hybridized to a radioactively labelled probe. The DNA fragment used as probe was isolated from an agarose gel as a 3.0Kb *EcoRI* fragment consisting entirely of *E.coli lacZ* sequence. An autoradiographic image of the filter is presented in Figure 4.9. Among the 30 p5'PGK-LACZ2 transformants 16 had incorporated *lacZ* sequence (Figure

Figure 4.8
The expression of *lacZ* fusion genes in transformed strains of *A.nidulans.*

Strains were tested for β-galactosidase production on Aspergillus minimal agar medium (buffered at pH6.0) with glucose as carbon source. The strains were incubated at 37°C for 3 days and were examined by including the chromogenic substrate X-gal (40μgml^{-1}) in the growth medium (Chapter 2).

The strains were:
1 -The untransformed strain QG716
2 -QG716 transformed with p5'PGK-LACZ1
3,5 -two strains of QG716 transformed with p5'PGK-LACZ2
4 -QG716 transformed with p5'PGK-LACZ3

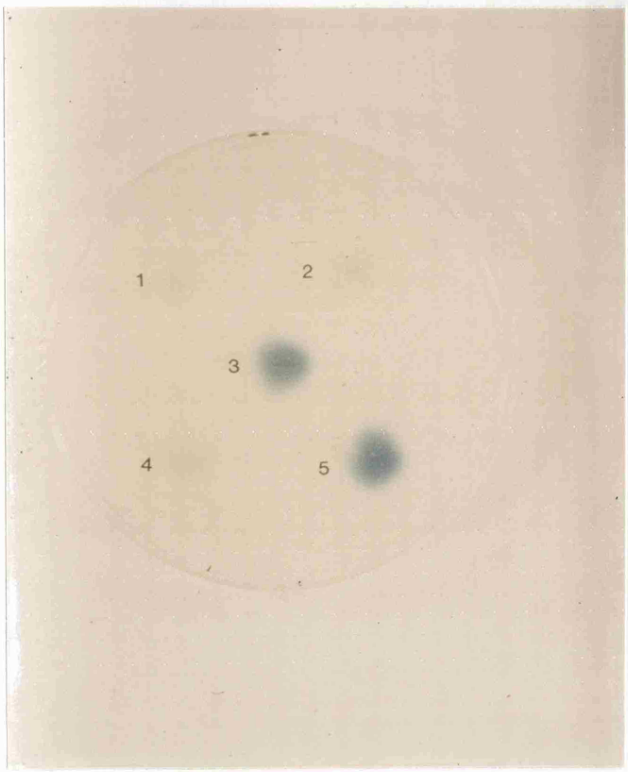
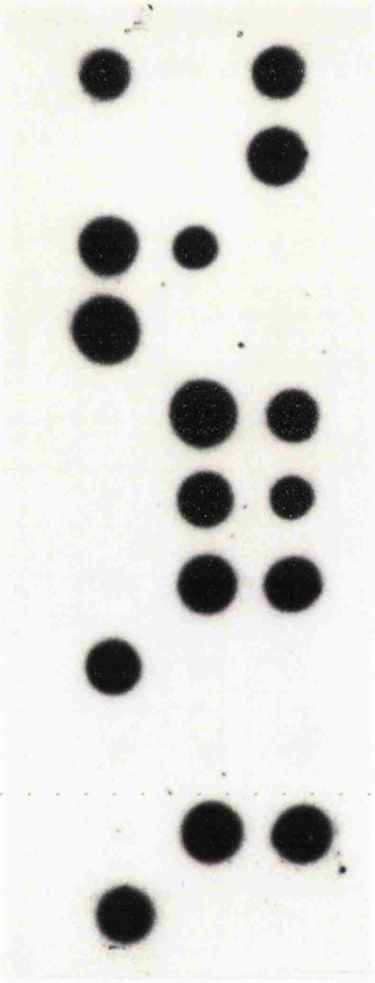

1
2
3
4
5

Figure 4.9

The incorporation of *lacZ* sequence into transformed strains of *A.nidulans.*

The autoradiograph shows the hybridization of an α^{32}P-dCTP labelled 3.0Kb *EcoRI* fragment consisting entirely of *E.coli lacZ* sequence (isolated from the vector p5'PGK-LACZ3), with genomic DNA of transformed strains of *A.nidulans*. The strains had been recovered from a transformation experiment of the *qutE*$^-$ mutant QG716 with the circular in phase fusion vector p5'PGK-LACZ2 by selection for growth on quinic acid, and were designated 5'PGK-LACZ2:C1 to C30. The *A.nidulans* recipient strain QG716 was included in the analysis as a control, and shows no sequence homology to the *lacZ* gene of *E.coli*.



C1	QG716	C2
C3	C4	C5
C6	C7	C8
C9	C10	C11
C12	C13	C14
C15	C16	C17
C18	C19	C20
C21	C22	C23
C24	C25	C26
C27	C28	C29
C30	-	-

4.9), whilst of the 8 p5'PGK-LACZ1 transformants 6 had incorporated *lacZ*, and similarly for 6 of the 8 p5'PGK-LACZ3 transformants (data not shown).

The untransformed recipient strain neither expressed β-galactosidase nor gave a positive signal when genomic DNA was hybridized to radioactivly labeled *lacZ* sequence. Furthermore, no transformant expressed β-galactosidase unless it had also incorporated *lacZ* sequence. However, not all transformants which had acquired *lacZ* necessarily expressed β-galactosidase. In all, 6 out of 8 of each of the p5'PGK-LACZ1 and p5'PGK-LACZ3 transformants contained *lacZ* sequence, and as expected none of these expressed β-galactosidase. Also, 16 out of the 30 p5'PGK-LACZ2 transformants had incorporated *lacZ* sequence, but only 12 of these 16 expressed β-galactosidase. Thus, although incorporation of gene fusion sequence in the correct translational reading frame appears to be a prequisite for β-galactosidase expression, it is not alone sufficient.

It is possible that the 4 transformants that have incorporated *lacZ* sequence from the in phase fusion vector but do not express β-galactosidase, may have vector integrated into the host genome in a transcriptionally silent region. Alternatively, the gene fusion sequence may have been damaged upon integration.

4.7 Summary.

Gene fusion vectors have been succesfully constructed, in all three translational reading frames, between sequence 5′ to the 3-phosphoglycerate kinase gene of *Aspergillus nidulans* and the *lacZ* gene of *Escherichia coli*. The *lacZ* gene was expressed when the in phase fusion vector was transformed into *E.coli*, and this expression was shown to depend upon the presence of 5′*pgk* sequence.

In transformation experiments with *A.nidulans*, 30 transformed strains were isolated with the in phase fusion vector. Among these transformants, 16 were shown to have incorporated *lacZ* sequence, and 12 of these expressed β-galactosidase. The immediate aim was to determine the fate of DNA sequences transformed into *A.nidulans* in order to see if this affected the expression of β-galactosidase. This analysis is described in Chapter 5.

Chapter 5.

Analysis of transformed strains of *Aspergillus nidulans*.

5.1 Introduction.

In order to analyse the *pgk* promoter of *Aspergillus nidulans* the gene fusion constructs described in Chapter 4 (and modified versions of these) had to be introduced into the host organism. The transformation procedure used was that developed by Ballance *et al.* (1983). Spheroplasts were generated using the enzyme preparation Novozym 234 to lyse hyphal walls, and polyethylene glycol and calcium ions were included to stimulate DNA uptake. Potassium chloride (0.6M) was used as an osmotic stabilizer during the production and transformation of spheroplasts and in the subsequent regeneration of mycelium. A suitable *bgaA*$^-$; *qutE*$^-$ recipient strain of *A.nidulans* (QG716) has been isolated (Chapter 4), and transformants were selected for growth on quinic acid medium. Transformed strains were routinely replated twice on selective medium to isolate single colonies, and were subsequently maintained on the same selective medium.

Stable transformed strains of Aspergillus can be obtained if vector DNA integrates into a chromosome. The nature and site(s) of the integration event(s) can be investigated by DNA hybridization, which provides a basis for typing transformed strains into one of three classes (Chapter 1, Figure 1.2). A single homologous recombination event involving the Aspergillus selective marker on the circular plasmid molecule would result in the incorporation of vector sequence at the corresponding host locus. This is referred to as a type I event following the classification of Hinnen *et al.* (1978). Integration of DNA elsewhere in the genome, possibly by recombination at a site of chance homology, is a type II event. A type III event refers to the replacement of a host gene with the corresponding sequence on the vector, possibly as a result of double homologous recombination, or by gene conversion. In this case no vector sequences are found in the chromosome.

5.2 Analysis of *Aspergillus nidulans* strains transformed with the circular in phase fusion vector.

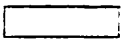
A total of 30 strains isolated by transforming the circular in-phase fusion vector into the *Aspergillus nidulans* recipient strain QG716 had been screened for β-galactosidase production on X-gal plates and for the incorporation of *lacZ* sequence by DNA hybridization (Chapter 4). Among the transformants, 16 had incorporated *lacZ* sequence, and 12 of these expressed β-galactosidase.

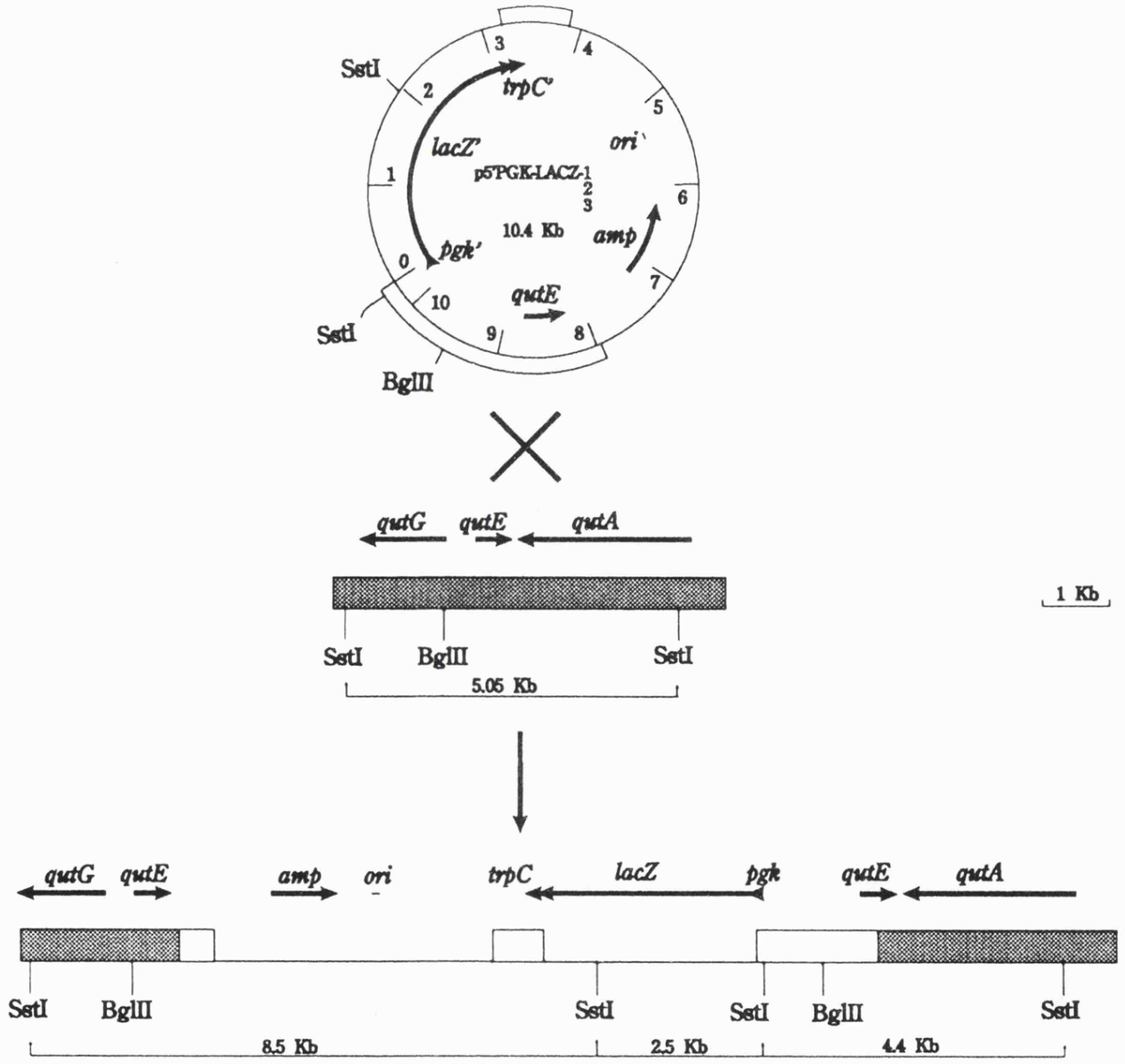
5.2.1 Analysis of genomic DNA isolated from transformed strains.

In order to establish a basis for the analysis of transformed strains, the sizes of the restriction enzyme digest fragments that would be predicted for the recipient strain and for transformed strains were calculated. From the known location of *SstI* restriction enzyme sites in the *qut* gene cluster, it can be predicted that the untransformed strain should give a 5.05Kb band on the autoradiograph. Since the enzyme *SstI* does not cut within *qutE* or neighbouring Aspergillus sequence on the vector molecule, a single copy type I integration event should replace this native band with two new bands of 8.5Kb and 4.4Kb. A multiple copy type I event is expected to generate both of these new bands plus an additional band of 7.9Kb, the intensity of which would correspond to the number of extra copies of the vector molecule that have integrated. A type II integration event is expected to give two or more bands: one corresponding to the undisturbed native *qutE* locus and other(s) representing the site(s) of vector integration. A type III, gene replacement, event would be indistinguishable from the untransformed strain. The type I integration of a single gene fusion construct into the *qutE* locus is outlined in Figure 5.1.

Genomic DNA prepared from each of the 30 transformed strains was digested with the restriction enzyme *SstI* and the products were separated by agarose gel electrophoresis.

Figure 5.1
Integration of the *pgk-lacZ* fusion construct at the *qutE* locus in the genome (a type I event).

In this figure:
sequence of bacterial origin is represented as ———
A.nidulans sequence located on the vector as 
and genomic *A.nidulans* sequence as 


SstI
trpC'
lacZ'
ori'
p5'PGK-LACZ-1
2
3
10.4 Kb
amp
pgk'
qutE
SstI
BglII
qutG
qutE
qutA
1 Kb
SstI
BglII
SstI
5.05 Kb
qutG
qutE
amp
ori
trpC
lacZ
pgk
qutE
qutA
SstI
BglII
SstI
SstI
BglII
SstI
8.5 Kb
2.5 Kb
4.4 Kb

The genomic digests were transferred onto Hybond-N filters by the method of Southern (1975). The filters were incubated in a pre-hybridization reaction at 65°C for 1 hour and then in a hybridization reaction with radioactively labeled probe at 65°C for 15 hours. The probe was a 1.2Kb *BglII-EcoRI* restriction digest fragment spanning the *qutE* gene. This fragment had been isolated from an agarose gel following electrophoresis of a restriction enzyme digest of the vector pPGK-LACZ1 (Chapter 4, Section 2). The filters were washed at high stringency; four 30 minute washes in 0.1 x SSC, 0.5% SDS at 65°C. Autoradiographic images of these filters were obtained by exposing them to X-ray film at -80°C (Figure 5.2). Fragment sizes calculated from the autoradiographs, together with the deduced transformant types are listed in Table 5.1.

Among the 30 transformed strains, 4 were characterized as type I integrants (C1, C5, C14 and C20), with C5 possibly in multiple copy; 12 as type II integrants, 6 of these at multiple copy; 13 as type III replacements. The transformant C7 gave a hybridization pattern of both the native undisturbed band and new bands corresponding to type I integration. This result may reflect analysis of an impure strain, and therefore it was replated twice on selective medium to reisolate single colonies and another sample of genomic DNA prepared.

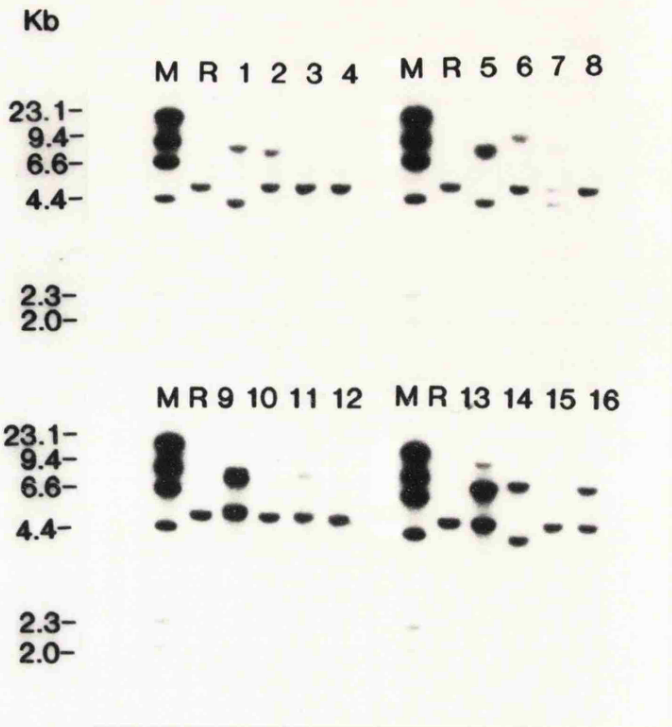
In order to resolve the status of the ambiguous strain C7, and to better characterize the possible multiple copy type I transformant C5 and the 12 type II transformants, the analysis described above was repeated but electrophoresis of the restriction digests was extended to obtain a better separation of fragments. The data are presented in Figure 5.3.

The ambiguous transformant C7 was recharacterized as a single copy type I integrant. Therefore the aberrant pattern observed before was probably due to ineffective strain purification resulting in a mixture of DNA being prepared from two transformants, a single copy type I

Figure 5.2

Analysis of strains transformed with the circular in phase fusion vector p5'PGK-LACZ2 to identify the nature of integration events.

The autoradiographs show the hybridization of an α^{32}P-dCTP labelled 1.2Kb *BglII-EcoRI* DNA fragment spanning the *qutE* gene, with *SstI* digests of genomic DNA prepared from transformed strains (numbered 1 to 30). The original recipient strain QG716 was included in this and in subsequent experiments as a control (symbol R).
Both in this experiment and in all later experiments, 10ng of phage λ DNA digested with *HindIII* provided the marker (symbol M). The marker was identified by including 1ng of an α^{32}P-dCTP labelled *HindIII* digest of λ DNA in the hybridization reaction.


Kb
M R 1 2 3 4
M R 5 6 7 8
23.1–
9.4–
6.6–
4.4–
2.3–
2.0–
M R 9 10 11 12
M R 13 14 15 16
23.1–
9.4–
6.6–
4.4–
2.3–
2.0–

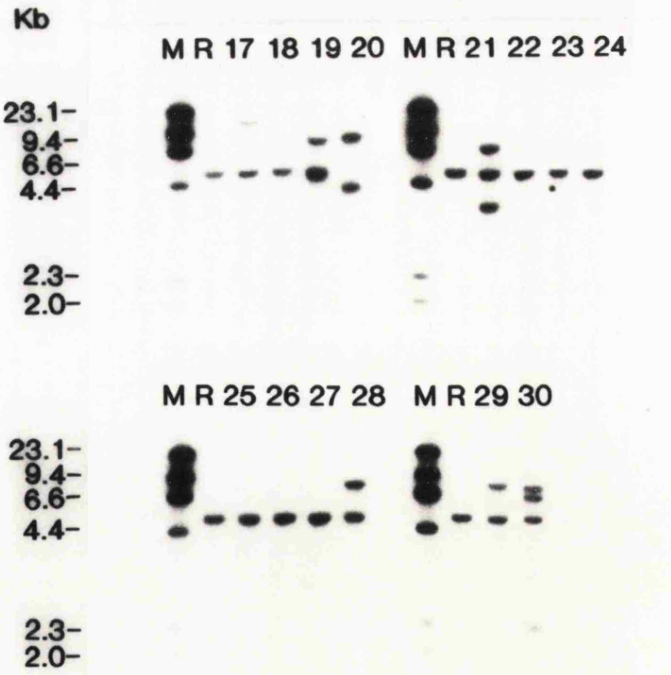

Kb
M R 17 18 19 20
M R 21 22 23 24
23.1–
9.4–
6.6–
4.4–
2.3–
2.0–
M R 25 26 27 28
M R 29 30
23.1–
9.4–
6.6–
4.4–
2.3–
2.0–

Table 5.1
Classification of transformed strains with the circular in phase fusion vector p5'PGK-LACZ2.

The *qutE*$^-$ mutant strain QG716 was transformed with the in phase fusion vector p5'PGK-LACZ2 to give the transformed strains 5'PGK-LACZ:C1 to C30, selected for growth on quinic acid.
The table shows the sizes of fragments that hybridized to an α^{32}P-dCTP labelled 1.2Kb *BglII-EcoRI qutE* DNA probe, when genomic DNA was digested with *SstI* and analysed by agarose gel electrophoresis and DNA hybridization. The third and fourth columns list the types of vector integration events that were deduced for the transformants, and whether single or multiple copies of the vector integrated.

The predicted fragment sizes (Kb) are given below.

Recipient strain		:5.05
Transformed strains:	Type I (single copy)	:8.5, 4.4
	Type I (multiple copy)	:8.5, 7.9, 4.4
	Type II	:5.05, other(s)
	Type III	:5.05

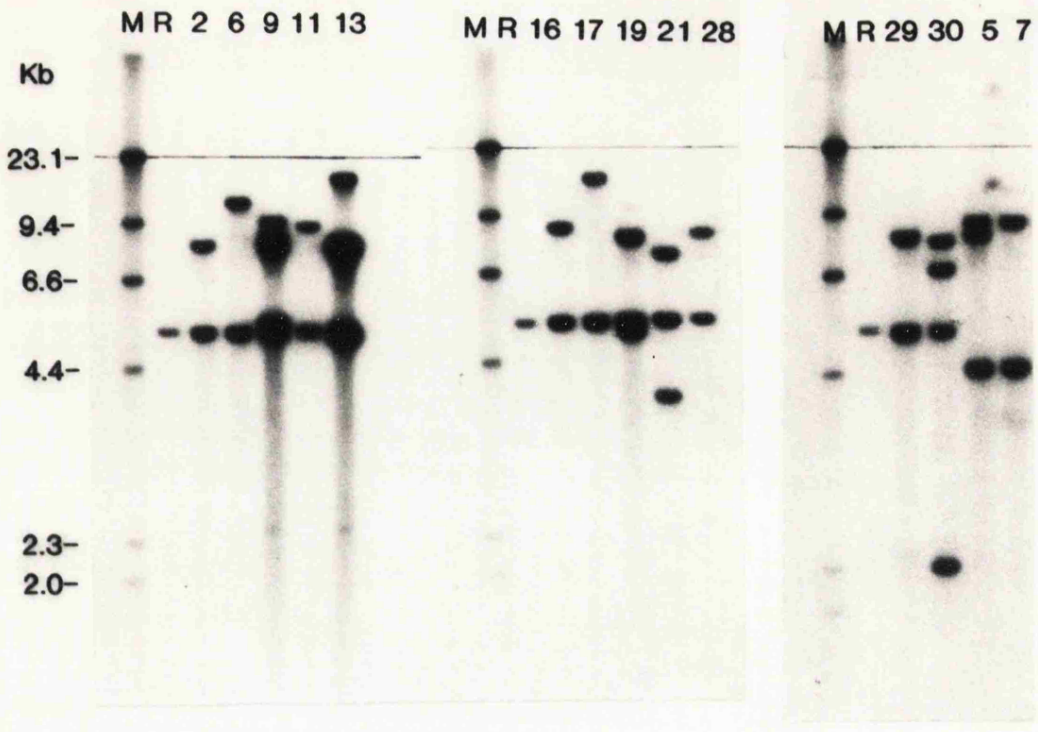
Strain	Observed fragment sizes (Kb)	Class of Transformant	Copies of vector
Recipient QG716	5.0	-	-
Transformants			
5'PGK-LACZ2:C1	8.6, 4.3	I	single
:C2	7.7, 5.0	II	single
:C3	5.0	III	-
:C4	5.0	III	-
:C5	8.1, 4.3	I	single/multiple
:C6	11.5, 5.0	II	single
:C7	8.5, 5.0, 4.3	Novel	unresolved
:C8	5.0	III	-
:C9	9.0, 8.1, 5.0	II	multiple
:C10	5.0	III	-
:C11	9.4, 8.6, 5.0	II	multiple
:C12	5.0	III	-
:C13	15.5, 7.7, 5.0	II	multiple
:C14	8.1, 4.3	I	single
:C15	5.0	III	-
:C16	8.1, 5.0	II	single
:C17	14.5, 5.0	II	single
:C18	5.0	III	-
:C19	7.9, 5.0, 4.9	II	multiple
:C20	8.3, 4.2	I	single
:C21	6.6, 5.0, 3.8	II	multiple
:C22	5.0	III	-
:C23	5.0	III	-
:C24	5.0	III	-
:C25	5.0	III	-
:C26	5.0	III	-
:C27	5.0	III	-
:C28	8.3, 5.0	II	single
:C29	8.3, 5.0	II	single
:C30	7.4, 6.1, 5.0, 2.3	II	multiple

Figure 5.3
Further analysis of genomic DNA isolated from strains transformed with the circular in phase fusion vector p5'PGK-LACZ2.

The autoradiographs show the hybridization of an α^{32}P-dCTP labelled 1.2Kb *BglII-EcoRI* DNA fragment spanning the *qutE* gene with *SstI* digests of genomic DNA prepared from transformed strains (numbered).

The table shows the sizes of the fragments along with the class of vector integration event that was deduced for each transformed strain.

Strain	Fragment sizes (Kb)	Class of Transformant	Copies of Vector
Recipient QG716	5.2	-	-
Transformants			
5'PGK-LACZ2:C2	8.2, 5.2	II	single
:C6	11.0, 5.2	II	single
:C9	9.5, 8.7, 8.1, 5.4, 5.2	II	multiple
:C11	9.3, 5.2	II	single
:C13	14.5, 8.1, 5.2	II	multiple
:C16	8.7, 5.2	II	single
:C17	13.1, 5.2	II	single
:C19	8.1, 5.2, 5.0	II	multiple
:C21	7.4, 5.2, 3.9	II	multiple
:C28	8.4, 5.3	II	single
:C29	8.0, 5.1	II	single
:C30	7.9, 6.7, 5.2, 2.3	II	multiple
:C5	8.7, 8.2, 4.4	I	multiple
:C7	8.7, 4.4	I	single


M R 2 6 9 11 13
M R 16 17 19 21 28
M R 29 30 5 7
Kb
23.1–
9.4–
6.6–
4.4–
2.3–
2.0–

integrant and a type III replacement. Transformant C5 was characterized as a multiple copy type I integrant. The type II transformants gave the same hybridization patterns as before, with the exception of C11, which appeared to be single rather than multiple copy. The only band which appears to be conserved across several of the type II transformants is an approximately 8Kb fragment which corresponds to the large SstI restriction digest fragment of the vector. Thus, there is no evidence for the preferential integration of vector sequence at any site other than the *qutE* locus.

In order to identify which strains express the *lacZ* gene, all 30 transformants had been screened on X-gal plates (Chapter 4, Section 6). The 5 type I transformants all have β-galactosidase activity. However of the 7 single copy type II transformants characterized above (C2, C6, C11, C16, C17, C28 and C29), only 3 (C6, C28 and C29) express *lacZ*. To establish if the absence of β-galactosidase activity reflected the failure of the entire *pgk-lacZ-trpC* fusion to be integrated into the genome, the 7 single copy type II transformants were subject to further analysis. In addition, 4 of the 5 multiple copy type II transformants exibited β-galactosidase activity, but it was not thought worthwhile attempting to further characterize these strains due to the complexity of their analysis.

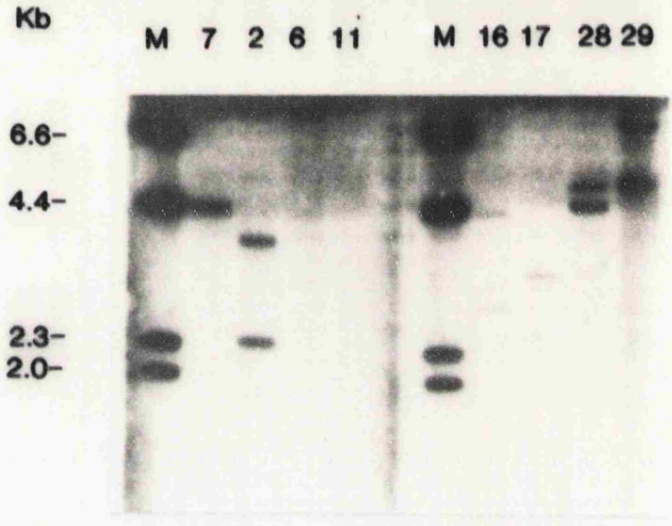
Genomic DNA from the 7 single copy type II transformants was digested with the restriction enzymes *XhoI* and *XbaI*, which cut on either side of the gene fusion sequence in the vector p5'PGK-LACZ2 (Chapter 4, Figure 4.6). The probe used was a 3.0Kb *EcoRI* restriction digest fragment consisting entirely of *lacZ* coding sequence and isolated from the construct p5'PGK-LACZ3 (Chapter 4, Figure 4.7). The reaction conditions for the hybridization and subsequent washes were as described above. Integration of the entire gene fusion sequence into the genome would be expected to result in the probe hybridizing to a 4.4Kb band on the filter. An autoradiographic image of the filter is presented in Figure 5.4. The single copy type I integrant

Figure 5.4
Analysis of single copy type II transformants to determine if the entire sequence of the fusion vector p5'PGK-LACZ2 had integrated into the genome.

The autoradiograph shows the hybridization of an α^{32}P-dCTP labelled 3.0Kb *EcoRI* DNA fragment consisting entirely of *lacZ* coding sequence with *XhoI-XbaI* digests of genomic DNA prepared from transformed strains (numbered). The single copy type I transformant C7 was included as a control and shows the predicted single 4.4Kb band. The portion of the autoradiograph corresponding to large fragments was completely blackened, indicating non-specific binding of probe DNA to a dirty filter. Nevertheless, fragments in the range of interest for this experiment (4.4 Kb) could be easily identified.

The table shows the sizes of the fragments observed.

Strain	Fragment sizes (Kb)
Transformant 5'PGK-LACZ2:C7	4.4
:C2	3.9, 2.4
:C6	7.7
:C11	-
:C16	4.4, 2.8
:C17	3.3, 2.4
:C28	5.1, 4.7
:C29	8.3, 7.4, 5.1


Kb
M 7 2 6 11 M 16 17 28 29
6.6-
4.4-
2.3-
2.0-

C7 was included as a control, and in this the *lacZ* probe hybridized to a 4.4Kb fragment as predicted. Amongst the 7 transformed strains that had been characterized as single copy type II integrants, only C16 generated this 4.4Kb *lacZ* fragment. Suprisingly, C16 had not stained blue on X-gal plates, whereas other transformed strains that do produce β-galactosidase (C6, C28 and C29) did not yield the 4.4Kb *lacZ* fragment. Therefore expression of the gene fusion does not appear to depend solely upon incorporation of the entire *pgk-lacZ-trpC* sequence. Also, it should be noted that although all 7 type II integrants examined here had been classified as single copy using the *qutE* probe, several of them generated more than one fragment that hybridized to the *lacZ* probe, implying that some integration events may have incorporated only a section of the vector molecule.

5.2.2 Assays of β-galactosidase activity in transformed strains.

All the type I and type II transformants identified above were then assayed for β-galactosidase activity. The strains were incubated in defined medium with glucose as carbon source under standard conditions. Cell free extracts were prepared from mycelium by grinding under liquid nitrogen and were assayed for β-galactosidase activity at 25^{o}C with ONPG as substrate (Wallenfels, 1962). Protein concentrations in the extracts were estimated by the method of Lowry *et al.* (1951). The data is given in Table 5.2.

The untransformed strain QG716 yields no β-galactosidase activity, and the single copy type I transformants (C1, C7, C14 and C20) give similar specific activities with a mean of 2600 units and a range of 310 units (12% of the mean). Suprisingly, the multiple copy type I transformant C5 gives a comparable specific activity to the single copy transformants, rather than an integral increase of this. This may be a consequence of additional copies of the gene fusion having been damaged whilst integrating into the chromosome.

Table 5.2

The β-galactosidase activities in strains transformed with the circular in phase fusion vector p5'PGK-LACZ2.

The strain QG716 was transformed with the in phase fusion vector p5'PGK-LACZ2, to give the transformed strains 5'PGK-LACZ2:C1 to C30.

In this and in subsequent experiments, cell free extracts prepared from mycelium were assayed for β-galactosidase activity at 25°C with ONPG as substrate (Wallenfels, 1962). Protein concentrations in the extracts were estimated by the method of Lowry *et al.* (1951). One unit of β-galactosidase activity was defined as that amount producing a change in optical density at 420nm of 0.001 per minute per mg protein.

Strain	Class of Transformant	Copies of vector	β-galactosidase specific activity ($\Delta OD_{420} x 10^3 min^{-1} mg^{-1}$)
Recipient QG716	-	-	0
Transformants			
5'PGK-LACZ2:C1	I	single	2760
:C7	I	single	2450
:C14	I	single	2630
:C20	I	single	2540
:C5	I	multiple	2850
:C2	II	single	0
:C6	II	single	2670
:C11	II	single	0
:C16	II	single	0
:C17	II	single	0
:C28	II	single	2750
:C29	II	single	490
:C9	II	multiple	11330
:C13	II	multiple	7490
:C19	II	multiple	1620
:C21	II	multiple	0
:C30	II	multiple	3100

Among the 7 single copy type II transformants, 4 (C2, C11, C16 and C17) do not show β-galactosidase activity, 2 (C6 and C28) have an activity comparable to the single copy type I transformants, and 1 (C29) has a reduced activity. Of the 5 multiple copy type II integrants, 4 (C9, C13, C19, and C30) have a β-galactosidase activity over a range of values from 1620 to 11330 units, and 1 (C21) does not show β-galactosidase activity.

5.3 Strategies to increase the transformation frequency and the proportion of targeted integrants.

5.3.1 Transformation and targeting frequencies with circular vector molecules.

A total of 6 transformation experiments had to be done with the circular in phase fusion vector to recover the 30 transformants described above. The transformation frequency never exceeded 0.5 colonies/μg DNA/ 10^7 spheroplasts, even though a spheroplast regeneration frequency of up to 45% was recorded. Furthermore, of the 30 transformants analysed, only 5 were characterized as desired type I integrants. As a consequence of the low frequencies of transformation and targeting, several transformation experiments would be necessary to isolate type I integrants for every new fusion construct made. Clearly, such an exercise would be exhausting of time and effort. Two strategies were followed in an attempt to increase the transformation frequency and the proportion of integrations at the *qutE* locus: cotransformation with the high frequency transforming vector pDJB2 (Ballance and Turner, 1985), and linearizing the vector molecule within Aspergillus sequence in the proximity of *qutE*.

