Transcription in the Yeast *Saccharomyces cerevisiae*

A thesis submitted for the degree of

Doctor of Philosophy

at the University of Glasgow

by

Gordon James Lithgow

Institute of Genetics

Church Street

Glasgow

September 1989

Dedicated to my Father and to
the memory of my Mother
and my Uncle Robert.

The work presented in this thesis is my own unless otherwise acknowledged. No part of this thesis has been previously submitted for examination leading to the award of a degree

Gordon J. Lithgow

September 1989

Summary

The aim of this study was to investigate aspects of transcription of highly expressed yeast genes. To investigate the sequence requirements for transcriptional initiation putative promoters were constructed which contained promoter elements subcloned from the 5'-flanking region of the phosphoglycerate kinase gene (PGK1). The analysis of these promoters was hindered by high levels of transcriptional readthrough from adjacent sequences in the promoter probe vector. A CT-rich block, found upstream of many highly expressed yeast genes has previously been shown to influence the position of the transcriptional initiation site (I-site). Preliminary data presented here is consistent with the existence of a yeast protein which specifically interacts with this region.

Sequences located within the coding region of some yeast genes have previously been shown to be required for maximal expression. Evidence presented here confirms that sequences located within the coding region of the pyruvate kinase gene (PYK1) activate transcription. A protein has been shown to interact selectively with a fragment of PYK1 coding region DNA previously shown to activate transcription. It would appear however, that this PYK1 transcriptional activator can function in trans. specifically, translation of the PYK1 coding region sequences results in the elevation of both PYK1 mRNA abundance but also PGK1 mRNA abundance. This apparent trans activation of PGK1 mRNA abundance is not due to a "Heat Shock"-type stress response and may be a previously unreported regulatory feature.

Acknowledgements

I wish to thank my Supervisor, Dr Alistair J.P. Brown, and the other members of the Genetics Department in Glasgow, for encouragement and guidance.
In particular, thanks go to Iain Purvis for pointing out that most experiments are improved by including controls and to Fiona Stuart for explaining what the word "controls" means.
This work has benefited from comments and advice from many people, in particular Al Brown, Jane Mellor, Ian Dawes, Iain Purvis, Joe Ross, Paul Moore, Jane McKenzie, Clive Stanway and Jon Swaffield.
I also wish to extend my gratitude to those people who have donated materials and technical help. Lynn "supertech" Loughlin, Ian Dawes, Joe Ross, Paul Moore, Jane McKenzie, Chris Hadfield, and Maggie Smith have all contributed in this respect.
The interpretation of parts of this work relies upon the unpublished work of many individuals. I wish to thank the following people for freely discussing there own data; Jane Mellor, Joy Rathjen, Paul Moore, Iain Purvis, Iain Dawes, Zaf Zaman, Chris Hadfield and Joyce Moore.
This work was made possible by the people, atmosphere and organisation of the Genetics Department, especially and the hard working, ever cheerful media ladies.
I am particularly grateful to Fiona Stuart, Iain Johnstone and Al Brown for critically reading this thesis. There comments and suggestions were absolutely vital. As was the proof reading undertaken by Alison, Ann, Rick, Arvind and Bret.
I am especially indebted to my family and to Fiona for encouragement, support and having the inexplicable confidence that everything would be fine.

I was supported during this work by the SERC.

CONTENTS

Summary
Declaration
Acknowledgements
Chapter 1 Regulation of Gene Expression in Saccharomyces cerevisiae
1.1 Non-Transcriptional Regulation of Gene Expression in Yeast
1.1.1 Translational Control
1.1.2 Regulation by RNA Splicing of Editing
1.1.3 Regulation by Modification of mRNA Stability
1.1.4 Regulation by DNA Rearrangements
1.2 Transcriptional Regulation of Gene Expression in Yeast
1.2.1 Yeast Nuclear RNA Polymerases
1.2.2 Yeast Promoter Elements
1.2.3 Yeast Regulatory Proteins
1.2.3.1 DNA Binding Domains
1.2.3.2 Further Consideration of Protein/DNA Recognition
1.2.3.3 Mechanism of Activation by Regulatory Proteins
1.3 Regulation of Glycolytic Enzyme Genes in Yeast
1.4 Aims of Study

Chapter 2 Materials and Methods
2.1 Chemicals and Enzymes
2.1.1 Enzymes
2.1.2 Chemicals and Membranes
2.1.3 Growth Media
2.2 Organisms and Growth Conditions
2.2.1 Escherichia coli
2.2.2 Saccharomyces cerevisiae
2.2.3 Random Spore Analysis
2.2.4 Yeast B-galactosidase plate assay
2.3 Transformation of E.coli
2.4 Transformation of S.cerevisiae
2.5 Isolation of RNA from S.cerevisiae

2.5.1 Procedure
2.5.2 Storage of Cells for RNA Isolation
2.6 Isolation of DNA from E.coli
2.6.1 Birnboim, Doly Procedure
2.6.2 Holmes and Quigley STET Prep Procedure
2.7 Gel Electrophoresis
2.7.1 Agarose gels
2.7.2 Formaldehyde gels
2.7.3 Polyacrylamide gels
2.8 Manipulation of DNA in vitro
2.8.1 Restriction enzyme Digest
2.8.2 CIP Treatment
2.8.3 Purification of DNA Fragments
2.8.4 Ethanol Precipitation of DNA
2.8.5 Blunting of Recessed 3'-ends
2.8.6 Ligation of DNA Fragments
2.9 Plasmid Sequencing
2.10 Radiolabelling of DNA
2.10.1 Nick Translation
2.10.2 End-labelling of Oligonucleotides
2.10.3 End-labelling 5'-Overhangs
2.10.4 Removal of Unincorporated Nucleotides
2.11 Filter Hybridisations
2.11.1 Northern Blotting
2.11.2 Dot Blotting
2.11.3 Hybridisation
2.12 Gel Retardation Assays
2.13 Footprinting of Protein/DNA Complexes

Chapter 3 Investigation of the Minimal Sequence Requirements for Transcription from the Yeast PGK1 Promoter; I-Site Determining Sequences
3.1 Introduction
3.2 Role of PGK1 CT Block in Transcriptional Initiation
3.2.1 Construction of pGJLP Series
3.2.2 Construction of pGJLG Series
3.2.3 Construction of pCJGAL12

3.2.4 Selection of yGJL661
3.2.5 Expression of CAT in Yeast
3.2.7 Summary and Discussion of Results
3.3 Search for Specific Protein Binding at a Yeast I-Site
3.3.1 Introduction
3.3.2 A 41 bp Oligonucleotide Containing the PGK I-site and CT Block Does Not Interact with Yeast Protein
3.3.3 A 100 bp Fragment Containing the PGK1 CT Block and I-Site Interacts With a Number of Yeast Protein Fractions
3.3.4 Discussion

Chapter 4 The PYK1 Coding Region Contains A Transcriptional Activator
4.1 Introduction
4.2 The Presence of PYK1 Coding Region Fragments Within the PYK/LacZ Transcription Unit Elevates mRNA Abundance by Transcriptional Activation
4.2.1 Hybrid Genes are Transcribed inYCpPKZ Series Transformants
4.2.2 Fragments of the PYK1 Coding are Associated with Elevated mRNA Abundance
4.2.3 Elevation of mRNA Abundance is Carbon Source Dependant
4.2.4 Elevation of PYK/LacZ mRNA by the PYK1 Coding Region is Independant of mRNA Stability
4.2.5 Discussion of Results

Chapter 5 The PYK1 Coding Region May Interact with a Yeast Protein
5.1 Introduction
5.2 PYK1 Coding Region/Protein Interactions
5.2.1 Preparation of PYK1 DNA Fragment for Gel Retardation Assays
5.2.2 Detection of Protein/DNA Interactions
5.2.3 A Factor of Protein Fraction 33 Interacts Selectively With a Fragment of PYK1 Coding Region DNA

5.2.4 Discussion
5.3 Footprinting Analysis of Complexes
5.3.1 Introduction
5.3.2 Preparation of Radiolabelled DNA for Footprinting
5.3.3 Preparative Gel Retardation
5.3.4 Protein Binding to the PYK1 Coding Region Fragment Does Not Protect a Specific Site From Cleavage
5.4 Conclusions

Chapter 6 The PYK1 Coding Region May Activate Transcription in Trans
6.1 Introduction
6.2 Activation of Transcription by the PYK1 Coding Region May Depend Upon Translation
6.2.1 Constructionof YCpPKZ6.1
6.2.2 PYK1 Coding Region Fragment Y Does Not Activate Transcription in YCpPKZ6.1
6.3 Cells Expressing PYK1/LacZ or Overexpressing PYK1 Have Elevated PGK1 mRNA Abundance
6.3.1 PGK1 mRNA Abundance is Elevated in Cells Expressing PYK1/LacZ Fusion Genes
6.3.2 Trans-Activation of PGK1 by the PYK1 Coding Region is Dependant upon Carbon Source
6.3.3 Cells Containing Multiple Copies of the PYK1 Gene Exhibit PGK1 mRNA Abundance
6.4 Expression of the PYK1 Coding Region May Affect the Abundance of a number of Glycolytic Enzyme mRNAs
6.5 Cells With Elevated PGK1 mRNA Abundance Are Not Undergoing a "Heat Shock" Stress Response
6.5.1 Growth Rates of YCpPKZ1, YCpPKZ3 and YCpPKZ4 Transformants are Identical
6.5.2 HSP90 and RPL3 mRNA Abundances Indicate That PGK1 Overexpressing Cells Are Not Undergoing a Stress Response
6.6 Summary of Results

Chapter 7 Transcription of the PYK1 Gene in Yeast
A Discussion
7.1 Transcriptional Activation by the PYK1 Coding Region
7.2 Carbon Source Regulation of PYK1
7.3 Proposals for Further Work

References

CHAPTER 1

REGULATION OF GENE EXPRESSION IN *SACCHAROMYCES CEREVISIAE*

The regulation of gene expression is fundamental to the development and successful existence of all organisms. Differentiation, efficient nutrient utilisation, response to disease and cell growth all require that genes are expressed at a suitable level and at the correct time and location. Consequently, complex processes have evolved which tightly regulate the rate at which a gene product is synthesised. These processes act at various stages in the transfer of information, from the synthesis of RNA, to the synthesis of a correctly folded polypeptide product.

The study of gene regulation in the yeast Saccharomyces cerevisiae (hereafter termed "yeast") has contributed to the understanding of gene regulation as a whole. Yeast has attracted a great deal of interest as it may be viewed as a basic model system for aspects of gene regulation in higher eukaryotes and is important as a host for the production of prophylactic and therapeutic proteins.

Yeast has proved to be an excellent organism for the study of gene regulation. The ability to obtain mutant strains by straightforward screening methods has allowed many regulatory components to be defined and their genes cloned by complementation. This, together with the ability to introduce in vitro mutated derivatives of these genes back in yeast, either on autonomously replicating plasmids or at normal chromosomal loci, has facilitated the elucidation of the structural/functional relationships of protein and nucleic acid components of regulatory circuits.

Yeast exhibits gene regulation in response to environmental conditions, such as nutrient source variation and limitation, and also to undertake developmental change, such as mating type switching, cell fusion and sporulation. The expression of some genes also varies throughout the cell cycle. To a great extent gene

regulation is exhibited at the level of transcription, but the synthesis of some proteins is controlled at other levels.
This study is mainly concerned with the transcriptional regulation of a yeast gene, PYK1, which codes for the glycolytic enzyme pyruvate kinase. However some results presented here elude to regulation of this gene at levels other than transcription. For this reason, there follows a brief outline of mostly post-transcriptional regulation in yeast and other eukaryotes. Consideration will then be given to mechanisms of transcriptional regulation in yeast and finally the control of glycolytic enzyme genes shall be outlined.

1.1 Non-Transcriptional Regulation of Gene Expression in Yeast

1.1.1 Translational Control

The steady state abundance of messenger RNA (mRNA), in most circumstances, determines the level of expression of a gene. However, the degree of gene expression is also dependent upon the rate at which the mRNA is translated. Total cellular translation can be influenced by growth conditions via the regulation of genes coding for components of the translational apparatus (Huet et al., 1985; Nagashima et al., 1986 and Huet and Sentenac, 1987).
The sequence environment of the translational initiation codon is important in determining the rate at which an individual mRNA species is translated (Cigan and Donahue, 1987; Donahue and Cigan, 1988). As a consequence, the level of expression is partly determined by sequences close to the AUG.
The rate of translation of an individual mRNA can also be regulated in response to different conditions. In yeast

all mRNAs which have been shown to be under translational regulation contain open reading frames upstream of the protein coding region. These are termed Upstream Reading Frames (URFs).

The yeast CPA1 mRNA, which codes for carbamoyl-phosphate synthetase required for arginine biosynthesis, is subject to translational repression in the presence of arginine (Messenguy et al., 1983). The mRNA extends 250 bp upstream of the enzyme coding region and contains an URF of 25 codons. Disrupting this URF causes derepression of CPA1 mRNA translation in the presence of arginine. Translational regulation of this mRNA may be mediated by a trans-acting factor, the product of the CPAR gene. It has been proposed that this trans-acting factor and the peptide sythesised from the upstream URF form a complex in the presence of arginine, which inhibits the 40S ribosomal subunit from translocating to the start codon of the CPA1 coding region.

A transcriptional regulator of CPA1, GCN4 is also regulated at the level of translation. Under amino acid starvation, translation of GCN4 is induced by the products of GCN1, GCN2 and GCN3 which are thought to act by inhibiting the repression mediated by GCD1 (Hinnebush, 1985). As with CPA1, translational regulation of GCN4 is dependent upon upstream URFs. In this case, 4 URFs are required. The presence of the two URFs most proximal to the GCN4 coding region inhibits translation, whilst the distal URFs negatively regulate the influence of the proximal URFs GCD1 negatively regulates GCN4 translation by repressing the influence of the distal URFs. (Reviewed by Hinnebush, 1988).

There is evidence to suggest that some yeast nuclear gene products may be required for the translation of specific mitochondrial genes. Translation of coxII mRNA may be dependent on the integrity of the pet3 gene. Nuclear pet3 mutants block accumulation of mitochondrial encoded

cytochrome c oxidase subunit II (coxII), but coxII mRNA levels are only three fold lower in the mutant than in PETIII cells (Poutre and Fox, 1987). Constanzo and Fox, (1986) have shown that PET494 is required to promote the translation of the mitochondrial mRNA encoding cytochrome c oxidase subunit III (coxIII) and it has been demonstrated recently that the expression of PET494 itself, is regulated at the level of translation in response to oxygen (Marykwas and Fox, 1989).

A number of other yeast mRNAs, such as SUC2, HTS1, LEU4, TRM1 and MOD5 contain upstream URFs (Brown, 1989). By analogy with GCN4 and CPA1, these genes may also be subject to translational control.

Translational control may not obligatorily require the presence of URFs. Data from this laboratory, (Moore et al., Submitted to Gene) is consistent with the existence of a specific factor required for the translation of the pyruvate kinase (PYK1) mRNA. There are no URFs in the PYK1 mRNA. This data will be discussed in Chapter 7.

1.1.2 Regulation by mRNA Splicing

Splicing of pre-mRNA is important for efficient expression of some genes, not only because recognition and precise excision of introns is required for accurate translation but also because differential splicing of a single transcript can yield diverse mRNAs coding for distinct proteins.

Although some yeast genes contain introns (Reviewed by Gallwitz et al., 1987) regulation of gene expression based upon differential splicing has not been observed in this organism. The most striking example of this type of regulation is exhibited in the sex determination of Drosophila melanogaster. Expression of different activities of genes responsible for somatic sexual dimorphism is a consequence of variable RNA splicing as

determined by sxl, tra and tra2. sxl and tra2 exhibit sequence similarity to genes coding for single-stranded RNA binding proteins (Nagoshi et al., 1988). Two distinct mechanisms appear to operate. sxl apparently blocks a male specific splice site thus indirectly promoting the use of a downstream female specific site (Sosnowski et al., 1989). In contrast, the combined activity of the products of tra and tra2 directly promote the use of an otherwise inactive splice site (Burtis and Baker, 1989)
Although regulation by differential splicing has not been observed in yeast, the expression of at least one yeast gene may be regulated by inhibition of splicing. Over-expression of the ribosomal protein (RP) L32 results in the accumulation of unspliced RPL32 transcript (Warner et al., 1985). This accumulation is apparently dependent upon the presence in the cell of excess RPL32. When rRNA synthesis is inhibited by depletion of RNA polymerase I, unspliced transcript accumulates (Warner, 1989) suggesting that "excess" RPL32 inhibits splicing of the RPL32 mRNA.

1.1.3 Regulation by Modification of mRNA Stability

The steady-state abundance of an RNA is dependent upon both the rate of synthesis and the rate of degradation. By influencing the rate at which specific RNAs are degraded, the cell can regulate the expression of specific genes. There are well documented examples of this type of regulation in higher eukaryotes. For example, oestrogen stabilises Xenopus liver vitellogenin mRNA, and human epidermal growth factor (hEGF) increases the stability of the hEGF receptor mRNA (Jinno et al.,1988; Brock and Shapiro, 1983).
The half-lives of yeast RNAs are inversely proportional to RNA length (Santiago et al., 1986) and range from approximately 1 min to over 100 min (Reviewed by Brown,

1989). Yeast RNAs fall into two broad groups, stable (half-lives > 30 min.) and unstable (half-lives < 25 min; Santiago et al., 1986; Brown, 1989). The reason for the existence of a stable and an unstable population is not apparent but may reflect general regulatory features. For example, mRNAs with short half-lives may correspond to genes which are tightly regulated in a temporal fashion. Rapid clearing of such mRNAs from the cytoplasm may be essential for such regulation. In contrast, mRNAs with long half-lives may correspond to genes which are less tightly regulated or constitutively expressed and can therefore be expressed with the efficiency that a low turnover affords (Shapiro et al., 1987).

The regulation of specific yeast genes at the level of mRNA stability, has also been described. The abundance of histone protein H2A and H2B mRNAs is regulated during the cell cycle (Hereford et al., 1981 and 1982); they accumulate during early S-phase and undergo a twenty-fold decrease in abundance during late S-phase. This decrease is associated with a reduced transcription rate but also with a three-fold decrease in histone mRNA stability (Lycan et al., 1987).

The decrease in ribosomal protein mRNA in response to heat shock is also mediated partly by a reduction in mRNA stability. In response to a 23^{o}C to 36^{o}C shift, the half-life of ribosomal protein L3 decreases from 10 to 2.5 min (Herruer et al., 1988).

The mechanism of mRNA stability regulation in yeast is essentially unknown. Specific RNases, trans-acting stability modulator proteins and antisense RNA may be components of such a mechanism (Brown, 1989).

1.1.4 Regulation of Gene Expression by DNA Rearrangements

Yeast chromosome III possesses 3 mating type loci. That is, an active MAT locus and the silent (unexpressed) HML

and HMR loci. Mating type is specified by the active MAT locus, which may be MAT-a or MAT-alpha . Conversion of the MAT locus from a to alpha or vice versa, is dependent upon the transposition of a mating type cassette from one of the HML or HMR loci to the MAT locus (Beach, 1983). This transposition is initiated by a double-strand cut by the HO endonuclease at the MAT locus (Strathern et al., 1982). Transcription is repressed at the HML and HMR loci. Transposition to the MAT locus results in expression.

1.2 Transcriptional Regulation of Gene Expression in Yeast

The yeast genome totals approximately 1.4 x 10^4 Kbp and specifies approximately 5000 protein coding genes. Under normal growth conditions 50% of the genome is transcribed (Hereford and Rosbash, 1977; Kaback et al., 1979). Unlike enteric bacterial genes, related yeast genes are not clustered into operons (exceptions include GAL1 /GAL10 and MAT-alpha I/MAT-alpha 2). Therefore, co-ordinate control of related genes cannot rely upon the synthesis of a polycistronic messenger RNA. The average yeast gene is transcribed 5-10 times during each cell cycle which results in a steady state abundance of 1-2 mRNA molecules per cell (St John and Davis, 1979). Although most genes are constitutively expressed, the transcription rate of some regulated genes can be varied by 2 to 3 orders of magnitude.

As the genes investigated in this study code for metabolic enzymes, the following discussion will centre upon transcription of protein coding genes. However, consideration will also be given to the transcription of 5S rRNA gene, as it contains a control region internal to the gene and is therefore pertinent to this study.

1.2.1 Yeast Nuclear RNA Polymerases

Three distinct forms of RNA polymerase have been described in yeast and in other eukaryotes. The yeast polymerases have been characterised biochemically and functionally. Each form has been purified and is shown to exhibit distinct properties and polypeptide composition (reviewed by Sentenac and Hall, 1982). As with higher eukaryotic RNA polymerases, yeast polymerases can be differentiated *in vitro* by their sensitivity to the mycotoxin alpha-amanitin (intact yeast cells are not permeable to alpha-amanitin). The B-form (II) is highly sensitive and the C-form (III) is relatively insensitive.

RNA polymerase III was clearly shown to initiate and terminate *in vitro* transcription of the 5S rRNA gene (Tekamp *et al.*, 1979) and tRNA genes (Klekamp and Weil, 1982). Transcription of these genes accounts for nearly 30% of total cellular mRNA. RNA polymerase I transcribes large rRNAs (aproximatly 70% of total cellular RNA). RNA polymerase II sythesises all RNA which serves as mRNA and other small RNA (Tekamp *et al.*, 1979). As more information pertaining to the structures of promoters recognised by RNA polymerases II and III becomes available, it is apparent that the mechanisms of transcriptional initiation by both polymerases are similar (Reviewed by Folk, 1988). It is therefore possible that a gene may be transcribed by both RNA polymerase II and III.

The functional relationship between the 9 to 14 polypeptides of eukaryotic RNA polymerases is not yet known. However, an unusual structural feature of an RNA polymerase II subunit (RPO21) has been implicated in the interaction of this polymerase with other transcriptional factors. RPO21, the largest subunit of RNA polymerase II, has an unusual heptapeptide sequence, tandemly repeated

26 times at its C-terminus (Allison et al., 1985), which forms a proteolytically sensitive "tail". This amino acid sequence, of consensus PRO THR SER PRO SER TYR SER, is conserved amongst eukaryotic RNA polymerase II sequences (Allison et al., 1985; Corden et al., 1985). It is not found at the C-terminus of the homologous yeast RNA polymerase III subunit, RPO31. The importance of this tail for correct polymerase function was illustrated by introducing deletion mutants of RPO21 into a rpo21 yeast strain. Only yeast synthesising RPO21 with 10 or more heptapeptide repeats were found to be viable (Nonet et al., 1987). Allison et al. (1985) illustrated that the yeast "tail" could be replaced with the longer homologous hamster RNA polymerase II tail without changing yeast phenotype. However, replacement with the more divergent D.melanogaster tail, resulted in a non-functional hybrid polymerase. A direct interaction between the repeat domain and transcription factors has been implicated from experiments in which variations in tail length were shown to suppress mutations in the transcriptional activator protein, GAL4. In a yeast strain in which the number of heptapeptide repeats had been increased, GAL4 deletion mutants were shown to activate transcription to a higher rate than in cells containing a wild type RPO21 polypeptide (Allison and Ingles, 1988).
As the repeat domain is not present in RNA polymerase III, activation of transcription by this polymerase may be dependent upon contacts distinct from those formed by RNA polymerase II.

1.2.2 Yeast Promoter Elements

For the purposes of this study, a yeast "promoter element" shall be considered as being any DNA element which contributes, positively or negatively to transcription, from a cis location relative to the RNA

initiation site (I-site). This definition may at first appear rather wide ranging as this includes elements which may be situated at large distances from the I-site. However, it is now almost impossible to draw a functional distinction between regulatory sequences and some promoter elements. Nevertheless, yeast promoter elements can, to a degree, be distinguished as either determining transcriptional start sites (I-sites), or contributing to the rate of initiation.

Those elements which have been implicated in I-site determination will be discussed at length in Chapter 3. Briefly, the mechanism of I-site determination in yeast appears to be distinct from that operating in bacteria and mammalian cells. As in higher eukaryotic genes, an element similar to the E.coli TATA box apparently influences I-sites. However, the relationship between the TATA element and the I-site is less constrained in yeast than in higher eukaryotic genes. Therefore, it is likely that additional specific I-site elements exist in yeast promoters.

Yeast transcription is influenced by elements which are located at distances as far as 2 kb from the mRNA I-site. These control elements are generally located upstream of the I-site in RNA polymerase II promoters (with the exception of the subject matter of this thesis) and will not function when placed at a downstream site (Guarente and Hour, 1984a; Struhl, 1984). Exceptions to this general finding are the mating type repressor (Miller et al., 1985) and Ty promoter elements (Fulton et al., 1988). An RNA polymerase I promoter element has also been shown to function at both an upstream and a downstream site (Elion and Warner, 1986). Upstream elements frequently comprise short DNA sequences, typically 10 to 30 bp in length, which are required for normal transcription rates and to mediate regulation of gene expression. Sets of genes which are co-ordinately

regulated, typically possess elements which are similar in sequence. Generally, upstream elements can be classified as upstream activation sequences (UAS) or upstream repression sequences (URS).

The role of UASs in regulation of transcription was clearly demonstrated by constructing hybrid promoters. The UAS from the galactose inducible GAl1/GAL10 divergent promoter (UAS_g) was shown to confer galactose inducibility upon the CYC1 TATA element and I-site element. (Guarente et al., 1982).