5.3.2 Cotransformation of the gene fusion construct with the high frequency transforming vector pDJB2.

There is evidence from the use of several different selective markers in *Aspergillus nidulans* for the integration of more than one molecule of transforming DNA into a single nucleus (Ballance *et al.*, 1983; Tilburn *et al.*, 1983; Yelton *et al.*, 1984; Wernars *et al.*, 1985; de Graff *et al.*, 1988). Incorporation of vector sequence in more than one copy is apparent in 6 of the 30 transformants characterized in Section 2. The integration of many copies of transforming DNA can be explained by assuming that spheroplasts are not all equally capable of taking up DNA, but those that are competent may take up several molecules simultaneously (Fincham, 1989).

The phenomenon of cotransformation and the integration of sequence from more than one different vector molecule has been used to introduce DNA into the genome when the single transformants cannot be easily recovered. For example, an acetate non-utilizing strain of *Neurospora crassa* has been transformed to acetate utilization by cotransforming the required gene with the glutamate dehydrogenase gene (Thomas *et al.*, 1988). This strategy was necessary because spheroplasts could not be regenerated on acetate growth medium, making direct selection impossible.

The technique of cotransformation has also been applied in *A.nidulans* (reviewed by Turner and Ballance, 1985). Cotransformation has been shown to reach 80% with *amdS* and *prnC* plasmids, but only 4% with a *pyr4* plasmid and the *oli31* (resistance) gene carried in phage λ. Wernars *et al.* (1987) applied the technique as a tool for replacing fungal genes, and demonstrated up to 95% cotransformation with *trpC* and *amdS* plasmids. This variation reported in cotransformation frequency using different selectable markers indicates that the success of the strategy cannot be predicted in any particular case.

My aim was to increase the frequency of transformation with the fusion vector carrying the *qutE* gene as a selectable marker. I proposed to do this by cotransforming with a high frequency transforming plasmid. Selection for the integration of DNA from the high frequency vector would be followed by screening for growth on quinic acid, indicating the cotransformation of the fusion construct.

In *A.nidulans*, whatever the selective marker employed, over 90% of the initially observed colonies remain very small and cannot be subcultured (Turner and Ballance, 1985). This effect has also been observed in *Saccharomyces cerevisiae* (Hicks *et al.*, 1979) and in *N.crassa* (Case, 1982). It is believed to result from the transient expression of plasmid sequence without integration into the genome.

The *ansl* sequence of *A.nidulans* is known to enhance transformation frequency by fifty to a hundred fold (Ballance and Turner, 1985). Although the sequence was originally identified as providing an autonomous replication (*ARS*) function in *S.cerevisiae*, it is believed to increase transformation frequency in *A.nidulans* by facilitating vector integration. The presence of several copies of *ansl* in the genome indicates that the high transformation frequency observed may reflect the number of sites available for homologous recombination, and it is indeed observed that *ansl* plasmids integrate widely over the genome (Ballance and Turner, 1985).

The vector pDJB2, which includes the *ansl* sequence of *A.nidulans* and the *pyr4* gene of *N.crassa* (Ballance and Turner, 1985), and the in phase gene fusion vector p5'PGK-LACZ2 were cotransformed into the *pyrG*$^{-}$; *bgaA*$^{-}$; *qutE*$^{-}$ *A.nidulans* strain QG716. Transformants which had integrated pDJB2 sequence were selected for growth on glucose medium in the absence of uracil. They were purified by replating twice on the same selective medium, before being tested for growth on quinic acid to identify cotransformants. As a comparison, the gene fusion vector alone was transformed into an aliquot of the same spheroplast preparation and was selected for directly on quinic acid medium including uracil.

Suprisingly, transformation with p5'PGK-LACZ2 alone gave a frequency of 12 colonies/μg DNA/10^7 protoplasts, despite a spheroplast regeneration frequency of only 4%. This is a 20-fold increase on the frequency recorded above. By contrast, pDJB2 gave a transformation frequency of 223 colonies/μg DNA/10^7 spheroplasts. A number of the transformants were replated once on selective medium and were then tested for growth on quinic acid. Out of a total of 73 transformants, 31 grew on quinic acid; implying a cotransformation frequency of 42% for p5'PGK-LACZ2 with pDJB2. Therefore, the transformation frequency of p5'PGK-LACZ2 in the cotransformation experiment was 223 x 0.42 = 94 colonies/μg DNA/10^7 spheroplasts. Thus, this

approach appears to have increased the integration of gene fusion vector DNA by a factor of about 8.

The 31 cotransformants were screened for the production of β-galactosidase on X-gal plates and all but one stained blue. This contrasts sharply with the 30 transformants previously screened (Chapter 4), when only 40% showed β-galactosidase activity. However, in the case of the cotransformants a wide range could be detected in the degree of staining, with many colonies only turning faintly blue (data not shown), whereas with the transformants characterized in Chapter 4, the intensity of staining was much more consistent. The weak activity of β-galactosidase observed with many of the cotransformants may reflect a large proportion of type II integrants, in which expression of *lacZ* may depend upon the site of vector integration into the genome. Thus, although cotransformation of the gene fusion construct with a high frequency transforming vector resulted in an 8 fold rise in the transformation frequency, the proportion of type I integration events may not have been significantly increased. To confirm whether or not this is the case, the cotransformants would have to be analysed further to determine the site of vector integration. It was not considered worthwhile to carry out this exercise.

5.3.3 Effect of linearizing vector molecule prior to transformation.

In *Saccharomyces cerevisiae* the frequency of type I transformation with integrating plasmids can be increased ten to a thousand fold by linearizing plasmid molecules within the selective marker (Orr-Weaver *et al.*, 1981). This is thought to be a consequence of the generation of free recombinogenic ends facilitating homologous recombination at the corresponding host locus. In *Neurospora crassa* a targeting effect was demonstrated when vector was linearized within the coding sequence of the selective marker (Kim and Marzluf, 1988). However, this effect was lost when plasmid was cut within flanking sequence with homology to chromosomal sequence, but outside the actual coding sequence

of the marker. When a plasmid carrying the *trpC* gene of *Aspergillus nidulans* was linearized within *trpC* coding sequence, there did not appear to be a significant increase, either in the transformation frequency, or in the proportion of type I integrants (Yelton *et al.*, 1984).

The fusion vector p5'PGK-LACZ2 was linearized with the restriction enzyme *BglII*, which cuts within sequence of *A.nidulans* origin, approximately 0.5kb upstream of the translation initiation codon of *qutE* (Figure 5.1). The linear vector molecule was transformed into the *A.nidulans* strain QG716, and transformants were selected for growth on quinic acid. To provide a comparison, the uncut plasmid was also transformed into an aliquot of the same preparation of spheroplasts. In this experiment 32 transformants were isolated from linear vector at a frequency of 9.9 colonies/μg DNA/ 10^7 spheroplasts. These were replated twice on selective medium to isolate single colonies. By contrast, circular vector gave a transformation frequency of 0.6 colonies/μg DNA/ 10^7 protoplasts. Thus, linearizing plasmid within *A.nidulans* sequence appears to have increased the transformation frequency 16.5 fold.

5.4 Analysis of *Aspergillus nidulans* strains transformed with linear in phase fusion vector.

5.4.1 Incorporation of *lacZ* sequence and β-galactosidase activity in transformants.

Strains isolated from transforming linear in phase fusion vector into *Aspergillus nidulans* were screened for the production of β-galactosidase on X-gal plates, and 21 of the 32 examined stained blue.

Genomic DNA was then prepared from each of the 32 transformed strains and was tested for the incorporation of lacZ sequence by "dot-blot" hybridization. The 3.0Kb *EcoRI* restriction digest fragment, consisting entirely of *lacZ* sequence, and isolated from the plasmid p5'PGK-LACZ3 (Chapter 4, Figure 4.7) was used as probe. An autoradiographic image of the filter obtained is shown in Figure 5.5. All 21 of the transformed strains which showed β-galactosidase activity gave a positive signal indicating incorporation of *lacZ* sequence from the vector into the genome, whereas none of the remaining 11 transformants that had not shown β-galactosidase activity had integrated *lacZ* sequence. In contrast, the same vector molecule in a circular form yeilded 16 transformants that had integrated *lacZ* sequence, but only 12 of these showed β-galactosidase activity (Chapter 4, Section 6).

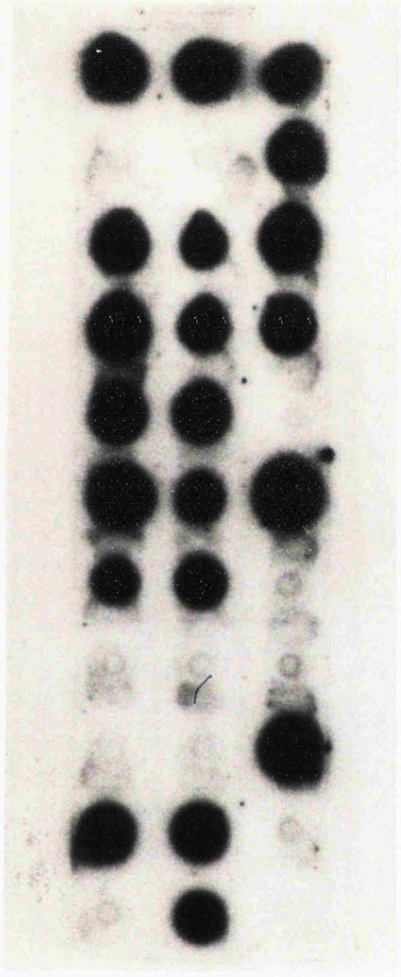
5.4.2 Analysis of DNA isolated from transformed strains.

Genomic DNA prepared from each of the 32 transformed strains (linear p5'PGK-LACZ2 into QG716) was subjected to "Southern blot" analysis exactly as above for the strains transformed with the circular vector (Section 2). The data is presented in Figure 5.6. In order to predict the hybridization patterns for the different transformant types, it was assumed that linear vector molecules would integrate into the genome to give the same structures as circular ones (Figure 5.1).

Figure 5.5

The incorporation of *lacZ* sequence in strains of A.*nidulans* transformed with the linearized gene fusion p5'PGK-LACZ2.

The autoradiograph shows the hybridization of an α^{32}P-dCTP labelled 3.0Kb *EcoRI* DNA fragment consisting entirely of *lacZ* sequence, with genomic DNA isolated from transformed strains of A.*nidulans*. The strains had been recovered from the *qutE* mediated transformation of QG716 with the *BglII* linearized in-phase gene fusion vector p5'PGK-LACZ2, and were designated 5'PGK-LACZ2:L1 to L32. The recipient strain QG716 was included in the analysis as a control.



L1	L2	L3
L4	L5	L6
L7	L8	L9
L10	L11	L12
L13	L14	L15
L16	L17	L18
L19	L20	L21
L22	L23	L24
L25	L26	L27
L28	L29	L30
L31	L32	QG716

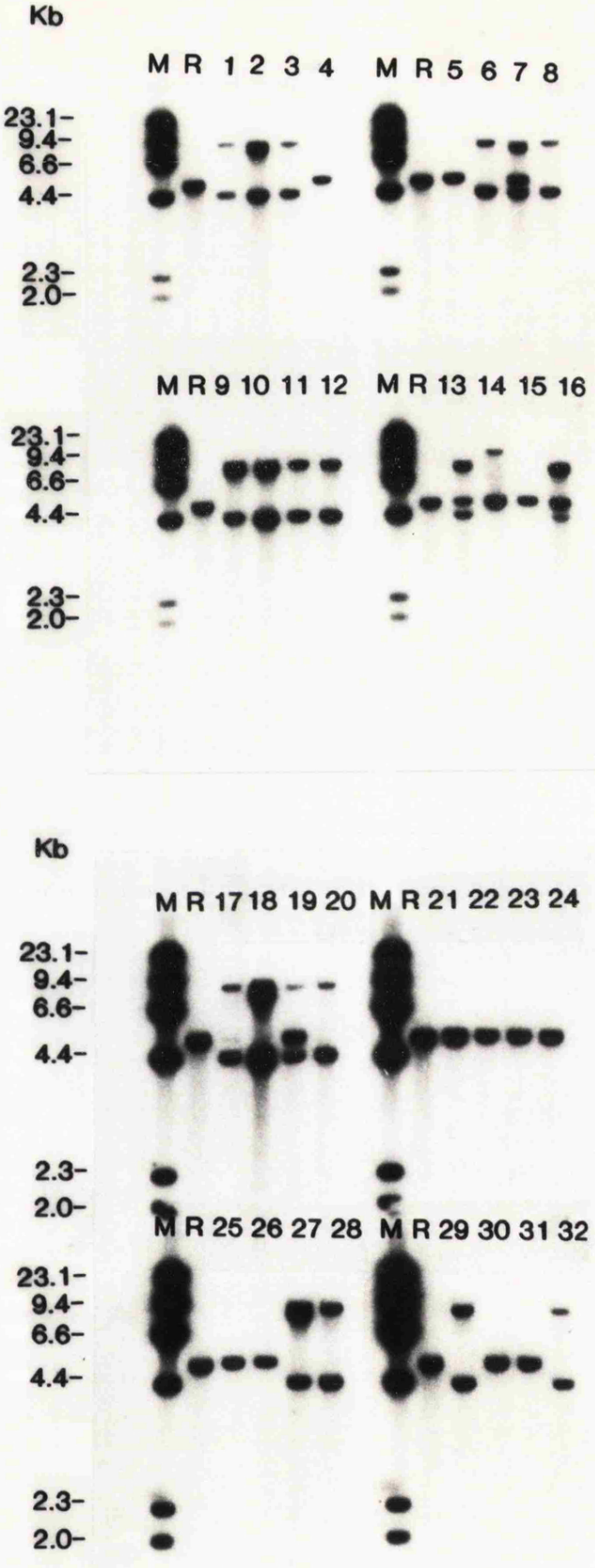
Figure 5.6

Analysis of strains transformed with the *BglII* linearized in phase fusion vector p5'PGK-LACZ2, to identify the nature of integration events.

The autoradiograph shows the hybridization of an α^{32}P-dCTP labelled 1.2Kb *BglII-EcoRI* DNA fragment spanning the *qutE* gene with *SstI* digests of genomic DNA prepared from transformed strains (numbered).

The table shows the sizes of the fragments along with the type of vector integration event that was deduced for each transformant.

Strain	Fragment sizes (Kb)	Class of Transformant	Copies of vector
Recipient QG716	5.0	-	-
Transformants			
5'PGK-LACZ2:L1	8.9, 4.5	I	single
:L2	8.8, 8.2, 4.4	I	multiple
:L3	8.9, 4.5	I	single
:L4	5.2	III	-
:L5	5.2	III	-
:L6	8.4, 4.4	I	single
:L7	8.6, 8.0, 5.0, 4.3	novel	-
:L8	8.8, 4.3	I	single
:L9	8.7, 8.1, 4.5	I	multiple
:L10	8.6, 4.4	I	single
:L11	8.9, 4.5	I	single
:L12	8.9, 4.5	I	single
:L13	8.6, 8.0, 5.2, 4.4	novel	-
:L14	11.0, 5.2	II	single
:L15	5.2	III	-
:L16	8.1, 5.0, 4.3	novel	-
:L17	8.9, 4.5	I	single
:L18	8.6, 8.0, 4.3	I	multiple
:L19	8.9, 5.1, 4.5	novel	-
:L20	8.9, 4.5	I	single
:L21	5.1	III	-
:L22	5.1	III	-
:L23	5.1	III	-
:L24	5.0	III	-
:L25	5.2	III	-
:L26	5.2	III	-
:L27	8.9, 8.0, 4.5	I	multiple
:L28	8.7, 4.4	I	single
:L29	8.7, 4.3	I	single
:L30	5.1	III	-
:L31	5.0	III	-
:L32	8.6, 4.3	I	single


Kb
M R 1 2 3 4 M R 5 6 7 8
23.1–
9.4–
6.6–
4.4–
2.3–
2.0–
M R 9 10 11 12 M R 13 14 15 16
23.1–
9.4–
6.6–
4.4–
2.3–
2.0–
Kb
M R 17 18 19 20 M R 21 22 23 24
23.1–
9.4–
6.6–
4.4–
2.3–
2.0–
M R 25 26 27 28 M R 29 30 31 32
23.1–
9.4–
6.6–
4.4–
2.3–
2.0–

Among the 32 transformants, 16 were characterised as type I, 4 of these at multiple copy number (L2, L9, L18 and L27), 1 as type II (L14), and 11 as type III. In addition, 4 transformants (L7, L13, L16 and L19) gave a hybridization pattern with new bands corresponding to type I integration, but also retaining the native band. To test whether this simply reflected heterogenous strains (see Section 2), these transformants were replated twice for single colony isolation, and new genomic DNA preparations were made. Surprisingly, analysis of these gave the same hybridization patterns as before (Figure 5.7), which indicates that a novel type of integration event has occurred.

Transformant L19 gives new bands corresponding to a single copy type I integration, but also retains the native band, whilst transformants L7, L13 and L16 give new bands corresponding to multiple copy type I events, but again retain the native band. It is not understood how vector molecules could integrate into the genome to produce these hybridization patterns.

The 4 multiple copy type I integrants (L2, L9, L18 and L27) were characterized further to determine the copy number of integrated vector sequence at the *qutE* locus. Genomic DNA prepared from the untransformed strain QG716, a single copy type I transformant (L8), and the 4 multiple copy transformants, was analysed by agarose gel electrophoresis and DNA hybridization. The DNA was digested with the restriction enzyme *BamHI* and probed with the 0.55Kb *XhoI-XhoII* restriction enzyme fragment spanning 5'*pgk* sequence and 13 N-terminal codons. The hybridization conditions and the stringency of the washes were as described above (Section 5.2.1). An autoradiographic image of the filter is presented in Figure 5.8, along with the predicted outcomes of single and multiple copy targeted integrations. The sequences hybridizing probe gave well separated fragments upon electrophoresis, and thus enabled the relevant regions on the filter to be easily isolated. The radioactivity of the probe hybridized to regions on the filter was quantified by radioisotope scintillation counting

Figure 5.7
Analysis of reisolated transformed strains.

The autoradiograph shows the hybridization of an α^{32}P-dCTP labelled 1.2Kb *BglII-EcoRI* DNA fragment spanning the *qutE* gene, with *SstI* digests of genomic DNA prepared from the reisolated transformed strains 5'PGK-LACZ2:L7, L13, L16, and L19.

The table shows the sizes of the observed fragments.

Strain	Fragment sizes (Kb)
Recipient QG716	5.2
Transformants	
5'PGK-LACZ2:L7	8.9, 8.2, 5.2, 4.4
:L13	8.9, 8.2, 5.2, 4.4
:L16	8.9, 8.1, 5.2, 4.4
:L19	8.9, 5.2, 4.4

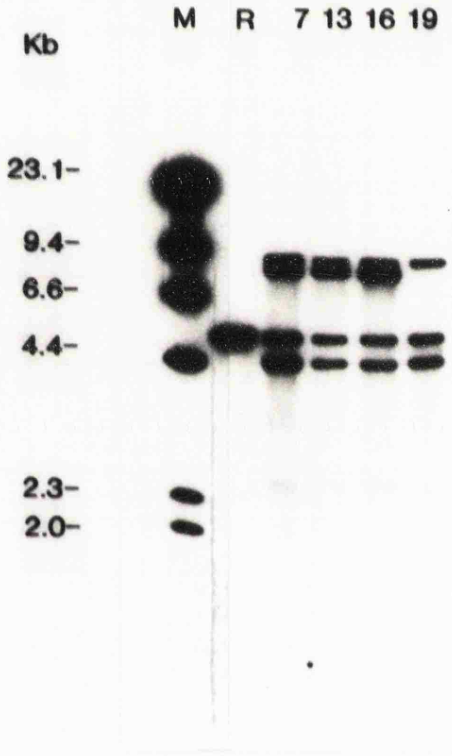

M R 7 13 16 19
Kb
23.1-
9.4-
6.6-
4.4-
2.3-
2.0-

Figure 5.8

Analysis of strains transformed with the *BglII* linearized in phase fusion vector p5'PGK-LACZ2, to determine the copy number of vector integration at the *gutE* locus.

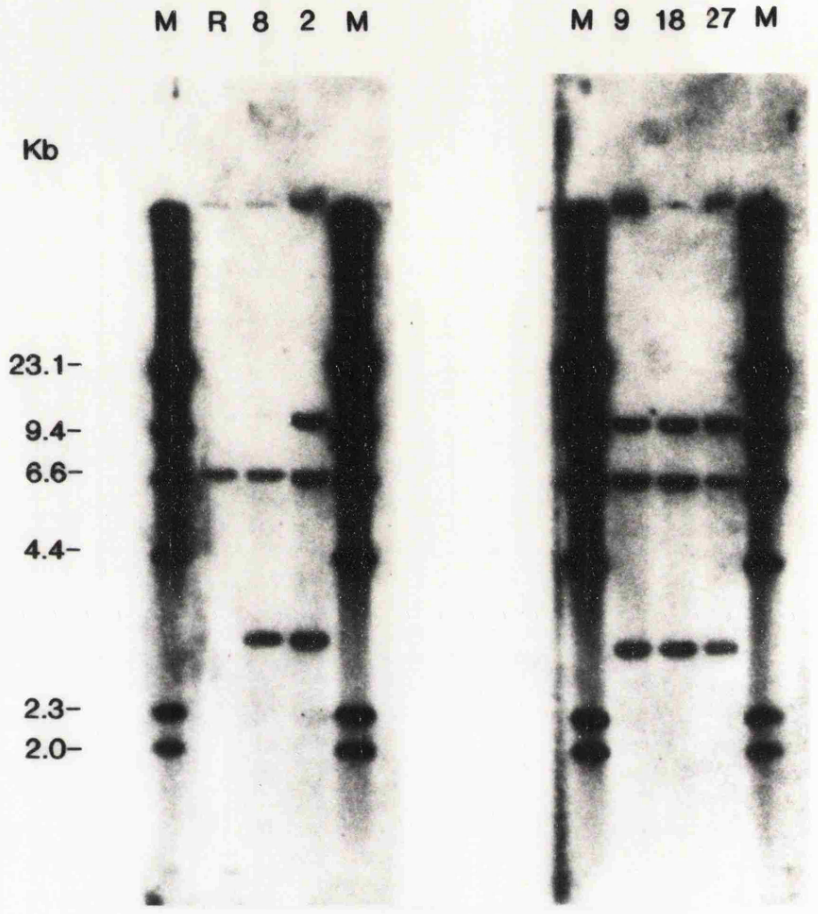
The autoradiograph shows the hybridization of an α^{32}P-dCTP labelled 0.55Kb *XhoI-XhoII* DNA fragment spanning 5'*pgk* sequence and 13 N-terminal codons of *pgk*, with *BamHI* digests of genomic DNA prepared from the transformed strains 5'PGK-LACZ2: L8, L2, L9, L18 and L27.

The table below shows the sizes of the fragments from the recipient strain, a single copy transformed strain and multiple copy transformed strains, that are predicted to hybridize to the DNA probe. The 6.8Kb fragment corresponds to the native *pgk* locus. The 3.05 Kb fragment corresponds to the integration of a single vector molecule at the *gutE* locus. The integration of further vector molecules results in a corresponding number of copies of the 10.4Kb fragment.

Class of strain	Predicted fragment sizes (Kb)
Recipient	6.8
Single copy, type I	6.8, 3.05
Double copy, type I	10.4, 6.8, 3.05
Triple copy, type I	(2)10.4, 6.8, 3.05
n copy, type I	(n-1)10.4, 6.8, 3.05

The table below shows the sizes of the fragments that were calculated from the autoradiograph, along with the deduced class of transformed strain.

Strain	Observed fragment sizes (Kb)	Class of Transformant	Copies of Vector
Recipient QG716	6.8	-	-
Transformants			
5'PGK-LACZ2:L8	6.8, 3.0	I	single
:L2	10.2, 6.7, 2.9	I	multiple
:L9	10.2, 6.7, 2.9	I	multiple
:L18	10.2, 6.7, 2.9	I	multiple
:L27	10.2, 6.7, 2.9	I	multiple


M R 8 2 M
M 9 18 27 M
Kb
23.1–
9.4–
6.6–
4.4–
2.3–
2.0–

(Table 5.3). The number of counts obtained from the resident *pgk* sequence (6.8Kb fragment) provides an internal control and can be compared to the counts obtained from the 10.4Kb and the 3.05Kb bands arising from the integration of vector molecule at the *qutE* locus. The ratio of counts obtained from the 6.8Kb band to that from the 10.4Kb band provides a good basis for estimating the copy number of vector molecule integration. From this data it was concluded that L2, L9 and L18 are each type I integrants with two copies of the vector p5'PGK-LACZ2, and that L27 has three copies.

The results of transforming the linear in phase fusion vector into *Aspergillus nidulans* are summarized in Table 5.4. Among the 32 transformants analysed, 16 are type I integrants (50%), 12 of these at single copy (37%). By contrast, transforming with circular vector gave type I integration in only 5 out of 30 cases (17%), four (13%) at single copy (Section 2). Thus, linearizing the vector molecule within *A.nidulans* sequence flanking the selective marker (*qutE*) increased the proportion of type I integrants almost 3 fold. In addition, the rate of transformation was itself increased by a factor of 16.5. Therefore, the overall frequency of type I integration using linear vector molecules is approximately 50-fold greater than with circular molecules, and in all later experiments with modified gene fusions the vector molecules were cut at the same unique *BglII* site prior to transforming into Aspergillus. This allowed type I integrants to be more readily isolated, so that β-galactosidase activity could be compared between transformants with varying extents of *pgk* promoter sequence upstream of *lacZ*, without concern over the effect upon expression of the site of integration.

In a previous experiment, the two out-of-phase fusion vectors, p5'PGK-LACZ1 and 3, had been linearized at the unique *BglII* restriction enzyme site within *A.nidulans* sequence and transformed into the *A.nidulans* strain QG716. More than 80 transformants were obtained with each construct, and 8 of each had been purified on selective

Table 5.3

The number of copies of the gene fusion vector p5'PGK-LACZ2 integrated at the *qutE* locus.

Equal areas of the hybond-N filter corresonding to bands shown in Figure 5.8 were analysed by recording α particle decay using a TRI-CARB 2000CA liquid scintillation analyser. Counts were recorded over the range of 5 to 1700keV for ten minutes. The recipient strain QG716 was included as a control.

* In each case, the ratio of cpm for a fragment to cpm for the approximately 6.8Kb fragment was internal to the particular strain.

Strain	Fragment sizes (Kb)	Radioisotope scintillation counts (cpm)	Ratio (*)	Deduced copy number
Recipient QG716	6.8	103	1.00	0
Transformants				
5'PGK-LACZ2:L8	6.8	113	1.00	1
	3.0	106	0.94	
:L2	10.2	129	0.81	2
	6.7	160	1.00	
	2.9	142	0.89	
:L9	10.2	159	0.83	2
	6.7	192	1.00	
	2.9	193	1.00	
:L18	10.2	198	0.88	2
	6.7	224	1.00	
	2.9	172	0.77	
:L27	10.2	201	1.81	3
	6.7	111	1.00	
	2.9	115	1.04	

Table 5.4
Classification of transformed strains carrying the linear in phase fusion vector p5'PGK-LACZ2.

The strain QG716 was transformed with the *BglII* linearized in phase fusion vector p5'PGK-LACZ2 to give the transformed strains 5'PGK-LACZ2:L1 to :L32.

Strain	Class of Transformant	Copies of Vector
5'PGK-LACZ2:L1	I	1
:L3	I	1
:L6	I	1
:L8	I	1
:L10	I	1
:L11	I	1
:L12	I	1
:L17	I	1
:L20	I	1
:L28	I	1
:L29	I	1
:L32	I	1
:L2	I	2
:L9	I	2
:L18	I	2
:L27	I	3
:L14	II	1
:L4	III	-
:L5	III	-
:L15	III	-
:L21	III	-
:L22	III	-
:L23	III	-
:L24	III	-
:L25	III	-
:L26	III	-
:L30	III	-
:L31	III	-
:L7	novel	-
:L13	novel	-
:L16	novel	-
:L19	novel	-

medium. They had then been screened for β-galactosidase production on X-gal plates, and for the presence of the *lacZ* gene by DNA hybridization (Chapter 4, Section 6). With neither vector did any transformed strain show β-galactosidase activity, even though 6 of each set of 8 transformants had incorporated *lacZ* sequence.

In order to check that this lack of β-galactosidase activity was not due to the incorporation of vector sequence into transcriptionally silent regions, the strains that had integrated *lacZ* sequence were futher investigated by agarose gel electrophoresis and DNA hybridization of *SstI* restriction enzyme digests (as above). Among the 12 transformed strains examined, 5 likely type I integrants were identified (3 with p5'PGK-LACZ1 and 2 with p5'PGK-LACZ3). In order to obtain a clearer image, the analysis was repeated with these strains. An autoradiographic image of the filter obtained is presented in Figure 5.9, along with the calculated fragment sizes. These patterns correspond to those predicted for single copy type I integrations (Section 2).

Thus, single copy integrants targeted to the *qutE* locus have been isolated from fusion vector in each translational reading frame, but only those with 5'*pgk* sequence fused to *lacZ* in the correct reading frame show β-galactosidase activity.

5.4.3 Assays of β-galactosidase activity in transformed strains.

To discover whether expression of fusion protein depended upon the state of growth of the mycelium, one of the single copy type I integrants (L28) was grown in liquid culture with glucose as carbon source and mycelium was harvested over a time course and stored frozen until all the samples had been collected. The β-galactosidase activities found in cell free extracts are given in Figure 5.10 along with a growth curve deduced from the wet weight of mycelium harvested at each time point. The specific activities,

Figure 5.9
Analysis of strains transformed with the *BglII* linearized out-of-phase fusion vectors p5'PGK-LACZ1 and 3, to confirm transformant types.

The autoradiograph shows the hybridization of an α^{32}P-dCTP labelled 1.2Kb *BglII-EcoRI* DNA fragment spanning the *qutE* gene with *SstI* digests of genomic DNA prepared from the transformed strains 5'PGK-LACZ1:L2, L5 and L7, and 5'PGK-LACZ3:L3 and L4.

The table shows the sizes of the fragments observed, along with the type of vector integration that was deduced for each transformant.

Strain	Fragment sizes (Kb)	Class of Transformant	Copies of vector
Recipient QG716	5.1	-	-
Transformants			
5'PGK-LACZ1:L2	8.6, 4.3	I	single
:L5	8.7, 4.4	I	single
:L7	8.7, 4.4	I	single
5'PGK-LACZ3:L3	8.7, 4.4	I	single
:L4	8.9, 4.4	I	single

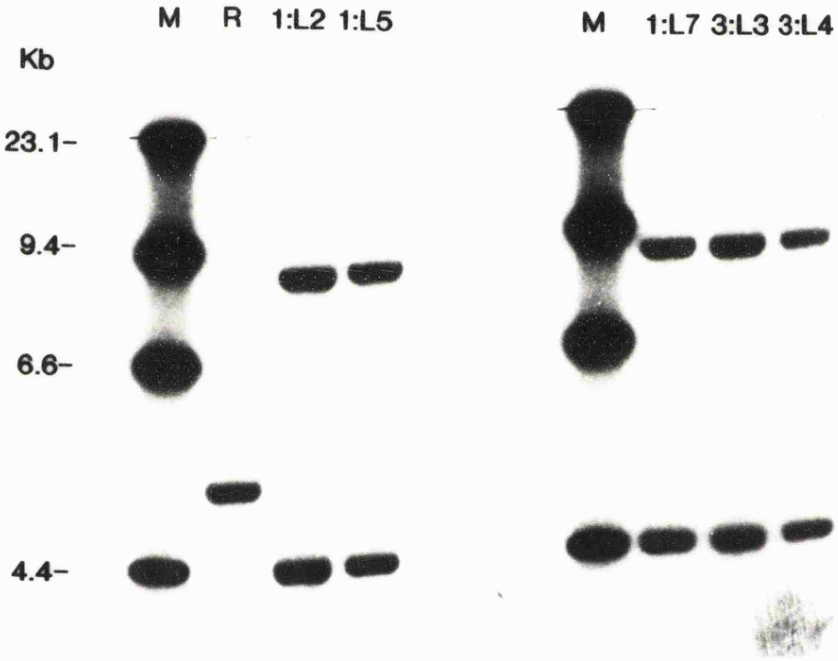
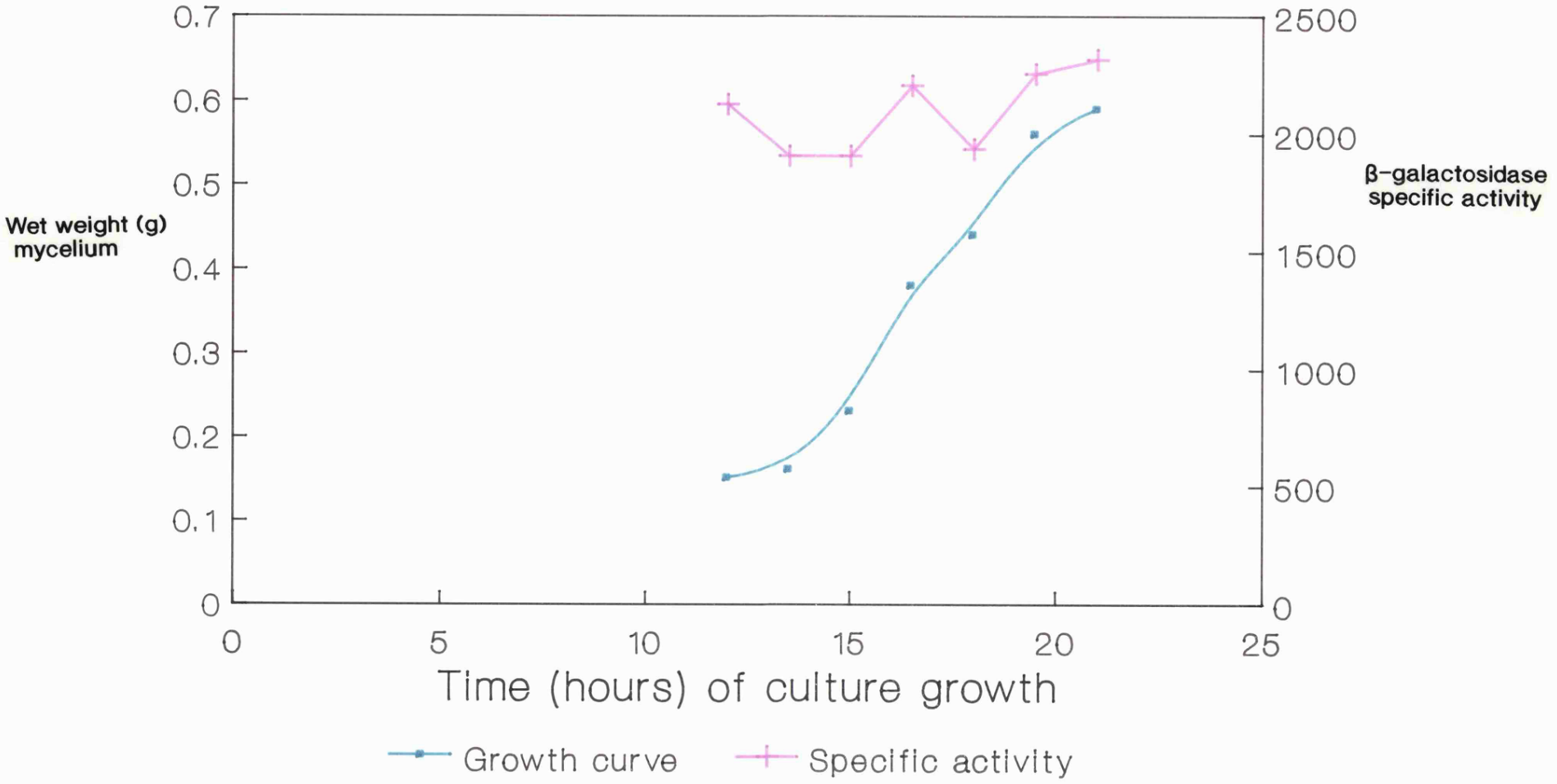

M
R
1:L2
1:L5
M
1:L7
3:L3
3:L4
Kb
23.1-
9.4-
6.6-
4.4-

Figure 5.10
The β-galactosidase activity of a strain transformed with the in phase fusion vector p5'PGK-LACZ2, harvested at different times from batch culture.