Complex, multifactorial regulation of transcription is mediated by tandem, distinctly regulated UASs. For example, CYC1 is subject to regulation by heme, carbon source and oxygen. In the absence of oxygen or heme, both UAS1 and UAS2 fail to activate transcription. Transcription is activated on glucose by UAS1, provided heme is present. On lactate, activation by UAS1 is derepressed 10-fold and UAS2 is derepressed 100-fold (Guarente and Mason, 1983; Guarente et al., 1984b).

Negative control is mediated by the presence of a URS (or operator) within the 5'-flanking regions of a number of yeast genes such as haploid specific HO (Miller et al., 1985), glycolytic enzyme isogene ENO1 (Cohen et al., 1987) and the catabolite repressed GAL1 and GAL10 (Struhl, 1985).

Elements which influence the transcription of yeast genes may also occur downstream of the I-site. Sequences required for the maximal expression of the phosphoglycerate kinase gene (PGK1) are apparently located within the gene coding region (Mellor et al., 1985; Chen et al., 1985). Similarly, the absence of the pyruvate kinase gene coding region diminishes expression from the PYK1 5'-flanking region (Purvis et al., 1987b). Evidence has been presented that these elements influence the rate of transcription (Mellor et al., 1987; Purvis et al., 1987b). Chapters 4 to 7 of this thesis outline a

study of transcriptional activation by the PYK1 coding region. The relevant literature will be discussed in detail in the introduction to Chapter 4.

The yeast retrotransposon Ty also contains elements within a transcribed region which contribute to wild-type levels of transcription. The TYA gene, which encodes the major structural proteins of the Ty- VLP's (Virus Like Particles), contains two elements which can activate transcription when placed upstream of a heterologous gene. One of these elements, which functions in an orientation dependent manner, confers cell-type regulation, that is 5 to 20-fold less expression is observed in a/alpha diploids, but does not exhibit sequence similarity to cell-type regulatory elements. This element may be required for transcription of the TYA element in a cell-type specific manner. The second element located within the Ty coding region, may be responsible for the synthesis of an mRNA initiating within the TYA gene. (Rathjen et al., 1987; Company and Errede, 1987; Fulton et al.; 1988). Therefore this second element may not activate transcription initiation from the upstream site.

Downstream elements contributing to transcription at an upstream site may be more common in higher eukaryotic genes. For example, human HPRT and mouse growth hormone genes contain intragenic sequences which determine tissue specific expression (Stout et al., 1985). Likewise, the immunoglobulin heavy chain genes contain enhancer sequences within the second intron of the activated gene (Queen and Baltimore, 1983) and there is clear evidence that the thymidine kinase genes of mouse, human and Herpes simplex virus all contain elements within the coding region which enhance expression (Kosche et al., 1985; Hofbauer et al., 1987; Koltunow et al., 1987).

The best studied examples of "downstream elements" are those in 5S RNA and tRNA genes. Transcription of these genes by RNA polymerase III is dependent upon internal control regions (ICR) which comprise of two discontinuous "boxes" (Reviewed by Geiduschek and Tocchini-Valentini, 1988). It is likely that elements located within the 5'-flanking regions are also required for normal transcription of these genes.

1.2.3 Yeast Regulatory Proteins

DNA control elements in yeast almost exclusively influence transcription by acting as binding sites for specific regulatory proteins. Eukaryotic transcriptional control is based, to a large extent, upon the interplay of specific DNA binding proteins and other components of the transcriptional apparatus, such as RNA polymerase. In higher eukaryotes, these proteins have been defined biochemically by their DNA binding activities. In contrast, yeast regulatory proteins were initially defined through mutations which affected the transcription of specific genes or sets of genes.

1.2.3.1 DNA Binding Domains of Regulatory Proteins

Regulatory proteins have a number of functional domains such that DNA binding function is distinct from transcriptional activation function. Consider the example of the GAL4 protein. This protein, which is necessary for the transcription of the family of genes required for galactose metabolism, binds to a 17 bp sequence of imperfect dyad symmetry situated in the upstream regions of GAL genes (Bram and Kornberg, 1985; Giniger et al., 1985). The GAL4 protein binds DNA via a protein domain localised at the N-terminal 73 amino acids (Keegen et al., 1986). Binding of GAL4 is necessary but not

sufficient for activation (Brent and Ptashne, 1985). Transcriptional activation is mediated by sequences localised at the carboxyl terminus (Ma and Ptashne, 1987b and 1987c). Likewise, the GCN4 protein, which activates the transcription of amino acid biosynthetic genes, has separate protein domains associated with DNA binding and transcriptional activation (Hope and Struhl, 1986).

DNA binding domains of eukaryotic regulatory proteins tend to exhibit one of three motifs. The first DNA binding motif to be described was the helix-turn-helix (HTH). The three-dimensional structures of bacteriophage lambda CRO and CI proteins and the E.coli CAP protein all revealed a distinctive succession of two alpha helices separated by a relatively sharp beta turn (Steitz et al., 1982). Collated evidence is consistent with the view that the second of these helices docks with the major groove of the DNA double helix whilst the first helix holds the second in the correct position (reviewed by Pabo and Sauer, 1984; Johnston and McKnight, 1989). Similarity has been observed in the primary amino acid sequences of these helices. This similarity facilitated the identification of the HTH motif in eukaryotic regulatory proteins. The two products of the yeast mating type locus, the MATa1 and MAT-alpha2 repressor proteins, which specifically bind DNA, exhibit a distribution of hydrophobic residues consistent with an HTH (Laughton and Scott, 1984). MATa1 and MAT-alpha2 share homology with eukaryotic homeotic gene products (Shepherd et al., 1984; Laughton and Scott, 1984) which also appear to contain HTH motifs.

The second DNA binding motif found in some yeast regulatory proteins is the "zinc finger". This motif is characterised by the sequential and ordered occurrence of cysteine and histidine residues. It is postulated that four of these residues, two cysteines and two histidines (cys2 his2), sequester a single zinc ion and that the

intervening 12 to 14 amino acid sequences loop out to form a structure which could interact directly with DNA (Fairall et al., 1986; Rhodes and Klug, 1986). This arrangement of cysteines and histidines is observed in the yeast ADR1 and SW15 proteins (Hartshorne et al., 1986; Stillman et al., 1988), and a related arrangement of cysteines (cys2 cys2) is present in GAL4, HAP1, PPR1, and ARGRII (Reviewed by Johnston and McKnight, 1989). There is genetic and biochemical evidence which suggest that GAL4 requires zinc ions in order to bind to DNA (Johnston, 1987b). In addition, of a number of mutations which inactivate GAL4, half occur within the potential zinc finger forming domain (Johnston and Dover, 1987). The zinc finger motif is not confined to regulators of RNA polymerase II transcription. The motif was originally observed within the eukaryotic transcription factor IIIA (TFIIIA) which is required for the transcription of 5S RNA by RNA polymerase III. Xenopus laevis TFIIIA contains 9 zinc finger repeats which interact with 50 bp of DNA (Engelke et al., 1980).

More recently, a third motif has been identified which may be characteristic of some DNA binding proteins. Primary amino acid sequence comparison of the products of several oncogenes, FOS, MYC and JUN, the yeast protein GCN4 and latterly the rat liver nuclear protein C/EBP revealed a commmon structural motif. Each protein contained strict heptad repeats of leucine residues and a high density of oppositely charged amino acids juxtaposed such that ion pairing could result in a stable helical domain (Landschulz et al., 1988). It is proposed that the leucines extending from one helix interdigitate with those of an analogous helix of a second polypeptide, forming an interlock termed a "leucine zipper". This structure is distinct from the HTH and zinc finger in that it is not thought to make contact with the DNA itself, rather it stabilises dimerisation of the DNA

binding protein. This dimerisation is necessary for specific interactions between basic regions of the protein and the DNA-recognition sequence.

1.2.3.2 Further Consideration on Protein/DNA Recognition

Sequence specificity of regulatory proteins is central to differential gene expression. The specificity of a protein for a particular binding site may not entirely be a consequence of the sequence of amino acid residues within the binding motifs described above. For example, replacing the entire zinc finger region of GAL4 with the analogous zinc finger region from PPR1, a *trans*-activator of pyrimidine biosynthetic genes, does not alter the specificity of the chimeric protein for UAS_G. Specific binding of GAL4 requires a 14 amino acid motif adjacent to the zinc finger region (Corton and Johnston, 1989).
The complexity of regulatory responses in yeast appears to have generated some intricate features. For example, multiple proteins can bind to identical or similar DNA sequences, suggesting that competition for binding sites may play an important role in the regulation of particular genes. UAS1 of the *CYC1* promoter interacts with both HAP1 and a distinct protein RC2, the role of which is unclear. The binding of these proteins appears to be mutually exclusive (Pfeifer *et al.*, 1987). Similarly, there is evidence that yeast contains a number of proteins which share DNA binding properties with GCN4 (Moye-Rowley *et al.*, 1989).
Single regulatory proteins may also bind at distinctly different binding sites. HAP1 for example, may bind to sequences within the *CYC7* UAS and a *CYC1* UAS which are not similar. Moreover, a HAP1 mutant fails to bind to *CYC1* UAS but retains unaltered specificity for the *CYC7* UAS (Verdiere *et al.*, 1986). It has been proposed that the HAP1 protein achieves promiscuity by the formation of

alternative and mutually exclusive zinc fingers (Verdiere et al., 1988).
In some cases specific binding requires the interaction of two or more distinct proteins. For example, HAP2 and HAP3 proteins interact and bind to the CCAAT box within the UAS2 of the CYC1 gene (Oleson et al., 1987; Hahn and Guarente, 1988). Recently the product of another locus HAP4 has been shown to be required for the binding of HAP2/3 to DNA in vitro (Forsburg and Guarente, 1989). There is also evidence that the MAT-alpha1 activator of alpha-specific genes will only bind to the appropriate UAS in the presence of a second protein factor, PRTF (Bender and Sprague, 1987). Co-operative interactions also occur between MAT-alpha2 and a second factor (Keleher et al., 1988), and between MAT-alpha2 and MATa1 in diploid cells (Millar et al., 1985).

1.2.3.3 Mechanism of Transcriptional Activation By Regulatory Proteins

As previously stated, binding of regulatory proteins to DNA is necessary but not sufficient for the activation of transcription (Brent and Ptashne, 1985; Hope and Struhl, 1986).
Although binding of a protein activator to DNA does not in itself activate transcription, activation of transcription through such proteins is associated with changes in chromatin structure. Almer et al. (1986) demonstrated that induction of the PHO5 gene results in the selective removal of two nucleosomes upstream and two nucleosomes downstream of the PHO5 UAS. In this way 600 bp of the PHO5 5'-flanking region becomes accessible. One could envisage that such nucleosome clearing would facilitate transcription by allowing RNA polymerase II and other transcription factors access to the PHO5 I-site.

Very little is known about how regulatory proteins bring about transcriptional activation. It appears however, that the activation function of these proteins is associated with short acidic regions and not with a specific amino acid sequence. This conclusion is based upon several observations. First, the activation regions of GCN4 and GAL4 are acidic but have no other similarities (Hope and Struhl, 1986; Ma and Ptashne, 1987a and 1987b). Second, dissimilar parts of the GCN4 protein are capable of activating transcription (Hope and Struhl, 1986). Third, mutations which increase or decrease the activation activity of GAL4 are almost always associated with an increase or decrease in the negative charge of the protein (Gill and Ptashne, 1987). Fourth, random E.coli polypeptide fragments which can complement deletions of an activator domain are acidic (Ma and Ptashne, 1987a). However, acidic character may not be solely responsible for activation since the extent of activation by derivatives of GCN4 only moderately correlates with net negative charge (Hope et al., 1988). Specific higher structure may be obligate and two have been proposed; an amphipathic helix (Giniger and Ptashne, 1987) and a dimer of intertwined alpha-helices (Hope et al., 1988).

The proposed structures of activation domains do not, in themselves, reveal how transcription by RNA polymerase is activated. Presumably, activator domains make contacts with either transcription factors bound close to the I-site, such as TATA box binding factor, or with RNA polymerase itself. One may envisage that acidic regions could make contacts with the heptapeptide repeat "tail" in the RPO21 subunit. Looping out of the intervening DNA could facilitate such contacts over large DNA duplex distances (Reviewed by Ptashne 1986).

The interactions between DNA-bound activator and other transcription factors may be regulated. For example, the

GAL80 protein is known to block the activation activity of GAL4 in cells grown on non-galactose medium (Torchia et al., 1984; Yocum and Johnston, 1984) perhaps via a direct protein/protein interaction close to the GAL4 acidic region (Lue et al., 1987). GAL80 does not inhibit activation by preventing GAL4 from binding to the UAS_G. In fact, GAL4 remains bound to UAS_G under non-inducing conditions (Lohr and Hopper, 1985; Giniger et al.,1985; Selleck and Majors, 1987a and b).

Activator proteins may also be regulated by phosphorylation resulting in the modification of interactions between the DNA bound protein and other transcription factors. Sorger and Pelham (1988) have hypothesised that the yeast Heat Shock Transcription Factor (HSF), which mediates the transcriptional activation of heat shock specific genes, is dependent upon phosphorylation for function. Phosphorylation of HSF may create an acidic activation domain analogous to that described for GAL4 and GCN4. Recently, phosphorylation has been implicated in modulating the activation activity of the GAL4 protein. Under galactose inducing conditions a phosphorylated form of GAL4 is present in yeast. The addition of glucose rapidly alters GAL4 protein in vivo in a manner consistent with dephosphorylation. The glucose affect does not require GAL80 protein, however GAL80 deletion results in the appearance of the phosphorylated form of GAL4 in the absence of galactose (Mylin et al., 1989). These results contrast with those of Cherry et al. (1989) who propose that phosphorylation of ADR1, a transcriptional activator, causes inactivation.

Repression of transcription in yeast may occur either by steric hindrance or by making contact with factors or surfaces normally contacted by activator domains. Consistent with this second explanation is the repression activity of MAT-alpha2 which has been shown to be

distinct from the DNA binding activity (Hall and Johnston, 1987).

Superficially, the mechanisms of yeast transcriptional regulation appear to be distinctly different from those of higher eukaryotic cells. The complexity of I-site determination in yeast (reviewed in Chapter 3) and the upstream location dependence of yeast activation sites, suggest that the mechanisms of transcription in yeast and other eukaryotes have diverged somewhat. However, this may not be the case. Considerable evidence now exists which supports the view that transcriptional mechanisms are highly conserved. It is now apparent that a number of yeast transcription proteins can also function in mammalian cells and can complement the function of mammalian transcription factors in *in vitro* mammalian transcription systems. The GAL4 activator can function in mammalian cells (Kakidani and Ptashne, 1987) and in *Drosophila* (Fisher *et al.*, 1988) when a binding site, UAS_G, is inserted upstream of a reporter gene. In mammalian cells, the GAL4 protein has been shown to function from a site downstream of the I-site which contrasts with the upstream requirement in yeast (Webster *et al.*, 1988). Compatibility is also observed when mammalian DNA binding proteins are expressed in yeast. For example, the human oestrogen receptor has been shown to direct transcription in yeast, imposing oestrogen regulation on a reporter gene (Metzger *et al.*, 1988) and FOS protein activates transcription in yeast (Lech *et al.*, 1988). Structural conservation between the yeast HAP2 and HAP3 proteins and the human CCAAT-binding protein is such that heterologous subunits are functionally interchangeable. Complementarity is not restricted to activator proteins. A yeast protein has been shown to exhibit the properties of the mammalian transcription factor IID, otherwise known as TATA box binding factor (Cavallini *et al.*, 1988; Buratowski *et*

al., 1988; Horikoshi et al., 1988 and 1989). Similarities between yeast and mammalian factors have uncovered previously unknown in vivo activities. For example, the similarity of the GCN4 DNA binding domain to domains of JUN and FOS oncogene products has led to experiments which demonstrate that these proteins have DNA binding activity (reviewed by Kouzarides and Ziff, 1989).

In summary, transcriptional regulation in yeast is dependent to a great extent upon the specific interaction of DNA elements with regulatory proteins and the subsequent interaction of these proteins with RNA polymerase. Genes which are co-ordinately expressed tend to be regulated by identical proteins. Most genes appear to be under the regulatory influence of a number of proteins. The exact molecular interplay which brings about activation or repression of transcription is, for the most part, unknown. The mechanisms of yeast and mammalian transcriptional activation appear to be highly conserved.

1.3 Regulation of Glycolytic Enzyme Genes in Yeast

Glycolysis is the sequence of reactions which converts glucose to pyruvate with the concomitant production of ATP. In fermenting yeast, pyruvate is subsequently transformed to ethanol and there is a net yield of two ATP molecules for every glucose molecule catabolised. The importance of the metabolic pathway in yeast is reflected in the fact that the enzymes catalysing glycolysis comprise 30 to 60% of total soluble protein. Recent interest in the expression of glycolytic genes is due to the desire to use highly transcribing control regions from these genes to direct high level synthesis of

heterologous proteins in yeast. Studying glycolysis will also reveal how the flux through a major metabolic pathway is controlled and how a cell can co-ordinately respond to specific environmental stimuli.

Maitra and Lobo (1971) demonstrated that the levels of many glycolytic enzymes could be induced by glucose after growth upon acetate-supplemented rich medium. Although some yeast strains do not exhibit such induction (Fraenkel, 1982; Baker, 1986), it is clear that the expression of some glycolytic genes, in certain strains and under particular growth conditions, is higher when cells are grown upon glucose as opposed to a non-fermentable carbon source (McAlister and Holland, 1982; Cohen et al., 1986; Nishizawa et al., 1989).

Mutants in glycolysis have been isolated by virtue of their ability to grow on pyruvate but not on glucose (Clifton et al., 1978; Ciriacy and Breitenbach, 1979). These mutants have facilitated the cloning of glycolytic enzyme genes by functional complementation and this in turn has prompted a number of studies, including this report, on the requirements for transcription and regulation of these genes.

Co-ordinate transcriptional regulation of glycolytic enzyme genes is purported by the existence of a mutation, gcr1, which results in a reduction in glycolytic enzyme levels to 1 to 10% of those found in wild type cells grown under gluconeogenic conditions (Clifton et al., 1978). This mutation was subsequently shown to affect mRNA levels (Clifton and Fraenkel, 1981). The GCR1 gene was shown to regulate the expression of many glycolytic genes (Baker, 1986). This regulation occurs at the transcriptional level for enolase (ENO) and glyceraldehyde 3-phosphate dehydrogenase (GDH) genes (Holland et al., 1987). The GCR1 gene has been cloned and sequenced (Baker, 1986; Holland et al., 1987).

The regulation of the two genes coding for yeast enolase

has been studied in detail. The two genes, ENO1 and ENO2, are expressed differently in cells grown on glycolytic (fermentative) or gluconeogenic (non-fermentative) carbon sources. The steady-state mRNA abundance of ENO2 mRNA, and the derived polypeptide, is greater than 20-fold higher in cells grown on glucose than on glycerol/ethanol. In contrast, the levels of ENO1 mRNA and the encoded polypeptide are identical on both carbon sources. (McAlister and Holland, 1982). Glucose dependent induction of ENO2 is mediated by a UAS located 460 bp upstream of the I-site (Cohen et al., 1986). Despite the fact that the ENO1 gene is not induced by growth on glucose it too is transcribed by virtue of a region containing two UASs which can mediate glucose regulation (Cohen et al., 1987). The constitutive nature of ENO1 is a consequence of a URS, located between the UAS and the I-site, which prevents glucose dependent induction.

Transcription of ENO2 is dependent upon the integrity of one of two UASs. The distal UAS is required for GCR1 regulated expression, although GCR1 protein does not bind directly to the DNA. The proximal UAS is known to interact with the protein RAP1 (Holland et al., 1989). RAP1 (Repressor/Activator binding Protein) was first characterised as a regulator of mating type, exhibiting both repressing and activating properties (Shore and Nasmyth, 1987; Brand et al., 1987). Subsequently, the protein has been shown to bind to UASs located within the upstream regions of PGK1 and PYK1 (Chambers et al., 1989). It is likely that the transcriptional activator TUF1, essential for the transcription of genes coding for components of the translational apparatus, is identical to RAP1. Therefore, the transcription of a number of glycolytic genes is dependent upon a factor which activates the transcription of a great range of genes. It is not clear how RAP1 mediated expression is related to GCR1 transcriptional control.

Since a large part of this study is concerned with the transcription of the pyruvate kinase gene (PYK1) consideration will now be given to the PYK enzyme and to its expression.

PYK phosphorylates ADP to ATP whilst converting phosphoenol pyruvate (PEP) to pyruvate. The in vitro properties of this enzyme, and of phosphofructokinase (PFK), are consistent with a role in physiological control (Reviewed by Fraenkel, 1982). However, recent studies are not consistent with this view (Schaaff et al., 1989). This particular publication will be discussed further in Chapter 7.

In yeast, pyk1 mutants were found to accumulate PEP and earlier glycolytic intermediates and to be rapidly depleted of ATP on addition of glucose (Ciriacy and Breitenbach, 1979). pyk1 mutants are the most commonly isolated glycolytic mutant (Fraenkel, 1982).

There is only a single PYK homotetrameric enzyme and a single PYK1 gene in yeast although there are four isoenzymes in adult vertebrate tissue. All four enzymes exhibit unique kinetic and structural properties, although the liver form and erythrocyte form are specified by a single gene. It has been proposed that these two forms of the enzyme result from differential tissue-specific pre-mRNA splicing (Marie et al., 1981).

Relatively little is known about the expression of yeast PYK1. PYK can potentially be induced by glucose (Maitra and Lobo, 1971; Nishizawa et al., 1989) but the mechanism by which this is achieved has not been investigated thoroughly. By analogy with ENO2, one may expect that induction occurs at the level of transcription. As with ENO2, transcription of PYK1 requires GCR1 (Baker, 1986) and the PYK1 upstream region contains a sequence which binds RAP1 (Chambers et al., 1989). Deletion analysis of the PYK1 5'-flanking region has revealed 3 regulatory regions: UAS1, between 600 and 700 bp upstream of the

coding region, UAS2, 800 bp upstream and a URS, approximately 350 bp upstream (Nishizawa et al., 1989) . These elements be discussed further in Chapter 7.

High levels of expression of, at least, some glycolytic enzyme genes, including PYK1, may require sequences outwith the 5' and 3' flanking regions (Mellor et al., 1987; Purvis et al., 1987). These particular publications are highly relavent to this study and will be discussed in full in Chapter 4.
The requirements for the initiation of transcription of glycolytic genes, in particular PGK1, will be discussed in Chapter 3.

1.4 Aims of This Study

There were two quite distinct aims in carrying out the work described in this study.

(a) To define the minimal cis-acting sequence requirements for high levels of trancription of the PGK1 gene and in particular to define I-site specific sequences in this gene.

(b) To delimit the sequences within the yeast PYK1 coding region which are required for maximal transcription and to determine by what mechanism these sequences function.

The background information pertinent to these aims is discussed at the begining of the relevant chapters. Chapter 3, which considers the first of these aims, is unlinked to the following three chapters which describe experiments undertaken in an attempt to achieve the second aim. The information presented in Chapters 4, 5

and 6 is collated and discussed in the Chapter 7.

CHAPTER 2

MATERIALS AND METHODS

2.1 CHEMICALS AND ENZYMES

2.1.1 Enzymes

DNA Restriction and modification enzymes were obtained from BRL, New England Biolabs, BCL, and Pharmacia.
Proteinase K, lysozyme, and B-glucuronidase were obtained from Sigma.

2.1.2 Chemicals and Membranes

All radiochemicals were purchased from NEN.
General laboratory chemicals were obtained from Sigma, Formachem, BDH Chemicals, Koch-Light, Pharmacia, Bio-Rad Laboratories, May and Baker, BCL, and Aldrich.
Nitrocellulose (0.45um) was obtained from Schleicher & Schuell.
PEG (3350 Mol. Wt.) and Sephadex-G50 were purchased from Sigma.
Biogel (100-200 mesh) and Dowex 50w-x8 were obtained from Bio-Rad Laboratories.
DE81 paper was obtained from Whatman Paper Ltd.
X-ray film was obtained from Kodak.
Geneclean kits were obtained from Stratech Scientific.
1,10-phenanthroline and neocuprone (2,9-Dimethyl-1,10-phenanthroline) were obtained from Sigma.

2.1.3 Growth Media and Additives

All media ingredients were supplied by Oxoid with the exception of Bacto-Agar which was obtained from Difco.
Amino acids and antibiotics were obtained from Sigma.
X-Gal (5-Bromo-4-chloro-3-indolyl-B-galactoside) and IPTG (Isopropylthio-B-galactoside) were obtained from BRL.

2.2 ORGANISMS AND GROWTH CONDITIONS

The following is an outline of growth conditions suitable for yeast and bacterial strains utilised in this study. All growth media was sterilised, without additives, for 15mins at 120^{o}C (% refers to weight/volume).

2.2.1 Escherichia coli :

Strains ; C1400: Ed8767, supE, supF, hsd5, met-, recA
JM103: deltalacpro, supE, thi, F' trad36, proAB, lacI^{q} ZdeltaM15

L-Broth : 1% Tryptone, 0.5% Yeast extract, 0.1% Glucose, 0.5% NaCl, pH 7.

Solid media : As L-broth but without glucose and addition of 2% Bacto-Agar to L-broth prior to autoclaving. In order to maintain selection of plasmids, ampicillin was added at a concentration of 50-100ug/ml.

D&M Medium : 7g K_2HPO_4, 2g KH_2PO_4, 4g NH_4SO_4, 0.25g trisodium citrate, 0.1 M $MgSO_4.7H_2O$, 17.5g agar, made up to one litre with 1.2% thiamine and 0.2% glucose in distilled water.

E.coli C1400 strain cultures were shaken at 37^{o}C for 3-4 hours in 2.5ml of L-Broth with an antibiotic supplement; 1.5 ml of this culture was then added to 300ul of glycerol and, after mixing, the strains were stored at -70^{o}C indefinitely. In order to use such strains, they were streaked out onto selective L-agar plates and incubated at 30^{o}C overnight. These plates could be stored at 4^{o}C for short periods of time. Liquid cultures were obtained by inoculating 5 ml of selective L-broth with a single colony from a storage plate and

shaking at 37°C for 6-8 hours. Often, 1 ml of this culture was added to 100-400ml of prewarmed, selective L-broth and shaken overnight at 37°C.

E.coli strain JM103 was invariably grown on D & M medium with appropriate selection. Screening of ligation products for insert containing pUC19 plasmids was carried out by transforming JM103 cells and plating onto D&M medium containing the substrate X-Gal (5-Bromo-4-chloro-3-indolyl-B-galactoside) and the inducer of B-galactosidase, IPTG (isopropylthio-B galactoside). Clones containing inserts are generally white; clones lacking inserts are blue. X-gal was stored at a concentration of 20 mg/ml in dimethylformamide (DMF) at -20°C and added to D&M agar plates to a concentration of 20 ug/ml.

2.2.2 Saccharomyces cerevisiae :

Strain ; DBY746: alpha, his3 delta, ura3-52, leu2-3, leu2-112, trp1-289

Complete medium (YPG) : Glucose 2%, Bactopeptone 2%, Yeast extract 1%

Minimal/Selective medium (GYNB) : Glucose 2%, Yeast Nitrogen Base (amino acid free) 0.65%, Amino Acids 50 ug/ml (where required by auxotrophic mutant).

LYNB : As GYNB replacing 2% Glucose with 2% Lactose

Solid media : 2% Bacto-Agar was added to medium before autoclaving.

Pre-sporulation medium : 0.8% Bacto yeast extract, 0.3% Bacto-peptone, 10% Glucose, 2% Bacto-agar.

Sporulation medium ; 1% Potassium acetate, 0.1% Bacto-yeast extract, 0.05% Glucose, 2% Bacto-agar.

B-Galactosidase Assay plates ; 0.1M KH_2PO_4, 0.015M $(NH_4)_2SO_4$, 0.075M KOH, 0.8mM $MgSO_4$, 2uM $Fe_2(SO_4)_3$, 2% Glucose, 40ug/ml thiamine, 40ug/ml pyridoxine, 40ug/ml pantothenic acid, 200ug/ml inositol, 2ug/ml biotin, 2%

agar and X-gal to a final concentration of 40ug/ml.

Auxotrophic yeast strains were grown for three days on GYNB plates with appropriate amino acid supplementation. The plates were sealed with parafilm and stored for up to three months at 4°C.
Liquid cultures were grown at 30°C with vigorous shaking (250 rpm). To obtain mid-logarithmic culture, a small volume (usually 20-100ul) of a stationary rich or minimal culture was added to the appropriate medium (prewarmed to 30°C) and grown at 30°C for 12-16 hours with shaking. Growth rate was determined by measuring the optical density (at 600nm) of 1ml samples taken every hour throughout logarithmic growth phase. Each culture was screened for yeast and bacterial contamination. To this end, culture was streaked onto selective GYNB agar in order to confirm the presence of auxotrophic markers. In addition, all cultures were examined microscopically to ensure no bacterial contamination was present.

2.2.3 Random spore analysis
This method allows the selection of haploid spores from a sporulated, or incompletely sporulated, culture. Strains with the desired genetic markers may be selected without the need to carry out ascus dissection. Diploids were streaked out, at high concentration on pre-sporulation medium, and grown at 30°C for 72 hours. These cells were then streaked at high concentration onto sporulation medium. Sporulation plates were incubated at 30°C for 48 hours. A small amount of sporulated culture was transferred, with a sterile toothpick, into 1ml 0.2 M potassium acetate buffer (pH 5.5); to this, was added 1ml of diethyl ether. The mixture was vortexed briefly and the phases allowed to separate. 0.1ml of the aqueous layer was spread onto the appropriate selective minimal medium plates and these selective plates were incubated

for 48 hours at 30°C. Colonies which appeared on these plates were picked and their genetic markers determined by replica plating.

2.2.4 Yeast B-galactosidase activity plate assay
Cells were patched onto assay plates and incubated for 72 hours at 30°C. Strains exhibiting B-galactosidase activity were blue and strains without activity were white.

2.3 TRANSFORMATION OF E.COLI

An overnight culture of the recipient strain was diluted 1 in 100 into L-broth and grown to a density of approximately 10^8 cells/ml (about 90 min). The cells were harvested at 12000 g, 4°C for 5 min and resuspended in 50ml of cold 50mM $CaCl_2$. The cells were pelleted again, resuspended in 5ml of cold 50mM $CaCl_2$ and kept on ice for at least 15 min before use. 200ul aliquots of the competent cells were added to the plasmid DNA, mixed gently and left on ice for up to 30 min. These cells were heat shocked at 42°C for 5 min and then transferred to 1ml of L-broth and incubated at 37°C for 30 min. The cells were then plated out on selective medium and incubated at 37°C for 16 hours. Unused competent cells were stored at -70°C in 20% (v/v) glycerol.

2.4 TRANSFORMATION OF S.CEREVISIAE

SE Solution : 1M Sorbitol, 25mM EDTA, pH 8.0
SEN Solution : 1.2M Sorbitol, 10mM EDTA, 100mM NaCitrate, pH 5.8.
SC Solution : 1.2M Sorbitol, 10mM $CaCl_2$.
CT Solution : 10mM$CaCl_2$, 10mM Tris.HCl, pH 7.5.
SY Solution : 2 volumes of 1.8M Sorbitol
1 volume 3 x YPG (2.2.2).

Bottom agar : 2% Glucose, 0.65% YNB (without amino acids), 2% Bacto-agar, 1.2M Sorbitol (amino acids added to 50ug/ml after autoclaving).
Top agar : 2% glucose, 0.65% YNB (without amino acids), 3% Bacto agar, 1.2M Sorbitol (amino acids added to 50ug/ml after melting, agar kept at 48^{o}C until required).

100ml of culture in mid-logarithmic growth (A_{600} = 0.6-0.7) was used for each transformation.
Cells were harvested by centrifugation at 960 g for 5 min at room temperature. The cells were resuspended in 50ml of SE solution containing 50mM DTT (added immediately before use and filter sterilised) and incubated with gentle shaking for 20 min at 30^{o}C. The cells were reharvested by centrifugation at 960 g for 5 min at room temperature. The cell pellet was resuspended in 50ml of SEN solution containing 2-3% (v/v) B-glucuronidase (prepared immediately before use and filter sterilised) and incubated with gentle shaking for 15-30 min at 30^{o}C. Spheroplast formation was monitored by light microscopy and was judged to be adequate when over 90% of cells lysed on addition of 35%(w/v)-lauryl sarcosine. The spheroplasts were harvested by centrifugation at 960 g for 3.5 min at room temperature and were washed twice in 50ml of 1.2M sorbitol, centrifuging each time at 960 g for 3.5 min at room temperature. The washed spheroplasts were resuspended very gently in 100ul of SC solution to give a thick suspension. DNA (5-15ug, in 10-15ul) was added to 50ul of spheroplasts in sterile eppendorf tubes and these were left for 15 min at room temperature. 500ul of 20% PEG in CT solution (made up freshly and filter sterilised) was then added and left for approximately 45 sec before centrifuging at 12000 g in an Eppendorf microfuge for 1 sec. The PEG was carefully removed, then the spheroplasts were resuspended in 100ul of solution SY and then incubated at 30^{o}C for 30 min. The spheroplasts

were diluted to 1ml by the addition of 1.2M Sorbitol, and 20ul of this suspension was added to 10ml of top agar (melted and cooled to 48^{o}C) before pouring onto bottom agar plates. The plates were then incubated at 30^{o}C and colonies usually appeared 4-7 days after transformation. Putative transformants were then streaked onto selective plates (GYNB agar containing the appropriate amino acids) to check that they exhibited the predicted auxotrophic markers and were not due to contamination.

2.5 ISOLATION OF RNA FROM S.CEREVISIAE

Lysis Buffer : 0.1 M LiCl, 0.1M Tris-HCl, pH7.5, 0.01M DTT (added immediately before use).
Phenol : All phenol contained 0.1% (w/v) 8-hydroxyquinoline and was buffered against 0.5M Tris-HCl (pH 8.0).

2.5.1 Procedure

RNA was isolated according to the method of Lindquist (1981). The time taken to prepare the RNA was minimised wherever possible to reduce the significance of any ribonuclease activity. 100ml of YPG culture in mid-logarithmic growth phase (A_{600} = 0.5-0.7 read against a YPG blank) was harvested by centrifugation at 960 g for 5 min at 4^{o}C. The pellet was resuspended in 5ml lysis buffer and the resulting solution transferred to RNase free tubes containing ; 7g glass beads (450-500um mesh), 1ml 10% (w/v) SDS, 2.5ml phenol, 2.5ml chloroform, and then vortexed continuously for 5 min, followed by centrifugation at 12000 g for 5 min at 4^{o}C. The aqueous phase was removed and added to 5ml phenol/chloroform (1:1), vortexed for 15 sec and centrifuged at 1200 g for 5 mins at 4^{o}C. Phenol/choroform extraction was repeated and the aqueous phase was added to two volumes of diethyl

ether and vortexed for 30 sec. A second ether extraction was performed, followed by precipitation of material in the aqueous phase by the addition of sodium acetate to a final concentration of 0.1 M and two volumes of ethanol. RNA prepared in this fashion was stored under ethanol. The integrity of the RNA was assessed by non-denaturing gel electrophoresis, the gel being stained with ethidium bromide and the RNA visualised via UV fluorescence.

2.5.2 Storage of cells for RNA preparations

For convenience harvested cells were occasionally stored at -20^{o}C with the addition of two volumes of ethanol to the culture. RNA was usually extracted within 3 days. Cells stored under ethanol were harvested by centrifugation at 9000 g at 4^{o}C for 5 min. The resulting pellet was subjected to the procedure outlined above.

2.6 ISOLATION OF DNA FROM E.COLI

Birnboim Doly I (BD I) : 50mM glucose, 25mM Tris-HCl (pH8),10mM EDTA. Lysozyme added to 1mg/ml immediately before use.

Birnboim Doly II (BD II) : 0.2M NaOH, 1% (w/v) SDS.

Birnboim Doly III (BD III) : 5M KOAc (pH4.8)

STET Buffer : 8% (w/v) sucrose, 5% (v/v) Triton X-1000, 50mM EDTA, 50mM Tris-HCl (pH8.0).

2.6.1 Birnboim Doly Procedure

DNA was isolated by the method of Birnboim and Doly (1979). A 200ml culture of stationary phase cells was harvested by centrifugation at 12000 g for 5min at 4^{o}C. The pellet was resuspended in 4ml of BD I solution and incubated on ice for 5 min. 8ml of BD II solution were added and left on ice for 5 min and then 6ml of BD III solution were added, gently mixed and left on ice for a further 5 min. The cell debris and most of the

chromosomal DNA was removed by centrifugation at 35000 g, for 5 min at 4°C. The plasmid DNA was then precipitated by the addition of an equal volume of isopropanol followed by centrifugation at 39200 g, for 15 min. This DNA was further purified by banding on a CsCl/EtBr gradient. The DNA was resuspended in 1ml of TE and added to 4.5g of CsCl dissolved in 3.5 ml of TE. 250ul of EtBr (10mg/ml) was added to the DNA/CsCl solution which was then injected into Beckman VTi65 ultracentrifuge tubes. The tubes were sealed and centrifuged in a Beckman VTi65 vertical rotor at 28900 g for over 3.5 hours at 20°C. Two bands were visible, a lower supercoiled plasmid band and an upper chromosomal and relaxed plasmid DNA band. Covalently closed circular plasmid DNA was purified away from EtBr and CsCl by passing the plasmid DNA over a composite column containing 1ml of the ionic exchange resin, Dowex 50W-x8 and 10 ml of the gel filtration medium, Biogel (100-200 mesh). Plasmid DNA was detected in the 1.5ml fractions collected from the column by dotting 1ul of each sample onto an EtBr containing agarose plate (1ug/ml EtBr, 1% (w/v) agarose) and viewing with UV.

2.6.2 Holmes and Quigley STET Procedure

The method of Holmes and Quigley (1981) was used for mini DNA preparations. 1.5ml of an overnight culture, containing the plasmid of interest, was harvested by centrifugation in a 1.5ml eppendorf tube and resuspended in 350ul of STET buffer. 25ul of STET buffer, containing lysozyme at a concentration of 10mg/ml, was added to the tube and vortexed briefly. The tube was then immediately placed in boiling water for 45 sec and centrifuged in a eppendorf microfuge for 15 min at 4°C. The pellet was removed from the tube with a toothpick and 40ul of 3M NaOAc and 400ul of isopropanol were added. The nucleic acid was pelleted by centrifugation for 5min at 4°C. The

pellet was washed once with 70% (v/v) ethanol and then dried by spinning in a vacuum drier for up to 10 min. 50-200ul of dH_2O was added and the tube was placed in a 90°C water bath for 1-5 min, resulting in dissolution of the pellet. This DNA was suitable for restriction enzyme digestion and other *in vitro* manipulations.

2.7 GEL ELECTROPHORESIS

TBE : 89mM Tris-HCl pH 8.3, 89mM Boric acid, 2.5mM EDTA.
TAE : 89mM Tris-HCl pH 8.3, 40mM NaAcetate, 2.5mM EDTA.
Sequencing gel loading buffer : 0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanol, 10mM Na_2EDTA, 95% (v/v) formamide.
Agarose gel loading buffer : 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol, 25% (w/v) ficoll, 0.5% (w/v) SDS, 50mM EDTA.
Formaldehyde gel loading buffer : 0.4% (w/v) bromophenol blue, 0.4% (w/v) xylene cyanol, 50% (v/v) glycerol, 1mM EDTA.
Acrylamide stocks : Either 40% or 10% (w/v) acrylamide (acrylamide:bisacrylamide = 19:1).
MOPS 10 x : 0.2M Morpholinopropanesulfonic acid, 0.05M Sodium acetate, 0.01M EDTA. Adjusted to pH 7.0 with NaOH and stored in dark, without autoclaving.
MFF (RNase free) : 25% (v/v) formamide, 0.8% (v/v) formaldehyde, 0.1 x MOPS

2.7.1 Agarose gels

Electrophoresis of nucleic acids was most frequently carried out on 0.8-1.5% (w/v, in either TBE or TAE) agarose gels run at a voltage of up to 12 V cm^{-1}, until the bromophenol blue had migrated to the end of the gel. The running buffer was equivalent to the gel buffer. For many purposes gels could be run for much shorter distances. 3% (w/v) gels were used to electrophorese

fragments smaller than 150 bp. DNA or RNA was visualised by staining the gel either with 0.5ug/ml ethidium bromide, or by adding 0.5ug/ml ethidium bromide to the molten agarose solution prior to casting the gel. After electrophoresis the gel was placed on a Chromato VUE UV source, model TM36. Gels were photographed using a Polaroid camera with Wratten 3A red filter.

2.7.2 Formaldehyde gels

Formaldehyde gels (1.5% w/v agarose) were prepared and run according to the method of Maniatis and co-workers (1982). For each 100ml of agarose solution required, 1.5g of RNase free agarose was dissolved in 73ml of RNase free water by heating in a microwave oven. The solution was allowed to cool to about 60^{o}C whereupon 10ml of 10X MOPS solution and 16.2ml of 37% (v/v) formaldehyde were added and the gel cast immediately. RNA solutions were denatured by the addition of 8 volumes of MFF, followed by heating for 20 min at 60^{o}C. One volume of loading buffer was added before the gel was loaded. Gels were run at about 10 V cm^{-1} in 1 X MOPS (which was recirculated) until the bromophenol blue dye front approached the end of the gel.

2.7.3 Polyacrylamide Gels

Denaturing polyacrylamide gels of the type used for sequencing and footprinting typically contained 15ml of 40% (w/v) acrylamide stock solution, 54g urea and 10ml of 10 X TBE, made up to a final volume of 100ml with distilled water. The urea was dissolved by heating to 37^{o}C. Before pouring, 300ul of 10% (w/v) ammonium persulphate and 50ul of TEMED per 50ml of gel solution. Urea was omitted from non-denaturing gels, which were prepared in a similar manner.

The glass plates (40cm x 20cm) were thoroughly cleaned with ethanol and water and then assembled using three

spacers of Whatman 3MM paper which were clamped together (taping of the plates is not required with this method). The gel solution was run down one edge.
After insertion of the comb the gel was laid at an angle of about 10^{o} to the horizontal. Gels polymerised at room temperature within 30 min.
Gels of the type used in gel retardation experiments are described in section 2.13.
Sequencing gels were pre-electrophoresed for 30 min at a constant power of 40W. The samples were heated in loading buffer at $95^{o}C$ for 5 min and then placed on ice prior to loading on the gel.

2.8 MANIPULATION OF DNA IN VITRO

NEW Solution : 110mM NaCl, 10mM Tris-HCl pH 7.5, 0.5mM EDTA.
NaI Solution : Saturated solution of sodium iodide.
De81 Paper Elution Buffer : 1.5 NaCl, 0.2M NaOH.

2.8.1 Restriction enzyme digests

Restriction digests were carried out at a DNA concentration of 50ug/ml, in the presence of buffers supplied by BRL (REACT^{TM} buffers) and typically with 5-10 units of enzyme per 1ug DNA. Digestion was carried out at $37^{o}C$ (except Sma 1 digestion which was carried out at $30^{o}C$) for 1-2 hours. Digests of large amounts of DNA (typically 100ug) were incubated for up to 16 hours to ensure complete restriction. Small samples of digest mix were subjected to TBE gel electrophoresis to check that restriction was complete. In some cases the extent of linearisation of vector molecules was determined by comparing its ability to transform E.coli, against a circular vector molecule.

2.8.2 Calf intestinal phosphatase treatment (CIP)

To increase cloning efficiency, digested vector DNA was

subjected to treatment with CIP which removes the 5'-terminal phosphate groups and thus prevents recircularisation during ligation reactions. CIP was shown to function in many of the BRL restriction buffers at a concentration of 1-10 units/ug DNA. It was thus added directly to restriction digests for the last 15 min of incubation. When treating blunt ends with CIP this 15 min incubation was followed by further 15 min incubation at 45^{o}C, more enzyme was added and a further cycle of 15 min incubations carried out. The activity was heat killed by rapid heating to 70^{o}C for 10 min followed by rapid cooling on ice. In addition to this, vector DNA was subjected to gel purification to ensure no CIP contamination of subsequent ligations. The efficiency of the CIP treatment was always assessed by comparing the abilities of treated and non-treated vector DNA to self-ligate. Self-ligation was assessed by transformation of E.coli with equal quantites of the vectors in question.

2.8.3 Purification of DNA fragments : DNA fragments were purified by one of a number of methods, the choice of which depended on the size of the DNA fragment and on the quantity of fragment required.

Geneclean : This method was used for fragments of 150 bp to 10 kbp. Restriction digests were run on TAE agarose gels until DNA bands were resolved completely. DNA bands were located by UV light, cut out of the gel, trimmed, and placed in an eppendorf tube. 3 volumes of sodium iodide were then added and the tube incubated at 60^{o}C for 10 min or until the gel slice had completely dissolved. Glassmilk suspension was then added (5ul of glassmilk added to solutions containing 5ug or less DNA and an additional 1ul for every 1ug of DNA above 5ug) and evenly distributed throughout the tube by pippetting the solution and vortexing. The tube was then incubated on

ice for 10 mins. From this stage onwards, all efforts were made to ensure that the temperature of the solution did not rise far above 0°C. The glassbeads were pelleted by centrifugation in a microfuge for 5 sec and the supernatant removed. The glass beads were washed with ice cold NEW solution, again pelleted by centrifugation and the supernatant dicarded. This washing procedure was repeated 3-4 times. After the final wash all the NEW solution was removed and 30-100ul (depending on the amount of DNA) of distilled water was added, the glassbeads resuspended and the tube incubated at 65°C for 5-10 min. The tube was then centrifuged for 45 sec and the supernatant, containing the DNA, was removed to another tube. A second wash of the beads with distilled water sometimes yielded significant amounts of DNA. The yield of DNA was established by mini gel electrophoresis. DNA purified by this method could be used directly in ligations or labelling reactions.

DE 81 Paper : This method was commonly used to purify radio-labelled DNA for use as probes in gel retardations. Restriction digests were run out on an agarose gel and a section of gel was cut out from in front of the desired band. A strip of De81 paper was placed in the well and the agarose block replaced such that when the voltage was reapplied across the gel the DNA band migrated onto the paper and remained bound. The paper strip was then extracted from the well and placed in an eppendorf tube to which was added 300ul of the elution buffer. The paper was mashed with a hypodermic needle and the tube vortexed for a few seconds. The paper remnants were pelleted and the supernatant transferred to a fresh tube. The process of elution of DNA from the paper mash was repeated. The supernatants were pooled and spun to remove De81 paper fines, then the DNA was precipitated by the addition of 150 ul of 3M NaOAc, pH 5.5, and two volumes of cold

ethanol.

Electroelution apparatus : The gene trap apparatus, manufactured by Schleicher & Schuell, and the Unidirectional Electroeluter, manufactured by IBI, were used occasionally according to the manufacturers instructions.

2.8.4 Ethanol precipitation of DNA : The DNA solution was made 0.3M in NaOAc and 2 volumes of cold ethanol added. After mixing, the DNA was precipitated by cooling on ice or by placing at -20^{o}C for up to 2 hours and pelleted by centrifugation (27000g, 15 min, 4^{o}C for large volumes or 12000g, 15 min, 4^{o}C, for small volumes in eppendorf centrifuge tubes). The pellet was usually washed in 70% (v/v) ethanol and dried briefly in a vacuum drier.

2.8.5 Blunting of recessed 3'-ends : The basic reaction mix contained 1ug of DNA (with 5' protruding ends), 1ul of 1mM dNTP's (all four), 2ul of 10 X Klenow buffer, 1ul of 2-mercaptoethanol (1:40,dH_2O), 1ul of Klenow fragment (1U/ul) and distilled water in a total volume of 20ul. Reactions were incubated at 22^{o}C for 15 min.

2.8.6 Ligation of DNA fragments
Restriction fragments to be ligated were mixed such that the insert was in a 10 x molar excess over vector molecule. Typically, reactions were made up to 15ul by the addition of 1.5ul of 10 x ligation buffer, distilled water and 1ul of T4 DNA ligase (0.01 units/ug for overhanging end ligation and 1 unit/ug DNA for blunt end ligation) and the solution incubated at room temperature for 1 hour, for blunt ended ligation, or for 16 hours at 14^{o}C, for overhanging end ligations. The efficiency of the ligation reaction was assessed by comparing the transformation efficiencies of equal amounts of ligated

and non-ligated vector/insert mixes. Prior to any transformation procedure, ligation reactions were diluted 5-10 fold to minimise the inhibition of transformation by constituents of the ligation mix.

2.9 SEQUENCING

Denaturing buffer : 0.2M NaOH, 0.2 mM EDTA, pH 8.0.
Annealing buffer, 10 x : 70mM Tris-HCl pH 7.5, 70mM $MgCl_2$, 300mM NaCl, 100mM DTT, 1mM EDTA pH 8.0.
Nucleotide Mixes (dNTP/ddNTP mix) : Working concentrations; 0.5mM Deoxynucleotides, 2mM Dideoxynucleotides,
Chase solution : 0.125 mM dNTP (all four)
Formamide dye-mix : 98% (v/v) deionised formamide , 10mM EDTA, pH 8.0, 0.2% (w/v) bromophenol blue and 0.2% (w/v) xylene cyanol.

A volume corresponding to 2ug of plasmid DNA was dried down, the pellet resuspended in 40ul of denaturing buffer and allowed to stand for 5 min at room temperature. 4ul of 2M ammonium acetate (pH4.5) was added to neutralise the sample. The DNA was then precipitated by the addition of two volumes of cold ethanol and incubation of the tubes at -70^{o}C for 5 min, followed by centrifugation in a microfuge for 10 min at 4^{o}C. The DNA was washed with 70% (v/v) ethanol and vacuum dried. The plasmid pellet was resuspended in the following mix; 1ul oligonuleotide primer (2.5pmol/ul), 1.5ul 10 x Annealing buffer, 2ul [^{35}S]-dATP (8uCi/ul, 650 Ci/mmol) and 11.5 ul of distilled water, and incubated at 65^{o}C for 2 min. The mix was allowed to cool to room temperature, during which time the oligonucleotide primer hybridised to the denatured plasmid. 1ul of Klenow-fragment (diluted to 2U/ul) was added to the hybridisation sample, mixed, then 3ul was added to each of four tubes containing dNTP/ddNTP

mixes. The tubes were incubated at 30°C for 20 min. 1.5ul of chase solution was added to each of the four tubes and then incubated for a further 15 min. The samples were then dried under vacuum, resuspended in 4ul of formamide dye-mix, boiled for 3 min and cooled on ice before loading onto a sequencing gel.