Time of culture growth (hrs)	Wet weight (g) of mycelium harvested per 100ml culture	β-galactosidase specific activity ($\Delta OD_{420} x 10^3 min^{-1} mg^{-1}$)
12	0.15	2130
13.5	0.16	1910
15	0.23	1910
16.5	0.38	2210
18	0.44	1940
19.5	0.56	2260
21	0.59	2320

The strain analysed was the single copy type I transformant 5'PGK-LACZ2:L28.
Harvested mycelium was stored at $-20^{o}C$ until all the samples had been collected. Cell free extracts prepared from mycelium from all samples were assayed for β-galactosidase activity under standard conditions.


Wet weight (g) mycelium
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2500
2000
1500
1000
500
0
β-galactosidase specific activity
0
5
10
15
20
25
Time (hours) of culture growth
Growth curve
Specific activity

taken at 7 time points over a range of 9 hours, clearly demonstrate that the activity of the fusion protein does not vary with the age of the culture: mean specific activity of 2100 units, range of 410 units (19.5% of mean).

Cell free extracts prepared from the 16 type I integrants and the single type II integrant were then assayed for β-galactosidase activity (Table 5.5). The single copy type I integrants all gave a similar specific activity with a mean of 2580 units and a range of 820 units (32% of mean). The 4 multiple copy type I transformants gave significantly greater activities. L2, L9 and L18 all gave approximately twice the activity, and L27 gave approximately three times the activity. These specific activities are in proportion to the copy numbers determined above by radioisotope scintillation counting (Section 4.2). Thus, the β-galactosidase activity of the type I transformants appears to be directly proportional to the number of vector molecules that have integrated into the genome at the *qutE* locus. Since copy number has been shown to affect expression, in later experiments comparis ons of β-galactosidase activities between transformants carrying differently modified constructs were always done with single copy type I integrants.

The only type II transformant identified (L14) expressed β-galactosidase at about 1.6 times that of the mean calculated for single copy type I integrants. This is despite L14 giving a DNA hybridization pattern corresponding to a single copy type II transformant. Taken together with the variable activities recorded for single copy type II integrants when circular vector was transformed into *A.nidulans* (Section 2), this implies that the site of integration of vector sequence into the genome may affect expression of the fusion gene.

Five examples of single copy type I transformants had been identified with the two out-of-phase gene fusion constructs pPGK-LACZ1 and 3 (Section 4.2). Cell free extracts from one example of each of the two out of phase

Table 5.5
The β-galactosidase activities recorded from strains transformed with linearized fusion vector.

The strain QG716 was transformed with the *BglII* linearized in phase fusion vector p5'PGK-LACZ2, to give the transformed strains 5'PGK-LACZ2:L1 to L32.
Cell free extracts prepared from mycelium were assayed for β-galactosidase activity under standard conditions.

Strain	Class of Transformant	Copies of vector	β-galactosidase specific activity ($\Delta OD_{420} x 10^3 min^{-1} mg^{-1}$)
Recipient QG716	-	-	0
Transformants			
5'PGK-LACZ2:L1	I	1	3030
:L3	I	1	2660
:L6	I	1	2375
:L8	I	1	2710
:L10	I	1	2500
:L11	I	1	2860
:L12	I	1	2540
:L17	I	1	2500
:L20	I	1	2770
:L28	I	1	2210
:L29	I	1	2460
:L32	I	1	2360
:L2	I	2	4790
:L9	I	2	5050
:L18	I	2	4630
:L27	I	3	7420
:L14	II	1	4045

transformant types (PGK-LACZ1:L5 and PGK-LACZ3:L4) were assayed for β-galactosidase activity. In neither case was any activity detected; confirming the observation made above, when strains were screened on X-gal plates, that an in-phase fusion vector must be transformed into Aspergillus for the fusion gene to be expressed.

5.5 Summary.

Transformation of the circular in phase fusion vector p5'PGK-LACZ2 into *Aspergillus nidulans* generated 30 transformed strains (Chapter 4). Analysis of these transformants by agarose gel electrophoresis and DNA hybridization revealed 5 type I integrants, 12 type II integrants and 13 type III integrants. β-Galactosidase enzyme assays conducted on cell free extracts prepared from the transformed strains demonstrated that the gene fusion was expressed at an equivalent level in all type I integrants, but varied in type II integrants. β-Galactosidase activity could not be related directly to the incorporation of the entire *pgk-lacZ-trpC* fusion into the genome, and may therefore depend upon the site of vector integration.

The transformation frequency with circular plasmid molecules did not exceed 0.5 colonies/μg DNA/10^7 spheroplasts, and only 17% of the transformants analysed had vector sequence integrated at the *qutE* locus. In an attempt to increase the transformation and targeting frequencies, two strategies were followed: cotransformation with the high frequency transforming vector pDJB2, and linearization of the fusion construct within *A.nidulans* sequence flanking *qutE*. A cotransformation frequency of 42% was recorded with pDJB2 and the circular fusion vector, increasing the transformation frequency of the latter 8 fold. However, the β-galactosidase activity, as recorded on X-gal plates, appeared to vary considerably, possibly implying a high proportion of type II integration events.

Linearizing the fusion vector prior to transformation increased the transformation frequency by a factor of 16.5. When 32 of these transformants were analysed by agarose gel electrophoresis and DNA hybridization, half were shown to have the entire construct targeted to the *qutE* locus. Thus, the overall frequency of type I transformation was increased almost 50 fold over that obtained with circular vector molecules. With 4 of the transformants radioactively

labelled probe hybridized to fragments characteristic of both the untransformed recipient strain and of a type I integrant. Thus, it appears that integration at the *qutE* locus has simultaneously led to its duplication without disturbance. The precise nature of the integration event that has occurred in these transformed strains is unclear.

Expression of the fusion gene was shown to be directly proportional to the number of copies of vector molecule that had integrated at the *qutE* locus, and the only type II integrant that was isolated expressed the fusion gene at 1.6 times the level of single copy type I integrants. Thus, the copy number and site of vector integration do appear to affect expression, and in all subsequent experiments with differently modified fusion constructs it was decided to isolate single copy targeted integrants and to compare β-galactosidase activities between these strains only.

Chapter 6.

Deletion analysis of the promoter of the 3-phosphoglycerate kinase gene of *Aspergillus nidulans*.

6.1 Introduction.

The *pgk* gene of *Aspergillus nidulans* is constitutively and strongly expressed. The steady state level of *PGK* mRNA is over 200 fold higher than that of the constitutively expressed biosynthetic dehydroquinase *AROMA* gene (Clements and Roberts, 1986). Phosphoglycerate kinase activities in mycelium grown on a variety of different carbon sources do not show more than a 2 fold difference from the activity in glucose grown mycelium (Clements, 1986). Furthermore, the activity of a phosphoglycerate kinase/β-galactosidase fusion protein has been shown to be independent of the phase of growth of mycelium in batch culture (Chapter 5, Section 4).

In order to identify which sequence elements in the *pgk* promoter may be responsible for the high level of constitutive expression observed, sequence 5′ to *pgk* had been scanned for the presence of regions of homology to other strongly expressed fungal genes (Clements and Roberts, 1986). The sequence TATTTAT is present 32bp upstream of the transcription start site, and is thus similar in both position and sequence to the consensus TATA box (reviewed by Chambon *et al.*, 1983). Also, a sequence with homology to the consensus CCAAT box is positioned approximately 80bp upstream of the transcription start site, again in an expected position (Dierks *et al.*, 1983). Furthermore, a pyrimidine rich sequence is present between the TATA box and the start site. Such a CT block has been identified just upstream or around the site of initiation of transcription in several filamentous fungal genes (Gurr *et al.*, 1987) and *Saccharomyces cerevisiae* genes (Dobson *et al.*, 1982), particularly those that are highly expressed.

In the original in phase gene fusion p5′PGK-LACZ2 (Chapter 4, Section 2) sequence running 487bp upstream of

the *pgk* transcription initiation site is linked to *lacZ*. The three sequence elements described above, which were expected to be important in giving the high level of expression seen with *pgk*, are all included within this stretch of 5'*pgk* sequence. Following transformation into *A.nidulans* the promoter was shown to drive expression of *lacZ* (Chapter 4, Section 6). The mean specific activity of β-galactosidase for single copy integrants targeted to the *qutE* locus was determined as 2580 units (Chapter 5, Section 4).

The central aim of my work was to determine experimentally whether the promoter sequence elements described above and identified through homology to other highly expressed gene promoters, are responsible for the strong expression of the native *pgk* gene and of the *pgk-lacZ* fusion.

6.2 Construction of a series of vectors with a decreasing extent of *pgk* promoter sequence.

6.2.1 The nature of the promoter deletions.

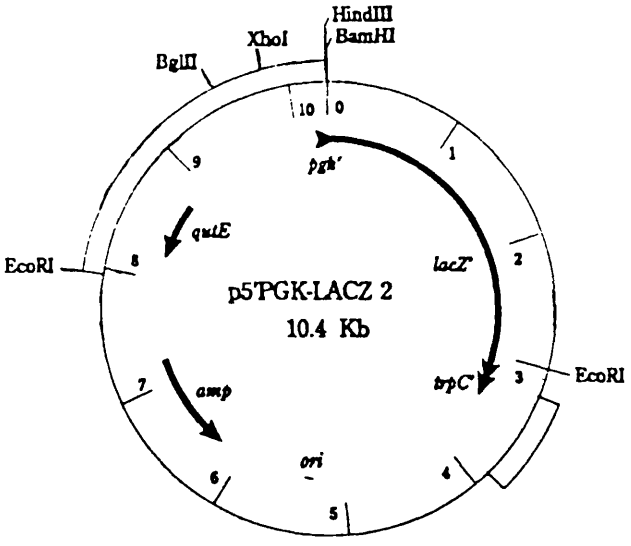
In order to determine the importance of sequence elements in the promoter for *pgk* gene expression, it was decided to delete such elements from the 5′ sequence of the gene. The two most straightforward ways of doing this are either to delete defined sequences from the promoter by excising specific restriction enzyme fragments, or to digest with exonuclease into the promoter starting from its 5′ extent on the vector. In the *pgk-lacZ* gene fusion p5′PGK-LACZ2 there are no restriction enzyme sites on either side of sequences thought to be important in transcriptional control, which do not also cut elsewhere in the vector. Thus, specific promoter sequences could not easily be deleted from the construct or replaced with synthetic sequences. However, during the construction of the vector, a unique *XhoI* restriction enzyme site had been located 5′ to the *pgk* sequence, and this was used as a start point for exonuclease digestion into the promoter, with the aim of determining its 5′ extent.

6.2.2 The generation of a *pgk* promoter deletion series.

The strategy for the formation of the *pgk* promoter deletion series of vectors is outlined in Figure 6.1. The in phase gene fusion construct p5′PGK-LACZ2 (487bp of 5′*pgk* sequence) was linearized by digestion with the restriction enzyme *XhoI*. Deletions into the *pgk* promoter sequence were obtained by incubating aliquots of linear vector molecule with exonuclease *Ba131* over a range of reaction times and stopping the reactions by adding EGTA. The extent of each exonuclease digestion was estimated by treating a sample taken from each reaction with the restriction enzyme *EcoRI* and comparing the mobility on an agarose gel of the two smaller fragments with those from a *XhoI-EcoRI* digest of the original vector. Reaction mixtures were identified in which exonuclease digestion was unlikely to have extended into

Figure 6.1
Construction of the *pgk* promoter deletion series of vectors.

The construction of the vector p5'PGK-LACZ2 is described in Chapter 4, Section 2.

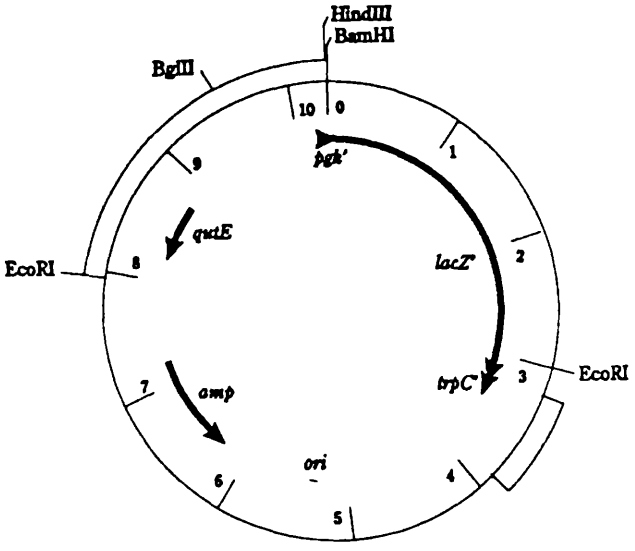


XhoI digest followed by *Bal31* exonuclease digest over a time course

Fill in reaction with Klenow fragment to give blunt ends followed by ligation reaction

Gives a series of vectors with reduced sequence 5' to *pgk* and 5' to *qutE*

p5'PGK-LACZ2-1 to 2-12



lacZ sequence, and in these the vector molecules were blunt ended by single strand synthesis using the Klenow fragment of DNA polymerase I. The populations of molecules were then recircularized in a ligation reaction and transformed into the *Escherichia coli* strain DH5 under selection for the plasmid borne *amp* gene. Plasmid DNA was prepared from a number of individual transformants and analysed by digesting with the restriction enzymes *BamHI* plus *EcoRI* and agarose gel electrophoresis. By comparing fragment mobilities with those of a *BamHI-EcoRI* digest of the original vector, the approximate extent of exonuclease digestion was determined for each clone (Figure 6.2).

6.2.3 Determining the precise 5′ extent of *pgk* promoter sequence in the modified constructs.

To determine the exact 5′ extent of *pgk* promoter sequence in particular derivatives, a fragment spanning the deleted region was subcloned into the M13 vector mp18 (Messing, 1983). Twelve of the modified gene fusion vectors were digested with the restriction enzymes *HindIII* plus *EcoRI*, and in each case the smallest of the three fragments generated was subcloned into *HindIII-EcoRI* cut M13mp18. Recombinants were identified by transforming ligation mixes into the *E.coli* strain JM101 and screening for white plaques on IPTG/X-gal plates. Single stranded DNA preparations from white plaques corresponding to each vector fragment were annealed to universal primer and used as templates in sequencing reactions (Sanger *et al.*, 1977). An autoradiographic image of sequences from two of the vectors is presented in Figure 6.3.

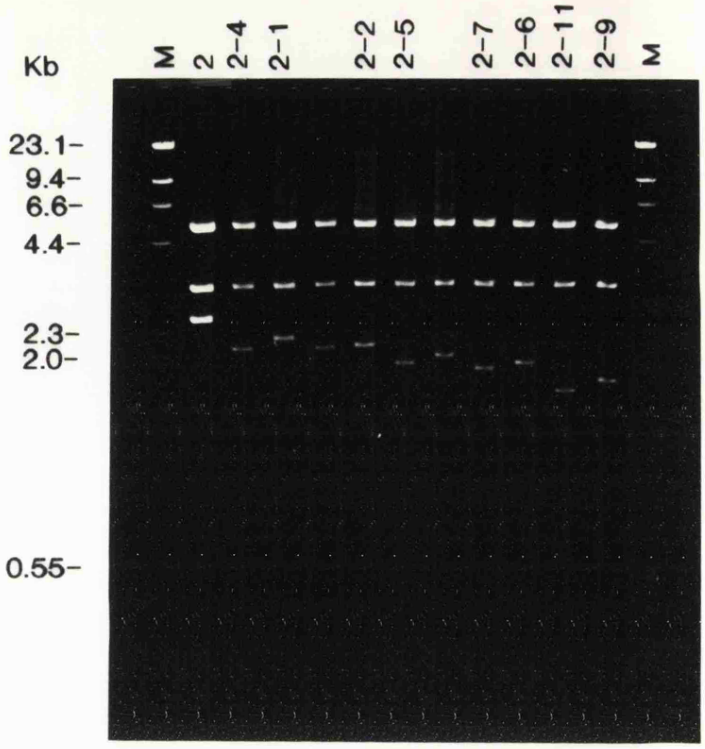
The nucleotide sequence that is predicted for the complete, untreated vector fragment is given in Figure 6.4. The sequence corresponds to the transcribed strand of *pgk*, and it runs from M13 primer through 5 N-terminal *pgk* codons, the 32nt leader sequence and 487nt of 5′ sequence, into sequence 1198nt upstream of the translation start codon of *qutE*. The precise extent of sequence deleted in each of the exonuclease digests described above was determined by

Figure 6.2
Restriction digest analysis of plasmids produced by 5'*pgk* sequence deletions.

Each of the exonuclease treated fusion vectors had been prepared by a small scale alkaline lysis procedure and incubated in a restiction digest reaction with the enzymes *BamHI* and *EcoRI*, and the products analysed by agarose gel electrophoresis.
The vectors p5'PGK-LACZ2-1 to 2-12, that were subsequently subcloned into M13mp18, are identified on the photographs. The original vector p5'PGK-LACZ2 was included as a control, and 0.1μg of phage λ DNA digested with *HindIII* was included as a molecular weight marker (symbol M).

The table shows the sizes of the smallest fragments that were generated by the digests and estimates of the extents of the exonuclease digests into 5'*pgk* sequence. These estimates were made by calculating the differences in the sizes of the smallest fragments derived from the vectors in the deletion series and that derived from the original vector p5'PGK-LACZ2. The difference was divided by 2, since exonuclease digestion procedes in both directions from the site of vector linearization, and it is assumed that the rate of digestion is approximately equal in each direction.

Vector	Size of smallest fragment generated by digestion (Kb)	Approximate extent of the exonuclease digest into 5'*pgk* sequence (Kb)
p5'PGK-LACZ2	2.50	0.00
2-1	2.20	0.15
2-2	2.15	0.17
2-3	2.10	0.20
2-4	2.05	0.22
2-5	1.95	0.25
2-6	1.90	0.30
2-7	1.85	0.32
2-8	1.85	0.32
2-9	1.80	0.35
2-10	1.75	0.37
2-11	1.65	0.42
2-12	1.60	0.45


Kb
M
2
2-4
2-1
2-2
2-5
2-7
2-6
2-11
2-9
M
23.1-
9.4-
6.6-
4.4-
2.3-
2.0-
0.55-

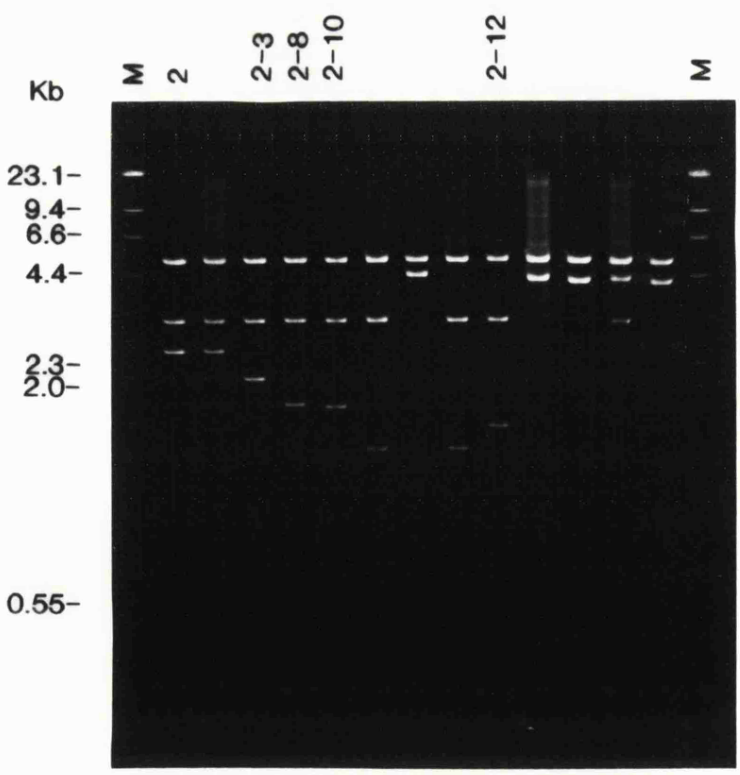
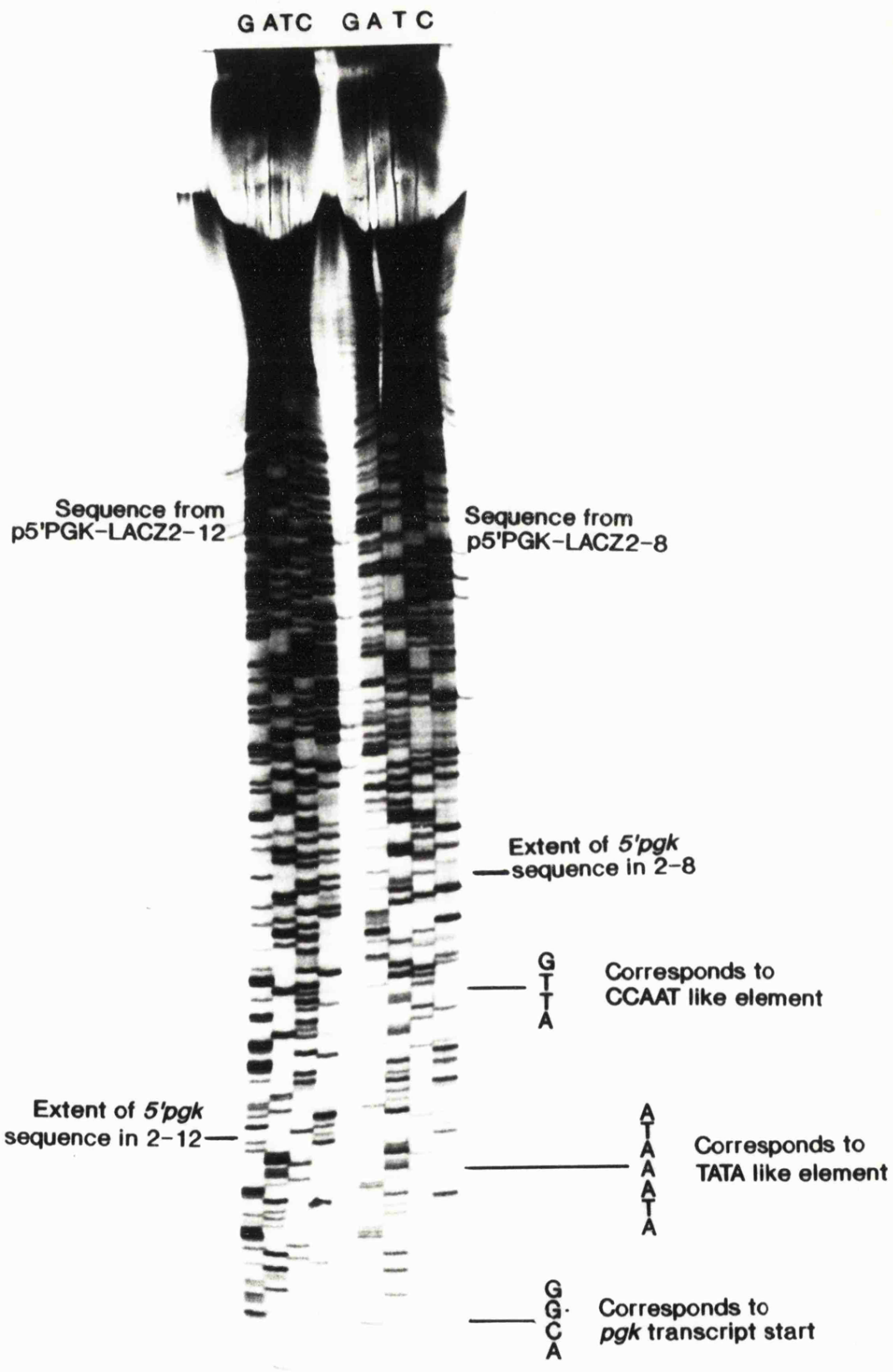

Kb
M
2
2-3
2-8
2-10
2-12
M
23.1-
9.4-
6.6-
4.4-
2.3-
2.0-
0.55-

Figure 6.3
Mapping the end points in the 5'*pgk* deletion series by DNA sequence analysis.

The autoradiograph shows the products of dideoxy chain termination sequencing reactions G, A, T and C (Sanger *et al.*, 1977) that have incorporated α^{35}S-dATP. Universal primer was annealed to the single stranded sequencing templates to initiate the reactions and the products were analysed on a 6% polyacrylamide gel. The sequences reading up the gel correspond to the transcribed strand of *pgk* originating from the fusion vectors p5'PGK-LACZ2-8 and 2-12.


G A T C
G A T C
Sequence from p5'PGK-LACZ2-12
Sequence from p5'PGK-LACZ2-8
Extent of 5'pgk sequence in 2-8
G T T A
Corresponds to CCAAT like element
Extent of 5'pgk sequence in 2-12
A T A A A T A
Corresponds to TATA like element
G G C A
Corresponds to pgk transcript start

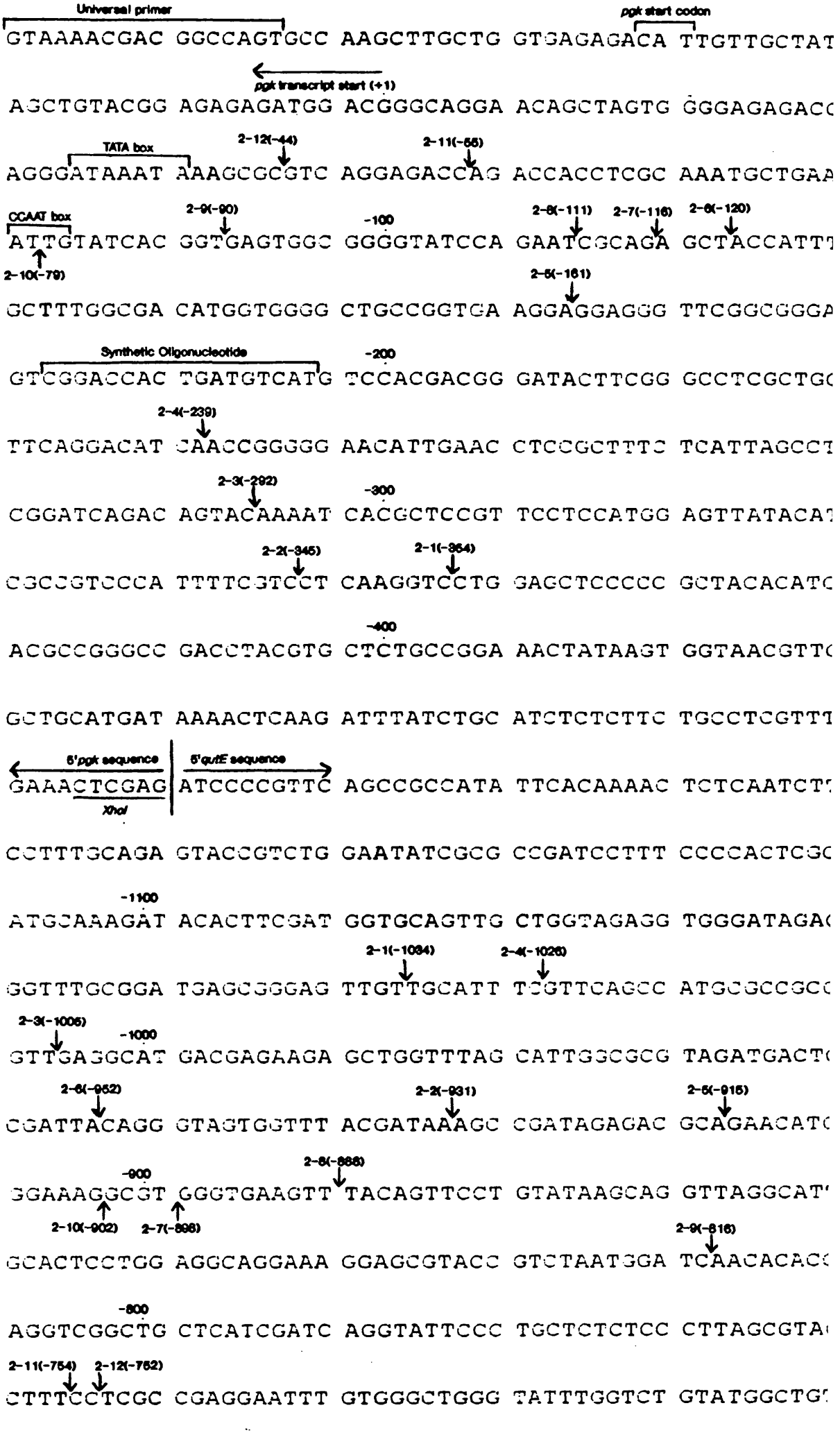
comparing the nucleotide sequence given in Figure 6.4 with those read from sequencing gels of vector fragments subcloned into M13.

The junction points between sequence upstream of *pgk* and sequence upstream of *qutE* was defined for each of the 12 modified promoters. The extent of 5'*pgk* and 5'*qutE* sequence for each construct is marked in Figure 6.4. In 8 of the modified vectors (p5'PGK-LACZ2-5 to 2-12) sequence could be read across the junction point by using universal primer in the sequencing reactions. However, with fragments from the remaining 4 vectors (p5'PGK-LACZ2-1 to 2-4) the junction was too far from the primer to be read from the sequencing gel. In these cases a synthetic oligonucleotide complementary to sequence upstream of *pgk* (-180 to -196 relative to the transcription start site) was used as primer in the sequencing reactions. The sequence of the oligonucleotide is indicated in Figure 6.4.

Thus, exonuclease digestion of the original in-phase gene fusion vector p5'PGK-LACZ2 generated 12 new constructs, in which the extent of sequences upstream of the *pgk-lacZ* fusions were precisely determined.

Figure 6.4
The end points in the 5'*pgk* deletion series of vectors.

The sequence of the transcribed strand of DNA (that generated in the sequencing reactions described) is given. The sequence runs from M13 universal primer through sequence complementary to 5 N-terminal *pgk* codons, the 32nt leader sequence and 487nt of 5' sequence, into sequence 1198nt upstream of the translation initiation codon of *qutE*. The numbering shown above the sequence is relative to the site of transcription initiation of *pgk* (+1) or the site of translation initiation of *qutE* (not shown), the transcript start site of *qutE* not yet having been mapped.
The universal primer and the synthetic oligonucleotide used to initiate sequencing reactions are indicated, as is sequence complementary to the specific 5'*pgk* consensus elements, the TATA box and the CCAAT box. The extent of 5'*pgk* sequence and 5'*qutE* sequence, as determined by the dideoxynucleotide-mediated chain-termination method of DNA sequencing, for each of the fusions in the exonuclease digest series p5'PGK-LACZ2-1 to 2-12 are marked. The exonuclease digests were initiated at the *XhoI* site.


Universal primer
pgk start codon
GTAAAACGAC GGCCAGTGCC AAGCTTGCTG GTGAGAGACA TTGTTGCTAT
pgk transcript start (+1)
AGCTGTACGG AGAGAGATGG ACGGGCAGGA ACAGCTAGTG GGGAGAGACC
TATA box
2-12(-44)
2-11(-55)
AGGGATAAAT AAAGCGCGTC AGGAGACCAG ACCACCTCGC AAATGCTGAA
CCAAT box
2-9(-90)
-100
2-8(-111)
2-7(-116)
2-6(-120)
ATTGTATCAC GGTGAGTGGC GGGGTATCCA GAATCGCAGA GCTACCATTT
2-10(-79)
2-5(-161)
GCTTTGGCGA CATGGTGGGG CTGCCGGTGA AGGAGGAGGG TTCGGCGGGA
Synthetic Oligonucleotide
-200
GTCGGACCAC TGATGTCATG TCCACGACGG GATACTTCGG GCCTCGCTGC
2-4(-239)
TTCAGGACAT CAACCGGGGG AACATTGAAC CTCCGCTTTC TCATTAGCCT
2-3(-292)
-300
CGGATCAGAC AGTACAAAAT CACGCTCCGT TCCTCCATGG AGTTATACAT
2-2(-345)
2-1(-354)
CGCCGTCCCA TTTTCGTCCT CAAGGTCCTG GAGCTCCCCC GCTACACATC
-400
ACGCCGGGCC GACCTACGTG CTCTGCCGGA AACTATAAGT GGTAACGTTC
GCTGCATGAT AAAACTCAAG ATTTATCTGC ATCTCTCTTC TGCCTCGTTT
5'pgk sequence
5'qutE sequence
GAAACTCGAG
XhoI
ATCCCCGTTC AGCCGCCATA TTCACAAAAC TCTCAATCTT
CCTTTGCAGA GTACCGTCTG GAATATCGCG CCGATCCTTT CCCCACTCGC
-1100
ATGCAAAGAT ACACTTCGAT GGTGCAGTTG CTGGTAGAGG TGGGATAGAC
2-1(-1034)
2-4(-1026)
GGTTTGCGGA TGAGCGGGAG TTGTTGCATT TCGTTCAGCC ATGCGCCGCC
2-3(-1005)
-1000
GTTGAGGCAT GACGAGAAGA GCTGGTTTAG CATTGGCGCG TAGATGACTC
2-6(-952)
2-2(-931)
2-5(-915)
CGATTACAGG GTAGTGGTTT ACGATAAAGC CGATAGAGAC GCAGAACATC
2-8(-868)
-900
GGAAAGGCGT GGGTGAAGTT TACAGTTCCT GTATAAGCAG GTTAGGCATT
2-10(-902)
2-7(-898)
2-9(-816)
GCACTCCTGG AGGCAGGAAA GGAGCGTACC GTCTAATGGA TCAACACACC
-800
AGGTCGGCTG CTCATCGATC AGGTATTCCC TGCTCTCTCC CTTAGCGTAC
2-11(-754)
2-12(-752)
CTTTCCTCGC CGAGGAATTT GTGGGCTGGG TATTTGGTCT GTATGGCTGT

6.3 Transformation of the pgk promoter deletion series of vectors into *Aspergillus nidulans*.

To evaluate the effect of *pgk* promoter deletions upon expression of the fused *lacZ* gene, 10 of the 12 modified constructs described in Section 2 were transformed into *Aspergillus nidulans*. In order that the level of β-galactosidase activity could be directly related to the extent of *pgk* promoter sequence remaining, it was necessary to isolate transformed strains with a single copy of vector integrated into the *qutE* locus. To ensure a relatively high transformation frequency and to facilitate targeting, the vectors were linearized at the unique *BglII* restriction enzyme site located upstream of *qutE* prior to transforming into A.*nidulans* (Figure 6.1).