2.10 RADIOLABELLING OF DNA

Kinase buffer : 100mM KCl, 70mM Tris-HCl pH7.5, 10mM $MgCl_2$, 5mM DTT.
TNE : 100mM NaCl, 10mM Tris-HCl pH7.5, 1mM EDTA.
Klenow buffer : 50mM Tris-HCl pH8.0, 10mM $MgCl_2$

2.10.1 Nick Translated probes : The following reaction mix was commonly used; 45ul Klenow buffer, 1ul dATP,dGTP,dTTP(10mM) , 1ul B-mercaptoethanol (1:40 with distilled water), 1ul DNase I (0.5ug/ml), 5ul distilled water, 5ul alpha [^{32}P]-dCTP (50uCi), 1ul E.coli DNA polymerase I, and 2ul of plasmid or fragment DNA (0.5-1.5ug). The reaction mixture was incubated for one hour at 14°C.

2.10.2 End-labelling of oligonucleotides : Labelling reactions contained the following; 4ul 10 x kinase buffer, 200-300ng oligonucleotide , 5ul gamma-^{32}PATP, 1ul T4 polynucleotide kinase and distilled water in a total volume of 40ul. The reaction was incubated at 37°C for 1 hour.

2.10.3 End-labelling of 5'-overhangs : As for filling in of recessed ends (2.8.5) but replacing dCTP's with 1-5ul (10-50 uCi) alpha [^{32}P]-dCTP.

2.10.4 Removal of unincorporated nucleotides : For most purposes it was advantageous to remove unincorporated

nuleotides from the labelled DNA. This was achieved by gel filtration on Sephadex-G50 columns (20cm in length, 0.5cm diameter). 5ul of Dextran Blue (50mg/ml) and 5ul of Phenol Red (50mg/ml) were added to the reactions. The samples were then loaded onto a, TNE equilibrated, column. The filtrate containing Dextran Blue was collected as it co-purified with nick translated probes. Oligonucleotide probes eluted from the column with, and just behind, the Dextran Blue.

2.11 FILTER HYBRIDISATIONS

SSC (Standard saline citrate) : 0.15M NaCl, 0.015M Na_3citrate.
SSCP (20X) : (All units are gl^{-1}). NaCl 193.2, Na_3citrate 88.4, Na_2HPO_4 38.0, $NaH_2PO_4.H_2O$ 19.2.
Denhardts solution (100x) : (All units are gl^{-1}). Bovine serum albumin 20, Ficoll 20, and Polyvinylpyrollidine 20.
Hybridisation solution : 10ml of hybridisation fluid was composed of the following; 2.7ml 20 x SSC, 4.5ml formamide, 0.5 ml 50% dextran sulphate (w/v), 0.45ml 100 x Denhardt's, 0.45ml 10%(w/v) SDS, 0.45ml 0.4M $NaPO_4$, 0.36ml 0.25M EDTA, 0.6ml distilled water.
N.B. Formamide concentration in oligonuceotide hybridisations varied from the above. These concentrations were optimised experimentally by Northern hybridisation (See oligonuleotide probes Figure 2.1).

2.11.1 Northern blotting

A formaldehyde gel was placed face down on three sheets of Whatman 3MM paper soaked, and in contact with, 20 x SSC. A piece of nitrocellulose (presoaked in distilled water, followed by 20 x SSC) was placed on the gel, followed by three pieces of Whatman 3MM paper cut to the size of the gel. Onto this were placed 3-4 layers of absorbent nappy pads, cut to size, followed by a 1.5kg

weight. The gel was blotted overnight with 20 X SSC as transfer buffer. After blotting, the nitrocellulose filter was baked for 3 hours at 80^0C.

2.11.2 Dot blotting

Ethanol precipitates of RNA were pelleted by centrifugation at 12000 g at 4^oC, vacuum dessicated, and resuspended in 200ul of RNase free water. To this, was added, 200ul of 6:4, 20 x SSC:37% (v/v) formaldehyde. The samples were vortexed and heated for 15 min at 60^oC. RNA samples were stored in this form at -70^oC. When required, the RNA samples were thawed, vortexed and diluted, if necessary, by the addition of SSC/formaldehyde mix. A sample volume of 60-100ul per dot was used. A nitrocellulose filter was placed in the 96-well BRL Hybridot manifold, a vacuum applied, and the samples added to the wells. After all the samples had been loaded onto the filter, 100ul of 15 x SSC was added to each well. When the SSC had been drawn through, and with the vacuum still applied, the manifold was dismantled and the filter removed. The filter was baked at 80^oC for 3 hours.

2.11.3 Hybridisations

Baked nitrocellulose filters were bathed in prehybrisation fluid in sealed plastic bags. These bags were incubated in a shaking water bath at the appropriate temperature for at least 2 hours. Probes, denatured by boiling for 10 min, were then added to the bags and the incubation continued at the appropriate hybridisation temperature for various lengths of time; 24 hours,for oligo probes and 16 hours for nick translated probes. The volume of denatured probe added was a function of the specific activity of the probe and the concentration of the probe DNA; routinely, for a nick translated probe with a specific activity of 10^7 cpm/ug a volume corresponding to 10^3 cpm/ml of hybridisation fluid was

added.
Following hybridisation with nick translated probes, the filters were washed two times for 15 min each with 2 x SSCP/ 0.1% (w/v) SDS at 62°C and three times for 15 min each with 0.56 x SSCP/0.1% (w/v) SDS at 55°C.
After hybridisations with oligonucleotide probes the filters were washed three times, for 15 min each, in 6 X SSC, at room temperature and twice, for 15 min each, in 6 X SSC, at 45°C.
The washed filters were sealed in plastic bags and exposed to Kodak X-Omat film, routinely for 16 hours in an autoradiography cassette usually with an intensifying screen. The film was processed using a Kodak X-Omat automatic processor.
After hybridisation and autoradiography, all dot blot filters were cut up into individual dots, placed into scintillation vials, 1ml of Ecoscint scintilation fluid added and the radioactivity counted in a Beckman LS1801 scintillation counter.

2.12 RETARDATION ASSAYS

Binding assay solution (1x). : 20mM Tris-HCl pH8.0, 70mM KCl, 5mM MgCl, 0.5mM $CaCl_2$, 0.5mM DTT, 0.1mM EDTA,

Binding activity was assayed by a method similar to that of Huet et al. (1987). Each 14ul binding assay contained binding assay solution, 7%(v/v) glycerol, 40pg-2.5ng of alpha-[^{32}P] end-labelled DNA fragment, 560ng of non specific competitor DNA (pSP65 or lambda DNA) and 0.35-16ug of protein. Specific competitor DNA, in addition to non-specific DNA, was included in some competition assays. The protein was added immediately before the assays were incubated at 22°C for 15 min. The samples were directly loaded onto a 20cm x 18cm, non-denaturing polyacrylamide vertical slab gel (4.94% (w/v) acrylamide,

0.06% (w/v) bisacrylamide, 5%(v/v) glycerol). The running buffer contained 20mM Tris-HCl pH8.0, 1mM EDTA and 5mM 2-mercaptoethanol. After polymerisation, the gel was run for at least 1 hour at constant voltage (250 V). The reservoir buffer was changed and the gel equilibrated overnight, at 4°C. The running buffer was exchanged again, for fresh, pre-cooled buffer, just prior to loading the samples. Electrophoresis was carried out at 4°C for 3 to 4 hours at constant voltage (250 V), with one buffer change after 1.5 hours. Loading buffer containing bromophenol blue and xylene cyanol was added to the side wells in order monitor migration. When the xylene cyanol was a few cm from the base of the gel plates, electrophoresis was terminated. The gels were then dried onto Whatman 3MM paper and autoradiographed.

2.14 Footprinting of Protein-DNA Complexes in Polyacrylamide Gel Slices

Solution A : Equal volumes of 40mM 1,10-phenanthroline [OP] (in 100% ethanol) mixed with 9.0mM copper sulphate (in water) then diluted 10-fold with water to 2.0mM OP/0.45mM $CuSO_4$.
Solution B : 58mM 3-mercaptopropionic acid (MPA). (neat MPA diluted 1/200).
Solution C : 0.5M ammonium acetate and 1mM EDTA

The method utilises the "nuclease" activity of the 1,10-phenanthroline-copper ion (Sigman, 1986), as illustrated by Kuwabara and Sigman (1987).
Gel retardations were conducted with 15 times the standard amounts of ^{32}P labelled DNA and protein. Binding assay mixes were loaded on a non-denaturing gel. Glycerol was omitted from the gel to prevent possible quenching of nuclease activity. After electrophoresis the wet gel was

autoradiographed for 30 mins. The autoradiograph was used to identify the position of the free and complexed DNA in the retardation gel. Gel slices corresponding to these positions were excised from the gel and immersed in 100ul of 50mM Tris-HCl pH8.0. 10ul of solution A, followed by 10ul of solution B was added. The digestion was allowed to proceed for 10 min at room temperature and then quenched by the addition of 28mM 2,9-dimethyl-1,10-phenanthroline (in 100% ethanol) and 270ul of solution C. The tubes were incubated at 37^{o}C overnight, during which time, the DNA eluted from the gel slices. This DNA was ethanol precipitated, pelleted, dried under vacuum, and resuspended in sequencing loading buffer. The samples were electrophoresed in a fashion identical to sequencing reactions.

Figure 2.1 Oligonucleotides

PGK1 (Phosphoglycerate kinase) CT Block and I-Site

5'-TTTCTTTTTCTCTTTTTTACAGATCATCAAGAAGAGCT-3'
3'-TCGAAAGAAAAAGAGAAAAAATGTCTAGTAGTTCTTC-5'

PGM (Phosphoglycerate mutase)

5'-CCAAAGTTAGTTTTAGTTAGACACGGTC-3'

ENO (Enolase)

5'-GCTGTCTCTAAAGTTTACGCTAGATCCGTC-3'

TPI (Triose Phosphate Isomerase)

5'-GCTAGAACTTTCTTTGTCGGTGGTAAC-3'

HEX A and B (Hexokinase)

5'-GGAGCAGTGATGACAACCTTCTTGGCACC-3'

PFK1 (Phosphofructokinase 1)

5 -CCACCTTGTTGAACATGGCCTGGGATAGC-3

PFK2 (Phosphofructokinase 2)

5 -CGGTATCAGAACCCAAAGAGTATTCAGTACC-3

RP1 (Ribosomal Protein I, also known as L3)

5'-CCTCTTGGGGTTTCGACGTAACCGACAACACC-3'

HSP90 (Heat Shock Protein 90)

5 -CACTAGCCATATCTTTGCGTGTTTGTTTGTGC-3

Figure 2.2 Hybridisation Conditions for Oligonucleotide Probes

Oligonucleotide	Formamide (%age)	Temperature (^{o}C)
PGM	20	37
ENO	20	37
TPI	20	37
HEX A and B	5	37
PFK1	10	37
PFK2	5	37
RP1 (RPL3)	20	37
HSP90	20	37

CHAPTER 3

INVESTIGATION OF THE MINIMAL SEQUENCE REQUIREMENTS FOR TRANSCRIPTION FROM THE YEAST PGK1 PROMOTER; I-SITE DETERMINING SEQUENCES

3.1 INTRODUCTION

Efficient transcription of DNA relies upon at least three types of information; that specifying the initiation position, the termination position and the specific rate at which initiation occurs. In this chapter experiments will be described which were designed to reveal the minimum sequence requirements for efficient transcription initiation in yeast, and specifically to investigate the role of a pyrimidine-rich sequence tract (CT block) in the definition of the transcriptional start site.

Prokaryotic genes (as represented by the genes of E.coli) have a highly conserved arrangement of cis-acting elements which constitute the transcriptional promoter. The sequences 5'-TTGACA -35 bp from the transcriptional start site, and 5'-TATAAT -10 bp (TATA or "Pribnow" box) are almost ubiquitous (Hawley and McClure, 1983). RNA polymerase interacts with both of these sequence motifs to initiate transcription (Schmitz and Galas, 1979). Together with a regulatory (operator) site, these sequences are all that is required for efficient, regulated transcription of most E.coli genes.

The constitution of the eukaryotic promoter appears far more complex. Many elements contributing to efficient transcription of eukaryotic genes can be found as multiple repeated units, sometimes overlapping, and frequently separated by large distances. Some elements even contribute to the expression of divergently transcribed genes (For example, see review by Johnston, 1989). Despite the complexity of eukaryotic promoters, most elements can be classified as either initiation site (I-site) determinants or initiation rate determinants. Initiation rate determining elements are usually associated with a specific regulatory response to an environmental factor or with tissue-specific gene expression.

In higher eukaryotes, initiation site (I-site) determination is, in most cases, strongly associated with sequences which have strong homology to the prokaryotic -10 TATA consensus. The AT-rich motif, TATAA/TA (abbreviated to TATA) is found 30 bp upstream from the I-site. The spacing between the TATA and the I-site is highly conserved in higher eukaryotic genes. The TATA element does not occur in all higher eukaryotic genes, but when it is present, the integrity of the sequence is essential for normal expression.

Higher eukaryotic genes which do not contain TATA-like elements are usually non-specialised, or "housekeeping" genes and these often have multiple I-sites (Dynan, 1986).

The TATA element is one of a number of elements found in close proximity to the I-site in cellular genes. Other elements include the CAAT box and Sp1 binding site (Sasson-Corsi and Bornelli, 1986). Taken together with the TATA-box, these elements constitute what may be termed the "core promoter". Outwith this core promoter, gene or gene family-specific regulatory elements may occur.

The exact mechanism by which the TATA element determines the I-site in higher eukaryotic genes is unknown. However, a specific protein has been shown to interact with the TATA element (Parker and Topol, 1984; Sawadogo and Roger, 1985 a and b). Therefore, it is possible that RNA polymerase II is induced to initiate transcription, at the correct position, by direct interaction with this protein.

The mechanism of I-site determination in yeast genes is poorly understood. The role of the TATA is not well defined in yeast and there is almost certainly a requirement for other I-site determining sequences.

General conclusions from experiments on yeast transcription may be invalid since, to date, the yeast

genes studied in detail may be atypical. In yeast, most genes are transcribed constitutively, at low levels (St. John and Davis, 1979; Struhl and Davis, 1981). However, studies on the mechanism of yeast transcription have been limited to genes that are highly expressed or tightly regulated by environmental stimuli. Many of the statements which follow in this chapter concerning I-site determination may only be relevant for a subset of yeast genes.

By analogy with higher eukaryotes the TATA element is a major candidate for a yeast promoter element. Sequences sharing homology with the higher eukaryotic TATA element are present in the 5'-flanking regions of most yeast genes. It is difficult to derive a consensus TATA-like sequence for yeast, as yeast 5'-flanking regions are highly AT-rich. Therefore many possible TATA-like sequences may occur. In contrast to the situation in higher eukaryotic genes, in yeast there is no tight relationship between I-site and TATA element position. On average, I-sites are found 60 bp downstream from a TATA-like sequence, but the actual distance may vary from 40 bp to 120 bp. This suggests that TATA-like sequences do not function in yeast genes exactly as they do in higher eukaryotic genes.

There is however, evidence that yeast TATA elements do contribute towards transcription and are related to I-site definition. For example, deletion of regions containing TATA elements, or mutation within TATA elements, is detrimental to the transcription of specific yeast genes (Struhl, 1982; Guarente and Mason, 1983; Siliciano and Tatchell, 1984). Reversal of the orientation of the TATA element also results in a reduction in transcription (Nagawa and Fink, 1985).

A role for TATA-like sequences in I-site determination has been demonstrated in the yeast TRP1 gene. TRP1 is transcribed to yield two mRNAs; transcript I and

transcript II. Two TATA elements are required for accurate initiation of transcript I. Deletion of one TATA element results in a change in the start site position and when both TATA elements are deleted multiple start sites are observed (Kim et al., 1986). The necessity of TATA-like sequences for accurate transcriptional initiation has been noted in other yeast genes, for example CYC1 (Guarente and Maso, 1983, Hahn et al., 1985), HIS4 (Nagawa and Fink, 1985),the MAT locus (Siliciano and Tatchell, 1984), and HIS3 (Struhl, 1982). Where multiple TATA elements are present within a single promoter each can exhibit distinct properties. For example, the HIS3 promoter contains one constitutive and one regulated upstream element. It has been proposed that each upstream element only functions in conjunction with its own distinct TATA element. Thus, the HIS3 promoter contains a TATA element which functions only with the regulated, rate determining upstream element ($TATA_R$) and a TATA element which functions solely with the constitutive rate determining upstream element ($TATA_C$). Four TATA elements also appear to function independently in the CYC1 gene, which has multiple I-sites. Each individual TATA element appears to be required for the initiation of a subset of mRNAs. For example, disruption of one TATA element affects the abundances of a subset of mRNAs with I-sites located 60-100 bp downstream (McNeil and Smith, 1986). Other I-sites outwith this "window" are preserved.

In summary, TATA elements in yeast genes are associated with I-sites but it is clear that the spacing between the TATA element and the I-site is not as critical as in mammalian systems. There is also evidence that individual TATA elements function independently within a single promoter. This suggests that an important difference exists between the TATA elements of higher eukaryotes and of yeast.

The requirement for a TATA element for the transcription of all yeast genes was questioned by Ogden et al, (1986). The highly expressed PGK1 gene contains two TATA-like sequences in the 5'-flanking region, TATAAA (-154 relative to the translational initiation codon) and TACATA (-117). Ogden et al (1986) demonstrated that deletion of one or both of these sequences does not effect PGK1 mRNA abundance as measured by Northern blotting. Further to this, these deletions were shown to have no affect on the I-site position. As this publication is of particular importance in the development of this project we should consider just how these conclusions were derived.

The I-site was identified in both wild type and deletion mutant promoters by riboprobe mapping procedures (Melton, et al, 1984; Zinn et al, 1983). PGK1 mRNA-complementary RNA probes were synthesised from a linearised pSP65 vector containing a specific PGK1 probe fragment. This BamHI/ClaI PGK1 fragment contained the entire PGK1 promoter region and 37 bp of the PGK1 coding region. Total yeast RNA was hybridised with excess probe and then RNA-RNA hybrids were treated with RNase as described by Zinn et al (1983). Protected fragments were resolved on an 8% denaturing polyacrylamide gel. In experiments carried out with RNA derived from yeast transformants carrying wild type or TATA-deletion PGK1, the size of protected fragment was found to be identical. Minor heterogeneity was observed in all samples. This result suggests that the deletion of TATA-like sequences does not prevent RNA synthesis initiating at the normal position and does not induce the initiation of a new I-site within the area encompassed by the riboprobe; that is, the entire promoter region and 37 bp of sequence downstream of the start of the coding region. This is consistent with TATA-like elements in the PGK1 promoter having no contribution to either the high rate of PGK1

transcription or to the position of transcription initiation. Contrary to the results of Ogden et al, (1986), Rathjen and Mellor have recently shown that the PGK1 TATA elements are essential for normal PGK1 transcription. This work is discussed in detail in 3.2.5. It is clear from the cumulative work on yeast promoters that TATA-like sequences are not tightly related to I-site position and the results of Ogden et al (1986) in particular illustrated that elements other than TATA are responsible for determining the position of I-sites. This contrasts with the role of TATA elements in higher eukaryotic genes, where the position of the TATA element is closely related to the I-site.

When speculating upon which sequences, other than TATA, may determine I-sites in yeast, one may consider sequences which occur frequently in yeast promoters. Sequences occurring close to the I-site of a number of highly expressed yeast genes have been observed by Dobson et al (1982). Figuring prominently in this survey is a CT-rich stretch, 20 to 60 bp in length, followed 8 to 12 base pairs downstream by a CAAG sequence. RNA initiation occurs at, or near, the CAAG seqence which is similar to a consensus initiation site sequence observed by Hahn et al (1985). Hahn et al showed that 55% of yeast promoters surveyed had either RRYRR (38%; R=purine, Y=pyrimidine) or TC(G/A)A (17%) at the I-site. CT-rich stretches or CT blocks are found in a number of highly expressed yeast genes such as SUC2, GAL1, ADR2 and CAR1, those coding for the glycolytic enzymes, PGK1, PYK1, ENO and GRD (Dobson et al., 1982; McNeil, 1988) and also in a number of filamentous fungal genes (Gurr et al., 1987).

The ubiquitous presence of CT blocks in highly expressed genes suggests a role in efficient transcription, but there is, to date, little evidence to support this supposition. However, an investigation into start site selection in the yeast CYCI gene indicated that CT

blocks inhibited initiation within the CT rich region itself, and also at immediately adjacent purines (McNeil and Smith, 1985). In this particular example, insertion of two G's and an A into a 10 bp CT block resulted in efficient transcription initiation at the most upstream of the inserted G's and minor initiation at both the other inserted purines. This indicates that CT blocks may fix a start site position by inhibiting initiation within a limited area thus promoting initiation at a site further downstream.

Sequence requirements at I-sites in yeast have also been examined experimentally by Hahn et al (1985) who demonstrated efficient initiation of transcription at the sequence TCGA when this sequence was introduced into the CYC1 promoter. Specific I- site sequences have also been revealed in HIS3 and DED1 (Chen and Struhl, 1985), PHO5 (Rudolf and Hinnen, 1987) and TRP1 (Mellor et al, 1988). For example, the PHO5 promoter contains two major I-sites which are of the form RRYRR. Disrupting normal initiation of transcription by sequential deletion of the PHO5 gene results in the utilisation of new I-sites. Each of these new I-sites are at RRYRR sequences (Rudolf and Hinnen, 1987).

In yeast, both I-site specific sequences, CT blocks and TATA-like sequences have been implicated in I-site determination. The relationship between TATA-like sequences and the I-site sequences is unknown. Ogden et al (1986) propose that deletion of the TATA-like sequences in PGK1 has no effect on the I-site position or level of transcription. One may conclude from this that the CT block and I-site specific sequence, together with an upstream activation site are sufficient for high level, accurate transcription of the PGK1 gene. The following experiments were undertaken to test this assumption, and with a view to defining the minimal sequence information required for high level

transcription in yeast.

3.2 Requirements For Efficient Transcription in Yeast and the Role of the PGK1 CT Block in Transcriptional Initiation.

It is of interest to know the minimal sequence requirements for transcription of the highly expressed gene PGK1. If the PGK1 CT block and I-site sequence were shown to be sufficient for accurate I-site determination in yeast, then the understanding of the initiation mechanism may be simplified. The role of TATA-like sequences in yeast may also be more apparent. Prompted by this goal, experiments have been carried out to answer the following questions;

i. Is the high level transcription of the PGK1 gene defined by the UAS_{PGK}, the Downstream Activation Site (see 4.1) the CT block and the I-site specific sequence (CAAT) only ?
ii. If the CT block functions in I-site determination, does it function unidirectionally (like TATA) or bidirectionally?

iii. Does the PGK1 CT block function in conjunction with a UAS from a promoter that requires a TATA-like sequence for function? The possibility exists that there are two very different mechanisms of transcription initiation in yeast; TATA-dependent and TATA-independent. It is possible that promoter elements from different types of promoter are incompatible?

iv. Does the CT block function by binding a specific protein?

In an attempt to answer questions i to iii, promoter

elements have been combined in novel arrangements and placed upstream of a reporter gene coding region linked to a transcriptional termination signal. These plasmid constructs were then introduced into yeast and the abundance of heterologous mRNA per gene copy was measured. It was intended that promoters which exhibited high levels of transcription would be analysed further to determine which sequences were responsible for I-site determination.
An experimental approach to answering question iv will be discussed in 3.3.

3.2.1 Construction of pGJLP Series

For convenience, yeast promoter elements were inserted into a bacterial plasmid polylinker. This approach allows modification of the putative promoters at a later date by insertion of other promoter elements (such as TATA) or the manipulation of distances between promoter elements by the insertion of "spacer" DNA. Putative promoter elements from the PGK1 5'-flanking region have been cloned into the polylinker of pUC19 and this polylinker was subsequently sub-cloned into the promoter-probe vector, pCJ19 (Chris Hadfield, Leicester Biocenter).
pCJ19 (figure 3.1b) is a 2u-based vector which propagates in yeast at high copy number. It contains the Ap^R gene, for selection in E.coli, and the TRP1 gene, for selection in yeast. It also contains the promoterless chloramphenicol acetyltransferase (CAT) gene from E.coli. When this gene is expressed, it confers upon the host resistance to the antibiotic, chloramphenicol. A polylinker is situated upstream of the CAT coding region, into which putative yeast promoters may be cloned. Downstream of the CAT coding region is a CYC1 transcriptional terminator.
There are advantages in using the CAT gene when studying

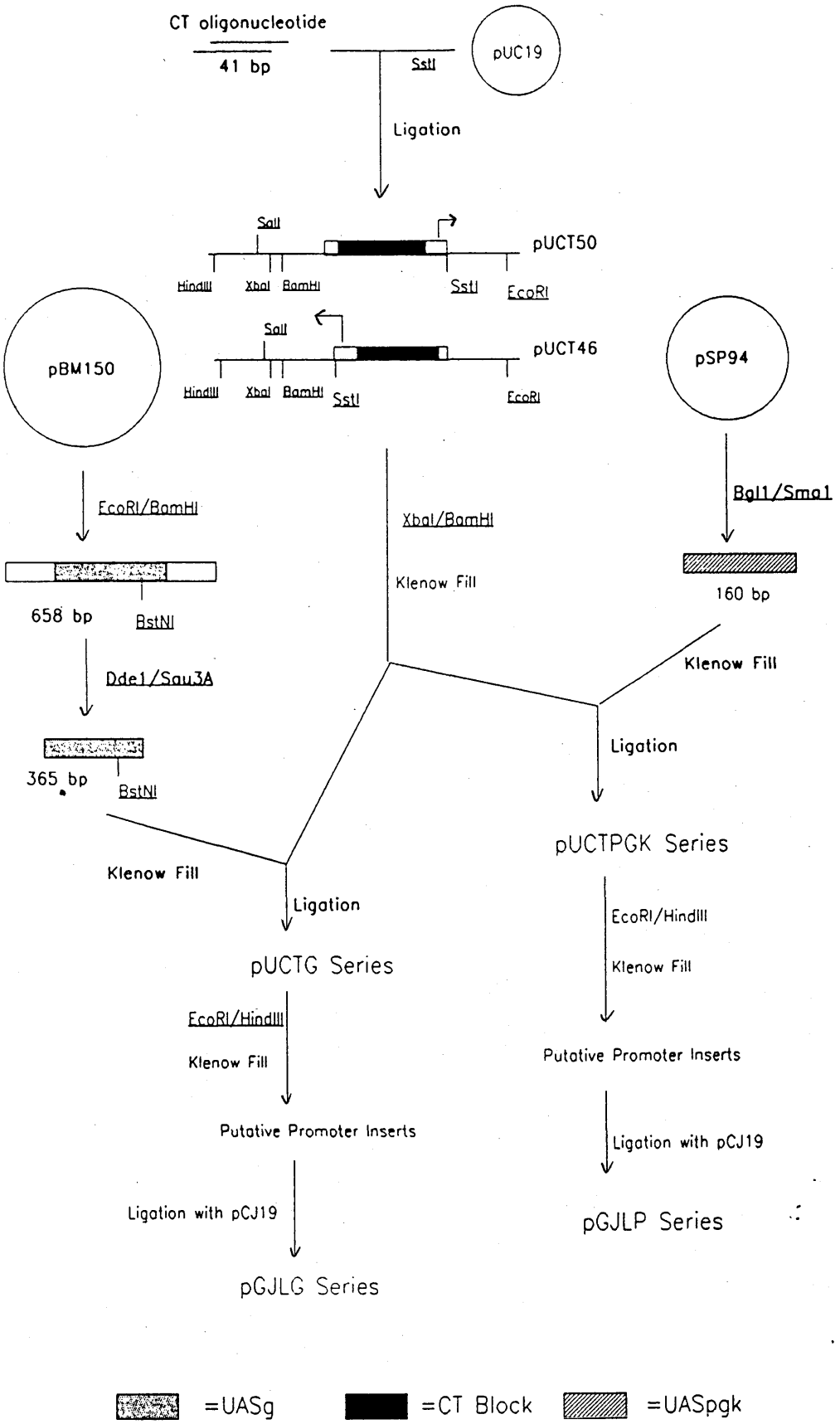


Figure 3.1 (a) Construction of pGJLG and pGJLP Series Plasmids
A narrative on the construction of these plasmids appears in 3.2.5. and 3.2.5.

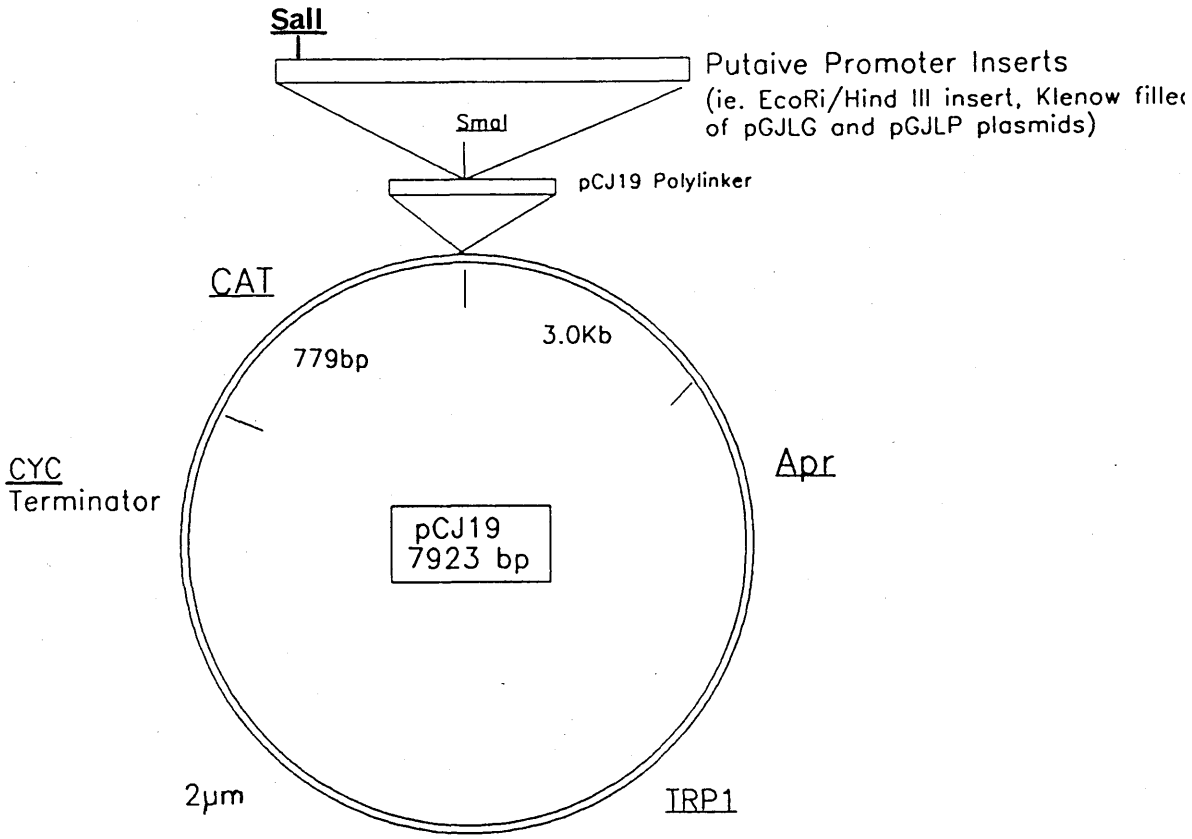


Figure 3.1 (b) Promoter Probe Vector pCJ19.
Supplied by Chris Hadfield (Leicester Biocentre). Putative yeast promoter elements were inserted into the SmaI site of this vector. The SalI site was used to determine the orientation of insertion. The CAT coding region, as indicated, is 779 bp long and the distance between the Ap^R and the promoter insertions is 3.0 Kb.

promoter efficacy. Because CAT is a heterologous gene in yeast, any CAT mRNA originates only from the plasmid copies of the gene. Thus, the wild type copy of the gene under investigation can be used as an internal control of promoter strength. In addition, as expression of the CAT gene confers drug resistance, preliminary expression studies can be carried out using a plate assay.

There is one possible disadvantage in using multicopy plasmids in these studies. It is possible that cells carrying many copies of the test promoter may suffer repression of the chromosomal locus of the test gene. This can occur by titration of a trans-acting factor by the multi-copy test promoter (Irani et al., 1987).

pCJ19 has been shown to function as a promoter probe vector (Chris Hadfield, personal communication). Promoterless vector does confer a very slight resistance to chloramphenicol: yeast colonies can be observed after an extended incubation period of 7 days on plates containing 500 ug/ml. Insertion of a yeast promoter into this vector confers high resistance to chloramphenicol (Chris Hadfield, personal communication).

A graphical representation of the construction strategy is presented in Figure 3.1 (a)

Two complementary 41 bp oligonucleotides were designed such that on hybridisation they recreated the PGK1 CT block and I-site (from nucleotide -33 to -67 bp upstream of the initiation codon) and, in addition, had ends complementary to those created by an SstI restriction digestion. On insertion into SstI digested vector an SstI site is recreated only at the 3'-end of the oligonucleotide insert (that is, the "CAAG" end).

The oligonucleotide was cloned into pUC19 as described below. Oligonucleotide concentrations were measured by the A_{260}, then equal amounts of each oligonucleotide were mixed in a solution of TE containing 100mM NaCl, to give a final DNA concentration of 300 ug/ml. This

mixture was heated to 80°C in a heating block and then cooled slowly to 0°C. To ensure that the oligonucleotides hybridised satisfactorily, the process was also carried out with gamma[^{32}P]-dATP polynucleotide kinase-labelled oligonucleotides which were then electrophoresed through 12% (w/v) non-denaturing polyacrylamide. Figure 3.2 illustrates an autoradiograph of a gel on which both single-stranded and annealed material have been electrophoresed. It is clear from this gel that almost all of the single-stranded material in the mixture has hybridised and consequently exhibits reduced electrophoretic mobility. Double-stranded oligonucleotide was then mixed, at 40-fold molar excess, with SstI cut pUC19 and ligated. Prior to this ligation, control transformations were carried out with undigested vector, digested vector and religated vector. These controls ensured that the vector DNA was sufficiently digested, and would subsequently religate. In addition, these controls also revealed whether transformation was inhibited by a factor introduced during vector preparation or subsequent ligation (for example, T4 ligase may have inhibited transformation). To reduce such inhibition, ligation mixes were diluted 5-fold with distilled water prior to transformation.

The cloning of DNA into the pUC19 polylinker was selected for on the basis of insertional inactivation of B-galactosidase activity as revealed by white colony colour on plates containing X-gal and IPTG. Insertion of a 41 bp oligonucleotide into the pUC19 polylinker should result in a frameshift mutation, abolishing B-galactosidase activity. Transformants containing products of the CT-oligo/pUC19 ligation were screened for inactivation of B-galactosidase. Plasmid DNA prepared from putative positives (white colonies) was restricted with EcoRI and HindIII or SstI and HindIII, end-labelled and electrophoresed through a 12% (w/v) polyacrylamide gel

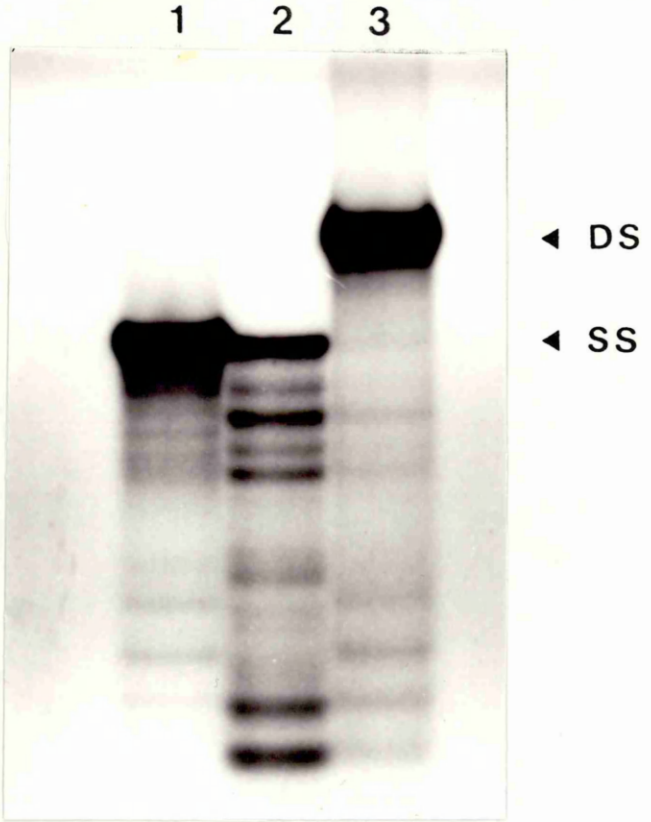


Figure 3.2 Solution Hybridisation of CT Block Oligonucleotides.
This autoradiograph of a 12% non-denaturing polyacrylamide gel illustrates the hybridisation of two complementary oligonucleotides which were subsequently cloned into pUC19. This gel is discussed in 3.2.1. Lane 1 and 2: Single stranded (SS) oligonucleotides. Lane 3: Double Stranded (DS) hybridised oligonucletides. The bands which appear below those indicated as SS correspond to oligonucleotides shorter than 38 bp in length.

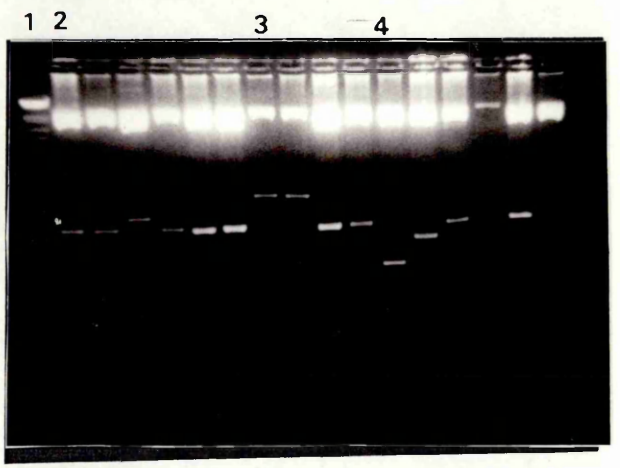


Figure 3.3 Screening Transformants for the Presence of UAS_G-Containing Plasmids.
This is an example of a ligation product screen in the construction of pGJLP series plasmids. Lane 1: lambda DNA EcoRI/HindIII size standards. All other lanes represent plasmid DNA isolated from individual transformants by the STET method (2.7.2) restricted with PvuII. Restriction of pUCT50 with PvuII yields a band of 322 bp (Lane 4). Lane 2: Insertion of a single UAS fragment, 482 bp band. Lane 3: Insertion of two UAS fragments, 642 bp band.

which was then autoradiographed. A plasmid containing the insert in the desired orientation (that is, CAAG proximal to the polylinker EcoR1 site therefore yielding a 100bp band upon restriction with HindIII and SstI) was identified and named pUCT50. The plasmid containing the insert in the reverse orientation was named pUCT46.
Each of these oligonucleotide containing plasmids were plasmid sequenced to ensure that no unpredicted insertions or deletions had occurred.
The PGK1 UAS (UAS_{PGK}) was then inserted into both pUCT46 and pUCT50. pSP94, containing the UAS_{PGK} (-402 to -538), was restricted with BglI and SmaI and termini were made flush using the Klenow fragment of DNA polymerase I. The 160 bp fragment was then gel purified by the geneclean method (2.8.3). pUCT50 and pUCT46 vectors were restriction digested with both Xba1 and BamHI and the termini similarly made flush. The terminal phosphates were removed with calf intestinal phosphatase (2.8.2), thus preventing re-circularisation of the vector molecule, without insertion of a non-phosphatase treated DNA fragment, upon ligation. The vector was gel purified by the geneclean method (2.8.3). The efficiency of digestion and phosphatase treatment was assessed by comparing the transformation efficiencies with those of ligated and non-ligated control vector taken from different stages of the vector preparation. Plasmid DNA was prepared from transformants containing the products of the pUCT46 or 50/UAS_{PGK} ligation, restriction digested with PvuI and electrophoresed through 1.5% (w/v) agarose gels (Figure 3.3). UAS_{PGK} containing plasmids were expected to contain a 482 bp fragment, or a 642 bp band if a double insertion occurred. The orientation of insertion was established by restricting plasmid thought to contain inserts with PstI. Five distinct types of ligation product were revealed by this analysis (pUCTPGK1 to 5). pUCTPGK1 to 4 contained single inserts. pUCTPGK5

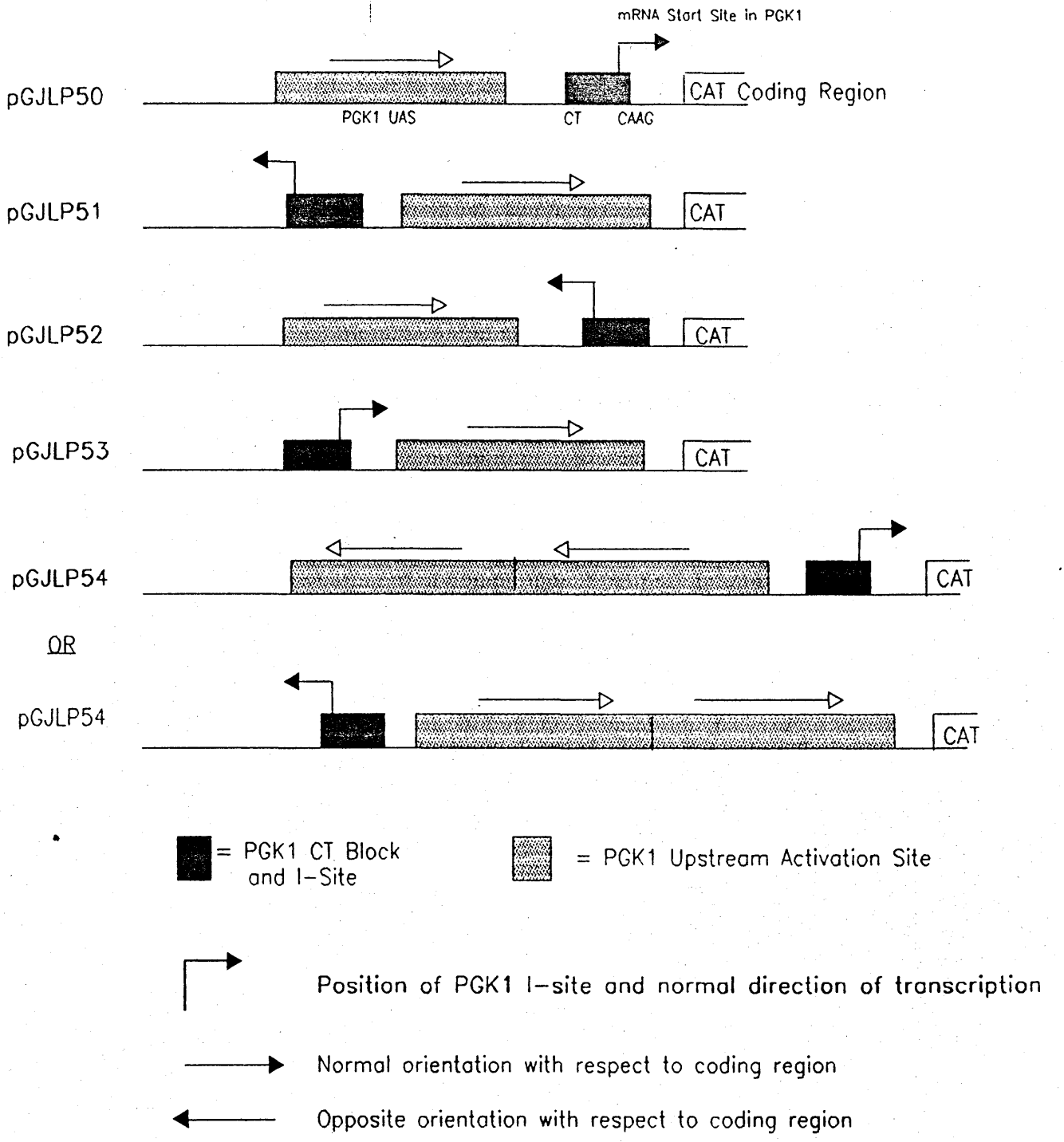


Figure 3.4 (a) Putative Yeast Promoters Containing PGK1 UAS, CT Block and I-Site.
The above is a representation of the putative promoters inserted into the promoter probe vector pCJ19. Note, it has not been determined whether transcription initiates at the sites indicated above. Rather, the position of the normal PGK1 I-site is indicated within the synthetic PGK1 insert. The orientation of the promoter insert in pGJLP54 has not been determined therefore both possibilities are illustrated. The elements are not drawn to scale.

contains a double insertion of the UAS_{PGK}, the orientation of which cannot be determined without sequencing through the insert. However, it is known that the UAS_{PGK} functions in both orientations relative to the I-site. Plasmids pUCTPGK1 to 5 contain putative yeast promoters within the polylinker.
These promoter-containing polylinkers were then subcloned into the promoter probe vector, pCJ19. pUCTPGK1 to 5 were restricted with EcoRI and HindIII, termini were made flush and the promoter-containing fragments were gel purified by the geneclean method. The pCJ19 promoter probe vector was restricted with SmaI and treated with calf intestinal phosphatase. The products of ligations of pCJ19 phosphatase-treated vector and putative promoter-containing inserts were screened by restricting plasmid DNA from transformants with SalI and BamHI. The SalI site within the promoter-containing polylinker was utilised to establish the insert orientation. Five of these products, which were named pGJLP50 to 54, are illustrated in Figure 3.4(a).

3.2.2 Construction of pGJLG1 to 4

To assess whether a non-PGK1 derived UAS would function in conjunction with the PGK1 CT block and I-site in determining efficient transcription pGJLG1, 2, 3 and 4 were constructed.
The pGJLG series of plasmids is identical to the pGJLP50 to 53 series except they lack the UAS_{PGK}, but have in its place the UAS derived from the divergent GAL1/10 promoter (UAS_G; Johnston, 1987a). pBM150 contains the entire GAL1/10 promoter on a 685 bp EcoRI/BamHI fragment. This fragment was purified and further restricted with DdeI and Sau3A. A 365 bp fragment was purified from this restriction digest. This fragment contains the UAS_G but does not contain TATA-like sequences required for normal

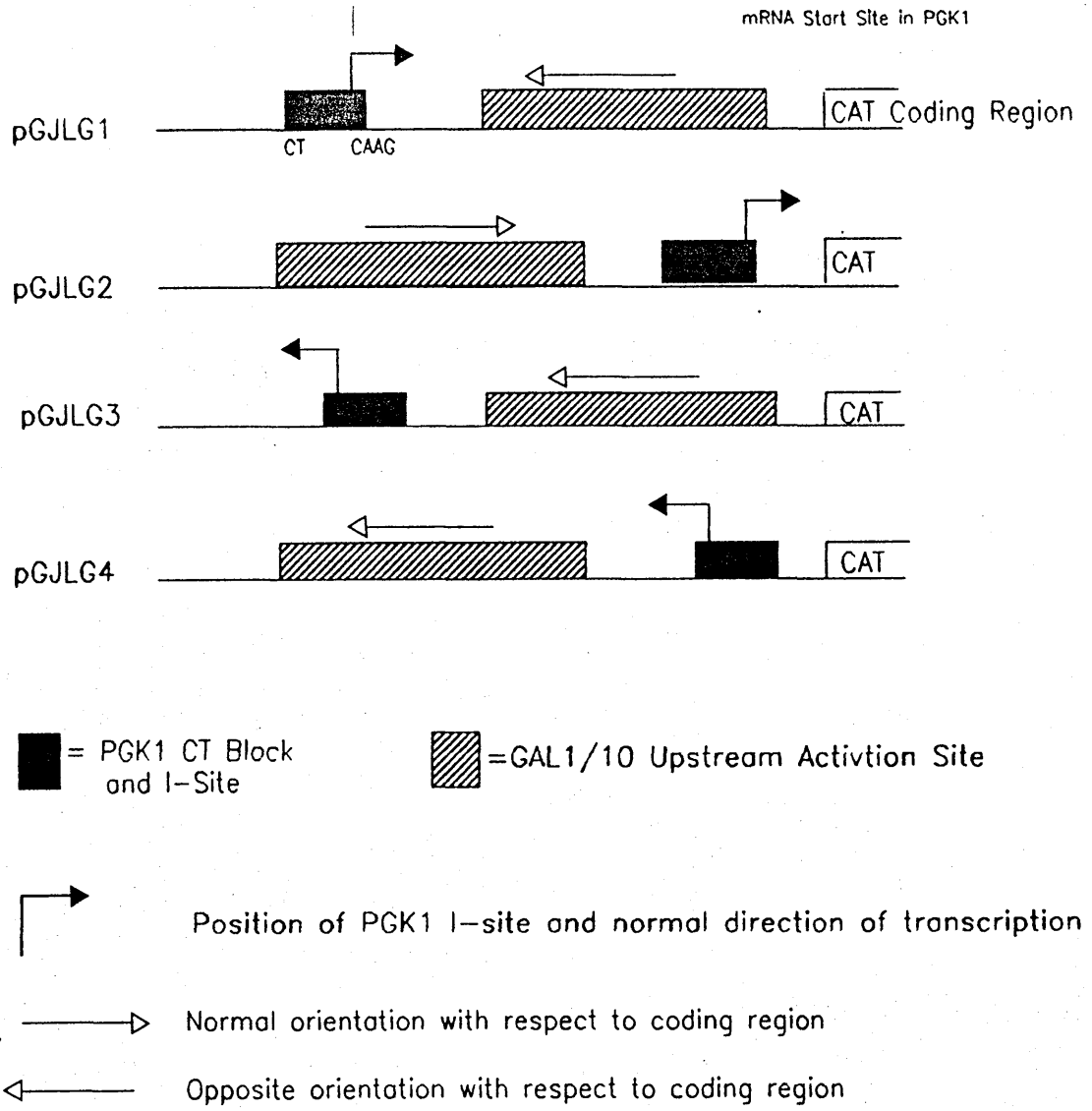


Figure 3.4 (B) Putative Yeast Promoters Containing GAL1/10 UAS and PGK1 CT Block and I-Site
See Figure 3.4 (a) for comments.

transcription of the GAL1/10 genes (West et al., 1984). The overhanging termini of the 365 bp fragment were made flush and then the fragment was gel purified. The fragment was then sub-cloned into pUCT46 and 50. The products of this ligation were called the pUCTG series. The orientation of the insertions was established by BstN1 restriction of the plasmid from transformants. All of the clones examined (70 in total) contained the UAS_G fragment inserted into pUCT 46 and 50 in a single orientation. All previous ligations involving these preparations of vector yielded clones with insertions in both orientations. Therefore, insertion of the UAS_G in the opposite orientation into this particular sequence may be detrimental to the E.coli cell or to plasmid maintenance. There is no obvious reason why this should be so, although it is possible that strong transcription from this fragment may read through into a locus essential for plasmid viability, such as the resistance Ap^R marker gene. The putative promoter containing polylinkers from the pUCTG series of plasmids were subcloned into the pCJ19 promoter probe vector as previously described (3.2.1). The orientation of the polylinker inserts in resultant plasmids was determined by restricting plasmid DNA with PstI and EcoRI. The products of these manipulations (pGJLG1 to 4) are illustrated in Figure 3.4 (b).