Initially, to obtain a preliminary set of transformants with a range of *pgk* promoter sequences fused to *lacZ*, the vectors p5'PGK-LACZ2-1 (with *pgk* promoter sequence extending -354bp relative to the transcript start site), 2-3 (-292bp), 2-8 (-111bp) and 2-12 (-44bp) were transformed into the A.*nidulans* strain QG716 and selected for growth on quinic acid. Eight transformants from each experiment with p5'PGK-LACZ2-1 and 2-3, 13 with 2-8 and 10 with 2-12 were replated twice on the same medium.

In order to determine whether the transformed strains had incorporated *lacZ* sequence, genomic DNA was prepared from each strain and was analysed by "dot blot" filter hybridization using the same radioactively labelled 3kb *EcoRI* probe described previously (Chapter 4, Section 6). The data revealed that six of the p5'PGK-LACZ2-1 transformants, four of the 2-3 transformants, seven of the 2-8 transformants and six of the 2-12 transformants had taken up *lacZ* sequence.

All of the transformed strains were then analysed further exactly as had been done previously by agarose gel electrophoresis of *SstI* digested genomic DNA, and DNA hybridization using the 1.2kb *BglII-EcoRI qutE* probe to

determine the nature of the vector integration event (Chapter 5, Section 2). The sizes of the fragments identified by the probe were compared to the fragment sizes predicted for type I integration events with each vector. This procedure identified two single copy type I transformants for p5'PGK-LACZ2-3, two for 2-8 and three for 2-12. However, none of the 2-1 transformants appeared to be single copy type I integrants. The distribution of types of integration events observed for each of the vectors is given in the Appendix.

At this point samples of the single copy integrants were assayed for β-galactosidase activity to obtain a first estimate for the extent of the *pgk* promoter. The data are described below (Section 6.4). Subsequently six of the eight remaining vectors in the deletion series [p5'PGK-LACZ2-2 (-345 bp), 2-4 (-239 bp), 2-5 (-161 bp), 2-6 (-120 bp), 2-7 (-116 bp) and 2-9 (-90 bp)] were transformed into *A.nidulans*. Also, a second transformation experiment was conducted with p5'PGK-LACZ2-1 (-354 bp) from which no single copy type I integrants had been previously isolated. In each case, transformed strains were selected on quinic acid and a sample of twelve transformants were analysed as described previously to determine the nature of vector integration events. Fragment sizes calculated from autoradiographs of hybridized filters were compared to the sizes that were predicted for type I integration events with each vector. In this way three single copy targeted integrants were identified for the vector p5'PGK-LACZ2-1, and one for each of the vectors 2-2, 2-4, 2-7 and 2-9. However, no transformants of this type were obtained for the vectors 2-5 or 2-6. The transformation experiment was repeated for these two vectors, and 20 transformants of each were analysed as before. Four single copy targeted integrants were identified for the vector 2-5 and two for the vector 2-6. The proportions of each type of integration event for all the vectors transformed into *A.nidulans* are given in the Appendix.

It is apparent, upon comparing the proportions of each transformant type between the three original unmodified vectors p5'PGK-LACZ1, 2 and 3 and the vectors resulting from exonuclease treatment of p5'PGK-LACZ2, that there is no decrease in the frequency of targeted integration with a decreasing extent of *A.nidulans qutE* sequence present on the vector. This is despite the fact that with the deletion series of vectors, the proportion of *A.nidulans qutE* sequence present on the vector molecule within which homologous recombination can occur, decreases as further *pgk* promoter sequence is removed (Figure 6.4). In the most extensive deletion construct, p5'PGK-LACZ2-12, as well as 443bp of 5'*pgk* sequence having been removed, 446bp of sequence upstream of the *qutE* gene has been removed from the original fusion vector. Thus, the extent of *A.nidulans* sequence in the construct has been reduced from 2.45Kb to 1.55Kb, and most pertinently, the extent of sequence flanking *qutE* has been reduced from 1.9Kb to 1.45Kb.

The gradual reduction in the extent of sequence flanking *qutE* present on the vector series with increasing size of deletion might have been expected to increasingly favour type II integration rather than events at the equivalent locus. Surprisingly, no such general trend is observed, even though with the unmodified in-phase fusion vector, p5'PGK-LACZ2, only one transformant out of 32 was characterized as being type II. However, it should be noted that in all the experiments described above, vector molecules were linearized within *A.nidulans* sequence flanking the *qutE* gene prior to transformation. The effect of linearizing vector molecules upon targeting has been described in Chapter 5, Section 3, where the proportion of type I integrants was three fold greater with linear vector rather than circular molecules. Thus, in all the transformation experiments described above, free homologous ends were available in *A.nidulans* sequence flanking *qutE* for homologous recombination. It is proposed that the presence of these free recombinogenic ends in all the transformations masks any effect upon the proportion of homologous to non-homologous recombination events that would have been

expected with a decreasing extent of sequence flanking *qutE* in the exonuclease digest series of vectors.

In the case of type II integration events, the only fragment hybridizing to *qutE* probe that appeared to be common to each set of transformants was that which corresponded in size to the relevant construct. Thus, as with the unmodified fusion vector p5'PGK-LACZ2 (Chapter 5, Section 2), there is no evidence for the preferential integration of vector sequence at any site other than the *qutE* locus.

In summary, single copy transformants of *A.nidulans*, with vector sequence targeted to the *qutE* locus, have been identified for ten constructs with a range of *pgk* promoter sequence upstream of a *pgk-lacZ* gene fusion. In the unmodified construct, p5'PGK-LACZ2, 487bp of 5'*pgk* sequence is present, whilst in the most extensive deletion, only 44bp of 5' sequence remains. In order to determine the 5' extent of the *pgk* promoter, cell free extracts derived from the *A.nidulans* transformants described above were assayed for β-galactosidase activity.

6.4 The effect of 5'*pgk* sequence deletions upon the expression of gene fusions.

6.4.1 Screening *Aspergillus nidulans* transformants for β-galactosidase activity.

All transformed strains that were isolated by introducing each fusion vector of the *pgk* promoter deletion series into *Aspergillus nidulans* were screened for β-galactosidase activity on plates containing X-gal. The proportion of each integrant type that expressed lacZ for each set of transformants is given in the Appendix.

Initially, to give an indication of the 5' extent of the *pgk* promoter, the vectors p5'PGK-LACZ2-1 (with *pgk* promoter sequence extending -354bp relative to the transcript start site), 2-3 (-292bp), 2-8 (-111bp) and 2-12 (-44bp) had been introduced into *A.nidulans* and the transformants subject to Southern blot analysis (Section 6.3). All type I transformants with the vectors 2-1 (-354bp) and 2-3 (-292bp) stained blue on X-gal plates, whilst those with the vectors 2-8 (-111bp)and 2-12 (-44bp) did not. Thus, sequences required for expression are located within 292bp of the transcript start site.

Subsequently a further six vectors in the *pgk* promoter deletion series p5'PGK-LACZ2-2 (-345bp), 2-4 (-239bp), 2-5 (-161bp), 2-6(-120bp), 2-7(-116bp) and 2-9 (-90bp) were introduced into *A.nidulans* and characterized (Section 6.3). All type I transformants with the vectors 2-2 (-345bp), 2-4 (-239bp) and 2-5 (-161bp) stained blue on X-gal plates, whilst those with the vectors 2-6 (-120bp), 2-7 (-116bp) and 2-9 (-90bp) did not stain blue.

Taken together these experiments have allowed promoter sequence elements required for expression of the *lacZ* gene fusion to be located within 161bp of the transcript start site. However, because the degree of staining apparent on X-gal plates is only qualitative, levels of *lacZ* expression obtained with constructs extending beyond 161bp upstream

cannot be accurately compared by this method. Therefore, sequence elements modulating expression from the larger constructs may lie beyond the point identified as the 5′ extent of the minimum promoter at this stage. In order that such putative promoter sequences could be investigated, β-galactosidase assays were conducted on cell free extracts prepared from the transformants described above (Section 4.2).

It is apparent from the data presented in the Appendix that all transformed strains with a particular construct targeted to the *qutE* locus either express β-galactosidase or do not. However, this uniformity is not observed with transformed strains in which the vector molecules have integrated elsewhere in the genome.

With the constructs p5′PGK-LACZ2-1, 2-2, 2-3 and 2-4, in which 5′ *pgk* sequence extends at least 239bp upstream of the transcript start site, all type II transformants tested had β-galactosidase activity. However, with the construct p5′PGK-LACZ2-5 (161bp), one out of fifteen type II transformants did not produce β-galactosidase, even though all type I transformants did. This single case could either reflect the damage or loss of gene fusion sequence upon integration, or the incorporation of vector molecule into a transcriptionally silent region of the genome.

In the cases of the constructs p5′PGK-LACZ2-6 (-120bp), 2-7 (-116bp) and 2-8 (-111bp), at least 50% of the type II transformants have β-galactosidase activity even though none of the type I transformants do. This may reflect integration of vector molecules into another promoter, or simply integration of vector downstream of sequence with chance homology to the *pgk* promoter sequence element that is missing in these constructs. If these *pgk-lacZ* fusions only lack a single element of the *pgk* promoter, then integration into the genome downstream of sequence with homology to this element may allow expression.

6.4.2 Levels of β-galactosidase expression with *Aspergillus nidulans* transformants.

In order to compare the effect of different extents of 5'*pgk* promoter sequence upon *lacZ* expression, enzyme assays were conducted on cell free extracts prepared from transformed strains of *Aspergillus nidulans* with the relevant vector sequence targeted to the *qutE* locus in a single copy. The consistency of β-galactosidase enzyme assays with single copy targeted integrants of a specific construct have been demonstrated with strains transformed with the unmodified vector p5'PGK-LACZ2 (Chapter 5, Section 4). Therefore, initially only one example of a single copy targeted integrant for each construct was assayed for β-galactosidase activity. However, for all transformants that were analysed, cell free extract was prepared and assayed independently on two occasions and a mean specific activity calculated. The variation in activity observed for a particular transformant did not exceed 9% of the mean.

In order to obtain an initial indication of β-galactosidase activities, cell free extracts were assayed from A.*nidulans* transformed strains that had incorporated the vectors p5'PGK-LACZ2-1 (-354bp), 2-3 (-292bp), 2-8 (-111bp) and 2-12 (-44bp) at a single copy at the *qutE* locus. As a control, a cell free extract was prepared and assayed from a single copy targeted integrant of the unmodified vector p5'PGK-LACZ2 (-487bp). Transformants with the vector 2-1 (-354bp) and 2-3 (-292bp) produced β-galactosidase at 91% and 84% respectively of the level obtained with the transformant with the unmodified construct (-487bp), and transformants with the vectors 2-8 (-111bp) and 2-12 (-44bp) did not produce β-galactosidase at all (Table 6.1). This confirms the results of screening transformed strains for β-galactosidase production on X-gal plates and clearly demonstrates that the 5' extent of the *pgk* promoter lies upstream of a point -111bp to the transcript start site.

Table 6.1

The effect of 5′*pgk* sequence deletions on the β-galactosidase activities of strains transformed with the *pgk-lacZ* fusion vectors.

The strain QG716 was transformed with the *BglII* linearized in phase fusion vectors in the 5′*pgk* exonuclease digest series, to give the transformed strains listed in the table. All of the transformants assayed had been characterized as being single copy type I integrants.
Cell free extracts prepared from mycelium were assayed for β-galactosidase activity under standard conditions.

* The 5′ extent of *pgk* sequence is relative to the transcript start site.

Strain	5′ extent of pgk sequence (bp) *	β-galactosidase specific activity ($\Delta OD_{420} x 10^3 min^{-1} mg^{-1}$)
Recipient QG716	-	0 (0.00)
Transformants		
5′PGK-LACZ2 :L8	487	2600 (1.00)
2-1 :L11	354	2370 (0.91)
2-2 :L1	345	2080 (0.80)
2-3 :L3	292	2180 (0.84)
2-4 :L4	239	1980 (0.76)
2-5 :L24	161	2570 (0.99)
2-6 :L25	120	0 (0.00)
2-7 :L6	116	0 (0.00)
2-8 :L6	111	0 (0.00)
2-9 :L10	90	0 (0.00)
2-12:L5	44	0 (0.00)

Subsequently cell free extracts were assayed from *A.nidulans* transformed strains with a single copy of the vectors p5'PGK-LACZ2-2 (-345bp), 2-4 (-239bp), 2-5 (-161bp), 2-6 (-120bp), 2-7 (-116bp) and 2-9 (-90bp) targeted to the *qutE* locus (Table 6.1). From the data it can be seen that when up to 120bp of sequence 5' to the *pgk* transcript start site is located upstream of *lacZ*, the fusion gene is not expressed. However, when 161bp of 5'*pgk* sequence is present upstream of the fusion, the β-galactosidase activity is 99% of that recorded with the most extensive construct extending 487bp upstream of the start site, whilst with 239bp of 5'*pgk* sequence, the level of expression is reduced to 76% of that with the largest construct. The β-galactosidase activities recorded for transformants with constructs including 292bp, 345bp and 354bp of 5'*pgk* sequence are intermediate between those recorded for constructs with 239bp and 487bp.

The most suprising aspect of these results is that sequence included within 120bp of the *pgk* transcript start site is not alone sufficient for the expression of the *lacZ* fusion gene. This is despite the presence of all the sequence elements which might encompass a "core promoter" predicted to be important in controlling gene expression by virtue of homology to "promoter boxes". The putative TATA and CCAAT boxes located upstream of many eukaryotic genes, as well as the pyrimidine block found upstream of many highly expressed fungal genes, are located within 81bp of the transcript start site.

The activity of β-galactosidase that was observed when 161bp of 5'*pgk* sequence is present upstream of the *lacZ* fusion implies that a positively acting promoter element is located within this sequence. Thus, since no β-galactosidase activity was recorded when 120bp of 5'*pgk* sequence were present, this promoter element can be localized between 120bp and 161bp upstream of the transcript start site; or possibly spanning position -120, when deletion of sequence upstream of this would cut into the element.

Comparison between the β-galactosidase activities recorded with transformed strains with more extensive 5'*pgk* sequence fused to *lacZ*, failed to reveal any further sequence elements likely to play a major role in the control of expression of the fusion gene. However, there may be two elements responsible for minor effects, since a strain with sequence extending 239bp upstream of the transcript start has only 76% of the activity of one with sequence extending either 487bp or 161bp upstream. The β-galactosidase activity is considerably less than that recorded with any of the original set of type I transformants, all with 487bp of 5'*pgk* sequence (Chapter 5, Section 4).

In summary, one major *pgk* promoter element has been identified. It is a positively acting element located between 120bp and 161bp upstream of the transcript start site. Two other elements playing a less significant role in the control of gene expression may lie further upstream; a negatively acting element between -161bp and -239bp relative to the transcript start site, and a positively acting one betweeen -239bp and -487bp.

6.5 Investigation of a putative glycolytic box sequence.

The original in phase fusion construct p5'PGK-LACZ2 contained 487bp of 5'*pgk* sequence upstream of the transcript start site, and this was expected to include all sequence elements that would be likely to affect the expression of the *lacZ* gene fusion. However, after the original construct had been made, another putative promoter element located further upstream than the 5'*pgk* sequence present in the vector was identified by virtue of sequence homology.

Three other *Aspergillus nidulans* genes encoding glycolytic enzymes in addition to *pgk* have been isolated and sequenced: the glyceraldehyde-3-phosphate dehydrogenase gene *gpd* (Punt *et al.*, 1988), the triose phosphate isomerase gene *tpi* (McKnight *et al.*, 1986) and the pyruvate kinase gene *pki* (de Graaff, 1989). Nucleotide sequences have been reported -712bp with reference to the transcript start site for *gpd*, -232bp for *tpi* and -678bp for *pki*. Comparison of sequences upstream of putative "core promoter" elements between *gpd* and *pgk* revealed a region of homology around 500 to 600bp 5' to the transcription start sites. There is a 16 out of 24 base pair match between sequence extending 592bp to 615bp upstream of the transcript start site of *gpd* and sequence extending 499bp to 522bp upstream of *pgk* (Punt *et al.*, 1988). Scanning sequence upstream of the transcript start sites of *tpi* and *pki* revealed sequence similar to this region which extends -182bp to -205bp for *tpi* and -457bp to -480bp for *pki*. These upstream regions of similar sequence in the four *A.nidulans* glycolytic genes are compared in Figure 6.5.

Within the block of similar sequence there is an eight base pair region that is well conserved, corresponding to sequence between 507bp and 514bp upstream of the transcript start site of *pgk*. This octomer has a seven nucleotide match to either *gpd* or *pki* sequence and a six nucleotide match to *tpi* sequence. The presence of this conserved sequence upstream of the four *A.nidulans* glycolytic genes indicates that it may be an important element involved in

Figure 6.5
Comparison of the putative glycolytic box identified upstream of the cloned *A.nidulans* glycolytic genes.

The numbering shown above the sequences is relative to the major site of transcription initiation of the genes (+1). An eight base pair region that is well conserved between the four sequences is indicated.

			-499
3-phosphoglycerate kinase	:	TGCTATTTTGAGGTGTAATGCATG	
			-592
glyceraldehyde-3-phosphate dehydrogenase	:	TGGCGCTCTGAGGTGCAGTGGATG	
			-182
triose phosphate isomerase	:	TCCTTTAATGTGGTATTTAGGTAG	
			-457
pyruvate kinase	:	ATTTATGTTGTGGTGTGACTGCCC	
		A GT	
Apparent consensus	:	TGTGGTAC	

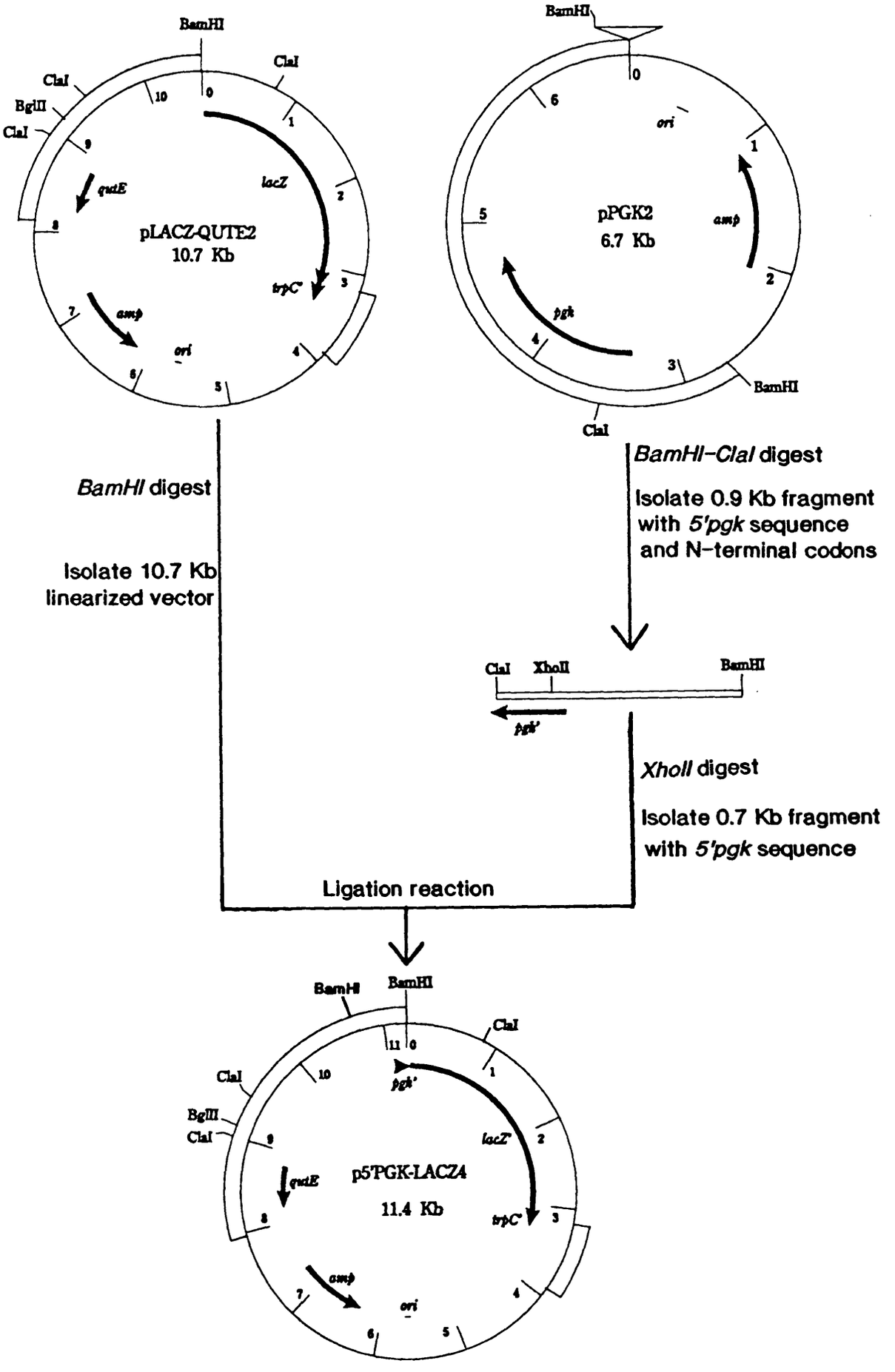
the regulation of glycolytic gene expression in this organism.

In order to investigate the effect of this newly identified element upon gene expression, another construct was made in which sequence extending further upstream of *pgk* was fused to *lacZ*. The route of vector construction is outlined in Figure 6.6. The aim was to introduce a 0.7Kb *BamHI-XhoII* restriction enzyme fragment consisting of *pgk* sequence extending from 638bp upstream of the transcription start site to 13 codons into the coding sequence, in the correct translational reading frame with respect to *lacZ*. Since this fragment could not easily be isolated in one step, the 0.9Kb *BamHI-ClaI* fragment spanning the above sequence was first isolated from the vector pPGK2. This was then digested with the restriction enzyme *XhoII* and the required fragment was isolated and introduced into the vector pLACZ-QUTE2 (Chapter 4, Section 2), which had been linearized with *BamHI* immediately upstream of *lacZ* sequence. The new construct has 638bp of 5' *pgk* sequence upstream of a *pgk-lacZ* fusion which should be in the correct reading frame. As before, the *qutE* gene is present on the vector to enable the construct to be selected for in *A.nidulans* and to facilitate targeting to the *qutE* locus. The construct was designated p5'PGK-LACZ4 and its structure was verified by restriction enzyme analysis.

This new 5'*pgk-lacZ* fusion was linearized at the unique *BglII* site (Chapter 5), and was transformed into the *A.nidulans* strain QG716. Transformants were selected for growth on quinic acid and 12 of those obtained were replated twice on selective media. The nature of the integration event in each transformed strain was determined as before by DNA hybridization using a *qutE* gene probe (Chapter 5). Comparison of the predicted fragment sizes with those observed allowed the transformed strains to be characterized. The proportions of each type of integration event are given in the Appendix. Amongst the 12 transformants 4 were identified as being type I integrants (3 at single copy), 4 as type II integrants, 2 as type III

Figure 6.6
Construction of the vector p5'PGK-LACZ4.

The vector pPGK2 consists of a 4.0Kb fragment spanning the *pgk* gene subcloned into the plasmid pUC13 (Clements and Roberts, 1985). The 0.7Kb *BamHI-XhoII* fragment with 5'*pgk* sequence could not be isolated directly because of numerous other *XhoII* sites in the vector pPGK2 (not shown in Figure). Therefore, the vector pPGK2 was first digested with the restriction enzymes *BamHI* and *ClaI*, and a 0.9Kb fragment isolated. This was then digested with *XhoII* to generate the required fragment. The construction of the vector pLACZ-QUTE2 is described in Chapter 4, Section 2.
The restriction enzymes *BamHI* and *XhoII* produce complementary single stranded ends. In the case of this ligation reaction, the *BamHI* site is reconstituted.


BamHI
ClaI
ClaI
BglII
ClaI
lacZ
qutE
pLACZ-QUTE2
10.7 Kb
trpC'
amp
ori
BamHI
ori
pPGK2
6.7 Kb
amp
pgk
BamHI
ClaI
BamHI-ClaI digest
Isolate 0.9 Kb fragment
with 5'pgk sequence
and N-terminal codons
BamHI digest
Isolate 10.7 Kb
linearized vector
ClaI
XhoII
BamHI
pgk'
XhoII digest
Isolate 0.7 Kb fragment
with 5'pgk sequence
Ligation reaction
BamHI
BamHI
ClaI
ClaI
BglII
ClaI
pgk'
lacZ'
qutE
p5'PGK-LACZ4
11.4 Kb
trpC'
amp
ori

replacements and 2 gave fragment sizes indicative of mixed type I integration and type III replacement events. The 12 transformed strains were screened for the production of β-galactosidase on X-gal plates. All of the strains stained blue, except for the 2 type III replacements. Thus, all transformants that had incorporated *lacZ* sequence expressed β-galactosidase.

To compare expression between transformants with or without the putative glycolytic box, cell free extract was prepared from two strains with a single copy of p5'PGK-LACZ4 (638bp of 5'*pgk* sequence) targeted to the *qutE* locus, and from one strain with the original *lacZ* fusion construct p5'PGK-LACZ2 (487bp). Assays for β-galactosidase activity were conducted on each of the cell free extracts prepared from mycelium grown in glucose media. The two strains transformed with 638bp of 5'*pgk* sequence upstream of *lacZ* recorded a β-galactosidase activity of 2944 units and 2673 units, whilst the strain with 487bp of 5'*pgk* sequence recorded an activity of 2700 units. Thus, the mean of the activities recorded for the cell free extracts with 638bp of 5'*pgk* sequence upstream of *lacZ* is 104% of that with 487bp of sequence. This is well within the range previously recorded with different type I transformants all having 487bp of 5'*pgk* sequence (Chapter 5, Section 4), and thus the difference can be discounted.

Therefore, it can be concluded that the putative glycolytic box does not play a role in controlling the level of expression of *pgk* when the organism is grown on glucose. However, to demonstrate whether the putative glycolytic box plays any role at all in modulating *pgk* gene expression, it was necessary to examine the effect of growing the strains on other carbon sources than glucose.

6.6 The effect of carbon source upon the expression of the gene fusion.

The activity of 3-phosphoglycerate kinase in *A.nidulans* after growth of mycelium on different carbon sources in liquid culture has been determined (Clements, 1986). Enzyme activity was demonstrated to be greater in extracts prepared from mycelium grown on acetate or alanine than in that grown on glucose. With glycerol as the carbon source, enzyme activity was less than on glucose. However, *PGK* activities were never more than two fold different from the activity in glucose grown mycelium, and it was considered that such a difference may be due to the mycelium being harvested at a different stage in the growth cycle for each carbon source (Clements, 1986). However, when cell free extract prepared from a transformant with 5'*pgk* sequence fused to *lacZ* (5'PGK-LACZ2:L28) was assayed for β-galactosidase activity over the time course of a growth cycle, no significant variation was detected (Chapter 5, Section 4).

To test whether expression of the fusion gene varied on different carbon sources in the way described above for the native phosphoglycerate kinase, two type I transformed strains (5'PGK-LACZ4:L1 and 5'PGK-LACZ2:L8) with 5'*pgk* sequence fused to *lacZ* were assayed for β-galactosidase activity following growth in glucose, glycerol, acetate, quinic acid or benzoic acid. The transformed strain 5'PGK-LACZ4:L1 has 638bp of 5'*pgk* sequence upstream of *lacZ*, including the putative glycolytic box identified upstream of the four cloned *A.nidulans* glycolytic genes, whilst the transformed strain 5'PGK-LACZ2:L8 has 487bp of 5'*pgk* sequence and lacks the putative glycolytic box. Cell free extracts were prepared from mycelium harvested from each of the cultures at a similar phase of growth (as determined by the weight of mycelium per unit volume of culture) and were assayed for β-galactosidase activity. The experiment was conducted on two separate occassions and the results are presented in Table 6.2. In the case of the transformed strain 5'PGK-LACZ4:L1, which includes the putative glycolytic box upstream of *lacZ*, following growth on

Table 6.2

Native 3-phosphoglycerate kinase and β-galactosidase fusion protein activities in A.*nidulans* strains grown on different carbon sources.

a) Activity of 3-phosphoglycerate kinase in the wild type strain R153 taken from Clements (1986). One unit of specific activity is defined as the amount of enzyme required for the formation of 1μmol of glyceraldehyde-3-phosphate dehydrogenase or NAD^- per minute per mg protein.

b) Activity of β-galactosidase in the transformed strains 5'PGK-LACZ4:L1 and 5'PGK-LACZ2:L8. The β-galactosidase assays were conducted under standard conditions, and one unit of specific activity is defined as the amount of enzyme required to produce an O.D. change of 0.001 per minute per mg protein.

With each set of assays, the experiment was conducted on two occassions and a mean taken.

c) Summary of data.

a)

Carbon source	Specific activity of *PGK* (μmol min^{-1} mg^{-1})		
	Experiment 1	Experiment 2	Mean
Wild type strain R153			
Glucose	1.03	1.11	1.07 (100%)
Glycerol	0.57	0.75	0.66 (62%)
Acetate	1.16	2.24	1.70 (159%)
Quinic acid	1.08	-	1.08 (101%)
Alanine	-	1.92	1.92 (179%)

b)

Carbon source	Specific activity of β-galactosidase ($\Delta OD_{420} x 10^3 min^{-1} mg^{-1}$)		
	Experiment 1	Experiment 2	Mean
Transformed strain 5'PGK-LACZ4:L1 (638bp of 5'*pgk* seq.)			
Glucose	2650	2850	2750 (100%)
Glycerol	1990	1880	1935 (70%)
Acetate	4740	4470	4605 (167%)
Quinic acid	5490	6300	5895 (214%)
Benzoic acid	-	4010	4010 (146%)
Transformed strain 5'PGK-LACZ2:L8 (487bp of 5'*pgk* seq.)			
Glucose	2500	2810	2655 (100%)
Glycerol	1220	-	1220 (46%)
Acetate	2500	3090	2795 (105%)
Quinic acid	3175	3935	3555 (134%)
Benzoic acid	-	2530	2530 (95%)

c)

Carbon source	Native PGK activity	β-galactosidase activity	
		With glycolytic sequence	Without glycolytic sequence
Glucose	100%	100%	100%
Glycerol	62%	70%	46%
Acetate	159%	167%	105%
Quinic acid	101%	214%	134%
Benzoic acid	-	146%	95%

glycerol the β-galactosidase activity was 70% of that recorded on glucose, whilst on acetate, quinic acid and benzoic acid, it was 167%, 214% and 140% respectively of that recorded on glucose.

The proportional enzyme activities recorded for growth on glycerol or acetate compared to glucose are very similar to those found by Clements, 1986 (Table 6.2). However, β-galactosidase activity in quinate grown mycelium was 214% the level of glucose grown mycelium, whereas the native *PGK* activity had been reported as being about the same. The greater level of activity observed with the fusion protein may result from the integration of vector sequence into the quinic acid gene cluster. The genes within this cluster are stongly induced by quinic acid, and it is suggested here that this may affect expression of the gene fusion when it has integrated at this locus and the transformed strain grown on quinate.

In the case of the transformed strain 5'PGK-LACZ2:L8, which lacks the putative glycolytic box upstream of *lacZ*, following growth on glycerol the β-galactosidase activity was 46% of that recorded with glucose, whilst on acetate, quinic acid and benzoic acid it was 105%, 134% and 95% respectively of that recorded on glucose. Thus, in the absence of sequence located between 487 and 638bp upstream of *pgk* (including the putative glycolytic box), expression of the *lacZ* fusion is very similar on acetate or benzoic acid to that recorded on glucose. Therefore, this stretch of 5'*pgk* sequence spanning the putative glycolytic box appears to modulate expression of the *lacZ* fusion (and by inference native pgk) in response to the carbon source for growth.

As shown in Chapter 3, Figure 1; glycerol enters the glycolytic pathway as glyceraldehyde-3-phosphate, but may not be as rapidly metabolised as glucose, as implied by the slower growth rate of *A.nidulans* on glycerol. Acetate and benzoic acid are metabolised through gluconeogenesis to generate glycolytic intermediates for biosynthesis. The

weaker β-galactosidase activity recorded for both transformed strains on glycerol compared to glucose was not surprising given the difference in growth rates, and is similar to that for native *PGK*. The greater enzyme activity recorded on acetate and benzoic acid when the putative glycolytic box is present upstream of *lacZ*, may reflect the gluconeogenic activity of *PGK*, since under these conditions *PGK* would be involved in glucose synthesis to provide a precursor for the pentose phosphate pathway and cell wall carbohydrates.

6.7 Summary.

Promoter elements likely to be involved in controlling *pgk* gene expression had been identified by sequence comparison with other fungal and higher eukaryotic genes (Clements and Roberts, 1986). In this way putative TATA and a CCAAT boxes were identified, as well as a pyrimidine rich sequence located just upstream of the transcription start site.

In order to investigate whether these sequences were required for *pgk* expression, a series of vectors were constructed with varying extents of 5'*pgk* sequence upstream of *lacZ*. These constructs were transformed into *Aspergillus nidulans* and single copy integrants with vector sequence targeted to the *qutE* locus were identified.

Assays for β-galactosidase activity on these transformed strains demonstrated that sequence within 120bp of the transcription start site has no effect upon expression of the gene fusion. This region includes the putative core promoter elements described above. However, since the mRNA start site has not been mapped in any of these fusions, it is possible that the putative core sequences play a role in positioning the initiation of transcription. In contrast, a sequence element located between 161bp and 120bp upstream (or possibly spanning 120bp) is essential for expression.