3.2.3 Construction of pCJGAL12

The construction of pGJLG and pGJLP series of plasmids has been described above and these plasmids are illustrated in Figure 3.4 and 3.5. For the purposes of comparative analysis an authentic yeast promoter, that from the intergenic GAL1/10 region, was cloned into the promoter probe vector pCJ19. Transcription from this promoter has previously been shown to be tightly

regulated by carbon source (Johnston, 1987a) via the UAS_G. The pGJLG series of constructs also contain the UAS_G, which may impose regulated transcription of the CAT gene. Therefore, it was appropriate to compare the intact GAL1/10 promoter with the hybrid UAS_G /PGK1 I-site putative promoters located in the pGJLG series.
The entire GAL1/10 divergent promoter was cloned into the pCJ19 polylinker. A 685 bp EcoRI/BamHI fragment of pBM150, containing the GAL1/10 promoter, was treated with Klenow fragment and ligated with SmaI restricted pCJ19. Insertion was confirmed by restricting the ligation products derived from subsequent transformants with SalI and BamHI. The orientation of this divergent promoter was not determined as it is known to function in both orientations (West et al., 1984)

3.2.4 Selection of Yeast Strain yGJL661

All plasmid constructs previously described (pGJLG's, pGJLP's and pCJGAL12) contain the TRP1 gene for selection and maintenance of plasmid in a trp1 auxotrophic host. Some of these plasmids also contain the galactose inducible UAS (UAS_G) from the GAL1/10 divergent promoter (pGJLG's and pCJGAL12). UAS_G will only activate transcription in cells expressing the GAL4 trans-activator protein, thus a gal4 mutant fails to grow on galactose as a sole carbon source.
Therefore, there were two requirements for the host strain used in this study. Firstly, it had to be an auxotrophic mutant for trp1, to facilitate selection of the plasmids, and secondly it had to be able to grow on galactose as a sole carbon source, which ensured the presence of GAL4 activator protein. Such a strain was not readily available. However a diploid strain of the following genotype was available;
a/alpha, ura3/ura3, trp1/TRP1, leu2/LEU2, pyk1/PYK1,

HIS4/his4, GAL/gal.
This diploid strain was sporulated and a haploid strain of the following genotype was isolated by the methods outlined in 2.2.3.;
ura3, trp1, leu2, GAL (yGJL661).
This strain (hereafter yGJL661) combined the desired characteristics of a requirement for trp1 complementation and the ability to grow on galactose.

3.2.5 Expression of CAT in Yeast

A number of novel putative yeast promoters have been inserted upstream of a CAT coding region in the promoter probe vector, pCJ19. The resultant plasmids were transformed into the yeast strain yGJL661. The expression of the CAT coding region in these transformants was assessed by measuring chloramphenicol resistance and by measuring the steady state abundances of RNA containing CAT sequence.

(a) Transformants Containing pCJ19 or Derivatives of pCJ19 Are Resistant To Chloramphenicol

Despite the absence of yeast promoter elements upstream of the reporter gene CAT, it has been shown previously that pCJ19 transformants are slightly more resistant to chloramphenicol than the untransformed host (Chris Hadfield, personal communication). This resistance allowed very slow growth (becoming visible only after a 7 day period) on minimal medium containing chloramphenicol at 100 ug/ml. However, Hadfield has clearly shown that insertion of a yeast promoter upstream of the CAT coding region results in a dramatic increase in chloramphenicol resistance. The simplest interpretation of these data was that the CAT coding region in pCJ19 is very inefficiently

transcribed in yeast and that insertion of an efficient yeast promoter greatly increases the rate at which the CAT gene is transcribed. Thus pCJ19 seemed a suitable candidate for use as a promoter probe.

If antibiotic resistance is to be considered an indicator of transcription then the untransformed host strain must be shown to be sensitive to the antibiotic activity. Untransformed yGJL661 was shown to be sensitive to choramphenicol at a concentration of 400 ug/ml by plating on minimal medium plates containing amino acid supplements and various concentrations of chloramphenicol up to 750 ug/ml. yGJL661 was then transformed with pCJ19 or one of the ten pCJ19 derivatives constructed as part of this study. Transformants were initially selected on the basis of complementation of the trp1 auxotrophic marker, not choramphenicol resistance.

Transformants were then patched onto minimal media plates (GYNB) containing the appropriate amino acid supplements and various concentrations of choramphenicol, from 100 ug/ml to 500 ug/ml. Surprisingly, all transformants, including the pCJ19 transformants were shown to be resistant to 500 ug/ml chloramphenicol. This result was unexpected because in the experiments of C.Hadfield (personal communication), pCJ19 conferred only very weak resistance to the antibiotic. In addition to this, transformants of pCJGAL12, which contains the entire GAL 1/10 promoter, proved to be equally resistant to chloramphenicol whether the cells were grown on glucose as the sole carbon source, or galactose. This is also an unexpected result as the GAL1/10 promoter is activated when cells are grown on galactose and is repressed when the cells are grown on glucose. One would therefore predict higher choramphenicol resistance when transformants are grown on galactose rather than glucose. Chloramphenicol concentrations in GYNB plates were increased in an effort to differentiate between the

resistance levels conferred by pCJ19 and those conferred by its derivatives. No differences were observed up to concentrations of 1mg/ml, at which concentration choramphenicol precipitated upon cooling of the agar. There are at least two interpretations of these results; low expression of the CAT gene from pCJ19 is sufficient to confer chloramphenicol resistance on the strain yGJL661, or CAT is expressed at high levels in pCJ19 transformants. To differentiate between these possibilities, the expression of CAT was assessed in transformants containing the unmodified promoter probe or derivatives.

(b) The CAT Gene of pCJ19 is Transcribed Independently of Insertion of Promoters Upstream of the Coding Region

The expression of the CAT gene in transformants containing pCJ19 or its derivatives could be assessed by one of two methods; either by measuring CAT specific activity or by measuring the steady-state abundance of the CAT mRNA. As the suitability of pCJ19 as a transcriptional promoter probe was being assessed at this stage, the steady-state mRNA abundance was measured. CAT assays were thought unsuitable as poor translation may result in the transcription of the CAT gene being underestimated.

Transformants containing pCJ19 or pGJLP series plasmids were grown on GYNB. Transformants containing pGJLG series plasmids were grown on GalYNB. Total RNA was prepared from exponentially growing cultures ($O.D._{600}$ = 0.5) and the integrity of the RNA was checked on non-denaturing gels and by probing Northern blots for specific mRNAs. The abundance of specific mRNAs was measured by dot blot analysis rather than by Northern blots. Dot blots have a

number of advantages over Northern blots in the quantitation of mRNA. For example, it is possible to analyse a wide range of RNA concentrations and RNA samples on a single filter. In addition, by probing a dilution series of RNA concentrations one can assess whether the amount of probe hybridised to a filter is proportional to the amount of RNA loaded. This is absolutely essential for quantitation. Furthermore, dot blotting does not require long transfer times as required in Northern blotting and circumvents errors due to uneven transfer of RNA onto nitro-cellulose membrane. It is also possible to cut out each dot from the filter after autoradiography, and measure the bound radioactivity directly by scintillation counting.

Routinely, 1ug of total RNA was diluted 2-, 3- and 4-fold and all concentrations were dotted out in duplicate onto nitro-cellulose. The filters were then probed for specific mRNAs and the hybridisation to each dot quantified by direct scintillation counting of the dots. Before quantitation, autoradiography was always performed to check for non-specific background radioactivity on each filter. The number of counts per minute (cpm) were then compared to the degree of dilution of RNA loaded on each dot. The example shown in figure 3.5 displays a linear relationship between the amount of probe bound and the amount of RNA loaded onto the filter, over the range studied. Therefore, the amount of radiolabelled DNA probe hybridised to a given dot is representative of the relative amount of specific mRNA in the original sample. However, this is only true if the DNA probe hybridises only to complementary mRNA. Thus, prior to dot blotting, each probe used was shown to hybridise to a single band on Northern blots (exceptions will be discussed individually). As large variations in RNA loading can occur, the abundance of a specific mRNA is always measured relative to a control RNA. A control RNA is one

whose abundance should not vary between samples. These samples are usually derived from different transformants. As one cannot always predict that the expression of a given gene will be identical between transformants, it is preferable to measure the abundance of the mRNA of interest relative to a number of control RNAs. In this study, and in the abundance measurements discussed in chapters 4 and 6, actin mRNA, ribosomal protein L3 (RP1) mRNA and 18s rRNA have been used as controls.

pCJ19 and all derivatives of this plasmid are 2um based. All yeast episomal-based (YEp) vectors vary in copy number between populations (for example, Futcher and Cox, 1984).

If the transcription rate induced by a given promoter is to be determined accurately, then the plasmid copy number must also be measured, in addition to mRNA abundance. This is usually achieved by measuring the number of copies of a plasmid-specific gene present in a total DNA sample, relative to a gene at single copy number on the yeast chromosome, by Southern blotting. This technique, although widely practised, relies upon a similar yield of plasmid, relative to chromosomal DNA, for each transformant.

It has been noted that the *E.coli* Ap^R gene is transcribed in yeast and that the abundance of its mRNA is proportional to the copy number, as measured by Southern analysis, of the Ap^r containing plasmid (IJP Purvis, Unpublished observation). For convenience, plasmid copy number in this study was measured indirectly by measuring the abundance of Ap^R mRNA in each transformant.

When quantifying mRNA by the dot blot method outlined above, it is essential that conditions are used in which the radiolabelled DNA probe hybridises only to the target mRNA. These conditions are usually determined by Northern

analysis, the rationale being that hybridisation to a single band indicates hybridisation to a single mRNA species. Where probing is carried out at high stringency, Northern analysis may also indicate that a number of individual RNA species contain sequences complementary to the probe.

RNA from yeast transformants carrying one of the following plasmids was electrophoresed on a denaturing gel and blotted onto a nitro-cellulose membrane, as described in 2.11.1 ; pGJLG1, pGJLG2, pCJ19 and pGJLP50 to 53. This filter was probed, as described in 2.11.3, with a nick-translated fragment containing only CAT sequences. The purity of this DNA fragment was assessed on agarose gels. The filter was then washed at high stringency (0.5 x SSCP at 55^{o}C) and autoradiographed. The autoradiograph is presented in Figure 3.6.

Perhaps the most striking feature of this Northern blot is that a number of bands are present in each lane. As this filter was washed at high stringency, it is likely that each of these bands represent species of RNA which contain CAT sequences. The approximate length of each RNA is given in the figure legend. Note that there has been no attempt to control for uneven loading of RNA onto the Northern gel, and therefore, the relative intensity of bands between lanes does not necessarily reflect the relative abundances of the mRNA in different transformants. It would appear from this Northern blot that there are a number of RNAs which contain CAT sequences. If we assume that each terminates at the CYC1 transcriptional terminator, then at least one of these transcripts initiates within the Ap^{R} gene. These results are consistent with those obtained by Chris Hadfield (personal communication). These results and the implications of the Northern will be discussed fully in 3.2.7.

Following the dot blotting procedures previously

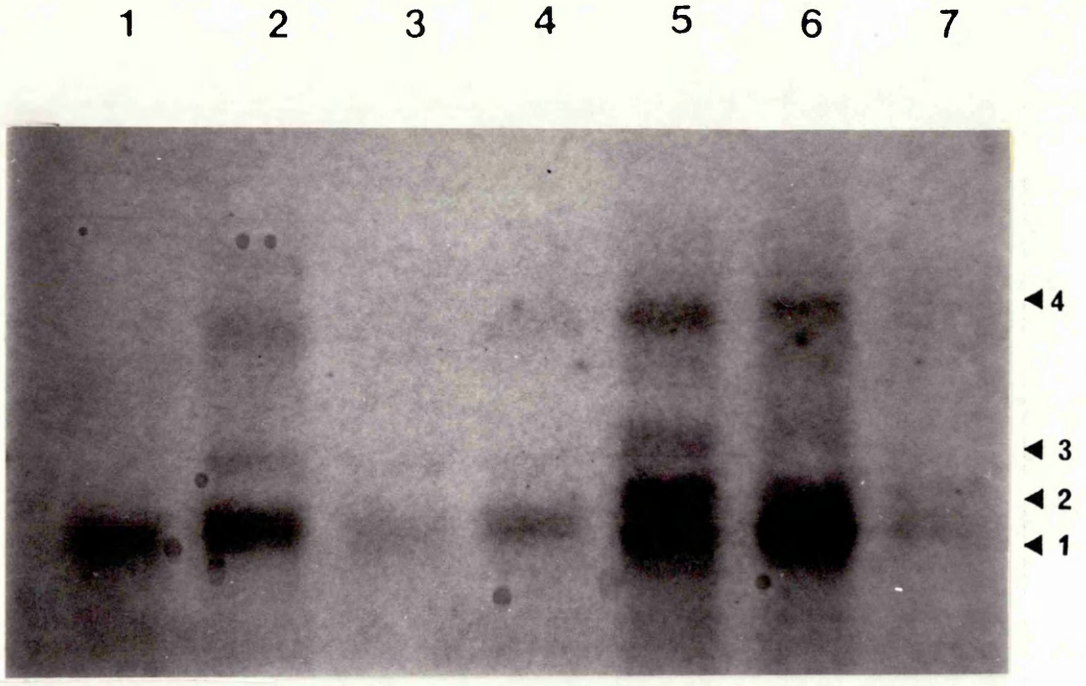


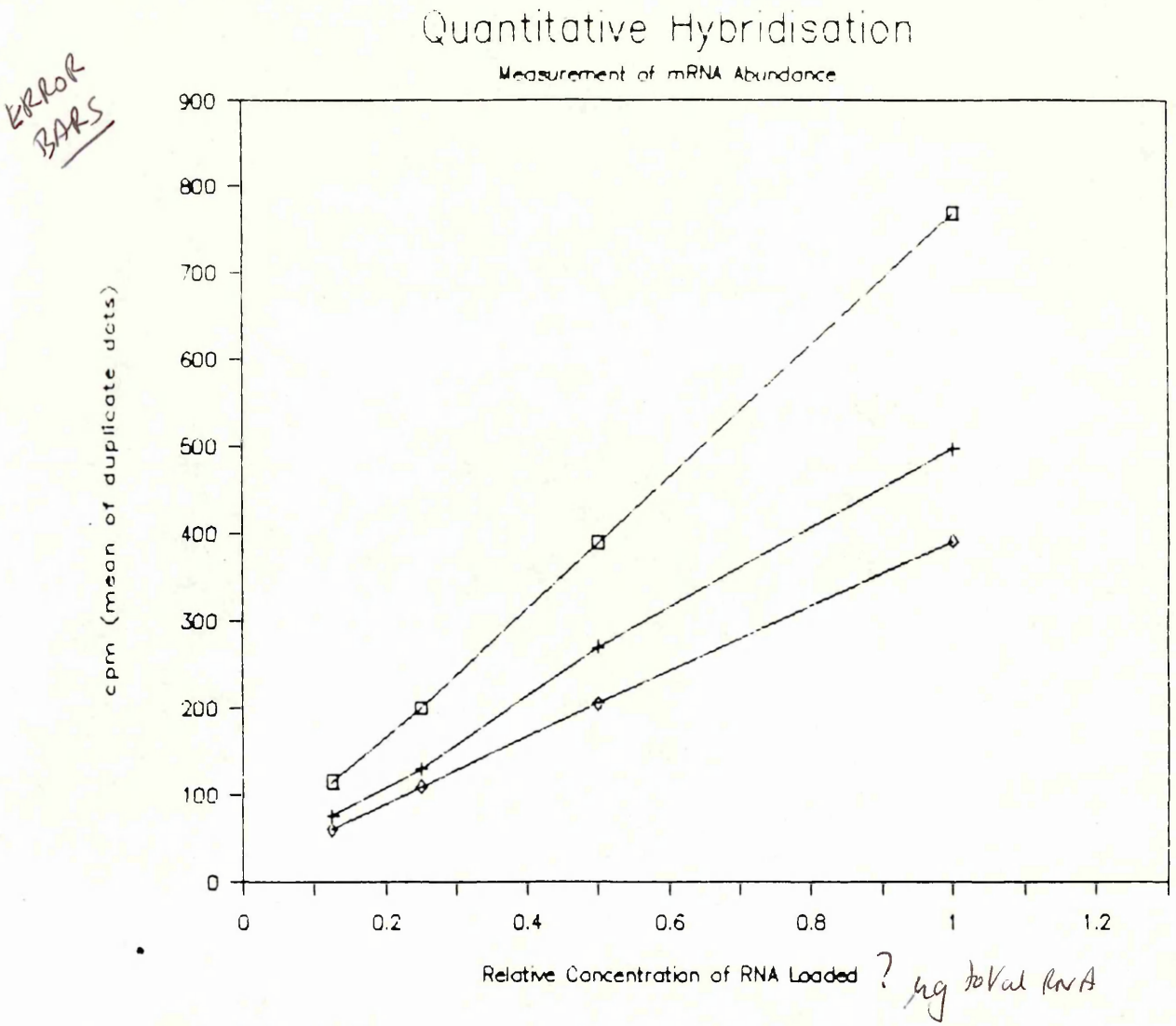
Figure 3.6 Northern Blot Analysis of the Expression of CAT in Yeast

RNA from the following transformants was probed at high stringency (0.5 X SSCP at 55°C) Lane 1: pGJLP50. Lane 2: pGJLP51. Lane 3: pGJLP52. Lane 4: pGJLP53. Lane 5: pCJ19. Lane 6: pGJLG1 Lane 7: pGJLG2.
Approximate sizes of indicated RNA species are as follows; A: 3.5 Kb. B: 2.0 Kb. C: 1.5 Kb. D: 1.0 Kb.

outlined, the abundance of CAT RNA was measured. RNA dot blot filters were probed specifically for CAT, actin and Ap^R RNAs (hybridisation conditions for use with actin and A^R probes as established empirically by IJP Purvis, personal communication). After autoradiography, the radioactivity associated with each dot was counted. The counts per minute (cpm) associated with duplicate dots were averaged and plotted against the dilution factor of the RNA (For example, 3.5). The cpm associated with actin mRNA were used as a loading control for both CAT and Ap^R RNAs. By dividing the Ap^R cpm with the actin cpm, the relative plasmid copy numbers were obtained. Similarly, by dividing the CAT cpm with the actin cpm the relative abundances of CAT RNA were obtained. These relative abundances were then adjusted for plasmid copy number. This generates the relative CAT RNA abundance per gene copy (Figure 3.7). These results indicate that insertion of putative yeast promoters into the promoter probe vector pCJ19 does result in an increase in the steady-state CAT mRNA abundance. However, the most important observation is the abundance relative to actin of CAT RNA in cells containing the unmodified pCJ19 promoter probe vector. This abundance is quantified by comparing the CAT cpms with the actin cpms and taking into account the specific activities and lengths of the probes. Actin mRNA constitutes approximately 0.25% of total mRNA (Purvis, Lithgow, Brown; unpublished observations). The abundance of the CAT RNA in pCJ19 transformants was calculated as >0.5% of total mRNA. A detailed appraisal of these results follows.

3.2.7 Summary and Discussion of Results

The ability of putative yeast promoters to direct transcription of a reporter gene in yeast has been measured. These putative promoters were inserted



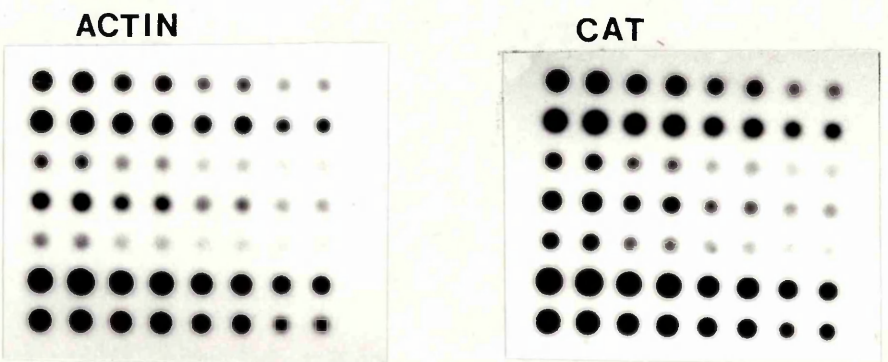


Figure 3.5 Quantifying Steady-State mRNA Abundance by the Dot Blot Method

The above illustrates two filters onto which were dotted identical serial dilutions of RNA isolated from pGJLG and pGJLP series transformants. The identical filters were then incubated with DNA probes specific for actin and CAT mRNAs (see text, 3.2.5). After autoradiography, the amount of probe bound to each dot was determined by scintillation counting. The cpm obtained were then plotted against the RNA dilution factor. This graph shows that hybridisation is quantitatively proportional to RNA loading. The average cpm obtained for each RNA dilution were averaged and plotted against the concentration of RNA loaded on the filter (Three examples are illustrated: Diamonds; PGJLP53, crosses; pJGLP50 and squares; pGJL51).

CONSTRUCT	mRNA Abundance/Gene Copy [Relative to pCJ19]
pCJ19	1.0
pGJLG1	5.1
pGJLG4	6.0
pGJLP50	6.2
pGJLP51	4.2
pGJLP52	5.3
pGJL53	2.6

Figure 3.7 Relative Abundances of *CAT* RNA
The relative abundances of *CAT* RNA were calculated as follows;

{(*CAT* cpm/actin cpm) x CN} / {(*CAT* cpm/actin cpm) x CN}
[Test plasmid] [pCJ19]

where CN is the transformant-specific copy number factor;

CN = Ap^R cpm/actin cpm

Constructs are illustrated in Figure 3.4. The derivation of steady-state RNA abundances is further discussed in 3.2.5.

upstream of a CAT coding region and the resultant RNA abundances were measured. When the CAT gene is expressed, transformants become resistant to the antibiotic chloramphenicol. Unfortunately, the "promoterless" control vector, pCJ19, conferred resistance upon the host strain yGJL661. In order to establish whether this resistance was a consequence of high expression of the CAT gene in pCJ19, the abundance of the CAT containing RNA was measured. Concurrently, the abundance of RNA containing CAT was assessed in transformants containing one of pGJLG1, pGJLG4, or pGJLP50-53. The abundance of CAT RNA in pCJ19 transformants was shown to be 0.5% of total mRNA. The CAT-containing RNA is therefore very abundant, comparative to the abundance of PGK1 mRNA (1%), which is considered to be a highly abundant yeast mRNA.
The CAT gene is transcribed in pCJ19 transformants at a high level despite the absence of known yeast promoter elements in the upstream region of the gene. When RNA from pCJ19 transformants is subjected to Northern blot analysis, at least 4 species of RNA which hybridise to a CAT specific probe at high stringency are revealed. If these RNAs terminate within the CYC1 transcriptional terminator, and no significant RNA processing has occurred, then transcription has initiated at a minimum of 4 sites within the pBR322 (bacterial plasmid) sequences. At least one of these RNAs initiated within the Ap^R gene.
Chris Hadfield (Leicester Biocentre) has recently examined the transcription in yeast of plasmid pCJ18. This plasmid is almost identical to pCJ19, the difference being that the CAT reporter gene has been replaced with APH1, which specifies G418 resistance. Hadfield has shown that the APH1 coding region is transcribed at high levels in yeast. This occurs without prior insertion of yeast promoter elements upstream of the coding region. 5 transcripts with APH1 sequences, of sizes 4.0 Kb, 3.2 Kb,

2.5 Kb, 1.6 kb and 1.5 Kb, are detected in pCJ18 transformants. The total abundance of these transcripts has been measured at 5% (Hadfield, personal communication) which is 5 times greater than the level of highly abundant PGK1. If these transcripts terminate at the CYC1 transcriptional terminator, then transcription initiates at five sites within the pBR322 plasmid sequences. This pattern of transcription is very similar to that described above for pCJ19. Both reporter genes are transcribed at high levels from multiple sites from within the pBR322 derived region and the sizes of the resultant RNAs are comparable (a number of RNAs approximatly 2 to 2.5 kb in length and a larger species of approximatly 4 kb).

Despite the existence of high background transcription, pCJ18 appears to be an efficient "promoter probe" vector, since insertion of a yeast promoter into the polylinker located upstream of the coding region results in a large increase in G418 resistance. Increased resistance does not depend upon increased transcription through the G418 coding region. Therefore, the success of pCJ18 as a promoter probe is likely to be based, not on the level of transcription through the resistance gene, but upon efficient translation of the coding region. The APH1 coding region within transcripts which initiate within pBR322 sequences in pCJ18, may be poorly translated, and so despite high APH1 RNA levels resistance to G418 is low. This could be due to the preponderance of yeast ATG codons 5' to the APH1 gene. When an authentic yeast promoter is cloned at a site immediately upstream of the coding region a transcriptional I-site is defined close to the coding region and within the yeast promoter. Transcripts initiating here may be translated efficiently (due to the absence of ATG codons prior to the coding region) and thus confer G418 resistance.

If the above explanation of pCJ18 behaviour is correct

and the similarities in transcription patterns between pCJ18 and pCJ19 are not coincidental, then we would expect that pCJ19 would have also functioned as a promoter probe in this study. However, pCJ19, without a yeast promoter insert, apparently confers chloramphenicol resistance on yGJL661 and thus cannot be used as a promoter probe. One explanation for this result could be the high intrinsic resistance of the host, yGJL661, to chloramphenicol (growth at 300 ug/ml). Even poorly translated CAT RNA may be sufficient to confer resistance to 2.5 mg/ml chloramphenicol.

The usefulness of the pGJLG and pGJLP series of plasmids in determining the minimal sequence information requirements for yeast transcription should now be reconsidered based on the data outlined in this chapter.

It is not possible from the above data to determine whether any of the novel hybrid promoter inserts elicit a high level of transcription, due to the influence of other elements, probably within pBR322, which promote transcription. The Northern blot implies that transcription of the CAT coding region initiates at sites within the pBR322 sequences. It is therefore difficult to establish whether transcription of pCJ19 derivatives results from putative promoter inserts or from non-yeast sequences.

There are differences in the abundances of CAT RNA per plasmid copy in transformants containing different pGJL-type plasmids (Figure 3.7), but against such a high background transcription rate it would not be prudent to attempt to relate these differences to structural features of the inserted promoters. Such mRNA abundance differences may occur as a consequence of interactions between the inserted yeast promoter elements and the sequences located within the pBR322 region which mediate the high transcription observed in yeast.

Insertion of putative promoters into pCJ19 may have some

affect on the transcription of the CAT gene. Differences occur in the ratio of Northern band intensities between different RNA samples. For example, the CAT transcripts 1, 2, and 4 are nearly all of equal abundance in pCJ19 transformants (see Northern Figure 3.6, Lane 5). Whereas, in all transformants containing pCJ19 derivative plasmids (Lanes 1, 2, 3, 4, 6 and 7) transcript 1 is more abundant than transcripts 2, 3 and 4. This suggests that insertion of putative yeast promoters into pCJ19 causes an increase in the abundance of transcripts initiating close to the CAT coding region relative to the abundance of transcripts initiating further upstream, assuming termination of all transcripts at the CYC1 terminator. Perhaps this reflects genuine promoter activity associated with the putative promoter inserts. However, one cannot infer whether transcription would be observed from these promoters in the absence of background transcription.

The Northern blot result may even have revealed an I-site definition function of the CT block. In pCJ19 there are at least three major start sites generating RNA's of 1.0, 1.5 and 3.5 kb in length. All pCJ19 derivatives containing CT blocks have only one major start site close to this position, as determined by Northerns. It is possible that the CT block (or some other component of the putative promoter inserts) is "focusing" 3 major I-sites into a single I-site. However, there are a number of problems in making such conclusions from Northern analysis. Firstly, Northerns do not indicate I-site position per se, only the length of a specific mRNA. Therefore, I-site changes can only be determined from a Northern if the termination site is constant for transcription from different constructs. Secondly, small alterations in RNA length cannot be detected by Northern blotting. More sensitive methods, such as S1 mapping, are required to detect closely clustered I-sites.

In summary, pGJLG and pGJLP constructs are unsuitable for the analysis of sequence requirements for transcription. Transcription promoted from hybrid yeast promoter inserts cannot be readily distinguished from very high background transcription originating from non-yeast sequences present within the promoter probe vector. Insertion of putative hybrid promoters into pCJ19 does affect the transcription pattern of the plasmid; specifically the number of major start sites is reduced from 4 to 2, as shown by Northern blotting. However, it is not possible to assign these changes to any particular structural features of the putative promoter inserts.
The plasmids used in this study may still provide useful data. To achieve this, the background transcription originating from the pBR322 sequences would have to be prevented from interfering with the expression of the CAT gene and upstream polylinker. This could be achieved by the insertion of a yeast transcriptional terminator upstream of the CAT gene. Convenient restriction sites exist for the insertion of such a terminator at the 5'-end of the polylinker in all pCJ19 derivatives.

During the course of this study, Joy Rathjen and Jane Mellor (Dept. of Biochemistry, University of Oxford) presented the results of experiments carried out to investigate transcriptional initiation at the yeast PGK1 gene (Rathjen and Mellor, personal communication; Rathjen and Mellor, 1988; Mellor, 1989). The consequences of the work, presented by the Mellor group, will now be discussed.
The rationale for the construction of yeast promoters deficient in TATA was discussed fully in 3.1 and in the early part of 3.2. Briefly, Ogden et al (1986) demonstrated that the TATA-like sequences in the PGK1

promoter were superfluous to PGK1 transcription. It was therefore concluded that sequences other than TATA were involved in I-site determination. Attention was focused upon the CT block followed by the sequence CAAG, which is conserved in many highly expressed yeast genes. To determine whether these sequences could facilitate transcription initiation, novel putative yeast promoters were then constructed in the course of this study which contained the PGK1 upstream activation site followed by the PGK1 CT block and CAAG sequence. Rathjen and Mellor have subsequently shown, contrary to the results of Ogden et al., that TATA elements are necessary for the normal transcription of the PGK1 gene. On deletion of TATA most transcripts do not map to the normal initiation site, but initiate heterogeneously at sites downstream of the ATG translational start codon. PGK protein levels (as determined by SDS/PAGE) are markedly reduced in these mutants. Mellor suggests that a cryptic TATA element, located at +30 with respect to the wild type I-site, is activated when the normal TATA element is deleted (Rathjen and Mellor personal communication; Mellor, 1989). It is likely that Ogden et al failed to observe these new start sites because they fall only 7 bp from the end of the probe used in the start site mapping procedures. Thus it is entirely possible that hybridisation of the mRNA to the probe over only 7 bp did not occur or that a 7 bp protected fragment was formed but it may not have been visualised on the polyacrylamide gel.

The TATA element of the PGK1 promoter is therefore involved in defining a single start site for transcription which is consistent with the role of TATA elements in other yeast genes. This knowledge tends to negate the rationale behind the construction of TATA-less promoters. However, further work by Rathjen and Mellor indicates that the PGK1 CT block, or sequences

immediately adjacent to it, may have a role in I-site determination. Experiments were carried out in which the spacing between the CT block and the I-site, and between the I-site and the TATA element, were varied. The results of these experiments indicate that sequences upstream of the I-site and downstream of the TATA element are involved in I-site determination. This sequence may be within, or adjacent to, the CT block. Insertion of "transcriptionally inert" DNA between the CT block and the TATA element results in heterologous I-sites. One interpretation of these data is that the CT block (or adjacent sequences) fix the unique I-site in PGK1, but only within the initiation "window" of a TATA element. The concept of a "window" of TATA element influence is consistent with the observations made on other yeast genes (for example, CYC1), as discussed in 2.1. In addition, the CT block originating from CYC1 has been shown to initiate transcriptional termination (McNeil, 1988). One interpretation of this data is that CT blocks function by preventing the transcriptional readthrough initiated from upstream of the normal I-site. Thus, the only mRNA which is detectable is that which has initiated downstream of the CT block. It is conceivable that the "focusing" activity of the CT block insertions in pCJ19 is a manifestation of transcriptional termination activity. However, the mechanism by which the CT block or associated sequences influence I-site position is unknown.

3.3 Search for Protein Binding at a Yeast I-Site

3.3.1 Introduction

It is clear from the work of Rathjen and Mellor described above, that either the pyrimidine-rich region (CT block) found in many highly expressed yeast genes, or a sequence immediately adjacent to and downstream of the CT block, has a role in I-site determination. The position of the I-site is also dependent upon the TATA element.

Most current models of transcriptional initiation incorporate a specific protein(s) binding close to the I-site. There is good evidence that specific proteins distinct from RNA pol II are indeed involved in I-site determination in higher eukaryotes and in yeast. This is based upon the fact that the TATA element binds a specific protein. A protein, which is present in HeLa cell types has been shown to bind specifically to cellular TATA elements (Parker & Topol, 1984; Sawadogo and Roeder, 1985a and b). In yeast, it is probable that a specific protein also binds to TATA elements within the GAL1/10 promoter (Selleck and Majors, 1987a and b). Recently, a TATA binding activity from yeast has been shown to exhibit properties of the mammalian transcription factor TFIID (Cavallini et al., 1988; Buratowski et al., 1988; Horikoshi et al., 1989).

A specific protein is also found to bind at the I-site of the TRP1 gene. The I-site in this gene is partly determined by two 6 bp direct repeats of the sequence, ACGTGA. The protein which specifically interacts with these sequences, thereby determining the mRNA start site, is thought to be CP1 which is known to also interact with the CDE1 region of yeast centromeres (Jane Mellor, personal communication). Intriguingly, this sequence also appears within the promoter region and coding region of the yeast lipoamide dehydrogenase (LPD1) gene. LPD1 fragments which contain this sequence have been shown to

specifically bind a yeast protein (Joe Ross, PhD. Thesis, Edinburgh, 1988).

The exact mechanism by which CP1 or the TATA element binding protein influence I-sites is unknown but may involve a direct interaction with RNA polymerase II. Alternatively, these proteins may have an influence on nucleosome positioning which in turn determines the accessibility of the DNA to RNA polymerase and other transcription factors.

In the case of PGK1, the CT block may be considered a candidate to bind a specific protein. Such a protein may bind to pyrimidine-rich regions of many highly expressed yeast genes. However no consensus sequence of the sort normally associated with specific protein binding sites is found within CT blocks. At least one precedent exists for a yeast protein which binds sequences of a restricted composition. Winter and Varshavsky (Late Poster, Helsinki; 1988) have identified a protein which specifically binds to (dA).(dT) sequences which had previously been implicated in the promotion of transcription (Chen et al., 1987).

Alternatively, as it is a formal possibility that sequences close to the CT block are responsible for I-site determination, (3.2.7) perhaps a specific protein binds to sequences which occur close to the authentic PGK1 I-site.

The experiments described below were carried out to determine whether the PGK1 CT block, or other possible I-site determining sequences which are positioned between the CT block and the I-site, interact with a specific yeast protein.

Protein/DNA interactions can be detected by a number of techniques. Different techniques yield different information about the nature of the interaction. Most interactions are therefore investigated in a number of ways. The most basic technique is the filter binding

assay, which is based on the principle that double stranded DNA does not bind to nitro-cellulose membrane but protein and protein/DNA complexes do become bound. The DNA is radiolabelled and then incubated with protein, and this mix is then passed through a membrane filter. Complex formation can be quantified by measuring the fraction of DNA bound to the filter.

Perhaps the most revealing technique is *in vitro* and *in vivo* footprinting. "Footprints" are areas of DNA protected from the action of an endonuclease or base modifying activity. Such an area of protection is usually presumed to be the result of binding of a protein to the DNA duplex. Footprinting will be discussed further in Chapter 5.3.

In this study, gel retardation or band shift assays were employed to search for specific protein binding to the CT block and I-site sequences of the *PGK1* promoter. By this method, interactions are revealed due to differences in electrophoretic mobility through non-denaturing polyacrylamide gel of uncomplexed ("free") DNA and protein complexed ("bound") DNA. Migration of a macromolecule through non-denaturing polyacrylamide gels is dependent upon both molecular weight and net charge. A DNA molecule is therefore more likely to migrate slowly if it is complexed to a protein, as a result of an increase in net molecular weight and usually a concurrent increase in positive charge. Migration of DNA is also dependent upon conformation. A bent, or non-linear DNA molecule will migrate more slowly than a linear DNA molecule. Many proteins, on binding DNA, induce bending of the molecule, which in turn causes retardation of migration.

Routinely, labelled or non-labelled specific DNA fragment is incubated, in suitable conditions, with total or nuclear protein extract, partially purified or purified protein. When purified protein is not available, a large

excess of non-specific DNA is included in the binding assay to titrate out many non-specific DNA binding proteins. The binding assay mixture is then electrophoresed under non-denaturing conditions and the gel is stained or autoradiographed to reveal the position of the DNA (see 2.12).

Gel retardation techniques are very suitable for investigating novel protein/DNA interactions as this method allows large numbers of different protein/DNA combinations to be assayed concurrently. Additionally, different types of interaction can be revealed within a single protein/DNA assay mixture. This is because different proteins can retard DNA to different extents. For example, if the DNA molecule in question bound two discrete proteins then two retarded bands may be observed.

The method used in this study closely followed that used by Huet et al. (1985) to reveal specific protein binding to sequences found upstream of genes coding for components of the yeast translational apparatus. This set of experiments was undertaken in collaboration with Dr.I.W.Dawes, of the University of New South Wales (formerly of the University of Edinburgh). Yeast protein fractions were prepared in the laboratory of Dr.Dawes as follows:

Yeast cells collected in the exponential phase of growth (25g) were broken by agitation with 40-mesh glass beads in a homogeniser in 45mls of a buffer containing 200 mM Tris-HCl pH 8; 10 mM $MgCl_2$; 10 mM 2-mercaptoethanol; 1 mM Na_2EDTA; 1 mM PMSF; 10% v/v glycerol and 0.3 M $(NH_4)_2SO_4$. The yeast extract was then centrifuged at 40K rpm for 30 min. The supernatant was removed, diluted with an equal volume of column buffer (20 mM Tris-HCl pH 8; 1 mM Na_2EDTA; 10 mM 2-mercaptoethanol) and applied to a 50ml heparin-Sepharose column equilibrated with column buffer containing 0.1 mM $(NH_4)_2SO_4$ at a flow rate of 1 ml min^{-1}.

The column was washed with column buffer containing 0.1 M $(NH_4)_2SO_4$ until unbound protein had eluted. Protein fractions (4ml) eluted by a salt gradient from 0.1 to 0.75 M $(NH_4)_2SO_4$ were dialysed against storage buffer (20mM Tris-HCl pH 8; 50mM KCl; 1 mM dithiothreitol; 0.2 mM EDTA; 10% v/v glycerol) overnight at 4^0C and were then stored at -70^0C in small aliquots. The protein concentration of the fractions was measured using the Biorad assay according to the manufacturer's instructions. Dr.Dawes kindly supplied protein fractions for use in this study.

By this method putative DNA binding proteins were selected and fractionated using heparin-Sepharose. It was thought that screening of these fractions for DNA binding activity would yield results more readily interpretable than those obtained with total crude protein extract. For example, a single type of protein molecule binding to a single DNA molecule, at a specific site, will result in the formation of a single retarded band. If however, the complexed protein subsequently interacts with other protein species, then a whole series of retarded bands may be formed. Such a pattern of retardation may be difficult to interpret, as perhaps only a single band within the series represents a specific protein/DNA interaction. When protein is eluted from the heparin-Sepharose column, individual proteins are separated, such that non-specific protein-protein interactions are less likely to occur. However, in some circumstances this may be a disadvantage. For example, the DNA binding protein of interest may be unstable when partially purified in this fashion. Alternatively, stable binding of the protein to the DNA may require the presence of a second protein; a stabilising factor. If such a stability factor is absent then very poor specific binding may be observed. It is important therefore to consider a number of different methods before dismissing

the possibility that a DNA element is a specific protein binding site.

The method of Huet et al. (1985) was chosen because it has been successfully applied to the detection of a yeast protein/DNA interaction. A search was made for protein binding to the PGK1 CT block and I-site sequences with this method and the specificity of these interactions was assessed.

3.3.1 A 41 bp Oligonucleotide Containing the PGK1 CT Block/I-Site Does Not Interact With Heparin-Sepharose Purified Yeast Protein.

A 41 bp oligonucleotide containing the PGK1 CT block and I-site (hereafter called CT41 and illustrated in Figure 2.1) was incubated with yeast heparin-Sepharose protein fractions and the resulting interactions were assessed by gel retardation methodologies, as described above and in Chapter 2.12.

In preparation for this experiment single stranded oligonucleotides were hybridised as described in 3.2.1., and then the resulting double-stranded DNA (CT41) was radiolabelled using the end-filling activity of the E.coli polymerase I Klenow fragment (2.10.3). Radiolabelled DNA was purified by phenol extraction and ethanol precipitation. A sample of the radiolabelled material was electrophoresed on a 5% (w/v) non-denaturing polyacrylamide gel. On autoradiography, a single band was observed. Thus, the labelled CT41 DNA was considered to be of sufficient purity for gel retardation assays.

It is highly likely that individual protein species will be present in more than a single protein fraction eluted

from the heparin-Sepharose column. Therefore, labelled CT41 was not assayed with every fraction in the first instance. 16 fractions were assayed for binding activity. Protein free assays were also carried out to control for non-protein fraction associated DNA retardation. All binding assays were carried out under conditions exactly as those described by Huet *et al.* (1985; 2.12). All binding assay mixes were electrophoresed through a 5% (w/v) non-denaturing polyacrylamide gel. This gel was then dried and autoradiographed. The autoradiograph revealed that each lane of the gel contains a single band only, and that this band migrated identically in each lane. There was therefore no evidence of DNA retardation (or "band shifting"). This suggests that under the conditions of this experiment CT41 does not bind a yeast protein.

From this experiment, there is no evidence to support the hypothesis that the *PGK1* CT block or I-site sequences interact with a specific yeast protein.

There are a number of possible reasons why a specific *in vivo* protein/DNA interaction may not be detected by the method outlined above. As discussed previously (3.3.1), a specific protein may be unstable upon extraction from yeast or stable binding may be dependent upon the presence of a second protein factor. Alternatively, the DNA fragment (CT41) may be unsuitable for the formation of a specific DNA/protein complex despite containing a specific binding site. Binding may require non-specific sequences flanking the specific binding site. If this were the case then CT41 may be less than the minimum length required for protein interaction. In an attempt to overcome this possible problem, the 100 bp *Eco*RI/*Hind*III fragment of pUCT50 (Figure 3.1a) was used in a subsequent screen for specific binding to the *PGK1* CT block and I-site.

3.3.3 A 100 bp DNA Fragment Containing The PGK1 CT Block and I-Site Sequences Interacts With Components of A Number of Heparin-Sepharose Protein Fractions

pUCT50 was restricted with EcoRI and HindIII. The products of this restriction were radiolabelled and purified such that only two bands, 2.68 kb and 100 bp, were observed on a non-denaturing polyacrylamide gel. The 100 bp (CT100) fragment comprises the pUC19 polylinker with CT41 inserted into the SstI site. CT41 is orientated such that the I-site sequence, CAAG, is close to the EcoRI site of the polylinker.

A screen for DNA binding activity was carried out in an identical fashion to that described above, substituting CT41 with labelled CT100. Figure 3.8 shows an autoradiograph of the resulting polyacrylamide gel. The contents of the individual binding assays represented by each lane are outlined in the figure legend. Lane 1 represents a protein free control binding assay. The labelled DNA in this binding assay was incubated in an identical fashion to that in all other binding assays. A single band is observed in this lane. Therefore any band retardation observed in other lanes was the result of incubation of the DNA in the presence of a protein fraction.

A number of protein fractions apparently induce retardation of the CT100 band. The pattern of retardation is not identical in all cases, therefore these protein/DNA interactions are specific with respect to the protein fraction. The interactions observed will now be discussed individually.

Band 1, as indicated in Figure 3.8, is common to almost all binding assays. This DNA retardation is therefore not specific to a single or subset of protein fractions. This

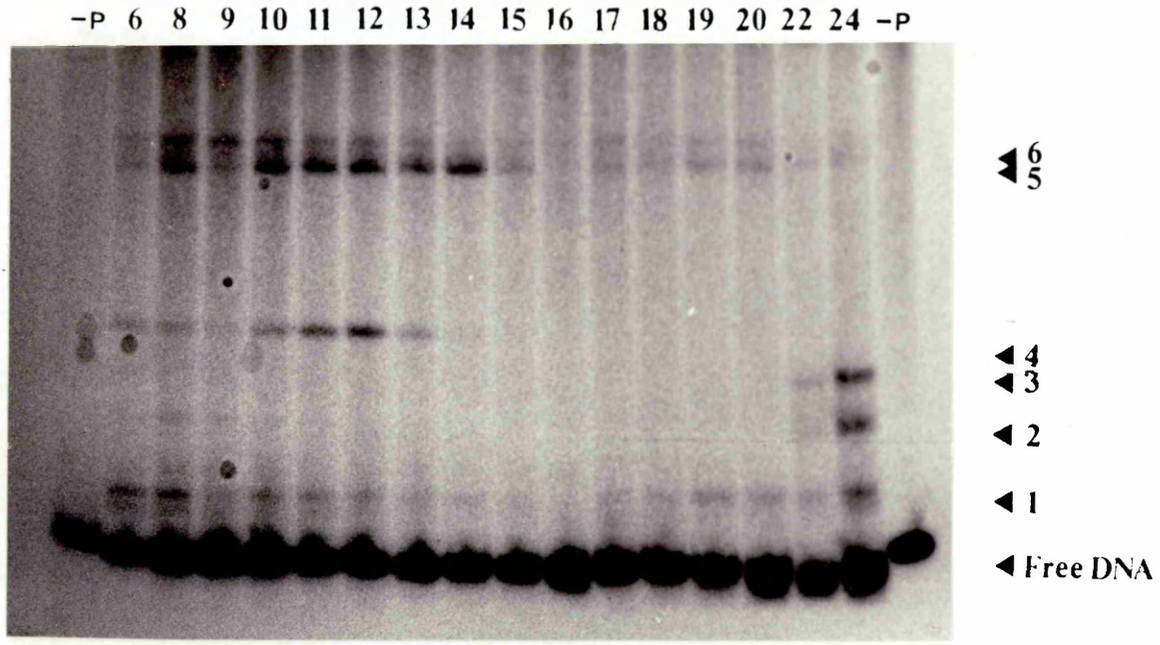


Figure 3.8 Screen for Binding Activity to PGK1 CT Block and I-Site

Unbound (free) CT100 fragment DNA (which contains the PGK1 CT Block and I-site) is indicated. The bands, indicated as 1 to 6, are discussed in 3.4.2. The contents of the binding assay mixes which were electrophoresed on this 5% non-denaturing polyacrylamide gel before autoradiography were as follows;

All Lanes : 110pg [^{32}P]-dCTP labelled CT100 DNA and 500ng of unlabelled phage lambda DNA and 0.7ug of protein. Heparin-Sepharose fraction numbers are indicated. "-p" represent a binding assay in which no protein was added.

suggests that it is caused, not by an interaction with a specific protein species, but by general, non-specific interactions with perhaps a component of all protein fractions. A specific binding protein is unlikely to be present in all heparin-Sepharose protein fractions unless multiple forms of the protein exist.

A component of Fractions 6 to 13 apparently interacts with CT100 generating Band 4. This component could be a protein which eluted from the heparin-Sepharose column at 0.25M $(NH_4)_2SO_4$ to 0.4M $(NH_4)_2SO_4$. The binding activity peaks at fraction 12.

A second retardant is observed using protein fractions 8 to 14. This is revealed by the formation of Band 5 and Band 6. Bands 5 and 6 can just be observed in assays containing protein fractions 17 to 22 but these bands are very faint. Complex formation peaks between fraction 10 and 14. This represents protein eluted from the heparin-Sepharose column at 0.3M $(NH_4)_2SO_4$ to 0.4M $(NH_4)_2SO_4$.

Complex formation was also observed when CT100 was incubated with protein fractions 22 and 24 (Band 2 and Band 3). These fractions contain protein eluted from the heparin-Sepharose column at 0.6M to 0.7M $(NH_4)_2SO_4$.

Complexes described above, represented by Bands 2 to 6, are apparently fraction-specific. That is to say, they are formed, not simply by the general interaction of CT100 with non-specific proteins. Rather, they are formed by the interaction of CT100 with specific proteins.

This fact does not in itself implicate involvement of these proteins with the <u>in vivo</u> function of the CT block and I-site. These specific proteins may be general DNA binding proteins with little or no sequence motif specificity. It is highly unlikely that such a general DNA binding protein would be involved in start site definition. To infer biological significance, a demonstration that one or more of these protein fractions contain a component which binds with a higher affinity to

CT100 than to non-specific DNA is required.

Relative binding affinities for specific and non-specific DNA can be assessed in binding assays where two types of DNA fragment "compete" for association with a specific protein. When interpreting such competition experiments it should be noted that at high concentrations any protein will retard the migration of DNA, and that non-specific DNA if present at sufficiently high concentrations will competitively inhibit the formation of specific interactions. In practice, the ability of specific and non-specific unlabelled DNA to inhibit complex formation between protein and specific labelled DNA is compared.