More recently a putative glycolytic box has been identified upstream of the four glycolytic genes that have been cloned from *A.nidulans*, and a further construct was made which included this sequence. Analysis of transformed strains of *A.nidulans* indicates that the putative glycolytic box is responsible for modulating expression on different carbon sources. When present upstream of *lacZ*, the β-galactosidase activity was shown to be increased on the gluconeogenic carbon sources of acetate, benzoic acid and quinic acid, compared to the glycolytic carbon source of glucose. In the light of this observation the glycolytic

box may be more appropriately named as the gluconeogenic box, since it apparently modulates a response to growth in gluconeogenic carbon sources.

The increased enzyme activity is most pronounced when quinic acid is the carbon source, and the positive effect of quinate upon expression may partly result from integration of the *pgk-lacZ* gene fusion into the quinic acid gene cluster: indicating that the location of a particular sequence in the genome may affect its expression (as discussed in Chapter 5).

The consensus sequence element that has been identified upstream of the cloned *A.nidulans* glycolytic genes has a different function to the carbon source modulation domain of the UAS of the *Saccharomyces cerevisiae PGK* gene. In contrast to the *A.nidulans* sequence, the yeast sequence is required for the induction of *PGK* expression by glucose (Stanway *et al.*, 1987). Furthermore, the *A.nidulans* and *S.cerevisiae* regulatory elements have no DNA sequence similarity.

Analysis of transformed strains in which the site of vector integration is elsewhere other than at the *qutE* locus (type II transformants), has provided evidence that the site of vector integration affects expression of the gene fusion. With constructs that included sufficient 5'*pgk* sequence to promote *lacZ* expression in targeted transformants, there was one example of a type II transformant that did not express the fusion gene. On the other hand, constructs that lacked the necessary sequence to promote expression when integrated at the *qutE* locus, sometimes regained expression of *lacZ* when located elsewhere in the genome.

Chapter 7.

The effect of *pgk* coding sequences on the expression of the *lacZ* gene fusions.

7.1 Introduction.

The *PGK* gene of *Saccharomyces cerevisiae* has been cloned and sequenced (Hitzeman *et al.*, 1982; Dobson *et al.*, 1982). The sequence 5′ to the gene includes two elements with homology to the consensus TATA box and one potential CCAAT box. However, these elements are positioned further upstream than the consensus positions for higher eukaryotic genes, and deletions within the 5′ sequence have demonstrated that neither the putative TATA boxes nor the CCAAT box are essential for either maintaining message level or for directing the position of the transcription start site (Ogden *et al.*, 1986). Rather, an upstream activation sequence (UAS) affecting transcription has been located further upstream. This element consists of two functionally distinct domains, one responsible for transcriptional activation and the other for carbon source dependent regulation of transcription (Stanway *et al.*, 1987). Although deletion of this element causes a fall in the rate of message synthesis, transcription is not totally abolished. From the deletion analysis of the remaining promoter, sequences responsible for this basal level of expression must lie within 55bp of the transcription start site (Ogden *et al.*, 1986). This region includes a pyrimidine rich sequence which has also been identified in other highly expressed yeast genes (Dobson *et al.*, 1982), and has been proposed as being an important element in positioning the transcriptional initiation site (McNeil and Smith, 1985).

In addition, results from gene fusion experiments suggest that sequences located within the *PGK* coding region itself may be more important in controlling transcription than 5′ promoter sequences. The *S.cerevisiae PGK* gene encodes one of the most abundant mRNA and protein species in

the cell, accounting for 1-5% of the total cellular message and protein (Holland and Holland, 1978). When present on a high copy number plasmid the gene can direct between 50 and 80% of total cellular protein synthesis, and this is reflected in the steady state level of the message (Mellor *et al.*, 1985). However, when the *PGK* coding sequence is replaced with that of other genes, protein synthesis is reduced at least 20-fold relative to *PGK* itself (Tuite *et al.*, 1982; Mellor *et al.*, 1983; Dobson *et al.*, 1984; Mellor *et al.*, 1985).

Suprisingly, fusions of the *S.cerevisiae PGK* promoter to the cDNA encoding human *PGK* do not show reduced message levels in yeast (Chen and Hitzeman, 1987). Although the amino acid sequences of the proteins from human and yeast are 65% conserved, codon usage in the two genes is very different. Thus, the drop in expression that occurs in the heterologous fusions described above cannot be explained by differences in codon usage. Furthermore, this reduced expression can only be partly explained by the relative instability of some of the heterologous proteins in yeast. Rather, the principal reason is a 5 to 10-fold decrease in the concentration of heterologous message compared to intact *PGK* message (Mellor *et al.*, 1983; Mellor *et al.*, 1985).

There is evidence implying that the *PGK* protein may stabilize its own message, since nonsense mutations within the yeast *PGK* gene result in a decreased message level (Chen *et al.*, 1984). However, the normal *PGK* expression system is unable to correct defective units present elsewhere on the same plasmid molecule, implying that the protein cannot act as a soluble factor to stabilize message produced at another transcription site.

Subsequent analysis indicated that the reduced message level seen with heterologous fusions correlates to a defect in the rate of synthesis of heterologous message, rather than decreased message stability (Mellor *et al.*, 1987). A positive transcriptional control element has been identified within the *PGK* coding region by comparing mRNA levels from

homologous and heterologous transcription units. This downstream activator sequence (DAS) has been located between 37 and 236bp downstream of the translation initiation codon, and its deletion results in a 6 to 10-fold reduction in the level of transcription.

There are several other examples of transcriptional control elements being located within the coding sequence of a gene. The best characterized are those of the 5S rRNA gene and tRNA genes transcribed by RNA polymerase III (Korn, 1982; Hall *et al.*, 1982), but intragenic transcriptional control regions have also been identified within RNA polymerase II transcribed genes. For example, β-lymphocyte specific enhancer sequences are present within the rearranged activated immunoglobulin heavy chain genes (Gilles *et al.*, 1983), and regulatory sequences are located within the β-globin gene (Kosche *et al.*, 1985). More pertinent to fungal glycolytic genes, a fusion of the *lacZ* gene of *Escherichia coli* to sequences upstream of the pyruvate kinase (*PYK*) gene of *S.cerevisiae* resulted in a 20-fold reduction in the steady state level of heterologous mRNA relative to the native *PYK* message (cited in Brown and Lithgow, 1987). Since these messages do not differ more than two-fold in stability, a DAS is again implicated in regulating transcription.

I intended to examine the coding region of the *pgk* gene of *A.nidulans* to discover whether such an intragenic activator sequence is present. I proposed to do this by constructing further fusion vectors with the *lacZ* reporter gene of *E.coli* positioned downstream of both 5'*pgk* sequence and extensive *pgk* coding sequence. Recovery of transformed strains with single copy targeted constructs would enable comparision of β-galactosidase activities with those previously observed with a *lacZ* fusion to 5'*pgk* sequence alone (Chapters 5 and 6), and would thus identify any effect upon expression of *pgk* coding sequence.

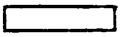
7.2 Construction of *lacZ* fusions that include *pgk* coding sequence, and introduction of these into *A.nidulans*.

The original 5'*pgk-lacZ* gene fusion has 487bp of 5' sequence, the 32bp untranslated leader sequence and the 13 N-terminal codons of *pgk* positioned immediately upstream of *lacZ* (Chapter 4). In order to precisely determine the effect upon expression of intragenic *pgk* sequences it was necessary to construct a vector resembling that described above, except that it should also include more *pgk* coding sequence. The putative downstream activator sequence of the *Saccharomyces cerevisiae pgk* gene has been located between codons 13 and 78, so it was decided to include at least an equivalent amount of N-terminal sequence of the *pgk* gene in the initial construct. The positions of the two *Aspergillus nidulans pgk* introns relative to the *S.cerevisiae* gene, which has no introns, are indicated in Figure 7.1.

The strategy chosen to construct the new *lacZ* fusion vector is outlined in Figure 7.2. The vector pLACZ-QUTE1 was digested with the restriction enzymes *XhoI* and *SalI*. The digestion products were separated by agarose gel electrophoresis and a 9.85Kb fragment, including the *lacZ* gene of *Escherichia coli* without its 7 N-terminal codons and the *qutE* gene of *A.nidulans* , was recovered from the gel. This fragment was incubated in a ligation reaction with a 1.2Kb *XhoI-SalI* restriction enzyme fragment isolated in a similar fashion from the vector pPGK2 (Clements and Roberts, 1985). This second fragment consisted of *pgk* sequence extending from -487bp to +694bp relative to the transcription start site. The structure of the new construct was verified by restriction digest analysis. It was designated pPGK-LACZ1 and included the 182 N-terminal codons and the two introns of *pgk* immediately upstream of *lacZ*. It was predicted to be in the correct translational reading frame.

Subsequently, a second fusion vector was constructed with a further 313bp of *pgk* coding sequence positioned between 5'*pgk* sequence and *lacZ*. The construction of this

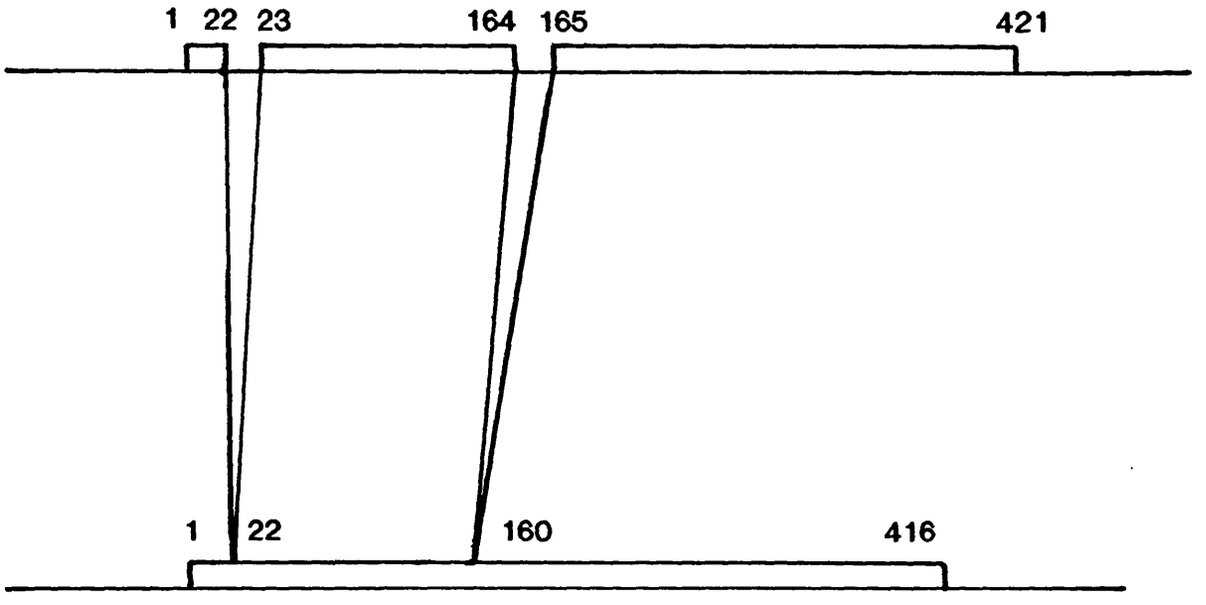
Figure 7.1
The positions of the *A.nidulans pgk* introns relative to the *S.cerevisiae pgk* gene.

In this figure: coding sequence is represented as 
and non-coding sequence as

The numbering shown above the sequences refers to the codons of the *A.nidulans* and *S.cerevisiae pgk* genes.

When aligned for maximum identity, the *A.nidulans* and *S.cerevisiae* sequences show 68% homology at the amino acid level (Clements and Roberts, 1986).

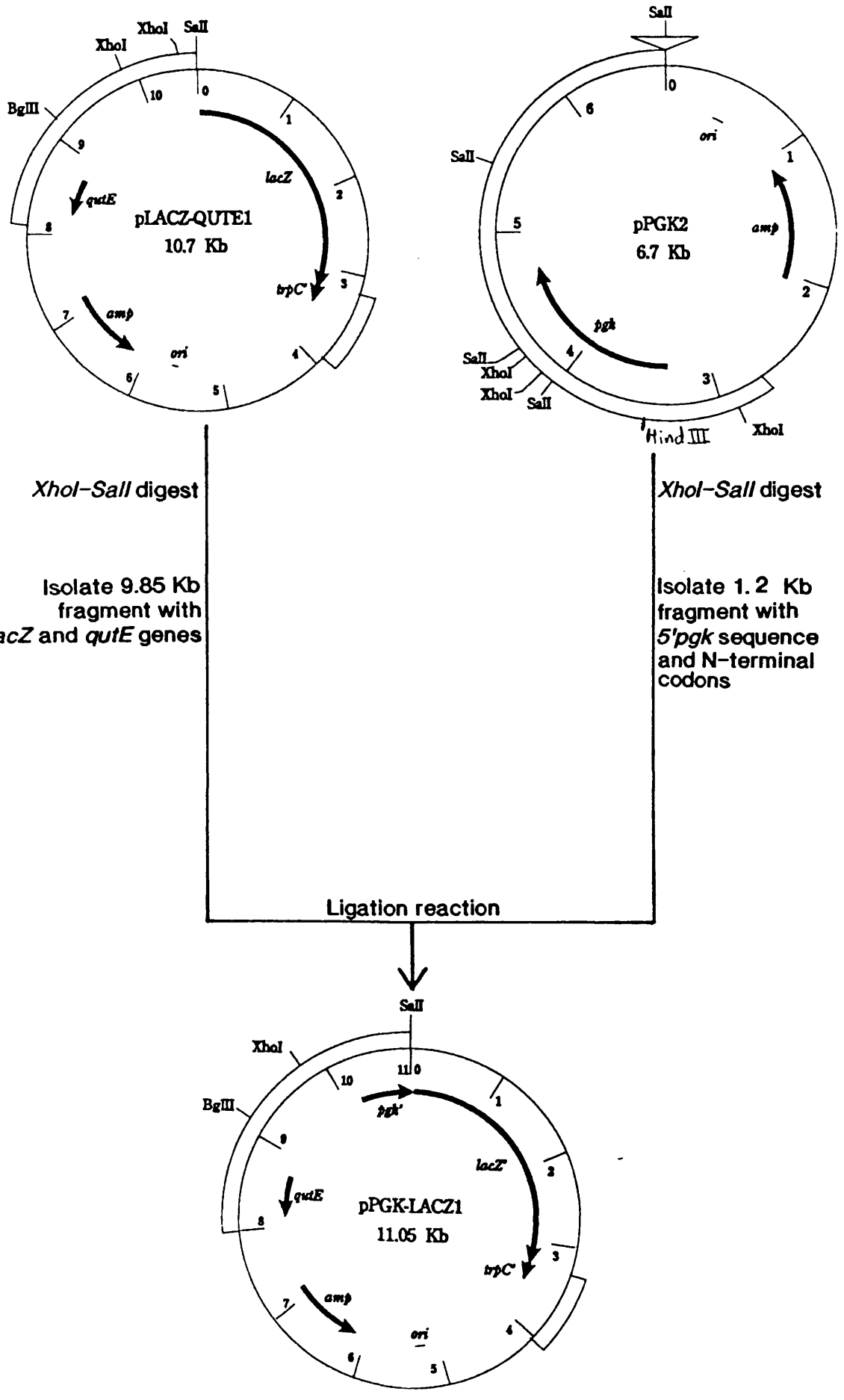
A.nidulans pgk



S.cerevisiae pgk

Figure 7.2
Construction of the vector pPGK-LACZ1.

The vector pPGK2 consists of a 4.0Kb fragment spanning the *A.nidulans pgk* gene subcloned into the plasmid pUC13 (Clements and Roberts, 1985). The construction of the vector pLACZ-QUTE1 is described in Chapter 4, Section 2.


XhoI
SalI
XhoI
BglII
lacZ
qutE
pLACZ-QUTE1
10.7 Kb
trpC
amp
ori
SalI
ori
pPGK2
6.7 Kb
amp
pgk
SalI
XhoI
XhoI
SalI
HindIII
XhoI
XhoI-SalI digest
Isolate 9.85 Kb fragment with lacZ and qutE genes
XhoI-SalI digest
Isolate 1.2 Kb fragment with 5'pgk sequence and N-terminal codons
Ligation reaction
SalI
XhoI
BglII
pgk'
lacZ'
qutE
pPGK-LACZ1
11.05 Kb
trpC
amp
ori

vector is given in Figure 7.3. A 0.3Kb *SalI* fragment, consisting of *pgk* coding sequence extending from +694 to +1007bp, was subcloned into the newly constructed vector pPGK-LACZ1, which had been linearised by digesting with *SalI*. As above, the structure of the construct was verified by restriction digest analysis. This vector, designated pPGK-LACZ2, was predicted to be in the correct translational reading frame and had the 286 N-terminal codons and the two introns of *pgk* fused to *lacZ*.

Finally, a third fusion vector was constructed to allow any far upstream regulatory elements to be identified. This vector had approximately 3150bp of 5'*pgk* sequence and the 182 N-terminal *pgk* codons, plus the two introns, upstream of *lacZ*. The construction of this final, most extensive fusion vector is outlined in Figure 7.4. A 3.85Kb *SphI-SalI* fragment consisting of sequence extending from approximately -3150 to +694bp was isolated from the phage λ clone λ-PGK3 (Clements and Roberts, 1985). This fragment was then incubated in a ligation reaction with a 9.75Kb *SphI-SalI* fragment isolated from the vector pLACZ-QUTE1 (Chapter 4, Section 2). As before, the structure of the desired construct was verified by restriction digest analysis. It was designated pPGK-LACZ3 and was predicted to have *pgk* sequence fused to *lacZ* in the correct reading frame.

The three new fusion vectors described above (pPGK-LACZ1, 2 and 3) were then transformed into the *A.nidulans bgaA*$^{-}$; *qutE*$^{-}$ recipient strain QG716 (Chapter 4, Section 4). Prior to transformation, each of the fusion vectors was linearised with the restriction enzyme *BglII*. This enzyme cuts the vectors at one site within *A.nidulans* sequence flanking the *qutE* gene. Both the transformation frequency and the degree of targeting to the *qut* gene cluster should be increased by linearising the vector at this point (Chapter 5). Transformed strains were selected for growth on quinic acid and were then replated twice on selective medium to isolate single colonies. A total of 33 transformed strains were isolated with the vector pPGK-LACZ1, 23 with pPGK-LACZ2 and 24 with pPGK-LACZ3.

Figure 7.3
Construction of the vector pPGK-LACZ2.

The vector pPGK2 consists of a 4.0Kb fragment spanning the *A.nidulans pgk* gene subcloned into the plasmid pUC13 (Clements and Roberts, 1985).

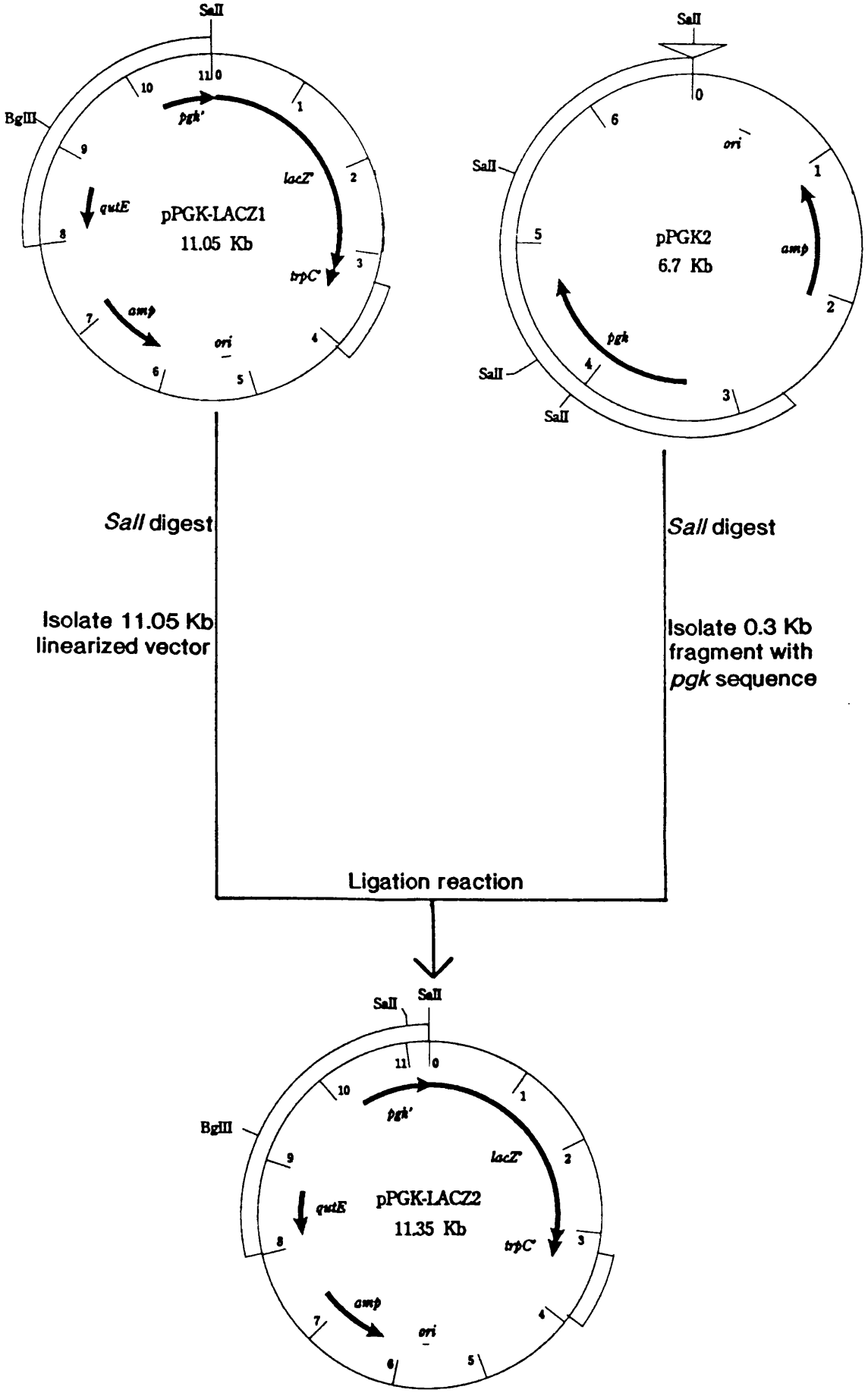
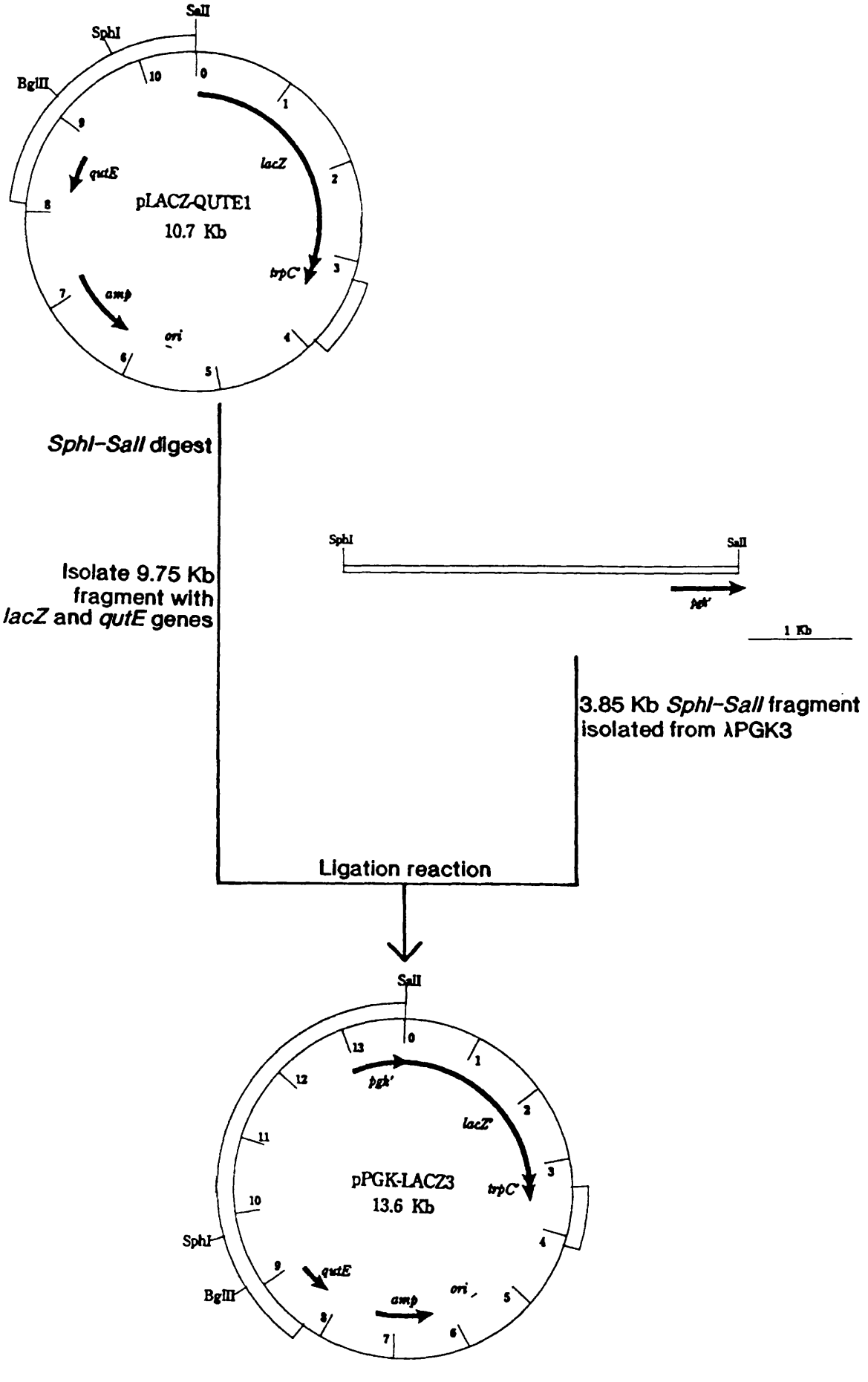

SalI
BglII
pgk'
lacZ'
qutE
pPGK-LACZ1
11.05 Kb
trpC'
amp
ori
SalI
ori
pPGK2
6.7 Kb
amp
pgk
SalI
SalI
Sall digest
Isolate 11.05 Kb
linearized vector
Sall digest
Isolate 0.3 Kb
fragment with
pgk sequence
Ligation reaction
SalI
SalI
BglII
pgk'
lacZ'
qutE
pPGK-LACZ2
11.35 Kb
trpC'
amp
ori

Figure 7.4
Construction of the vector pPGK-LACZ3.

The phage λ clone λ-PGK3 consists of approximately 12.3Kb of *A.nidulans* sequence spanning the *pgk* gene cloned into phage λ (Clements and Roberts, 1985). The construction of the vector pLACZ-QUTE1 is described in Chapter 4, Section 2.


SalI
SphI
BglII
10
0
1
9
lacZ
qutE
2
pLACZ-QUTE1
10.7 Kb
8
trpC
3
7
amp
ori
4
6
5
SphI-SalI digest
Isolate 9.75 Kb
fragment with
lacZ and qutE genes
SphI
SalI
pgk'
1 Kb
3.85 Kb SphI-SalI fragment
isolated from λPGK3
Ligation reaction
SalI
13
0
1
12
pgk'
2
lacZ'
11
3
pPGK-LACZ3
13.6 Kb
trpC'
10
4
SphI
9
qutE
ori
5
BglII
amp
8
7
6

In order to compare the level gene expression from the different gene fusions and relate this to the extent of *pgk* sequence present, it was necessary to identify single copy integrants targeted to the *qutE* locus. These were expected to express β-galactosidase, whereas some of the other types of transformants were not, in particular type III gene replacements. To save time all the transformed strains isolated were first screened for the production of β-galactosidase on X-gal plates, and genomic DNA was only prepared from those that were positive. The result was that 24 of the 33 strains transformed with the vector pPGK-LACZ1, 17 of the 23 transformed with pPGK-LACZ2, and 19 of the 24 transformed with pPGK-LACZ3 expressed the *lacZ* gene fusion.

In order to identify single copy type I integrants genomic DNA was then prepared from a number of the transformed strains that had β-galactosidase activity. In all, genomic DNA was prepared from 5 strains transformed with pPGK-LACZ1, 10 with pPGK-LACZ2 and 18 with pPGK-LACZ3. To distinguish single copy type I transformants, DNA prepared from the pPGK-LACZ1 and 2 transformants was digested with the restriction enzyme *SstI*, whereas that prepared from the pPGK-LACZ3 transformants was digested with *XhoI*. All of the restriction digests were analysed by agarose gel electrophoresis and DNA hybridization using the radioactively labelled 1.2Kb *BglII-EcoRI qutE* fragment as probe (Chapter 5). As a result of this analysis, 3 single copy type I transformants were identified with the vector pPGK-LACZ1, 1 with pPGK-LACZ2 and 9 with pPGK-LACZ3.

With the type II transformed strains the only fragment found to hybridize to the *qutE* probe in all cases within each group of transformants was that which corresponded in size to the vector sequence alone, and would be expected to result from all integrations. Thus, as with previously constructed vectors (Chapters 5 and 6), there is no evidence for the preferential integration of vector sequence at any site other than the *qutE* locus.

7.3 Expression of *lacZ* from gene fusions which include *pgk* coding sequence.

Having identified single copy type I transformants with each of the three *pgk-lacZ* fusion vectors described in Section 2, one strain of each was grown in glucose media and β-galactosidase assays conducted on cell free extracts prepared by grinding mycelium under liquid nitrogen as previously described (Chapter 5). A single copy type I transformed strain that had incorporated the original *lacZ* fusion vector was included in the analysis as a reference to previous data (Chapter 5, Section 4).

7.3.1 Instability of the fusion proteins.

Initially the assays were done with cell free extracts prepared immediately from freshly harvested mycelium. When the assays were repeated on new extracts prepared from the same mycelium that had been stored at -20°C, it was discovered that the β-galactosidase activities had, in some cases, fallen significantly (Table 7.1, Experiment 1). After four days of storage the β-galactosidase activities recorded for strains transformed with fusions including extensive *pgk* coding sequence were only approximately 50% of their original value. By contrast, the activity recorded with the strain transformed with the fusion including only 13 N-terminal *pgk* codons had decreased by less than 20%. After another four days storage the β-galactosidase activity in mycelium of this latter strain had not decreased any further, whilst that of the strains transformed with the more extensive *pgk-lacZ* fusions had again fallen. Thus, it appears that the larger fusion proteins are unstable when mycelium is stored frozen.

In order to confirm the loss of enzyme activity in frozen mycelium, cell free extracts were prepared again and assayed from each of the four transformed strains analysed above after storage for increasing periods of time. The results of these enzyme assays are given in Table 7.1, Experiment 2. Although the β-galactosidase activity

Table 7.1

The effect of storage of harvested mycelium at -20°C upon the β-galactosidase activities observed in transformed strains.

The strain QG716 was transformed with the *BglII* linearized in phase fusion vectors p5'PGK-LACZ2 and pPGK-LACZ1, 2 and 3, to give the transformed strains 5'PGK-LACZ2:L8 and PGK-LACZ1:L14, 2:L7 and 3:L2. The table shows the results of two separate experiments with these strains. The numbering is relative to the *pgk* transcript start site (+1). Cell free extracts prepared from mycelium were assayed for β-galactosidase activity under standard conditions.

Strain	Extent of *pgk* sequence	Time of storage (days)	β-galactosidase specific activity ($\Delta OD_{420} x 10^3 min^{-1} mg^{-1}$)
Experiment 1			
5'PGK-LACZ2:L8	-487 to +72	0	2700 (1.00)
		4	2240 (0.83)
		8	2190 (0.81)
PGK-LACZ1:L14	-487 to +694	0	6050 (1.00)
		4	3000 (0.50)
		8	2650 (0.44)
PGK-LACZ2:L7	-487 to +1007	0	3320 (1.00)
		4	1680 (0.51)
		8	1420 (0.43)
PGK-LACZ3:L2	-3150 to +694	0	8290 (1.00)
		4	4320 (0.52)
		8	2910 (0.35)
Experiment 2			
5'PGK-LACZ2:L8	-487 to +72	0	2830 (1.00)
		1	2600 (0.92)
		3	1870 (0.66)
		7	1870 (0.66)
PGK-LACZ1:L14	-487 to +694	0	5620 (1.00)
		1	4390 (0.78)
		3	3290 (0.59)
		7	2810 (0.50)
PGK-LACZ2:L7	-487 to +1007	0	3410 (1.00)
		1	2270 (0.66)
		3	1610 (0.47)
		7	1410 (0.41)
PGK-LACZ3:L2	-3150 to +694	0	6970 (1.00)
		1	6630 (0.95)
		3	3950 (0.57)
		7	3910 (0.56)

recorded with the strain transformed with the fusion including only 13 N-terminal codons had decreased by 34% after storing mycelium frozen for seven days, the activity recorded with the strains transformed with fusions including more extensive *pgk* coding regions decreased to a greater extent. Most strikingly, the transformed strain with the greatest extent of the *pgk* coding region (286 N-terminal codons plus both introns) fused upstream of *lacZ* had a β-galactosidase activity which decayed by 59% over seven days. The decay in enzyme activity for each of the transformants is plotted in Figure 7.5.

The native β-galactosidase protein of *Escherichia coli* is a tetramer of four identical 1021 amino acid polypeptides (Goldberg, 1969; Wallenfels and Weil, 1972). Although fusion proteins generally have β-galactosidase activity (Casadaban et *al.*, 1983), the large fusions constructed here may be affecting the stability of the tertiary and quaternary structure, and thus the enzyme activity.

7.3.2 The effect of *pgk* coding sequence upon expression of the *lacZ* gene fusion.

In order to relate the β-galactosidase activities recorded for the transformed strains directly to the extent of *pgk* sequence upstream of *lacZ* and to minimize the problem of protein instability all further enzyme assays were done on extracts prepared from mycelium immediately after harvesting (as in Chapters 5 and 6). Mycelium was harvested from cultures of the strains on four separate occasions, and the results of these assays, along with those recorded with extracts prepared from freshly harvested mycelium at the outset of the time course experiments described above, are presented in Table 7.2.

Sequence between codons 14 and 182 of *pgk*, which includes the two introns, increases β-galactosidase activity approximately two fold, whilst sequence beyond this region between codons 183 and 286 reduces activity to approximately 60% of this higher level. Moreover, sequence located far

Figure 7.5

The decay of β-galactosidase activity during frozen storage in transformed strains.

The strain QG716 was transformed with the *BglII* linearized in phase fusion vectors p5'PGK-LACZ2 and pPGK-LACZ1, 2 and 3, to give the transformed strains 5'PGK-LACZ2:L8 and PGK-LACZ1:L14, 2:L7 and 3:L2. The extent of *A.nidulans pgk* sequence is given in brackets.

Cell free extracts prepared from mycelium were assayed for β-galactosidase activity under standard conditions. Units of β-galactosidase activity are $\Delta OD_{420} x 10^3 min^{-1} mg^{-1}$.

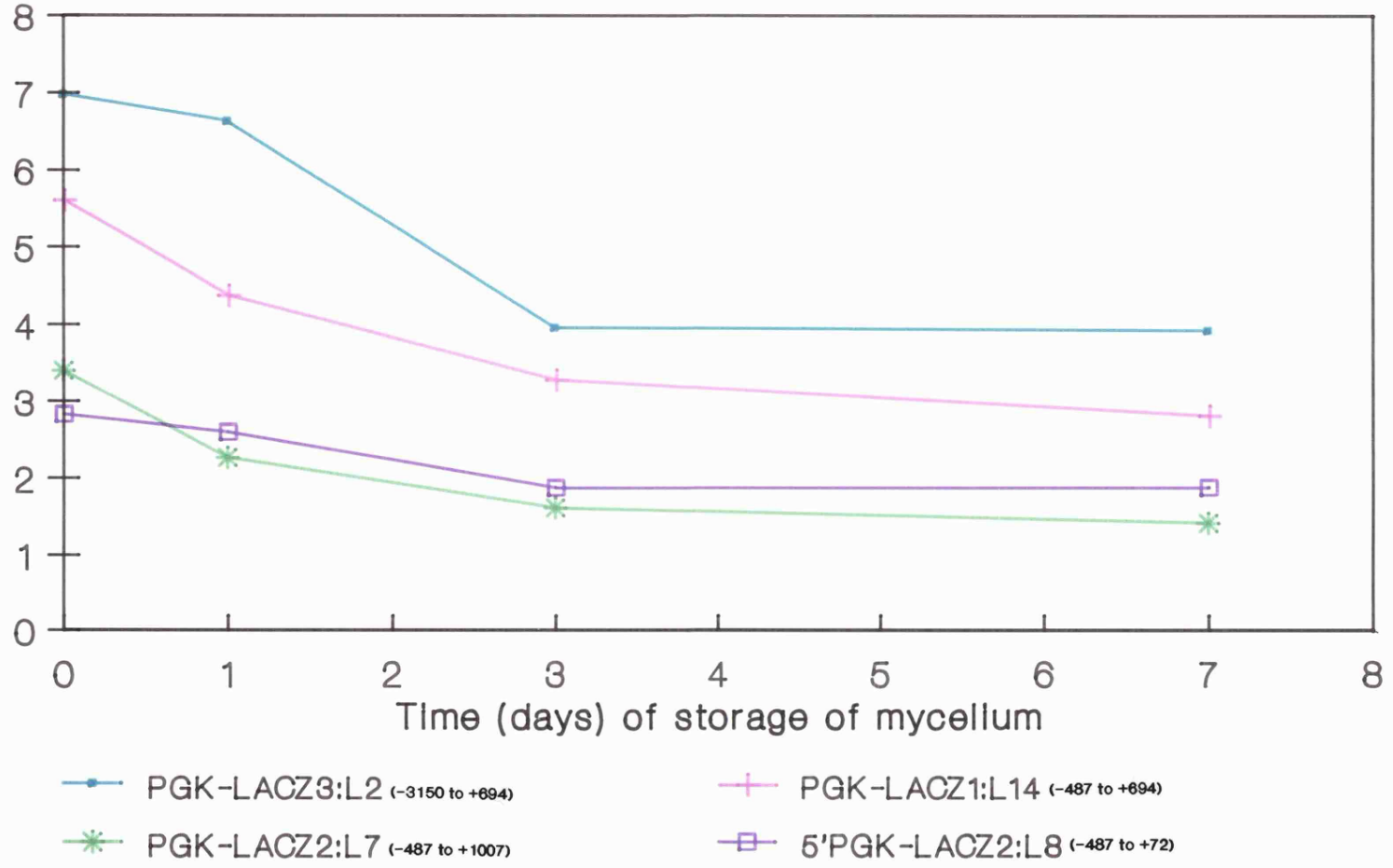

8
7
6
5
4
3
2
1
0
0
1
2
3
4
5
6
7
8
Time (days) of storage of mycelium
PGK-LACZ3:L2 (-3150 to +694)
PGK-LACZ1:L14 (-487 to +694)
PGK-LACZ2:L7 (-487 to +1007)
5'PGK-LACZ2:L8 (-487 to +72)

Table 7.2

The β-galactosidase activities of strains transformed with *pgk-lacZ* fusion vectors.

The strain QG716 was transformed with the *BglII* linearized in phase fusion vectors p5'PGK-LACZ2 and pPGK-LACZ1, 2 and 3, to give the transformed strains 5'PGK-LACZ2:L8 and PGK-LACZ1:L14 and :L8, 2:L7, and 3:L2 and L4.
Cell free extracts prepared from mycelium were assayed for β-galactosidase activity under standard conditions.
The upper part of the table summarizes the results of six separate experiments conducted on cell free extracts prepared from freshly harvested mycelium. The lower part of the table shows the results of a single experiment conducted on further transformed strains.

Strain	Extent of *pgk* sequence	β-galactosidase specific activity ($\Delta OD_{420} x 10^3 min^{-1} mg^{-1}$)	
		Mean	Range
5'PGK-LACZ2:L8	-487 to +72	2650 (0.37)	130
PGK-LACZ1:L14	-487 to +694	5130 (0.71)	1650
PGK-LACZ2:L7	-487 to +1007	3130 (0.43)	1320
PGK-LACZ3:L2	-3150 to +694	7210 (1.00)	2520
PGK-LACZ1:L8	-487 to +694	5980	
PGK-LACZ3:L4	-3150 to +694	7020	

upstream, between approximately -3150bp and -488bp relative to the transcript start, increases activity a further 40% over the higher level. Thus, there appear to be two sequence elements that activate expression of the *lacZ* fusion, one located between 72bp and 694bp downstream of the transcription start site of *pgk* and one located between approximately 3150bp and 488bp upstream. There may also be a sequence element that damps expression located between 694bp and 1007bp downstream.

However, although the transformed strain with -3150 to +694 of *pgk* sequence fused to *lacZ* always had the greatest β-galactosidase activity, there was considerable variation in the actual activities observed for the strains with extensive *pgk* coding sequence upstream of *lacZ*. Whilst the β-galactosidase activities recorded for the strain with only 13 N-terminal codons upstream of *lacZ* only had a range in activity of about 5% of the mean, strains transformed with the larger fusions had a range of at least 30%. The range in β-galactosidase activities observed with these latter strains is most likely to reflect instability of the large fusion proteins. Although cell free extracts were always prepared from freshly harvested mycelium and were stored on ice for not more than two hours prior to conducting assays, the precise treatment of mycelium and extract may have varied considerably, leading to different degrees of degradation of the fusion protein in separate samples of a particular strain.

The loss of β-galactosidase activity observed when mycelium was stored frozen before preparing cell free extract emphasizes that the large fusion proteins may be unstable. This is particularly likely with the largest fusion protein, which includes the N-terminal 286 amino acids of *PGK*. Thus, the weaker β-galactosidase activity recorded from a strain synthesizing this protein, compared to one synthesizing a protein with only the N-terminal 182 amino acids of *PGK*, could equally well be due to instability of the larger fusion protein as to the presence of a transcriptional silencer within *pgk* coding sequence.

However, the two fold difference in β-galactosidase activity recorded between strains with 13 or 182 N-terminal *pgk* codons fused upstream of *lacZ* implies that there is a real effect of the coding region upon overall expression. This region includes both coding sequence and the two introns. Although the effect is much less in degree than that observed with the *pgk* gene of *Saccharomyces cerevisiae* (Mellor *et al.*, 1987), the decreased stability that is probable for the larger fusion protein may be masking the full extent of any internal activating sequence.

Only a single type I transformant representing each fusion construct was examined in the above analysis, since different isolates of type I integrants with a particular vector had previously shown little variation in expression (Chapter 5). However, to confirm the effects recorded above, further type I transformants were examined that had incorporated the vectors pPGK-LACZ1 (-487 to +694 of *pgk* sequence) and pPGK-LACZ3 (approximately -3150 to +694). No further type I transformants with the vector pPGK-LACZ2 (-487 to +1007) had been recovered. When β-galactosidase assays were conducted on cell free extracts prepared from the transformed strains, enzyme activities correlated closely to those previously recorded. The data is given in Table 7.2.

7.4 Relative mRNA concentrations in transformed strains.

In all previous experiments comparisons have been made between the expression of the gene fusions by recording β-galactosidase activities. Although this approach monitors overall expression and does not distinguish between potential controls in transcription, message stability and translation, it has allowed likely activator sequences to be identified. However, it is still not known if all control elements have been located, and this can only be confirmed by comparing expression from the *pgk-lacZ* fusions with that from the native *pgk* gene. Since the specific activity of *Aspergillus nidulans* 3-phosphoglycerate kinase is not known, levels of expression could not be compared by assaying the activities of *PGK* and β-galactosidase. Moreover, a specific antibody is not available to Aspergillus *PGK* to allow comparison of the protein concentration with β-galactosidase. However, the steady state message levels transcribed from the fusion gene and the native *pgk* gene can be determined by quantitative "Northern blot" analysis.

Total cellular RNA was prepared from the type I transformed strain PGK-LACZ3:L2 (Section 2), which has 5'*pgk* sequence extending approximately 3150bp upstream of the transcription start site and has coding sequence extending 182 codons into the gene, and from the untransformed recipient strain QG716 (Chapter 4). In order to detect the presence of fusion message, aliquots of the RNA preparations were spotted onto a nitrocellulose filter and were incubated in a hybridization reaction with radioactively labelled probe at 42^{o}C for 15 hrs. A 3.0Kb *EcoRI* DNA fragment, consisting entirely of *lacZ* sequence (isolated from the vector pLACZ-QUTE3: Chapter 4), was used to synthesize probe. The filter was given four 30 minute washes in 0.1x SSC/0.1% SDS at 50^{o}C, and an autoradiographic image was obtained by exposing it to X-ray film. This is presented in Figure 7.6. As expected, RNA prepared from the transformed strain hybridized the *lacZ* probe, whereas message from the untransformed strain did not.

Figure 7.6

The synthesis of β-galactosidase fusion message in a strain transformed with the fusion vector pPGK-LACZ3.

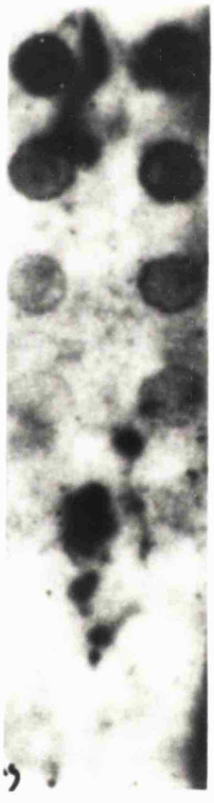
a) The autoradiograph shows the hybridization of an α^{32}P-dCTP labelled 3.0Kb *EcoRI* DNA fragment consisting entirely of *lacZ* sequence, with 5μg of cellular RNA prepared from the single copy type I transformed strain PGK-LACZ3:L2, and the recipient strain QG716.

b) The autoradiograph shows the hybridization of an α^{32}P-dCTP labelled 0.65Kb *HindIII-SalI* DNA fragment consisting entirely of *pgk* sequence, with a dilution series (μg) of total cellular RNA prepared from the single copy type I transformed strain PGK-LACZ3:L2, and the recipient strain QG716.

a)

QG716 PGK-LACZ3:L2

b)



QG716	PGK-LACZ3:L2
8	8
4	4
2	2
1	1
0.5	0.5
0.25	0.25
0.125	0.125

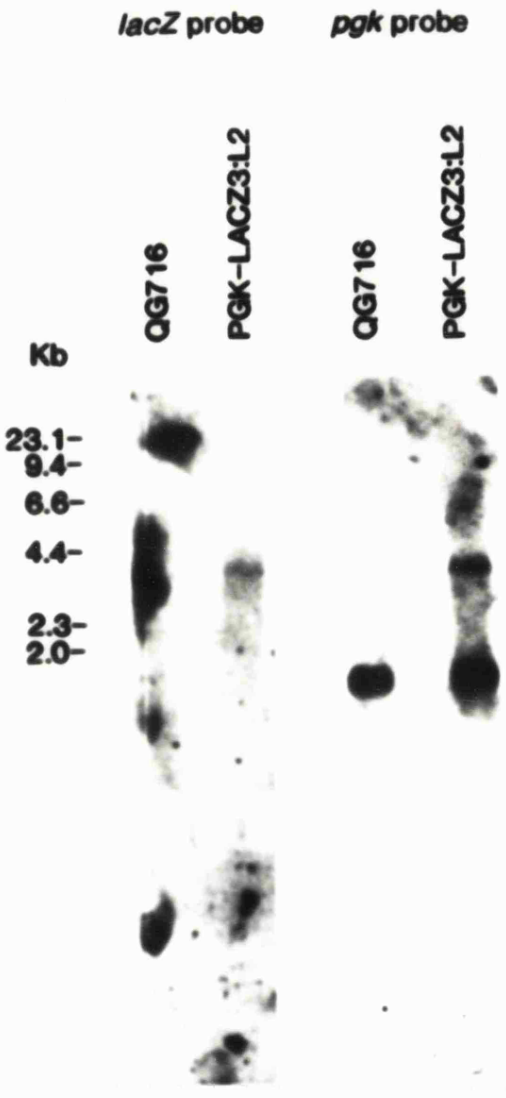
To provide an initial comparison between the steady state levels of fusion message and *PGK* message, a "dot blot" hybridization experiment was conducted in which the two RNA samples analysed above were incubated with a radioactively labelled 0.65Kb *HindIII-SalI* DNA fragment consisting entirely of *pgk* seqence (isolated from the vector pPGK2; Clements and Roberts, 1985). This probe was predicted to hybridise to native *PGK* message in both strains and to the β-galactosidase fusion message in the transformed strain alone, in which *pgk* sequence located 5′ of *lacZ* completely spans the DNA sequence of the probe. Thus, the hybridisation signal was expected to be more intense with the transformed strain than with the recipient strain. A first approximation to the relative levels of the fusion and native *pgk* messages was made by spotting a dilution series of each RNA preparation onto the filter. An autoradiographic image of this filter is presented in Figure 7.6. A positive hybridisation signal was still obtained with the transformed strain when there was only half the amount of RNA present required to give a positive signal with the untransformed strain. Thus, RNA prepared from the transformed strain appears to be hybridising to about twice as much probe DNA as that from the untransformed strain, indicating similar steady state levels of fusion message and *PGK* message in the transformed strain.

So that message levels could be more accurately compared between the transformed and untransformed strains, equal quantities of each RNA preparation were subject to agarose gel electrophoresis followed by transfer of samples onto a nitrocellulose filter, which was incubated in hybridisation reactions to either the *lacZ* or the *pgk* DNA probes exactly as in the "dot blot" experiments described above. Autoradiographic images of the filter are presented in Figure 7.7. The *lacZ* probe hybridizes to a single 4.05Kb message prepared from the transformed strain; the size predicted for the fusion message. As expected, the *lacZ* probe did not hybridize to message prepared from the untransformed strain. The *pgk* probe also hybridized to the 4.05Kb fusion message prepared form the transformed strain,

Figure 7.7

Analysis of a strain transformed with the fusion vector pPGK-LACZ3, to compare the synthesis of β-galactosidase fusion message with native *PGK* message.

The autoradiographs show the hybridization of α^{32}P-dCTP labelled *lacZ* and *pgk* DNA fragments, with 5μg of total cellular RNA prepared from the single copy type I transformed strain PGK-LACZ3:L2, and the recipient strain QG716, which had been size separated by agarose gel electrophoresis. The *lacZ* probe was a 3.0Kb *EcoRI* fragment, and the *pgk* probe was a 0.65Kb *HindIII-SalI* fragment (Figure 7.2). Phage λ DNA digested with *HindIII* and single stranded provided the marker (not shown).


lacZ probe
pgk probe
QG716
PGK-LACZ3:L2
QG716
PGK-LACZ3:L2
Kb
23.1
9.4
6.6
4.4
2.3
2.0

and in addition it hybridized to a 1.7Kb message prepared from both strains and corresponding in size to the native *PGK* message (Clements and Roberts, 1986). The intensity of signal is much greater for the *pgk* probe hybridizing the native *PGK* message than it is hybridizing the fusion message. Thus, contrary to the "dot blot" result above, the steady state level of the native *PGK* message in the transformed strain appears to be considerably greater than that of the fusion message.

In order to quantify the difference in message levels, autoradiographic images like those presented in Figure 7.7 were scanned using a laser densitometer. In this case the X-ray film was made sensitive by low level prior exposure to light to give a linear response to radioactive decay. The traces obtained are shown in Figure 7.8. RNA prepared from the untransformed strain and probed with *pgk* sequence gave a single clearly defined peak corresponding to the native *PGK* message, whilst RNA prepared from the transformed strain gave two peaks. As for the untransformed strain, one of these was clearly defined and corresponded to the native *PGK* message, but the other, corresponding to the fusion message, was neither as sharp nor as strong. Comparing the areas under the two peaks obtained with RNA prepared from the transformed strain, implies that the steady state level of the fusion message is approximately 0.2 of that of the native *PGK* message. Thus, transcription of the *lacZ* fusion gene driven by *pgk* promoter sequence appears to be much weaker than transcription of the native *pgk* gene. This implies that *pgk* sequence elements not present on the most extensive construct (approximately -3150 to +694bp relative to the transcript start site) are required for efficient transcription. An alternative explanation is that the large fusion message is unstable. The lack of definition of the peak corresponding to this message on the tracing obtained from the densitometer scan (Figure 7.8), indicates that this may be the case. However, to distinguish between the effects of transcriptional control and message stability, the rates of turnover of the fusion and native *PGK* messages would have to be calculated and compared.

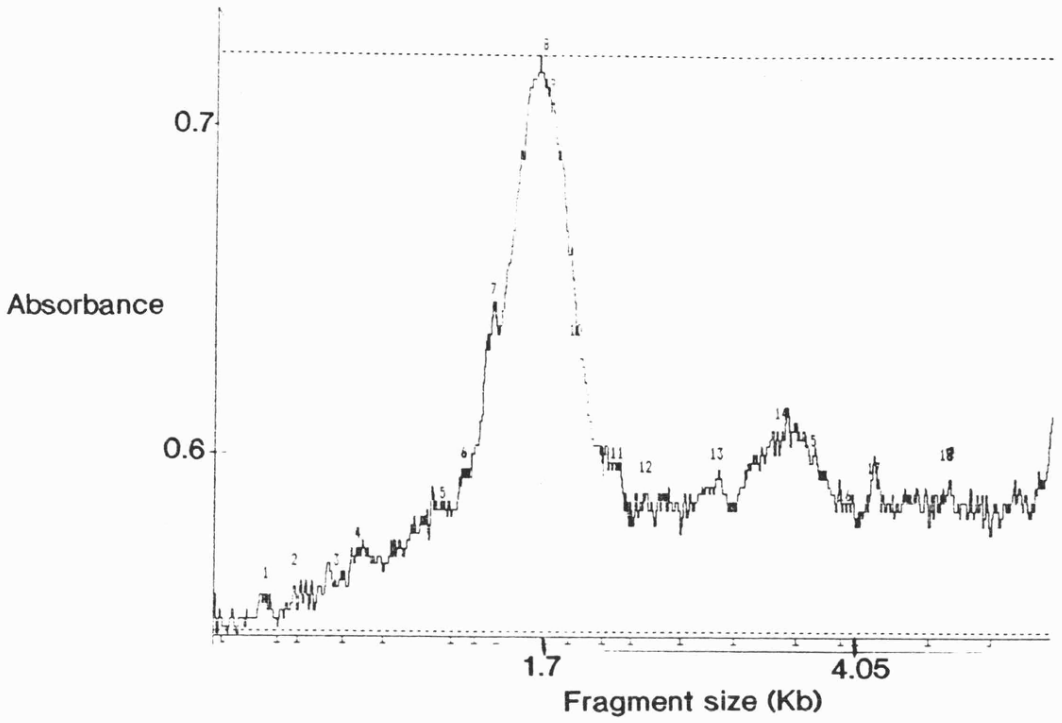
Figure 7.8

Comparison of the synthesis of β-galactosidase fusion message with native *PGK* message, in a strain transformed with the fusion vector pPGK-LACZ3, by densitometric analysis.

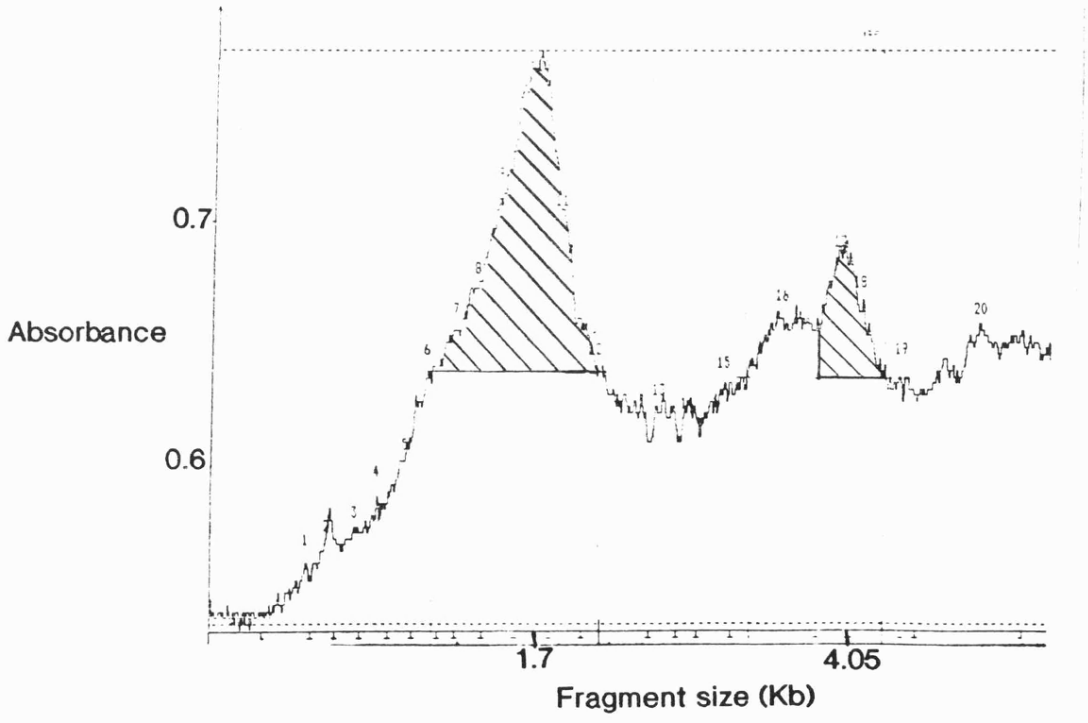
The traces were obtained using a laser densitometer from autoradiographs of total cellular RNA prepared from the strains QG716 and PGK-LACZ3:L2, that had been separated by agarose gel electrophoresis and hybridized to a *pgk* DNA probe (Figure 7.7). To give a linear response to radioactive decay, the X-ray film was made sensitive by low level prior exposure to light.

The areas compared to estimate relative mRNA levels are indicated by shading in the Figure. The base level was taken as that of the background absorbance in the region between the and peaks.

Recipient strain QG716



Transformed strain PGK-LACZ3:L2



7.5 Summary.

Experiments with gene fusions in *Saccharomyces cerevisiae* have identified sequence located between the 12th and 79th codons as being essential for efficient high level expression of the *PGK* gene (Mellor *et al.*, 1987). The presence of a downstream activator sequence has been proposed to explain this effect. To discover whether such an element is located within the coding sequence of the *Aspergillus nidulans pgk* gene, further fusion vectors were constructed with up to 286 N-terminal *pgk* codons and both introns immediately upstream of *lacZ*. These were transformed into *A.nidulans* and single copy integrants with vector sequence targeted to the *qutE* locus were identified. Cell free extracts were prepared from representative type I transformants and were assayed for β-galactosidase activity. Comparison between the activities recorded with strains transformed with the different gene fusion vectors identified sequence extending from codon 14 across the two introns to codon 182 as having a positive effect on the expression of the gene fusion. Also, sequence lying between codons 183 and 286 had an apparent negative effect. However, the larger fusion proteins have been shown to be unstable, and this latter negative effect of coding sequence upon expression may simply reflect this instability.

The construction of a further gene fusion vector, including sequence far upstream of the *pgk* gene, has allowed an additional activator sequence exerting a relatively minor effect to be located between 487bp and approximately 3150bp upstream of the transcriptional start site. Any far upstream sequence element must lie beyond -638 bp, because a construct including the putative glycolytic box has previously been shown to have no effect upon expression when, as here, glucose is the carbon source (Chapter 6, Section 5).

In order to discover if this most highly expressed *lacZ* fusion included all the promoter elements of *pgk*, steady state message levels were compared between the fusion and

native *PGK*. The fusion message was shown to be present at only about 0.2 the level of the native *PGK* message, indicating that the most extensive fusion does not include all the *pgk* promoter elements necessary to ensure efficient expression. Alternatively, the large fusion message may be unstable, a proposal consistent with the absence of a sharply defined peak on a densitometer trace of an autoradiographic image of the message. One further possibility is that the integration of the gene fusion construct into the *qutE* locus in the *A.nidulans* genome results in reduced activity compared to the *pgk* locus. Such an effect of the site of vector integration upon gene expression has been demonstrated in Chapters 5 and 6.

Chapter 8.

Discussion.

Several conserved sequences postulated to bind regulatory proteins have been identified upstream of genes cloned from filamentous fungi, but these putative promoter elements are only now being investigated functionally. The major thrust of my work has been to attempt to elucidate the role of some of these potential control sequences using the 3-phosphoglycerate kinase (*pgk*) gene of *Aspergillus nidulans* as a model.

Analysis of sequence upstream of the *pgk* gene identified sequence elements with homology to the consensus TATA and CCAAT boxes characteristic of higher eukaryotic genes, and a pyrimidine rich region common to several filamentous fungal genes and the more highly expressed genes of *Saccharomyces cerevisiae* (Clements and Roberts, 1986). These elements are located in approximate consensus positions, and in addition an eight base pair direct repeat has been identified within the core promoter that is not common to other fungal genes. In order to analyse the function of these putative transcriptional control elements, I devised a strategy to fuse *A.nidulans* 5'*pgk* sequence to the *lacZ* gene of *Escherichia coli* (to provide a "reporter" system), to delete regions of the *pgk* promoter and to introduce these various constructs into *A.nidulans*. The effect of specific sequence elements upon gene expression could then be investigated by monitoring β-galactosidase expression in the transformed *A.nidulans* strains with defined 5'*pgk* sequence driving expression of the *lacZ* gene.

Furthermore, since no mutant strain of *A.nidulans* deficient in 3-phosphoglycerate kinase (*PGK*) activity has been identified, the location of the *pgk* gene on the *A.nidulans* genetic map is unknown. A secondary aim of my project has been to isolate and characterize a *pgk*$^-$ mutant strain of *A.nidulans* and to assign the *pgk* locus to a particular chromosome.

8.1 Disruption of the 3-phosphoglycerate kinase (*pgk*) gene of *Aspergillus nidulans*.

Mutant strains of *Escherichia coli* (Irani and Maitra, 1974), *Saccharomyces cerevisiae* (Lam and Marmur, 1977) and *Pseudomonas aeruginosa* (Banerjee *et al.*, 1987) have been identified that are deficient in 3-phosphoglycerate kinase (*PGK*) activity, and in each case a single gene locus has been identified. The *E.coli* and *S.cerevisiae pgk* mutants are incapable of growth on either glycolytic or gluconeogenic precursors alone, but will grow on a combination of the two. This requirement for both a glycolytic and a gluconeogenic carbon source demonstrates the essential role of *PGK* in both pathways. Suprisingly, although the *pgk* mutant of *P.aeruginosa* is also unable to grow on gluconeogenic precursors, it will grow on glycolytic precursors. These anomalous growth requirements of the *P.aeruginosa pgk* mutant result from the ability of the organism to utilize the unidirectional Entner-Doudoroff pathway which bypasses glycolysis, providing the citric acid cycle precursor pyruvate from glycolytic carbon sources. The Entner-Doudoroff pathway has been reported in *Aspergillus niger* (Elzaing *et al.*, 1973), and if it were functional in *A.nidulans*, then any one of four acetate non-utilizing strains of *A.nidulans* for which no enzyme lesions had been previously attributed (Armitt *et al.*, 1976), could correspond to a deficiency in *PGK*. However, none of the four strains proved defective for *PGK* activity.

Previous attempts to isolate a *pgk*$^-$ mutant strain of *A.nidulans* by classical mutation techniques had proved unsuccessful because the precise carbon source requirements of such a mutant were not known (Roberts, personal communication). Therefore, a gene disruption strategy was devised in which the native *pgk* gene would be replaced with the homologous cloned sequence interrupted by a selectable marker, thus allowing the single *pgk* gene locus identified by Clements and Roberts (1985) to be assigned to a particular chromosome. A direct gene disruption strategy designed by Rothstein (1983) in *S.cerevisiae*, and applied by

Miller *et al.* (1985) in *A.nidulans* was used. The orotidine decarboxylase (*pyr-4*) gene of *N.crassa* was introduced into the isolated *pgk* gene of *A.nidulans*, and this disrupted sequence was transformed into a corresponding mutant strain (*pyrG*$^-$) of *A.nidulans* on a vector molecule in both circular and linear forms. Since the precise phenotype of a *pgk*$^-$ transformed strain of *A.nidulans* was unknown, a *pyrG*$^-$ diploid strain of *A.nidulans* was transformed with the expectation of disrupting one of the two resident *pgk* genes, which would then be maintained by dominance of the wild type allele.

I demonstrated that the frequency of regeneration of uracil independent colonies was increased approximately 10-fold by linearizing the vector prior to transformation. This observation is discussed below in Section 8.2 in relation to similar results obtained with 5'*pgk-lacZ* fusion vectors.

The diploid uracil independent transformed strains were analysed by classical genetic techniques to determine their nature with respect to the disrupted *pgk* sequence that had been introduced. Chromosome segregation from transformed heterozygous diploid strains was induced by use of the mitotic spindle inhibitor benomyl. Initially, diploid transformants were forced to segregate haploid strains on media containing benomyl and supplemented with uracil and glucose to select for the wild type *pgk*$^+$ allele. Analysis of the segregants demonstrated that uracil dependence (*pyr-4*$^-$, *pgk*$^+$) and the genetic marker on chromosome VIII (*qutE*) had co-segregated in 60% of the cases that could be characterized, indicating that the *pgk* locus is on chromosome VIII. In the case of one particular transformed strain, all of the haploid sectors isolated were uracil requiring, indicating disruption of the native *pgk* locus on one chromosome of the homologous pair. Subsequently, this particular transformed diploid strain was induced to segregate on media containing benomyl and lacking uracil. By carefully varying combinations of carbon sources, a *pgk::pyr-4* segregant was finally isolated and clearly

demonstrated to be deficient in *PGK* activity. This *pgk*$^-$ mutant strain requires both acetate (0.1M) and glycerol (0.05M) for growth, when it conidiates poorly. The mutant strain is poisoned by low concentrations of hexoses, probably due to the accumulation of toxic levels of the glycolytic intermediate 1,3-diphosphoglycerate, which is the substrate for *PGK*. Thus, the *pgk*$^-$ mutant strain of *A.nidulans* resembles that of *E.coli* and *S.cerevisiae*, but not that of *P.aeruginosa*, indicating that the Entner-Doudoroff pathway does not operate in *A.nidulans*. As with the *pgk*$^+$ segregants, the genetic marker on linkage group VIII (*qutE*) segregated with the disrupted *pgk* allele (*pgk::pyr-4*$^-$), thus confirming the location of the single *pgk* gene of *A.nidulans* on chromosome VIII. However, the position of the *pgk* gene has not been mapped through the sexual cycle, an exercise that may be difficult due to the likely meiotic instability of the *pgk*$^-$ mutant strain.

Molecular analysis revealed that although the resident *pgk* locus had indeed been disrupted, neither a straightforward gene replacement, nor any other standard recombination event had occurred. The complex nature of the gene disruption event at the *pgk* locus may result from the original vector molecules having been linearized within sequence of polylinker rather than *A.nidulans* origin prior to transformation into the diploid strain. In an attempt to promote a classical gene replacement through two homologous recombination events on either side of the disrupted sequence, the original recipient diploid strain was transformed with either a fragment consisting only of the disrupted *pgk* sequence, or of the whole vector linearized within 5'*pgk* sequence. However, transforming with molecules with free, potentially recombinogenic ends originating from *A.nidulans* sequence, also failed to generate any transformed strains in which the natve gene had been replaced by gene disruption. Moreover, the transformants isolated did not correspond to standard integrative recombination events at the resident gene locus.

Comparison of the above transformation experiments with the direct gene replacement strategy deployed by Miller *et al.* (1985), in which the native *argB*$^+$ allele of *A.nidulans* was replaced with a disrupted *argB::trpC* allele, indicates that the type of vector integration events recorded depend upon the specific sequence under study. In the case of the *argB* gene disruption, 20% of the transformed strains recovered had the native gene precisely replaced with the disrupted sequence. The difference in the proportion of classical gene replacements between this experiment and the *pgk* gene replacement described above may reflect differences in the extent of *A.nidulans* sequence flanking the disrupted genes on the linear molecules to be transformed. With the *argB* gene disruption experiment there were at least 1.5Kb of *argB* coding and flanking sequence on either side of the disrupting *trpC* sequence, whereas with the *pgk* gene disruption experiment there was only about 0.75Kb of *pgk* coding and flanking sequence on the 5'side and 1.15Kb on the 3' side of the disrupting *pyr-4* sequence. Thus, the extent of flanking sequence within which homologous recombination events could occur was considerably less in the case of the *pgk* gene disruption. However, one-step gene disruption can be achieved in *S.cerevisiae* with as little as 0.3Kb of targeting sequence flanking the selectable marker (Rothstein, 1983). It is possible that locus specific effects may be responsible for the difference in the proportion of gene replacements at the *argB* and *pgk* locus in *A.nidulans*, with the local environment of the genes affecting the frequency and type of recombination events. The importance of locus specific effects is discussed further in relation to the integration of 5'*pgk-lacZ* fusion vectors into the A.nidulans genome in Section 8.2.

In summary, one of the pair of *pgk* genes in a diploid strain of *A.nidulans* has been disrupted. The analysis of *pgk*$^+$ and *pgk*$^-$ segregants has located the *pgk* locus on chromosome VIII. A *pgk*$^-$ mutant strain has been isolated that requires both acetate and glycerol for growth, and conidiates poorly. The mutant strain is poisoned by even moderate concentrations of hexoses.

8.2 Transformation of *Aspergillus nidulans* with 5'*pgk-lacZ* fusion vectors.

In order to analyse the putative function of consensus eukaryotic 5' sequence elements identified upstream of the transcription initiation site of the *Aspergillus nidulans* 3-phosphoglycerate kinase (*pgk*) gene, 5' sequence extending almost 0.5Kb upstream of the transcription initiation site of *pgk* (and well beyond these consensus sequences) was fused to the *lacZ* reporter gene of *Escherichia coli* in each translational reading frame.

When the three fusion vectors were transformed into *Escherichia coli* the *lacZ* gene was expressed only with the translationally in frame vector. Transcriptional read-through from prokaryotic vector sequence is unlikely, since heterologous expression was shown to specifically depend upon the presence of *A.nidulans* 5'*pgk* sequence upstream of *lacZ*. Thus a eukaryotic promoter appears to be driving gene expression in a prokaryote, despite differences in prokaryotic and eukaryotic consensus control elements. However, mapping the transcription start site by S1 nuclease analysis would be necessary to confirm that a filamentous fungal promoter was acting in *E.coli*.

Incorporation of vector DNA into the genome is the only way in which sequences introduced into *A.nidulans* can be stably maintained (reviewed by Turner and Ballance, 1985). Inclusion of the wild type catabolic dehydroquinase (*qutE*) gene of *A.nidulans* on the gene fusion constructs allowed selection of quinic acid utilizing transformants of a recipient *qutE*$^-$ mutant strain of *A.nidulans*, and some of the transformants had the *pgk-lacZ* fusion integrated at the *qutE* locus. This fixed genomic location provided a reference point from which differently modified promoter fusions were analysed for *lacZ* expression. Thus, the effect upon gene expression of specific consensus sequence elements in the *A.nidulans pgk* promoter could be directly determined, and position effects arising from integration at other sites in the genome (discussed below) discounted.

Transformation of a *qutE*$^-$ mutant strain of *A.nidulans* with the in frame fusion vector linearized at a site within *A.nidulans* sequence approximately 0.5Kb upstream of the translational start codon of the cloned *qutE* gene, increased the frequency of transformation 16.5-fold compared to the circular vector. This result presumably arises because free potentially recombinogenic ends of *A.nidulans* sequence were available to promote the integration of vector sequence into the genome. Furthermore, with circular vector only 17% of the characterized integration events had the entire vector integrated at the *qutE* locus, whilst with linear vector the proportion of type I integrants was 50%. Thus, the proportion of type I transformants was three-fold greater with linear vector and the overall frequency of type I transformation was increased about fifty-fold.

I have described in Section 8.1, that linearizing the vector used in the *pgk* gene disruption strategy (prior to transformation of *A.nidulans*) also increased the transformation frequency, though only about ten-fold. However, with the gene disruption vector the plasmid was linearized within sequence of polylinker origin, making direct comparison of the increase in transformation frequency observed when linearizing the 5'*pgk-lacZ* fusion vector within *qutE* sequence difficult, even though the polylinker region of the gene disruption vector was immediately adjacent to cloned *A.nidulans* sequence.

The result with the 5'*pgk-lacZ* fusion vector is similar in fashion, though less in degree to the results of transformation experiments conducted in *S.cerevisiae*, in which the linearization of the vector molecule can increase the frequency of type I transformation up to some thousand-fold. However, the degree to which the type I transformation frequency in *S.cerevisiae* is increased upon linearizing vector molecules depends upon the specific sequence that is cut. Whilst making a double stranded break within the *HIS-3* gene stimulates targeted transformation 2000- to 3000-fold, cutting within the *LEU-2* gene shows only a 50- to 100-fold effect (Orr-Weaver *et al.*, 1981).

In a study with *Neurospora crassa*, all of the transformants recovered were type I integrants when the vector was linearized within the coding sequence of the selected marker, the tryptophan biosynthetic gene *trp-1*, (Kim and Marzluf, 1988), probably because homologous recombination was the only way in which the lesion in coding sequence could be repaired. However, when the vector was linearized within *N.crassa* sequence flanking the the coding region, no increase in the proportion of homologous recombination events was observed. Thus, the effect of linearizing within sequence approximately 0.5Kb upstream of the cloned *qutE* gene on the *lacZ* fusion vector described above in *A.nidulans*, is not reflected by a similar experiment in *N.crassa*. Furthermore, experiments in which the tryptophan biosynthetic gene *trpC* of *A.nidulans* present on a vector was linearized within coding sequence, did not affect the proportion of type I integrants isolated following transformation into the host organism (Yelton *et al.*, 1984). Thus, it appears that just as the extent of the increase in the frequency of type I transformation in *S.cerevisiae* after linearizing vector molecules within the selected marker, is dependent upon that specific marker (Orr-Weaver *et al.*, 1981), so the effects of linearizing vector molecules prior to transformation of filamentous fungi also depends upon the specific cloned fungal sequence that is cut in the vector.

In the course of subsequent experiments to analyse the *pgk* promoter, exonuclease digests were made into sequence 5′ of the *pgk* gene. Due to the nature of the *lacZ* fusion vector and the bidirectional action of the exonuclease, sequence 5′ of the *qutE* gene was also deleted, thus reducing the extent of *qut* gene cluster sequence on the vector (Chapter 6, Figure 6.4). It was possible that such deletion of sequence upstream of *qutE* might affect the degree of targeting of the fusion vectors to the *qutE* locus in the genome. However, no decrease in homologous recombination was observed. The site at which vectors were linearized is approximately 0.5Kb upstream of *qutE* and proximal to all the exonuclease generated deletions. Thus, the generation of

free potentially recombinogenic ends by linearizing each of the vectors within 5'*qutE* sequence prior to transformation is proposed to have an overriding influence upon the degree of targeting. In *S.cerevisiae* directing vector sequence to a particular locus is difficult when the targeting sequence within which the vector is linearized is less than about 0.2Kb (Orr-Weaver *et al.*, 1983). This may reflect either insufficient sequence to allow recombination to occur, or the intracellular degradation of DNA past the end of the targeting sequence on the vector following linearization. However, in the case of the 5'*pgk-lacZ* fusions transformed into *A.nidulans*, even with the most extensive deletion construct, the extent of *qut* cluster sequence was only reduced from 1.9Kb to 1.45Kb, and thus should still be sufficient to promote targeted integration.

The importance of the site of vector integration for the expression of vector sequences was demonstrated by recording β-galactosidase activities in transformed strains of *A.nidulans*. Considering first the original fusion vector with almost 0.5Kb of 5'*pgk* sequence upstream of *lacZ*; all transformed strains in which a single copy of the vector molecule had been targeted to the *qutE* locus expressed *lacZ* to a similar degree, with a range of 32% of the mean. Furthermore, the β-galactosidase activity in transformed strains with multiple copies of the vector integrated at the *qutE* locus was directly proportional to the copy number. However, when the vector integrated elsewhere in the *A.nidulans* genome, expression of *lacZ* was no longer uniform. One half of the transformed strains with a single copy of the vector integrated elsewhere in the genome did not express *lacZ* at all, and the remainder expressed *lacZ* over a range of between approximately 20% and 160% of the single copy targeted integrants. Subsequently, fusion vectors with reduced 5'*pgk* sequence upstream of *lacZ* were transformed into *A.nidulans*. When screened on X-gal plates for β-galactosidase activity, strains with a single copy of a particular construct targeted to the *qutE* locus gave consistent results, either never expressing *lacZ* or all showing the same level of expression. However, amongst

transformed strains with vector sequence integrated at a site other than the *qutE* locus, expression of *lacZ* was more variable. There were several examples of transformed strains which regained β-galactosidase activity when a particular vector was integrated at a non-specific site, although the same vector integrated at the *qutE* locus had no β-galactosidase activity. In one case the reverse was also true; one transformed strain lost β-galactosidase activity with a particular vector integrated at a non-specific site, whilst transformed strains with the same vector targeted to the *qutE* locus expressed *lacZ*.

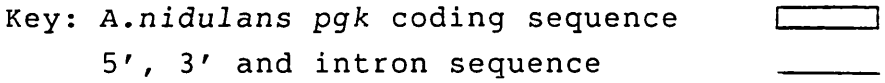
In summary: as predicted at the outset of the project, the site of vector integration affects the expression of the fusion gene. Consequently, comparisons of β-galactosidase activity between transformed strains carrying *lacZ* fusions with different extents of 5'*pgk* sequence in order to determine any requirement for consensus *pgk* promoter sequences, were made between strains with a single copy of the vector molecules integrated at the *qutE* locus.

8.3 Identification of the 5′ extent of the *Aspergillus nidulans* 3-phosphoglycerate kinase (*pgk*) gene.

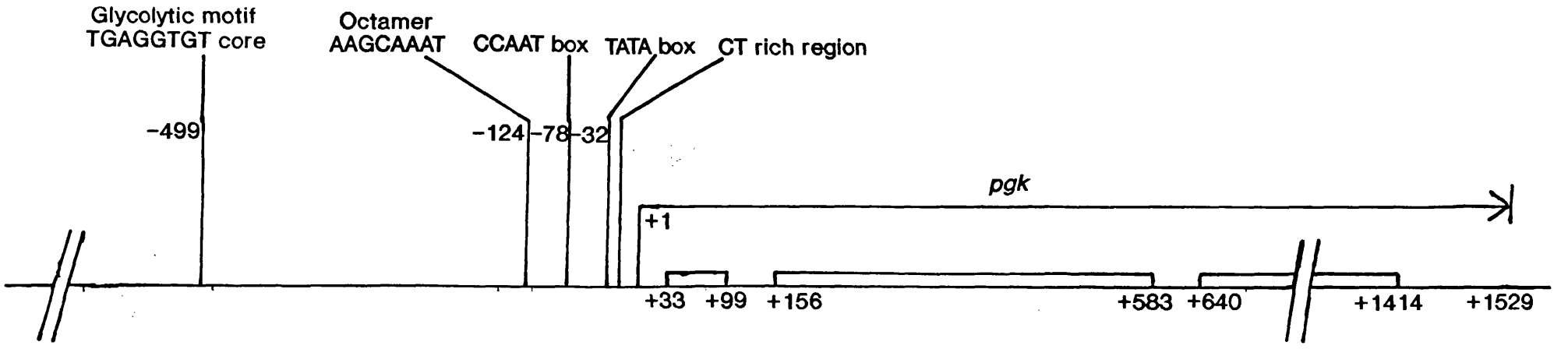
In order to identify the functional promoter elements of the 3-phosphoglycerate kinase (*pgk*) gene of *Aspergillus nidulans*, deletions were made in the original *lacZ* fusion vector that diminished the extent of 5′*pgk* promoter sequence. A series of constructs were generated by these deletions with between 44 and 487bp of 5′*pgk* sequence fused to *lacZ*. These constructs were introduced into *A.nidulans*, and transformed strains with a single copy of the vectors integrated at the *qutE* locus were assayed for β-galactosidase activity. The results of the *pgk* promoter analysis are summarized in Figure 8.1.

Suprisingly, sequences within 120bp of the transcription initiation site are not sufficient alone to drive the expression of the fusion gene, even though these include all of the apparent consensus promoter sequence elements identified 5′ of the gene: TATA and CCAAT like elements, and a pyrimidine rich sequence. This region also includes an eight base pair direct repeat which, it was postulated, might affect expression. Although these conserved sequences alone have been shown insufficient to promote gene expression, it would be necessary to specifically delete them whilst leaving further 5′ sequence intact in order to see if they are required for efficient expression in conjunction with the -161 to -120 region. Furthermore, the transcription start sites of the fusion messages were not mapped, so it is not known whether these upstream elements are required to position the site of transcript initiation. In particular, the pyrimidine rich sequence, which is a feature of most strongly expressed *S.cerevisiae* and filamentous fungal genes, especially those lacking a canonical TATA sequence (Dobson *et al.*, 1982; Brown and Lithgow, 1987; Gurr *et al.*, 1987), has been shown to be important in positioning transcript initiation sites (McNeil and Smith, 1985).

Figure 8.1
Regulatory regions affecting the expression of the *A.nidulans pgk* gene.



The position of sequence elements is relative to the transcription initiation site of *pgk* (+1).



approx.
-3150 -638 -488

Positively acting region 29% total expression

Carbon source dependent regulatory region

-161 -120

Essential region 37% total expression

+72 +694

Positively acting region 34% total expression

Although a canonical TATA sequence (TATAAA) is often located in approximately the consensus higher eukaryotic position upstream of filamentous fungal genes (Gurr *et al.*, 1987), the *pgk* gene of *A.nidulans* is an example of a second group of genes that have a less rigidly conserved AT rich sequence with a less precise location. Such an AT rich element can be removed from the *A.nidulans* tryptophan biosynthetic gene, *trpC*, without affecting expression (Hamer and Timberlake, 1987). Thus, an AT rich sequence positioned 5′ of *pgk* is insufficient to promote the expression of a fusion gene, and a similar sequence upstream of *trpC* is not required for expression. These results indicate that AT rich sequences located 5′ to filamentous fungal genes are more likely to reflect chance homology than consensus promoter elements.

In contrast to the effect of the deletion of the proximal 120bp of 5′ *pgk* sequence, sequence located between 161bp and 120bp upstream of the transcription initiation site is essential for the expression of the *lacZ* fusion gene, indicating that an essential promoter sequence element is located within this region (or possibly spanning -120bp). Comparison of this sequence with that of other eukaryotic genes reveals the presence of a previously unnoticed element (AAGCAAAT) between positions -124 and -131 with a seven out of eight base pair match to the higher eukaryotic octamer (ATGCAAAT). This octamer motif has been identified upstream of several eukaryotic genes, and even 5′ to prokaryotic genes (Falkner *et al.*, 1986). Although originally identified as a distal upstream element, it can also be an important motif in enhancers, such as that of SV40 (Nomiyama *et al.*, 1987), where multimers of this sequence define cell specificity (Ondek *et al.*, 1987). Furthermore, this octamer sequence is located in reverse orientation in histone H2B promoters, where it is required to stimulate transcription following a shift from G1 to S phase (Harvey *et al.*, 1982; La Bella *et al.*, 1988). It is also present in the reverse orientation in the promoters of the immunoglobulin heavy and light chain genes, where it is essential for promoter activity (Eaton and Calame, 1987;

Falkner and Zachau, 1984), and in the IgH enhancer, where its deletion reduces, but does not abolish, gene expression (Gerster *et al.*, 1987), implying that it may play a role in determining B-cell specificity. Both ubiquitous and cell specific DNA binding proteins have been identified that specifically bind to the octamer motif, and are reviewed by La Thangue and Rigby (1988) and by Calame (1989).

Although the sequence with a seven base pair match to this well characterised octamer is the only element located between 161bp and 120bp upstream of the *A.nidulans pgk* transcript start site with homology to any consensus eukaryotic promoter sequence, it cannot be assumed that this is responsible for the expression of the *lacZ* fusions described above. First, this octamer sequence has not been identified upstream of any other fungal genes, and second, in all the other promoters in which it has been identified, it is absolutely conserved, whereas for *A.nidulans pgk* it is deviant in the second nucleotide (A rather than the consensus T). A detailed functional analysis of this element, either by its precise deletion or by site directed mutagenesis, is required to ascertain its role in *pgk* expression in *A.nidulans*.

The comparison of β-galactosidase activities between transformed strains with more extensive stretches of 5'*pgk* sequence (up to 487bp) fused to *lacZ*, failed to identify any further sequence elements likely to play a major role in the control of gene expression. However, a minor negative effect appears to be exerted by sequence located between 161bp and 239bp upstream, and a corresponding positive effect by sequence located between 239bp and 487bp upstream.

Subsequent to the construction of the deletion series of 5'*pgk-lacZ* fusion vectors, an additional conserved element of sequence was identified upstream of the four glycolytic genes that have been cloned from *A.nidulans*: triose phosphate isomerase (McKnight *et al.*, 1986), glyceraldehyde-3-phosphate dehydrogenase (Punt *et al.*, 1988), 3-phosphoglycerate kinase (Clements and Roberts,

1986) and pyruvate kinase (de Graaff, 1989). This element (termed the glycolytic box) consists of a block of sequence located between 499bp and 522bp upstream of *pgk* (Figure 8.1), and within this is an eight base pair region (TGAGGTGT) that is particularly well conserved between the four *A.nidulans* glycolytic genes. In order to investigate the effect of this putative promoter element, a further *lacZ* fusion vector was constructed that included 5′ *pgk* sequence extending beyond this region of homology. When a transformed strain of *A.nidulans* carrying this more extensive construct was grown on various glycolytic and gluconeogenic carbon sources, and its β-galactosidase activity compared to that of a transformed strain carrying fusions lacking the conserved upstream sequence, it was evident that this element modulates gene expression in response to the carbon source. Strains transformed with the more extensive construct expressed the gene fusion more strongly on acetate (167%), benzoic acid (146%) and quinic acid (214%) than on glucose, whilst those transformed with the shorter construct showed no increase in expression on acetate and benzoic acid and a less pronounced increase on quinic acid (134%). The results with the longer fusion (including the putative glycolytic box) are similar to those of Clements (1986), who compared native *PGK* activity in *A.nidulans* following growth on different carbon sources. Thus, sequence located between 488 and 638bp upstream of the *pgk* transcription start site appears to increase expression on gluconeogenic carbon sources compared to a glycolytic one. This increase in expression may reflect a need for greater *PGK* activity when the organism is grown on gluconeogenic carbon sources. Under these conditions the gluconeogenic activity of *PGK* would be required to synthesize glucose, which would provide a precursor for the pentose phosphate pathway and cell wall carbohydrates. In light of its effect on *pgk* gene expression, this element might be more appropriately termed a gluconeogenic box. In order to assign the modulation of *pgk* expression specifically to this sequence, it would have to be precisely deleted or mutagenized.

The *pgk-lacZ* fusion was more strongly expressed on quinic acid than on other gluconeogenic carbon sources whether or not the glycolytic box was present, but this effect was not observed with the native *pgk* gene (Clements, 1986). The increased expression of the gene fusion on quinic acid may be a consequence of the targeting of the constructs to the *qutE* gene cluster in the genome. The structural genes required for quinic acid metabolism are induced in the presence of quinic acid but are repressed on other carbon sources. In the transformed strains the regulation of the *qut* gene cluster may be affecting expression of the *pgk-lacZ* fusion integrated at the *qutE* locus. This effect is discussed further below (Section 4), in relation to mRNA levels of the gene fusion compared to native *pgk*.

Subsequently, in order to investigate the effect of far upstream seqeuences upon expression initiated at the *A.nidulans pgk* start site, a further *lacZ* fusion vector was constructed including extensive 5′ *pgk* sequence. The analysis of transformed strains of *A.nidulans* revealed that sequences located between approximately 3150 and 639bp upstream of the *pgk* gene have a positive effect upon the expression of the *lacZ* fusion, accounting for about 30% of the maximum β-galactosidase activity recorded (Figure 8.1). Since the DNA sequence of this far upstream region has not been determined, putative cis-acting control elements cannot yet be identified by virtue of homology to consensus eukaryotic promoter sequences, and it therefore remains unexplored.

8.4 The effect of 3-phosphoglycerate kinase (*pgk*) coding sequence upon gene expression.

Analysis of the 3-phosphoglycerate kinase (*PGK*) gene of *Saccharomyces cerevisiae* by constructing gene fusions has identified sequence between codons 12 and 79 as necessary for the high level of expression observed with the native gene (Mellor *et al.*, 1987). A positive cis-acting control sequence element, termed a downstream activator (DAS), has been proposed to be located within this stretch of sequence. In order to investigate whether such an element plays a role in promoting the expression of the *Aspergillus nidulans pgk* gene, *lacZ* fusion vectors were constructed which included at least an equivalent extent of *A.nidulans pgk* coding sequence to that required to ensure the efficient expression of genes fused to yeast *PGK*. When these constructs were introduced into *A.nidulans*, and the transformed strains assayed for β-galactosidase activity, sequence extending from codon 14 across the two introns to codon 182 of *pgk* was identified as promoting the expression of the *lacZ* fusion (Figure 8.1). However, this positive effect of *A.nidulans pgk* coding sequence is not as pronounced as that recorded with the yeast gene. Whilst the DAS of the yeast gene can increase the level of heterologous mRNA and protein encoded by a reporter gene six to ten-fold (Mellor *et al.*, 1987), the positively acting sequence element located within the coding region of the Aspergillus gene is responsible for only about 30% of the maximum β-galactosidase activity recorded.

However, the large *lacZ* fusions that included extensive *A.nidulans pgk* coding sequence were shown to encode unstable proteins. There was a significant loss in catalytic activity upon storage of mycelium at -20°C, and extracts prepared from freshly harvested mycelium of any one transformed strain showed considerable variation in activity. In contrast, the original *lacZ* construct including only the N-terminal 13 codons of *pgk* encoded a protein, the activity of which decayed to a lesser extent upon storage at -20°C, and gave much more consistent levels of activity when prepared from freshly grown mycelium.

Thus, the extensive *A.nidulans pgk* sequence in the fusion proteins may be affecting the stability of the tertiary and quaternary structure normally adopted by β-galactosidase, so reducing the catalytic activity of the proteins and masking the full effect of a DAS located within the coding region of *A.nidulans pgk*.

Intragenic transcriptional control sequences have been identified within other RNA polymerase II transcribed genes, including the immunoglobulin heavy chain genes (Gilles *et al.*, 1983), the β-globin gene (Kosche *et al.*, 1985), and most pertinant to fungal gene expression, the *S.cerevisiae* pyruvate kinase gene (cited in Brown and Lithgow, 1987).

In all of the work described in this Section and in Section 8.3, sequences were identified as being *pgk* promoter elements by assaying for "reporter" function. However, since the specific activity of *A.nidulans PGK* is not known, the expression of the reporter gene could not be related to the expression of native *pgk*. Thus, from the above analysis, it is unclear whether all *pgk* promoter sequences are present within the most extensive *lacZ* fusion construct (*pgk* sequence from approximately -3150 to +694). In order to compare the levels of expression between native *pgk* and the most extensive *lacZ* reporter fusion, steady state message levels were compared from the relevant transformed strain of *A.nidulans*. The fusion message was shown to be present at only about 20% of the level of the native message. However, densitometric analysis indicated that the fusion message may be unstable, and it is therefore unclear whether its reduced steady state level relative to the native *pgk* message is due to decreased transcription resulting from a missing activator sequence, or fusion message instability. These alternatives could be distinguished by estimating the half lives of the two messages. In *S.cerevisiae*, a gene fusion in which the human interferon gene was located immediately downstream of the entire yeast *PGK* gene produced a transcript which was of approximately equal stability to the native yeast *PGK* message (Mellor *et al.*, 1987).

An alternative explanation for the reduced expression of the gene fusion relative to native *pgk* is the effect of the genomic location of the relevant sequences. The transformed strain from which the message levels were analysed had the *pgk-lacZ* fusion integrated into the *qut* gene cluster. The gene fusion has been shown to be more strongly induced than native *pgk* by quinic acid (Section 3). Thus, the reduced level of fusion message compared to native *PGK* message may reflect a damping of the expression of the gene fusion integrated at the *qutE* locus when *A.nidulans* is grown on glucose. As discussed in Section 2, the site of integration of vector in the genome may be affecting gene expression.

8.5 Summary of *pgk* promoter analysis and future work.

The functional analysis of putative promoter sequences described in this thesis has identified three regions of sequence that appear to constitutively activate *Aspergillus nidulans* 3-phosphoglycerate kinase (*pgk*) gene expression (Figure 8.1). Firstly, sequence located between -161 and -120bp relative to the *pgk* transcription initiation site is essential for expression and contributes to almost 40% of the maximum activity recorded. The only putative promoter element that has been identified within this sequence is a seven out of eight base pair match to the octamer sequence ATGCAAAT, located upstream of several eukaryotic genes. The disruption of the homologous sequence upstream of *A.nidulans pgk* by site directed mutagenesis would allow its importance in stimulating *pgk* expression to be investigated. Constructs in which the modified 5'*pgk* sequences were fused to the *lacZ* reporter gene could be introduced into *A.nidulans*, and transformed strains with single copies of the vector molecules targeted to the *qutE* locus assayed for β-galactosidase activity.

A second element identified as important in stimulating *pgk* expression is located between codons 14 and 182, and is responsible for over 30% of the total activity recorded. In order to position this element more accurately, further *lacZ* fusion constructs would need to be made with decreasing amounts of *pgk* coding sequence. Also, a third region, located between 639 and approximately 3150bp upstream of the transcript start site has been identified as being responsible for about 30% of the maximum activity recorded. However, this region has not yet been mapped in detail or sequenced, and its analysis to identify any far upstream enhancers of the *pgk* gene would therefore be a lengthy exercise.

Furthermore, sequence located between -638 and -488bp, and including a consensus element located upstream of the four cloned *A.nidulans* glycolytic genes, has been shown to modulate gene expression on different carbon sources. This

sequence is responsible for increasing expression about 1.5-fold on gluconeogenic relative to glycolytic carbon sources. The investigation of this consensus sequence element by its precise deletion or by site directed mutagenesis would be required to confirm its role in modulating expression.

Although the regions of sequence discussed above have been identified as being necessary for efficient *pgk* expression, the proximal and distal upstream sequences originally proposed by Clements and Roberts (1986) to be control elements (CCAAT box, TATA box and CT rich sequence) have been shown to be insufficient to promote expression. However, in order to investigate whether they are required for expression, it would be necessary to construct a *lacZ* fusion vector in which only sequence spanning these elements had been deleted.

The analysis of putative *pgk* promoter sequences described in this thesis has been limited to determining the quantitative effect upon overall gene expression. In order to evaluate the effect of specific sequences upon the positioning of the message start site, it would be necessary to S1 nuclease map with RNA prepared from the transformed strains of *A.nidulans*. In addition, by comparing fusion message levels prepared from transformed strains, the effect of particular sequences upon expression could be related specifically to transcriptional control, although it would also be necessary to determine the stability of the fusion messages when conducting these studies.

The lines of investigation described above will allow the further elucidation of functional elements in the *pgk* promoter of *A.nidulans*. The identification of specific sequence elements responsible for the constitutively high level of expression of *pgk* might allow putative promoter elements to be identified upstream of other fungal genes by virtue of sequence similarity.

Appendix

Characterization of transformed strains of *A.nidulans*.

The recipient *A.nidulans* strain QG716 was transformed with the vectors listed in the Table. The vectors p5'PGK-LACZ1, 2 and 3 consist of 487bp of 5'*pgk* sequence upstream of *lacZ*. With p5'PGK-LACZ2 N-terminal *pgk* sequence is fused to *lacZ* in the predicted translational reading frame, whereas with p5'PGK-LACZ1 and 3 the fusion is out-of-phase (Chapter 4). The vectors p5'PGK-LACZ2-1 to 2-12 constitute an exonuclease digest series with a decreasing extent of 5'*pgk* sequence fused upstream of *lacZ*. The fusions are in the predicted translational reading frame (Chapter 6). The vector p5'PGK-LACZ4 is a more extensive construct with 638bp of 5'*pgk* sequence upstream of *lacZ*. It is also in the predicted translational reading frame (Chapter 6). The vector p5'PGK-LACZ2 was initially transformed into *A.nidulans* as a circular plasmid (as indicated in the Table). In a later experiment this vector was linearized prior to transformation, and subsequently in all transformation experiments with other vectors (p5'PGK-LACZ1 and all vectors below this in the Table), the constructs were linearized prior to being introduced into *A.nidulans*.

Transformed strains were classified by DNA hybridization analysis with regard to the type of integration (Type I or II) or replacement (Type III) event that had occurred between vector and genomic sequences (Chapters 5 and 6). Transformed strains classified as Novel gave a hybridization pattern with new bands corresponding to a Type I integration event but also retaining the native band (Chapter 5).

The number of strains of each type that express *lacZ*, as determined by colony staining on X-gal plates is given in parenthesis (see also Chapters 4, 5 and 6).

Vector	Extent of 5'pgk sequence (bp)	Number of Transformed Strains Analysed	Class of Transformant					
			Type I		Type II	Type III	Novel	Unclear
			Single copy	Multiple copy				
p5'PGK-LACZ2 (circular)	487	30 (12)	4 (4)	1 (1)	12 (7)	13 (0)	0	0
p5'PGK-LACZ2 (linear)	487	32 (21)	12 (12)	4 (4)	1 (1)	11 (0)	4 (4)	0
p5'PGK-LACZ1	487	8 (0)	3 (0)	0	3 (0)	2 (0)	0	0
p5'PGK-LACZ3	487	8 (0)	2 (0)	0	4 (0)	2 (0)	0	0
p5'PGK-LACZ2-1	354	20 (15)	3 (3)	3 (3)	8 (8)	5 (0)	1 (1)	0
2-2	345	12 (10)	1 (1)	3 (3)	2 (2)	2 (0)	3 (3)	1 (1)
2-3	292	8 (4)	2 (2)	0	2 (2)	4 (0)	0	0
2-4	239	12 (8)	1 (1)	0	5 (5)	4 (0)	2 (2)	0
2-5	161	32 (19)	4 (4)	1 (1)	15 (14)	12 (0)	0	0
2-6	120	32 (9)	2 (0)	0	13 (9)	14 (0)	3 (0)	0
2-7	116	12 (4)	1 (0)	1 (0)	7 (4)	1 (0)	1 (0)	1 (0)
2-8	111	13 (2)	2 (0)	0	4 (2)	6 (0)	0	1 (0)
2-9	90	12 (0)	1 (0)	1 (0)	5 (0)	3 (0)	2 (0)	0
2-12	44	10 (0)	3 (0)	0	3 (0)	4 (0)	0	0
p5'PGK-LACZ4	638	12 (10)	3 (3)	1 (1)	4 (4)	2 (0)	2 (2)	0

References.

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

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

Arst,H.N. and Bailey,C.R. (1977) The regulation of carbon metabolism in *Aspergillus nidulans*. In: Smith,J.E. and Pateman,J.A. (eds), "Genetics and physiology of Aspergillus." Academic Press, London and New York. pp 97-130

Arst,H.N.,Jr. and Scazzocchio,C. (1985) Formal genetics and molecular biology of the control of gene expression in *Aspergillus nidulans*. In: Bennett,J.W. and Lasure,L.L. (eds.), "Gene manipulation in fungi", Academic Press, London and New York. pp 309-34

Atkins,R.A. and Lambowitz,A.M. (1985) A general method for cloning *Neurospora crassa* nuclear genes by complementation of mutants. Mol. Cell. Biol., 5, 2272-2278

Ballance,D.J., Buxton,F.P. and Turner,G. (1983) Transformation of *Aspergillus nidulans* by the orotidine-5'-phosphate decarboxylase gene of *Neurospora crassa*. Biochem. Biophys. Res. Commun., 112, 284-289

Ballance,D.J. and Turner,G. (1985) Development of a high frequency transforming vector for *Aspergillus nidulans*. Gene, 36, 321-331

Banks,R.D., Blake,C.C.F., Evans,P.R., Haser,R., Rice,D.W., Hardy,G.W., Merret,M. and Phillips,A.W. (1979) Sequence, structure and activity of phosphoglycerate kinase; a possible hinge-bending enzyme. Nature, 279, 773-777

Bannerjee,P.C., Darzins,A. and Maitra,P.K. (1987)

Gluconeogenic mutations in *Pseudomonas aeruginosa*: genetic linkage between fructose-biphosphate aldolase and phosphoglycerate kinase. J. Gen. Microbiol., 133, 1099-1107

Beri,R.K., Grant,S., Roberts,C.F., Smith,M. and Hawkins,A.R. (1990) Selective overexpression of the *QUTE* gene encoding catabolic 3-dehydroquinase in multicopy transformants of *Aspergillus nidulans*. Biochem. J., 265, 337-342

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

Bigelis,R., Keesey,J. and Fink,G.R. (1977) The *his4* fungal gene cluster is not polycistronic. In: Wilcox,G., Abelson,J. and Fox,C.F. (eds.), "Molecular Approaches to Eukaryotic Genetic Systems", 8, ICN-UCLA Symposia on Molecular and Cellular Biology. Academic Press, London and New York. pp 179-187

Biggin,M.D., Gibson,T.J. and Hong,H.F. (1983) Buffer gradient gels and 35S label as an aid to rapid DNA sequencing. Proc. Natl. Acad. Sci. USA, 80, 3963-3965

Birboim,H.C. and Doly,J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res., 7, 1513-1523

Boguslawski,G. (1985) Yeast transformation. In: Bennett,J.W and Lasure,L.L. (eds.), "Gene Manipulations in Fungi", Academic Press, London and New York. pp 161-195

Bohmann,D., Keller,W., Dale,T., Scholer,H.R., Tebb,G. and Mattaj,I.W. (1987) A transcriptional factor which binds to the enhancers of SV40, immunoglobulin heavy chain and U2 snRNA genes. Nature, 325, 268-272

Bradford,M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein, utilizing

the principle of protein-dye binding. Anal. Biochem., 72, 248-254

Broach,J.R., Strathern,J.N. and Hicks,J.B. (1979) Transformation in yeast: development of a hybrid cloning vector and isolation of the *CAN1* gene. Gene, 8, 121-133

Brown,A.J.P. and Lithgow,G.J. (1987) The structure and expression of nuclear genes in *Saccaromyces cerevisiae*. In: Kinghorn,J.R. (ed.), "Gene structure in Eukaryotic Microbes", IRL Press, Oxford. pp 1-26

Bull, J.H., Smith,D.J. and Turner,G. (1988) Transformation of *Penicillium chrysogenum* with a dominant selectable marker. Curr. Genet., 13, 377-382

Buxton,F.P. and Radford,A. (1984) The transformation of mycelial spheroplasts of *Neurospora crassa* and the attempted isolation of an autonomous replicator. Mol. Gen. Genet., 196, 337-344

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

Calame,K.L. (1989) Immunoglobulin gene transcription: molecular mechanisms. Trends Genet. 5, 395-399

Casadaban,M.J., Martinez-Arias,A., Shapira,S.K. and Chou,J. (1983) β-Galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. Meth. Enzymol., 100, 293-308

Case,M.E., Schweizer,M., Kushner,S.R. and Giles,N.H. (1979) Efficient transformation of *Neurospora crassa* by utilizing hybrid plasmid DNA. Proc. Natl. Acad. Sci. USA, 76, 5259-5263

Case,M.E. (1982) Transformation of *Neurospora crassa*. In:

Hollaender,A., DeMoss,R.D., Koninsky,J., Savage,D. and Wolfe,R.S. (eds.), "Engineering of Microorganisms for Chemicals", Basic Life Sciences, 19. Plenum, New York. pp 87-100

Chamberlin,M.J. (1976) RNA polymerase: an overview. In: Losick,R and Chamberlin,M.J. (eds.), "RNA polymerase", Cold Spring Harbour Laboratory. pp 17-67

Chambon,P., Dierich,A., Gaub,M-P., Jakowler,S., Jongstra,J., Krust,A., LePennec,J-P., Oudet,P. and Keudelhuber,T. (1983) Promoter elements of genes coding for proteins and modulation of transcription by oestrogens and progesterone. Rec. Prog. Horm. Res., 40, 1-42

Chan,C.S.M. and Tye,B-K. (1980) Autonomously replicating sequences in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA, 77, 6329-6332

Charles,I.G., Keyte,J.W., Brammer,W.J., Smith,M. and Hawkins,A.R. (1986) The isolation and nucleotide sequence of the complex *AROM* locus of *Aspergillus nidulans*. Nucl. Acids Res., 14, 2201-2213

Chen,C.Y. and Hitzeman,R.A. (1987) Human, yeast and hybrid *PGK* gene expression in yeast. Nucl. Acids Res., 15, 643-660

Chen,C.Y., Opperman,H. and Hitzeman,R.A. (1984) Homologous versus heterologous gene expression in the yeast *S.cerevisiae*. Nucl. Acids Res., 12, 8951-8970

Chen,W. and Struhl,K. (1985) Yeast mRNA initiation sites are determined primarily by specific sequences, not by the distance from the TATA element. EMBO J., 4, 3273-3280

Choi,O-R and Engel,J.D. (1986) A 3′ enhancer is required for temporal and tissue-specific transcriptional activation of the chicken adult β-globin gene. Nature, 323, 731-734

Chomczynski,P. and Sacchi,N. (1987) Single-step method of

RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Analytical Biochemistry, 162, 156-159

Ciliberto,G., Castagnoli,L. and Cortese,R. (1983) Transcription by RNA polymerase III. Curr. Topics Dev. Biol., 18, 59-88

Clarke,L. and Carbon,K. (1980) Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature, 287, 504-509

Clements,J.M. (1986) Molecular characterization of the 3-phosphoglycerate kinase gene from *Aspergillus nidulans*. Ph.D. Thesis, University of Leicester

Clements,J.M. and Roberts,C.F. (1985) Molecular cloning of the 3-phosphoglycerate kinase (*PGK*) gene from *Aspergillus nidulans*. Curr. Genet., 9, 293-298

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

Clutterbuck,A.J. (1982) *Aspergillus nidulans*. Genet. Maps, 2, 209-217

Cockerill,P.N. and Garrard,W.T. (1986) Chromosomal loop anchorage of the kappa immunoglobulin gene locus next to the enhancer in a region containing topoisomerase II sites. Cell, 44, 273-282

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

Daboussi,M.J., Djeballi,A., Gerlinger,C., Blaiseau,P.L. Bouvier,I., Cassan,M., Lebrun,M.H., Parisot,D. and Brygoo,Y. (1989) Transformation of seven species of filamentous fungi using the nitrate reductase gene of *Aspergillus nidulans*. Curr. Genet., 15, 453-456

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

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

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

deGraff,L., van den Broek,H. and Visser,J. (1988) Isolation and transformation of the pyruvate kinase gene of *Aspergillus nidulans*. Curr. Genet., 13, 315-321

deGraaff,L.H. (1989) The structure and expression of the pyruvate kinase gene of *Aspergillus nidulans* and *Aspergillus niger*. Ph.D. Thesis, Wageningen Agricultural University, The Netherlands

Dierks,P., van Ooyen,A., Cochran,M., Dobkin,C., Reiser,J. and Weissmann,C. (1983) Three regions upstream from the capsite are required for the efficient and accurate transcription of the rabbit β-globin gene in mouse 3T6 cells. Cell, 32, 695-706

Dobson,M.J., Mellor,J., Fulton,A.M., Roberts,N.A., Bowen,B.A., Kingsman,S.M. and Kingsman,A.J. (1984) The identification and high level expression of a protein encoded by the yeast *Ty* element. EMBO J., 3, 1115-1119

Dobson,M.J., Tuite,M.F., Roberts,N.A., Kingsman,A.J. and Kingsman,S.M. (1982) Conservation of high efficiency promoter sequences in *Saccharomyces cerevisiae*. Nucl. Acids Res., 10, 2625-2637

Dower,W.J., Miller,J.F. and Ragsdale,C.W. (1988) High efficiency transformation of *E.coli* by high voltage

electroporation. Nucl. Acids Res., 16, 6127

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

Dynan,W.S. and Tjian,R. (1983) The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell, 35, 79-87

Eaton,S. and Calme, K. (1987) Multiple DNA sequence elements are necessary for the function of an immunoglobulin heavy chain promoter. Proc. Natl. Acad. Sci. USA, 84, 7634-7638

Eicher,E.M., Cherry,M. and Flaherty,L. (1978) Autosomal phosphoglycerate kinase linked to mouse major histocompatibility complex. Mol. Gen. Genet., 158, 225-228

Elzaing,T.A., Massan,M.M. and Allan,A.M. (1973) New pathway for non-phosphorylated degradation of glucanate in *Aspergillus niger*. J. Bacteriol., 114, 457-459

Everret,R.D., Baty,D. and Chambon,P. (1983) The repeated GC-rich motifs upstream from the TATA box are important elements of the SV40 early promoter. Nucl. Acids Res., 11, 2447-2464

Falkner,F.G., Mocikat,R. and Zachau, H.G. (1986) Sequences closely related to an immunoglobulin gene promoter/enhancer element occur also upstream of other eukaryotic and of prokaryotic genes. Nucl. Acids Res., 14, 8819-8827

Falkner,F.G. and Zachau,H.G. (1984) Correct transcription of an immunoglobulin κ gene requires an upstream fragment containing conserved sequence elements. Nature, 310, 71-74

Fantes,P.A. (1972) Studies on the β-galactosidase system in *Aspergillus nidulans*. Ph.D. Thesis, University of Leicester

Fantes,P.A. and Roberts,C.F. (1973) β-Galactosidase activity and lactose utilization in *Aspergillus nidulans*. J. Gen. Microbiol., 77, 471-486

Feinberg,A.P. and Vogelstein,B. (1984) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Analytical Biochemistry, 137, 266-267

Fincham,J.R.S (1989) Transformation in fungi. Microbiol. Rev., 53, 148-170

Fitzgerald-Hayes,M., Buhler,J-M., Cooper,T.G. and Carbon,J. (1982) Isolation and subcloning analysis of functional centromere DNA (*CEN1*) from *Saccaromyces cerevisiae* chromosome XI. Mol. Cell. Biol., 2, 82-87

Frederick,G.D., Asch,D.K. and Kinsey,J.A. (1989) Use of transformation to make targeted sequence alterations at the *am* (*GDH*) locus of Neurospora. Mol. Gen. Genet., 217, 294-300

Free,S.J., Rice,P.W. and Metzenberg,R.L. (1979) Arrangement of genes coding for ribosomal ribonucleic acids in *Neurospora crassa*. J. Bacteriol., 137, 1219-1226

Gasser,S.M. and Laemmli,U.K. (1986) Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D.melanogaster*. Cell, 46, 521-530

Gerster,T., Matthias,P., Thali,M., Jiricny,J. and Schaffner,W. (1987) Cell type-specificity elements of the immunoglobulin heavy chain gene enhancer. EMBO J., 6, 1323-1330

Gidoni,D., Kadonaga,J.T., Barrera-Saldana,H., Takahashi,K., Chambon,P. and Tjian,R. (1985) Bidirectional SV40 transcription mediated by tandem Sp1 binding interactions. Science, 230, 511-517

Giles,N.H., Case,M.E., Baum,J., Geever,R., Huiet,L., Patel,V. and Tyler,B. (1985) Gene organisation and regulation of the *qa* (Quinic acid) gene cluster of *Neurospora crassa*. Microbiol. Rev., 49, 338-358

Giles,N.H., Case,M.E., Partridge,W.H. and Ahmed,S.I. (1967) A gene cluster in *Neurospora crassa* coding for an aggregate of five aromatic synthetic enzymes. Proc. Natl. Acad. Sci. USA, 58, 1453-1460

Gillies,S.D., Morrison,S.L., Oi,V.T. and Tonegawa,S. (1983) A tissue specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell, 33, 717-728

Giniger,E. and Ptashne,M. (1987) Transcription in yeast activated by a putative amphipathic α-helix linked to a DNA binding unit. Nature, 330, 670-672

Giniger,E., Varnum,S.M. and Ptashne,M. (1985) Specific DNA binding of *GAL4*, a positive regulatory protein of yeast. Cell, 40, 767-774

Goldberg,M.E. (1969) Tertiary structure of *Escherichia coli* β-D-galactosidase. J. Molec. Biol., 46, 441-446

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

Graves,B.J., Johnson,P.F. and McKnight,S.L. (1986) Homologous recognition of a promoter domain common to the MSV LTR and the HSV *tk* gene. Cell, 44, 565-576

Greenberg,M.E., Siegfried,Z. and Ziff,E.B. (1987) Mutation of the *c-fos* gene dyad symmetry element inhibits serum inducibility of transcription *in vivo* and the nuclear regulatory factor binding *in vitro*. Mol. Cell. Biol., 7, 1217-1225

Guarente,L. and Hoar,E. (1984) Upstream activation sites of the *CYC1* gene of *S.cerevisiae* are active when inverted but not when placed downstream of the TATA box. Proc. Natl. Acad. Sci. USA, 81, 7860-7864

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

Gwynee,D.I., Buxton,F.P., Silbey,S., Davis,R.N., Lockington,R.A., Scazzochio,C. and Sealey-Lewis,H.M. (1987) Comparison of the cis-acting control regions of two coordinately controlled genes involved in ethanol utilization in *Aspergillus nidulans*. Gene, 51, 205-216

Hall,B., Clarkson,S.G. and Tocchini-Valentini,G. (1982) Transcriptional initiation of eukaryotic transfer RNA genes. Cell, 29, 3-5

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

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

Hansen,U. and Sharp,P. (1983) Sequences controlling *in vitro* transcription of SV40 promoters. EMBO J., 2, 2293-2303

Harvey,R.P., Robins,A.J. and Wells,J.R.E. (1982) Independently evolving chicken histone H2B genes: identification of a ubiquitous H2B-specific 5' element. Nucl. Acids Res., 10, 7851-7863

Hastie,A.C. (1970) Benelate-induced unstability of Aspergillus. Nature, 226, 771

Hawkins,A.R. (1987) The complex *Arom* locus of *Aspergillus*

nidulans. Evidence for multiple gene fusions and convergent evolution. Curr. Genet., 11, 491-498

Hawkins,A.R., DaSilva,A.J.F. and Roberts,C.F. (1985) Cloning and characterization of the three enzyme structural genes *QUTB*, *QUTC* and *QUTE* from the quinic acid utilization gene cluster in *Aspergillus nidulans*. Curr. Genet., 9, 305-311

Hawkins,A.R., Lamb,H.K., Smith,M., Keyte,J.W. and Roberts,C.F. (1988) Molecular organisation of the quinic acid utilization (*QUT*) gene cluster in *Aspergillus nidulans*. Mol. Gen. Genet., 214, 226-231

Hatzopoulos,A.K., Schlokat,U. and Gruss,P. (1988) Enhancers and other cis-acting regulatory sequences. In: Hames,B.D. and Glover,D.M. (eds.), "Transcription and splicing", IRL Press Ltd, Oxford. pp 43-96

Heuchel,R., Matthias,P. and Schaffner,W. (1989) Two closely separated promoters are equally activated by a remote enhancer: evidence against a scanning model for enhancer action. Nucl. Acids Res., 17, 8931-8947

Hicks,J.B., Hinnen,A. and Fink,G.R. (1979) Properties of yeast transformation. Cold Spring Harbour Symp. Quant. Biol., 43, 1305-1313

Hinnen,A., Hicks,J.B. and Fink,G.R. (1978) Transformation of yeast chimaeric *ColE1* plasmid carrying *LEU2*. Proc. Natl. Acad. Sci. USA, 75, 1929-1933

Hitzeman,R.A., Hagie,F.E., Hayflick,J.S., Chen,C.Y., Seeburg,P.H. and Derynck,R. (1982) The primary structure of the *Saccharomyces cerevisiae* gene for 3-phosphoglycerate kinase. Nucl. Acids Res., 10, 7791-7808

Holland,J.P. and Holland M.J. (1978) Isolation and identification of yeast messenger ribonucleic acids coding for enolase, glyceraldehyde 3-phosphate dehydrogenase and

phosphoglycerate kinase. Biochemistry, 17, 4900-4907

Horikoshi,M., Wang,C.K., Fujii,H., Cromlish,J.A., Weil,P.A. and Roeder,R.G. (1989) Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. Nature, 341, 299-303

Huang,I.Y., Welsh,C.D. and Yoshida,A. (1980) Complete amino acid sequence of human phosphoglycerate kinase. J. Biol. Chem., 255, 6412-6420

Huiet,L. and Giles,N.H. (1986) The *qa* repressor gene of *Neurospora crassa*: wild type and mutant nucleotide sequences. Proc. Natl. Acad. Sci. USA, 83, 3381-3385

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

Hynes,M.J. and Davis,A. (1986) The *amdS* gene of *Aspergillus nidulans*: Control by multiple regulatory signals. BioEssays, 5, 123-128

Ilgen,C., Farabaugh,P.J., Hinnen,A., Walsh,J.M. and Fink,G.R. (1979) Transformation in yeast. In Setlow,J.K. and Hollaender,A. (eds.), "Genetic Engineering", Vol 1. Plenum, New York. pp 117-132

Irani,M. and Maitra,P.K. (1974) Isolation and characterization of *Escherichia coli* mutants defective in enzymes of glycolysis. Biochem. Biophys. Res. Commun., 56, 127-133

Johnston,A.D. and Herskowitz,I. (1985) A repressor (*MAT* α2 product) and its operator control expression of a set of cell-type specific genes in yeast. Cell, 42, 237-247

Jones,K.A., Yamamoto,K.R. and Tjian,R. (1985) Two distinct transcription factors bind to the HSV thymidine kinase

promoter *in vitro*. Cell, 42, 559-572

Karin,M., Haslinger,A., Holtgreve,H., Richards,R., Krauter,P., Westphal,H. and Beats,M. (1984) Characterization of DNA sequences through which cadmium and glucocorticoid hormones induce human metallothionein-II_A gene. Nature, 308, 513-519

Kaster,K.R., Burgett,S.G. and Ingolia,T.D. (1984) Hygromycin B resistance as dominant selectable marker in yeast. Curr. Genet., 8, 355-358

Kim,S.Y. and Marzluf,G.A. (1988) Transformation of *Neurospora crassa* with the *trp-1* gene and the effect of the host strain upon the fate of the transforming DNA. Curr. Genet., 13, 65-70

Kinghorn,J.R. and Pateman,J.A. (1977) The regulation of nitrogen assimilation. In: Smith,J.E. and Pateman,J.A. (eds), "Genetics and physiology of Aspergillus", Academic Press, London and New York. pp 203-241

Kingsman,A.J., Gimlich,R.L., Clarke,L. Chinault,A.C. and Carbon,J. (1981) Sequence variation in dispersed repetitive sequences in *Saccharomyces cerivisiae*. J. Mol. Biol., 145, 619-632

Kinsey,J.A. and Rambosek,J.A. (1984) Transformation of *Neurospora crassa* with the cloned *am* (glutamate dehydrogenase) gene. Mol. Cell. Biol., 4, 117-122

Korn,L.J. (1982) Transcription of Xenopus 5S ribosomal RNA genes. Nature, 295, 101-105

Kosche,K.A., Dobkin,C. and Bank,A (1985) DNA sequences regulating human β-globin gene expression. Nucl. Acids Res., 13, 7781-7793

Kozak,M. (1981) Composition and analysis of sequence upstream from the translational start site in eukaryotic

mRNAs. Nucl. Acids Res., 12, 857-872

Krumlauf,R. and Marzluf,G.A. (1979) Characterization of the sequence complexity and organization of the *Neurospora crassa* genome. Biochemistry, 18, 3705-3713

Krumlauf,R. and Marzluf,G.A. (1980) Genome organization and characterization of the repititive and inverted repeat DNA sequences in *Neurospora crassa*. J. Biol. Chem., 225, 1138-1145

Kulbe,K.D. and Bojanovski,M. (1982) 3-Phosphoglycerate kinase from bovine liver and yeast. Meth. Enzymol., 90, 115-120

La Bella,F., Sive,H.L., Roeder,R.G. and Heintz,N. (1988) Cell-cycle regulation of a human histone H2b gene is mediated by the H2b subtype-specific consensus sequence.. Genes Dev., 2, 32-39

Lam,K.B. and Marmur,J. (1977) Isolation and characterization of *Saccharomyces cerevisiae* glycolytic pathway mutants. J. Bacteriol., 130, 746-749

La Thangue,N.B. and Rigby,P.W.J. (1988) Trans-acting protein factors and the regulation of eukaryotic transcription. In: Hames,B.D. and Glover,D.M. (eds.), "Transcription and splicing", IRL Press Ltd, Oxford. pp 1-42

Lee,W., Haslinger,A., Karin,M. and Tjian,R. (1987) Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. Nature, 325, 368-372

Legerton,T.L. and Yanofsly,C. (1985) Cloning and characterization of the multifunctional *his-3* gene of *Neurospora crassa*. Gene, 39, 129-140

Lis,J.T., Simon,J.A. and Sutton,C.A. (1983) New heat shock puffs and β-galactosidase activity resulting from

transformation of Drosophila with an *hsp70-lacZ* hybrid gene. Cell, 35, 403-410

Lowry,O.H., Rosenbrough,N.J., Farr,A.L. and Randall,R.J. (1951) Protein estimation with the Folin phenol reagent. J. Biol. Chem., 193, 265-275

Lusky,M., Berg,L., Weiher,H. and Botchan,M. (1983) Bovine papilloma virus contains an activator of gene expression at the distal end of the early transcription unit. Mol. Cell. Biol., 3, 1108-1122

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

Matthes,H.W.D., Zenke,W.M., Grundstrom,T., Staub,A., Winterzerith,M. and Chambon,P. (1984) Simultaneous rapid chemical synthesis of over one hundred oligonucleotides on a microscale. EMBO J., 3, 801-805

McCullough,W., Payton,M.A. and Roberts,C.F.(1977) Carbon metabolism in *Aspergillus nidulans*. In: Smith,J.E. and Pateman,J.A. (eds), "Genetics and physiology of Aspergillus." Academic Press, London and New York. pp 131-146

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

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

McNeil,J.B. and Smith,M. (1985) *S.cerevisiae* mRNA 5'-end positioning: analysis by *in vitro* mutagenesis, using synthetic duplexes with random mismatch base pairs. Mol.

Cell. Biol., 5, 3545-3551

McNeil,J.B. and Smith,M. (1986) Transcription initiation of the *S.cerevisiae* iso-1-cytochrome c gene. Multiple, independent TATA sequences. J. Mol. Biol., 187, 363-378

Mellor,J., Dobson,M.J., Kingsman,A.J. and Kingsman,S.M. (1987) A transcriptional activator is located in the coding region of the yeast *PGK* gene. Nucl. Acids Res., 15, 6243-6259

Mellor,J., Dobson,M.J., Roberts,N.A., Kingsman,A.J. and Kingsman,S.M. (1985) Factors affecting heterologous gene expression in *Saccharomyces cerevisiae*. Gene, 33, 215-226

Mellor,J., Dobson,M.J., Roberts,N.A., Tuite,M.F., Emtage,J.S., White,S., Low,P.A., Patel,T., Kingsman,A.J. and Kingsman,S.M. (1983) Efficient synthesis of enzymatically active calf chymosin in *S.cerevisiae*. Gene, 24, 1-14

Mercola,M., Goverman,J., Mirell,C. and Calame,K. (1985) Immunoglobulin heavy-chain enhancer requires one or more tissue specific factors. Science, 227, 266-270

Messing,J. (1983) New M13 vectors for cloning. Meth. Enzymol., 101, 20-78

Michelson,A., Markham,A. and Orkin,S. (1983) Isolation and DNA sequence of a full length cDNA clone for human X-chromosome-encoded phosphoglycerate kinase. Proc. Natl. Acad. Sci. USA, 80, 472-476

Miller,B.L., Miller,K.Y. and Timberlake,W.E. (1985) Direct and indirect gene replacements in *Aspergillus nidulans*. Mol. Cell. Biol., 5, 1714-1721

Mills,F.C., Fisher,L.M., Kuroda,R., Ford,A.M. and Gould,H.J. (1983) DNaseI hypersensitive sites in the chromatin of human μ immunoglobulin heavy-chain genes. Nature, 306, 809-812

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

Morris,N.R. (1978) Preparation of large molecular weight DNA from the fungus *Aspergillus nidulans*. J. Gen. Microbiol., 106, 387-389

Mount,S.M. (1982) A catalogue of splice junction sequences. Nucl. Acids Res., 10, 459-472

Muller,H-P., Sogo,J.M. and Schaffner,W. (1989) An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge.
Cell, 58, 767-777

Murray,A.W. and Szostak,J.W. (1983a) Pedigree analysis of plasmid segregation in yeast. Cell, 34, 961-970

Murray,A.W. and Szostak,J.W. (1983b) Construction of artificial chromosomes in yeast. Nature, 305, 189-193

Newbury,S.F., Glazebrook,J.A. and Radford,A. (1986) Sequence analysis of the pyr-4 (orotidine 5'-P decarboxylase) gene of *Neurospora crassa*. Gene, 43, 51-58

Nielsen,J.T. and Chapman,V.M. (1977) Electrophoretic variation for X-chromosome linked phosphoglycerate kinase (*PGK-1*) in the mouse. Genetics, 87, 319-325

Nielson,D.A., Chou,J., MacKrell,A.J., Casadaban,M.J. and Steiner,D.F. (1983) Expression of a preproinsulin-β-galactosidase gene fusion in mammalian cells. Proc. Natl. Acad. Sci. USA, 80, 5198-5202

Noll,M. (1976) Differences and similarities in chromatin structure of *Neurospora crassa* and higher eukaryotes. Cell, 8, 349-355

Nomiyama,H., Fromental,C., Xiao,J.H. and Chambon,P. (1987) Cell-specific activity of the constituent elements of the simian virus 40 enhancer. Proc. Natl. Acad. Sci. USA., 84, 7881-7885

Oakley,B.R., Rinehart,J.E., Mitchell,B.L., Oakley,C.E., Carmona,C., Gray,G.L. and May,G.S. (1987) Cloning mapping and molecular analysis of the *PYR-G* orotidine-5'-phosphate decarboxylase gene of *Aspergillus nidulans*. Gene, 61, 385-400

Ogden,J.E., Stanway,C., Kim,S., Mellor,J., Kingsman,A.J. and Kingsman,S.M. (1986) Efficient expression of the *Saccharomyces cerevisiae PGK* gene depends on an upstream activation sequence but does not require TATA sequences. Mol. Cell. Biol., 6, 4335-4343

Ondek,B., Shepard,A. and Herr,W. (1987) Discrete elements within the SV40 enhancer region display different cell-specific enhancer activities. EMBO J., 6, 1017-1025

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

Orr-Weaver,T.L., Szostak,J.W. and Rothstein,R.J. (1981) Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. USA, 78, 6354-6358

Orr-Weaver,T.L. and Szostak,J.W. (1983) Multiple tandem plasmid integration in *Saccharomyces cerevisiae*. Mol. Cell. Biol., 3, 747-749

Orr-Weaver,T.L., Szostak,J.W. and Rothstein,R.J. (1983) Genetic applications of yeast transformation with linear and gapped plasmids. Meth. Enzymol., 101, 228-245

Paietta,J.V. and Marzluf,G.A. (1985) Gene disruption by transformation in *Neurospora crassa*. Mol. Cell. Biol., 5,

1554-1557

Parslow,T.G., Jones,S.D., Bond,B. and Yamamoto,K.R. (1987) The immunoglobulin octanucleotide: independent activity activity and selective interaction with enhancers. Science, 235, 1498-1501

Pateman,J.A., Doy,C.H., Olsen,J.E., Kane,H.J. and Creasor,E.M. (1985) Molecular analysis of alcohol metabolism in Aspergillus. In: Timberlake,W.E. (ed.), "Molecular Genetics of Filamentous Fungi", A.R.Liss, New York. pp 171-184

Pelham,H. (1982) A regulatory upstream promoter element in the Drosophila *Hsp 70* heat shock gene. Cell, 30, 517-528

Picard,D. and Schaffner,W. (1984) A lymphocyte-specific enhancer in the mouse immunoglobulin κ gene. Nature, 307, 80-82

Plon,S.E. and Wang,J.C. (1986) Transcription of the human β-globin gene is stimulated by an SV40 enhancer to which it is physically linked but topologically uncoupled. Cell, 45, 575-580

Pontecorvo,G., Roper,J.A., Hemmons,J.M., MacDonald,K.D. and Bufton,A.W.J. (1953) The genetics of *Aspergillus nidulans*. Adv. Genet., 5, 141-238

Proudfoot,N.J. and Whitelaw,E. (1988) Termination and 3' end processing of eukaryotic RNA. In Hames,P.D. and Glover,D.M. (eds.) "Transcription and Splicing", IRL Press Ltd., Oxford. pp 97-129

Ptashne,M. (1986a) A Genetic Switch. Cell Press and Blackwell Scientific Publ., Cambridge, Palo Alto.

Ptashne,M. (1986b) Gene regulation by proteins acting nearby and at a distance. Nature, 322, 697-701

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

Rine,J. and Carlson,M. (1985) *Saccharomyces cerevisiae* as a paradigm for modern molecular genetics of fungi. In: Bennett,J.W. and Lasure,L.L. (eds.), "Gene manipulation in fungi", Academic Press, London and New York. pp 125-160

Roberts,C.F. (1969) Silica gel stock cultures of *Aspergillus nidulans*. Aspergillus News Letters, 10, 29

Roper,J.A. (1952). Production of heterozygous diploids in filamentous fungi. Experimentia, 8, 14-15

Rothstein,R.J. (1983) One-step gene disruption in yeast. Meth. Enzymol., 101, 202-211

Russel,P.J., Rodland,K.D., Cutter,J.E., Racklin,E.M. and McClosky,J.W. (1985) DNA methylation in Neurospora: Chromatographic and isoscizomer evidence for changes during development. In: Timberlake,W.E. (ed.), "Molecular Genetics of Filamentous Fungi", Liss,A.R., New York. pp 321-333

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

Scherer,S. and Davis,R.W. (1979) Replacement of chromosome segments with altered DNA sequences constructed *in vitro*. Proc. Natl. Acad. Sci. USA, 76, 4951-4955

Scopes,A.K. (1973) 3-Phosphoglycerate kinase. In: Boyer,P.D. (ed), "The Enzymes", Vol ^{8}, Academic Press, London and New York. pp 335-351

Seguin,C. and Hamer,D.H. (1987) Regulation *in vitro* of metallothionein gene binding factors. Science, 235, 1383-1387

Selker,E.U., Cambareri,E.B., Jensen,B.C. and Haack,K.R. (1987) Rearrangement of duplicated DNA in specialized cells of Neurospora. Cell, 51, 741-752

Shaul,Y., Rutter,W.J. and Laub,O. (1985) A human hepatitis B viral enhancer element. EMBO J., 4, 427-430

Shortle,D., Haber,J.E. and Botstein,D. (1982) Lethal disruption of the yeast actin gene by integrative DNA transformation. Science, 217, 371-373

Sigler,P.B. (1988) Transcriptional activation: Acid blobs and negative noodles. Nature, 333, 210-212

Sollner-Webb,B. and Tower,J. (1986) Transcription of cloned eukaryotic ribosomal RNA genes. Annu. Rev. Biochem., 55, 801-830

Southern,E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Molec. Biol., 98, 503-517

Sproat,B.S. and Gait,M.J. (1985) Chemical synthesis of a gene for somatomedin C. Nucl. Acids Res., 13, 2959-2977

Stanway,C., Mellor,J., Ogden,J.E., Kingsman,S.M. and Kingsman,A.J. (1987) The UAS of the yeast *PGK* gene contains functionally distinct domains. Nucl. Acids Res., 15, 6855-6873

Stinchcomb,D.T., Struhl,K. and Davis,R.W. (1979) Isolation and characterization of the yeast chromosomal replicator. Nature, 282, 39-43

Stohl,L.L. and Lambowitz,A.M. (1983) Construction of a shuttle vector for the filamentous fungus *Neurospora crassa*. Proc. Natl. Acad. Sci. USA, 80, 1058-1062

Stohl,L.L., Atkins,R.A. and Lambowitz,A.M. (1984) Characterization of deletion derivatives of an autonomously

replicating Neurospora plasmid. Nucl. Acids Res., 12, 6169-6178
Stout,J.T., Chen,H.Y., Brennand,J., Caskey,C.T. and Brinster,R.L.(1985). Expression of human HPRT in the central nervous system of transgenic mice. Nature 317, 250-252
Struhl,K. (1983) The new yeast genetics. Nature, 305, 391-397

Struhl,K. (1984) Genetic properties and chromatin structure of the yeast GAL regulatory element: an enhancer-like sequence. Proc. Natl. Acad. Sci. USA, 81, 7865-7869

Takahashi,K., Vigneron,M., Matthes,H., Wildeman,A., Zenke,M. and Chambon,P. (1986) Requirement of sterospecific alignments for initiation from the simian virus 40 early promoter. Nature, 319, 121-126

Thomas,G.H., Connerton.I.F. and Fincham,J.R.S. (1988) Molecular cloning, identification and transcriptional analysis of genes involved in acetate utilization in *Neurospora crassa*. Molec. Microbiol., 2, 599-606

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

Timberlake,W.E. and Barnard,E.C. (1981) Organisation of a gene cluster expressed specifically in the asexual spores of *Aspergillus nidulans*. Cell, 26, 29-37

Timberlake,W.E., Boylan,M.T., Cooley,M.B., Mirabito,P.M., O'Hare,E.B. and Willet,C.E. (1985) Rapid identification of mutation-complementing restriction fragments from *Aspergillus nidulans* cosmids. Exp. Mycol., 9, 351-355

Timberlake,W.E. (1978) Low repititive DNA content in *Aspergillus nidulans*. Science, 202, 973-974

Turner,G. and Ballance,D.J. (1985) Cloning and transformation in Aspergillus. In: Bennett,J.W. and

Lasure,L.L. (eds.), "Gene manipulation in fungi", Academic Press, London and New York. pp 259-278

Turner,G. and Ballance,D.J. (1986) Cloning vectors for *Aspergillus nidulans* based on *ansl*, a chromosomal fragment enhancing transformation frequency. In: Alacevic,M., Hranueli,D. and Toman,Z. (eds.), Proceedings of the Fifth International Symposium on the Genetics of Industrial Microorganisms. Pliva, Zagreb, Yugoslavia. pp 121-127

Tuite,M.F., Dobson,M.J., Roberts,N.A., King,R.M., Burke,D.C., Kingsman,S.M. and Kingsman,A.J. (1982) Regulated high efficiency expression of human interferon-alpha in *S.cerevisiae*. EMBO J., 1, 603-608

van Gorcom,R.F.M., Pouwels,P.H., Goosen,T., Visser,J., van den Broek,H.W.J., Hamer,J.E., Timberlake,W.E. and van den Hondel,C.A.M.J.J. (1985) Expression of an *Escherichia coli* β-galactosidase fusion gene in *Aspergillus nidualns*. Gene, 40, 99-106

van Gorcom,R.F.M., Punt,P.J., Pouwels,P.H. and van den Hondel,C.A.M.J.J. (1986) A system for the analysis of expression signals in Aspergillus. Gene, 48, 211-217

Vanhanen,S., Penttila,M., Lehtovaara,P. and Knowles,J. (1989) Isolation and characterization of the 3-phosphoglycerate kinase (*pgk*) gene of *Trichoderma reesei*. Curr. Genet., 15, 181-186

van Solingen,P., Muurling,H., Koekman,B. and van den Berg,J. (1988) Sequence of the *Penicillium chrysogenum* phosphoglycerate kinase gene. Nucl. Acids Res., 16, 11823

Vieira,J. and Messing,J. (1982) The pUC plasmids, an M13mp7 derived system for the insertion mutagenesis and sequencing with synthetic universal primers. Gene, 19, 259-268

Vollmer,S.J. and Yanofsky,C. (1986) Efficient cloning of genes of *Neurospora crassa*. Proc. Natl. Acad. Sci. USA, 83,

4867-4873

Voss,S.D., Schlokat,U. and Gruss,P. (1986) The role of enhancers in the regulation of cell-type-specific transcriptional control. Trends in Biochem. Sci., 11, 287-289

Wallenfels,K. (1962) β-Galactosidase (Crystalline). Meth. Enzymol., 5, 212-219

Wallenfels,K. and Weil,R. (1972) β-Galactosidase In: Boyer,P.D. (ed.),"The Enzymes", 3rd edition, Vol 7, Academic Press, London and New York. pp 617-663

Ward,M., Wilkinson,B. and Turner,G. (1986) Transformation of *Aspergillus nidulans* with a cloned oligomycin-resistant ATP synthase subunit 9 gene. Mol. Gen. Genet., 202, 265-270

Watson,H.C., Walker,N.P.C., Shaw,P.J., Bryant,T.N., Wendell,P.L., Fothergill,L.A., Perkins,R.E., Conroy,S.C., Dobson,M.J., Tuite,M.F., Kingsman,A.J. and Kingsman,S.M. (1982) Sequence and structure of yeast phosphoglycerate kinase. EMBO J., 1, 1635-1640

Wernars,K., Goosen,T., Sewart,K. and van den Broek,H.W.J. (1986) Genetic analysis of *Aspergillus nidulans amdS*$^{+}$ transformants. Mol. Gen. Genet., 205, 312-317

Wernars,K., Goosen,T., Wennekes,B.M.J., Sewart,K., van den Hondel,C.A.M.J.J. and van den Broek,H.W.J. (1987) Cotransformation of *Aspergillus nidulans*: a tool for replacing fungal genes. Mol. Gen. Genet., 209, 71-77

Wernars,K., Goosen,T., Wennekes,B.M.J., Visser,J., Bos,C.J., van den Broek,H.W.J., van Gorcom,R.F.M., van den Hondel,C.A.M.J.J. and Pouwels,P.H. (1985) Gene amplification in *Aspergillus nidulans* by transformation with vectors containing the amdS gene. Curr. Genet., 9, 361-368

Yang,L., Rowe,T.C., Nelson,E.M. and Liu,L.F. (1985) *In vivo*

mapping of DNA topoisomerase II-specific cleavage sites on SV40 chromatin. Cell, 41, 127-132

Yelton,M.M., Hamer,J.E. and Timberlake,W.E. (1984) Transformation of *Aspergillus nidulans* using a *trpC* plasmid. Proc. Natl. Acad. Sci. USA, 81, 1470-1474

Yelton,M.M., Timberlake,W.E. and van der Hondel,C.A.M.J.J. (1985) A cosmid for selecting genes by complementation in *Aspergillus nidulans*: Selection of the developmentally regulated *yA* locus. Proc. Natl. Acad. Sci. USA., 82, 834-838