Figure 3.9 illustrates a "competition" experiment carried out to assess the specificity of the interaction between CT100 and a component of protein fraction 12. The contents of the individual binding assays are described in the figure legend. The possible identity of each individual band is also indicated. Lane 8 represents a protein-free, control binding assay. Lane 7 illustrates the binding activity previously identified in protein fraction 12. Complex I is equivalent to Band 4 indicated in Figure 3.8. Lanes 3 to 6 represent binding assays containing decreasing concentrations of unlabelled CT100 DNA fragment. As little as 10ng of unlabelled CT100 is sufficient to completely abolish formation of complex I. Lanes 7 and 8 represent binding assays in which linear pSP65 plasmid was added as a cold competitor. Complex I formation is not inhibited by the addition of up to 1ug of this plasmid. This indicates that CT100 and pSP65 competitively inhibit complex formation with different efficiencies. This suggests that a component of protein fraction 12 binds to CT100 DNA fragment with a higher affinity (greater than 100-fold) than to non-specific DNA. One could construe this as a reflection of specificity. Thus a DNA fragment containing the PGK1 CT

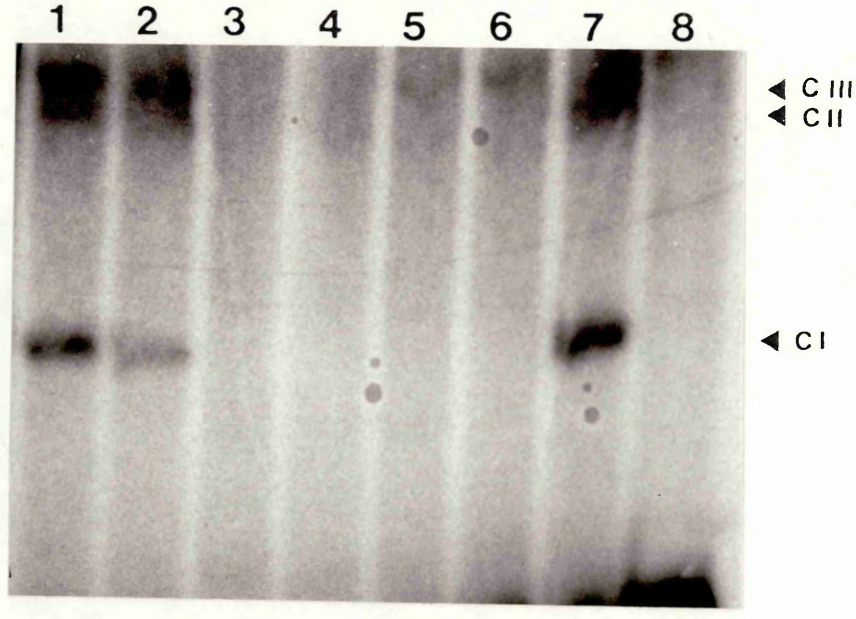


Figure 3.9 Specificity of the Interaction Between a Yeast Factor and a DNA Fragment Containing the PGK1 CT Block and I-Site

Unbound CT100 fragment DNA runs at the very foot of the autoradiograph. Complex I (CI) corresponds to band 4 in Figure 3.8. The contents of the binding assays electrophoresed on this 5% non-denaturing polyacrylamide gel were as follows;

All lanes : 110pg radiolabelled CT100 DNA, 500ng of unlabelled phage lambda DNA and (except lane 8) 2ug of protein (heparin-Sepharose fraction 12).

Lane 1 : 1ug linear pSP65 plasmid.
Lane 2 : 100ng linear pSP65 plasmid.
Lane 3 : 1ug unlabelled CT100.
Lane 4 : 500ng unlabelled CT100.
Lane 5 : 100ng unlabelled CT100.
Lane 6 : 10ng unlabelled CT100.
Lane 7 : No additional DNA.
Lane 8 : No protein.

block and I-site interacts specifically with a yeast protein. This is discussed further below.

3.3.4 Discussion 4

A yeast protein(s) has been shown to interact with a DNA fragment (CT100) which contains sequences found at the PGK1 RNA initiation site. There is evidence that this protein binds with a higher affinity to this DNA than to non-specific DNA. Before biological significance can be assigned to this interaction the exact conditions in which these results were obtained should be considered carefully.

From the competition experiment described above, it is clear that a factor of protein fraction 12 binds to unlabelled CT100 fragment with over a hundred-fold higher affinity than to non-specific plasmid DNA. By this criterion one may suggest that this protein fraction component specifically binds to DNA containing sequences originating from close to the PGK1 I-site. However, the specificity observed may not be the result of selective binding to a specific sequence. CT100 differs from pSP65 in more than just sequence content. Linear pSP65 is 30x the length of CT100, therefore there are 30x more molecules of CT100 than pSP64 in an equal weight.

Consider then a protein which specifically binds to the termini of a DNA duplex. In a competition assay designed to examine the binding specificity of this protein, equal microgram quantities of different types of DNA may be assayed for their ability to competitively inhibit an interaction of this protein with a particular DNA molecule. If one of the competitor DNA molecules is 10x smaller than the other then it will be present in the binding assay in ten-fold greater numbers than its counterpart. As the protein binds selectively to DNA termini, the smaller competitor will be more successful than the larger because the number of DNA termini

associated with a microgram of the smaller DNA will be greater. As a result, the protein in question may be thought to be exhibiting "sequence" specificity when, in fact, the protein binds non-specifically to duplex sequence, but selectively to duplex ends.

A similar set of circumstances may have occurred in the competition experiments designed to assess the specificity of the factor of protein fraction 12 for CT100. Although the protein does clearly exhibit selective binding to this DNA fragment, the basis of selectivity may not be sequence specificity. Regarding a biological role for this component, only sequence specificity is significant.

It is rather straightforward to overcome these difficulties. Firstly, competitors of equal length and if possible, similar conformation (e.g. linear) should be employed. Secondly, the site of protein binding within the specific DNA fragment should be determined. This could be accomplished by footprinting procedures.

Further experiments on the binding of yeast proteins to I-site sequences were hampered by poor stability of the binding activity. Aliquots of protein fractions were stored at -70^{o}C. Samples were thawed immediately before use and then the remainder of the sample was discarded. Despite this, subsequent aliquots exhibited highly diminished binding activity. Faint complexes migrating at the expected rates were observed but were considered unsuitable for competition assay purposes. The same was found to be true for binding activities from protein fractions 22-24, which were previously shown to initiate complex formation with CT100 (Figure 3.8). Perhaps binding activities are more stable in crude extract and therefore the above experiments should be repeated with crude extract replacing heparin-Sepharose protein fractions. For the above reasons, further work is required to confirm specific protein binding to PGK1 I-

site sequences.

Direct footprinting of the PGK1 I-site region may provide a rapid means of revealing whether specific interactions occur in this region. Alternatively, a search could be made for a yeast protein which binds CT rich sequences. (following the procedures for the isolation of a (dA).(dT)- specific binding protein; Winter and Varshavsky, Late Poster Helsinki, 1988).

Establishment of specific binding would provide an assay for the purification of any specific trans-acting factors. Ultimately, the role of such a protein in I-site determination would be revealed by structural studies, perhaps of reconstituted transcriptional initiation complexes.

In summary, components of various yeast protein fractions interact with a DNA fragment containing the PGK1 CT block and I-site sequences. At least one of these components binds specifically to the DNA fragment; however, this may not reflect sequence specificity. Further experiments in this area were undermined by poor stability of binding activities. The results outlined here may represent preliminary evidence that sequences within, or close to the CT block, determine transcriptional start sites by interacting with a specific protein. The identity of this protein is not obvious; It may be a component of RNA polymerase II or an unknown transcription factor.

CHAPTER 4

THE *PYK1* CODING REGION CONTAINS A TRANSCRIPTIONAL ACTIVATOR

4.1 Introduction

A number of reports suggest that, in yeast, not all the elements which determine the rate of transcription, are located upstream of the RNA Initation Site (I-site) (Mellor et al., 1987; Purvis et al., 1987b; Fulton et al., 1988; Company and Errede, 1987). The conclusions of these reports are outlined briefly (Chapter 1.2 and 1.3). It is unclear how, mechanistically, intragenic elements contribute to transcription, or how they may be incorporated into current models of yeast transcriptional regulation. In the following three chapters, experiments are described which were designed to investigate a putative transcriptional activation site located within the coding region of the pyruvate kinase (PYK1) gene. The existence of this element was proposed by Purvis et al., (1987). Before a description of the evidence for the existence of a transcriptional activator within the PYK1 coding region is presented, published evidence relating to a similar system in the PGK1 gene, will be discussed. 3-phosphoglycerate kinase (encoded by PGK1) is a highly abundant protein in yeast, constituting 1 to 2% of total cellular protein. For this reason, the promoter of the PGK1 gene has been utilised in an expression system for the synthesis of heterologous polypeptides in yeast (Hitzeman et al., 1981; Hitzeman et al., 1982; Derynck et al., 1983; Mellor et al., 1983; Tuite et al., 1983; Mellor et al., 1985).

When the PGK1 coding region is replaced with that of the desired protein product, and introduced into yeast on a multicopy vector, heterologous protein yields of around 1 to 2% of total cellular protein are obtained (Mellor et al., 1985). This is a rather poor yield when one considers that when the PGK1 gene is present in yeast on a high copy number vector, PGK constitutes 40 to 80% of total cellular protein (Mellor et al., 1985). Two

independent studies have shown that poor expression of heterologous genes from the PGK1 promoter is due to an unexpectedly low steady-state mRNA abundance (Chen et al., 1984; Mellor et al., 1987). Low mRNA abundance can be due to either low transcription rate or to low mRNA stability or to a combination of both. In the first of these studies, replacement of the PGK1 coding region with that of human interferon-alpha1 (IFN-alpha1) resulted in a reduction in the abundance of the mRNA species transcribed from the PGK1 5'-flanking region. In addition, insertion of the IFN-alpha1 coding region into an intact PGK1 coding region was also shown to cause a dramatic decrease in the steady state mRNA abundance. The authors concluded that a cis- or trans-acting element had been deleted or inactivated in the construction of these heterologous transcription units. To determine whether this element functioned in cis or in trans a plasmid was constructed which contained both an uninterrupted PGK1 gene and a PGK1/IFN-alpha1 hybrid gene, and the relative steady state mRNA abundances were measured. PGK1 mRNA was found to be 20 to 40 times more abundant than PGK1/IFN-alpha1 mRNA, showing that the defect in hybrid gene expression is not alleviated by the over-expression of the PGK1 gene in trans. Although mRNA stability data was not presented in this publication, a model was proposed in which the PGK1 protein enhances steady state mRNA abundance by stabilising, in cis, the PGK1 mRNA. In the second of these studies, Mellor et al., (1987) demonstrated that replacement of the PGK1 coding region with that of IFN resulted in a 6-fold reduction in the steady state mRNA abundance associated with the PGK1 "promoter". Insertion of IFN into the PGK1 transcription unit, but outwith the coding region, did not reduce mRNA abundance. This suggests that the presence of the IFN gene did not, in itself, lead to decreased mRNA abundance, and indicates a reduced rate of transcription.

To map the location of the element required for normal PGK1 mRNA abundance, deletion analysis of the PGK1 coding region was then carried out. Measurement of transcription rates and mRNA stabilities indicated that sequences located between 37 bp and 236 bp of the PGK1 coding region were required for a high rate of transcription. Therefore, Mellor, et al. proposed that the PGK1 coding region contains a "Downstream Activator Sequence (DAS)" which contributes to wild type rates of transcription. The results of Mellor et al. are incompatable with the model proposed by Chen et al. based on stabilisation of the PGK1 mRNA by "protein feedback" as Mellor et al. failed to observe any gross changes in mRNA stability when the PGK1 coding region is replaced.

In 1987 Chen et al. reported that replacing the yeast PGK1 coding region with the human PGK1 coding region resulted in a relatively minor decrease in mRNA abundance (30% reduction). Human and yeast PGK1 share 35% similarity at the nucleotide level, 65% similarity at the amino acid level and a conserved three-dimensional protein structure. The authors suggested that this observation supports the hypothesis that the PGK protein itself stabilises its own mRNA. Presumably, the human PGK1 coding region contains information which contributes to transcription in yeast. As protein structure is more highly conserved than the DNA sequence it was thought mRNA stability was conferred by information contained within the protein. However, no mRNA stability data was presented to substantiate the "protein-feedback" model.

In summary, prior to this study, two models had been proposed to explain how information present within the PGK1 coding region contributes to PGK1 mRNA abundance. Direct measurements of transcription rate and mRNA stability strongly suggest that the coding region activates transcription rate, and therefore, may contain a DAS. However, it is surprising that the human PGK1

coding region, apparently contains information which also activates transcription, despite the fact that no significant sequence motifs appear to be shared between the human and the yeast gene.

Results obtained in this laboratory indicate that the PYK1 coding region may also contain information pertinent to mRNA abundance. The experiments leading to this conclusion are now described.

Ian Purvis et al. (1987), whilst investigating heterologous gene expression in this laboratory, replaced the PYK1 coding region with the coding region of the bacterial B-galactosidase gene (LacZ) and introduced this hybrid transcription unit into yeast, on a centromeric(CEN)-based plasmid (YCpPKZ1; called YCpPKG2 in Purvis et al., 1987). This construct is illustrated in Figure 4.1

The steady-state abundance of the resulting PYK/LacZ fusion mRNA, was observed to be thirty-fold lower than that of the chromosomally derived PYK1 mRNA, within the same cells. In addition, Purvis measured mRNA stabilities on total RNA and found that the chromosomally derived PYK mRNA was twice as stable as the plasmid derived PYK/LacZ mRNA (Figure 4.2). This result shows that the observed thirty-fold reduction in mRNA abundance is not wholly due to a reduction in mRNA stability.

There are a number of interpretations of the above data. Replacement of the PYK1 coding region with that of LacZ apparently causes a reduction in PYK1 mRNA abundance. mRNA stability measurements suggest that this reduction in mRNA abundance is the consequence of something other than just a reduction in mRNA stability. It is conceivable that the PYK1 coding region contains information essential for the efficient transport of the mRNA from the nucleus to the cytoplasm. The absence of this information may result in sequesteration of the mRNA within the nucleus. This in turn, may lead to the mRNA

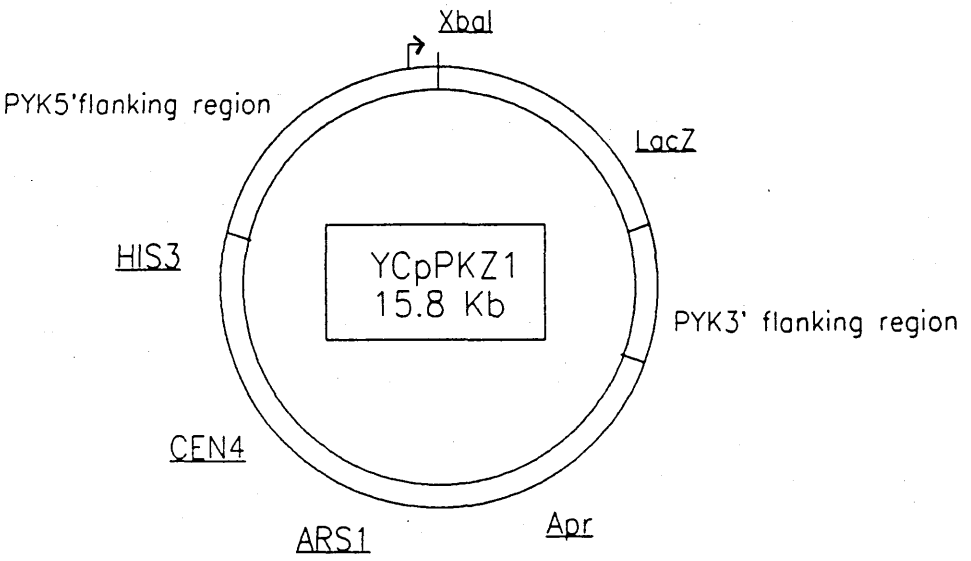


Figure 4.1 Control Plasmid YCpPKZ1
The construction of this plasmid has been previously described (Purvis et al., 1987; this plasmid is called YCpPKG2 in this publication).

	mRNA Abundance	mRNA Half-Life
PYK1/LacZ [YCpPKZ1]	0.04	27.4
PYK1 [chromosomal]	0.61	57

Figure 4.2 Steady-State Abundance and Stability of Chromosomally Derived PYK1 and Plasmid Derived PYK1/LacZ mRNA.
This data is taken from Purvis et al. (1987).

being degraded at a higher rate. However, as Purvis et al., isolated total yeast RNA, such differential decay would be revealed by mRNA stability measurements. In addition, YCpPKZ1 transformants exhibit B-galactosidase activity and hybrid PYK/LacZ mRNA has been found to be associated with polysomes (Purvis et al., 1987), suggesting the mRNA is exported from the nucleus and translated.

Alternatively, the mRNA abundance and stability meaurements may indicate that the rate of transcription from the PYK1 promoter decreases when the PYK1 coding region is replaced with a heterologous coding region. This could be due to either the introduction of sequences which repress transcription or the removal of sequences which, in some way, enhance or activate transcription.

Purvis et al., (1987), inferred from the above results that replacement of the PYK1 coding region with the LacZ coding region results in a fifteen-fold reduction in transcription rate. These results are consistent with those of Mellor et al., (1987) on the transcription of PGK1. Replacement of the PYK1 coding region causes a reduction in mRNA abundance which is not associated with a reduction of mRNA stability. PYK1 may therefore contain a DAS.

Although the terminology "Downstream Activation Sequence" (DAS) is useful in describing the information contained within the PYK1 coding region contributing to transcription, it is perhaps unsuitable for the following reason; this termnology may imply that a DAS is analogous to a UAS (Upstream Activation Sequences) which function by binding specific regulatory proteins. At the outset of this study there was no evidence that PYK1 coding region sequences functioned in this manner. Additionally, a "Downstream promoter element" may be construed as being an element located close to the I-site: Some authors have classified TATA elements as downstream sites. For this

reason, the term "transcriptional activator" will be commonly used in this study in preference to "DAS". Even this terminology is not ideal, because it implies that the transcriptional activation is carried out directly by the PYK1 coding region. Formally, "Transcriptional activator" should be taken to mean, "that information which contributes, directly or indirectly, to transcription".
The following experiments were carried out to confirm that the PYK1 coding region contains a transcriptional activator and to locate sequences associated with this activation.

4.2 The Presence of PYK1 Coding Region Fragments Within The PYK/LacZ Transcription Unit Elevates mRNA Abundance By Transcriptional Activation

With a view to delimiting the sequences required for transcriptional activation, Purvis constructed a series of plasmids of the form shown in Figure 4.3. These constructs contain the LacZ coding region inserted between the PYK 5'-untranslated region and the PYK 3'-untranslated transcriptional termination region (as described in Purvis et al. 1987) . All sequences up to 3Kb upstream of the PYK1 coding region, including the proposed UAS's at -600 to -900, are present. This "control" hybrid transcription unit was modified by the insertion of PYK1 coding region fragments of progressively smaller size. These insertions are in frame, such that translation of the hybrid mRNA should produce a PYK/B-galactosidase fusion protein.
Yeast transformants containing these transcription units exhibit B-galactosidase activity on plates containing X-gal (colonies appear blue). The resulting transcription units were introduced into yeast on a yeast/E.coli shuttle vector containing the CEN4 and ARS1 sites for

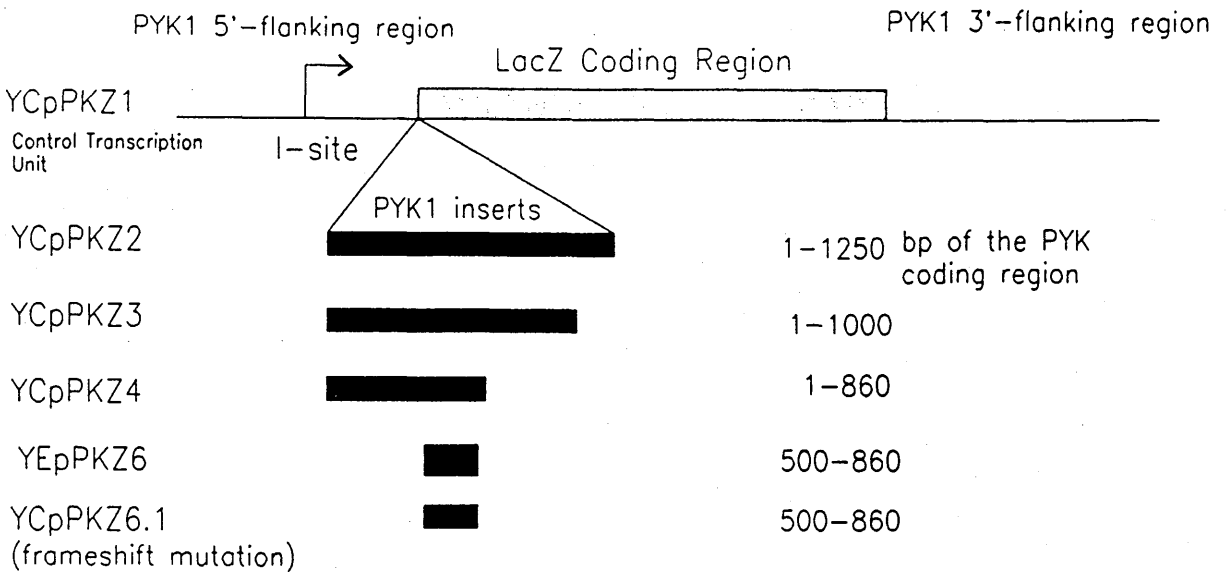


Figure 4.3 PYK1 Coding Region Deletion Series.
Fragments of the PYK1 coding region were inserted into the control plasmid YCpPKZ1 (Figure 4.1). All insertions are in frame with the exception of YCpPKZ6.1. The construction of YCpPKZ1 is discussed in Chapter 6.

propagation in yeast at approximately single copy, and the HIS3 gene for selection of transformants on minimal media. The resulting plasmids were called YCpPKZ1 to 4 and are illustrated in Figure 4.3.

There are two advantages of introducing these chimeric transcription units into yeast on centromeric (CEN) based plasmids. Firstly, to assess the influence of PYK1 coding region fragments on transcription, the mRNA abundance "per gene copy" should be measured. As CEN based plasmids propagate in yeast at 1 to 2 copies per cell (Clark and Carbon, 1980; Clark and Carbon, 1985), errors inherent in the measurement of multiple plasmid copies can be eliminated. Secondly, transformants containing episomal based plasmids exhibit copy number heterogeneity between transformants (Futcher and Cox, 1984). Multiple copies of a regulatory region may titrate out trans-acting transcription factors (Irani et al., 1987), therefore different plasmid copy numbers may sequester transcription factors to different extents. Thus the observed transcription rate per gene copy may appear to vary between transformants containing identical plasmids. CEN based plasmids, being approximately single copy, should give consistent results in different transformants.

The methods by which mRNA abundance was measured in this study, are outlined in detail in 3.2.5.

Briefly, total RNA was prepared from exponentially growing cultures (A_{600} = 0.5) and it's integrity was checked by electrophoresis through non-denaturing agarose gels and by probing Northern blots for specific mRNAs. Northern blots were also carried out to ensure that, under the conditions used, the radiolabelled DNA probes hybridised to the appropriate specific RNA species. The abundance of specific mRNAs was measured by dot blot analysis. Routinely, 1ug of total RNA was diluted 2, 3 and 4 fold and all concentrations were dotted out in

duplicate onto nitrocellulose membrane. The filters were then probed for specific mRNAs and the radioactivity in each dot measured by autoradiography and direct scintillation counting of the dots. The amount of hybridisation, as represented by cpm, was compared to the amount of RNA loaded. The data was used in mRNA quantitation only if the cpm were directly proportional to the amount of RNA loaded. As large variations in RNA loading occur, the abundance of a specific mRNA was always measured relative to internal loading controls such as actin, ribosomal protein L3 and 18s rRNA.

4.2.1 Hybrid Genes Are Transcribed In The YCpPKZ-Containing Series Transformants

Yeast strain DBY746 was transformed with YCpPKZ1, YCpPKZ2, YCpPKZ3 or YCpPKZ4 (Figure 4.3).
The B-galactosidase coding region on YCpPKZ1 has previously been shown to be expressed in yeast (Purvis *et al.*, 1987). Similarily, transformants containing YCpPKZ2 -3 and -4 exhibited B-galactosidase activity on minimal media plates containing X-gal (colonies appeared blue) thus confirming both transcription and translation of the hybrid gene.
Transcription of the *PYK/LacZ* genes was further confirmed by probing Northern blotted RNA (Figure 4.5) with a nick-translated plasmid (pMC1856, Casadaban *et al.*, 1983) which contains the *LacZ* gene. The expected hybrid mRNA lengths estimated from *PYK1* insert size are as follows: PKZ1 - 3.2Kb, PKZ2 - 4.5Kb, PKZ3 - 4.25Kb, PKZ4 - 4.0Kb mRNA species of the predicted lengths were observed and are illustrated. Minor hybridisation to another two mRNA species can be observed. At least one of these RNAs almost certainly contains *Ap*^{R} sequences (the probe, pMC1856 also contains *Ap*^{R}).
It should be noted that the differences in relative band

intensities between lanes, are not due to differences in hybrid mRNA abundance between transformants (as quantified by dot blot analysis Section 4.2.2) but are due to RNA loading differences.
This Northern blot confirms that the RNA extracted from YCpPKZ transformants is intact and that the probe used hybridises predominantly with the hybrid PYK/LacZ mRNA. When PYK1/LacZ mRNA was quantified, dot blots were probed with a DNA fragment containing LacZ sequences only.

4.2.2 Fragments of PYK1 Coding Region Are Associated With Elevated mRNA Abundance

Figure 4.4 (a and b) illustrates the results of an experiment designed to delimit the sequences located within the PYK1 coding region required to activate transcription. In this experiment the steady state mRNA abundance derived from each plasmid was measured in cells grown exponentially on 2% glucose medium.
To ensure that the observed variations in mRNA abundance are related to structural differences between constructs, measurements were made on three separate transformants for each plasmid.
The abundance of the fusion PYK/LacZ mRNA was assessed in each transformant by probing dot blot filters with a probe specific for the LacZ-containing mRNA (the probe was a LacZ coding region fragment purified from pMC1856). Hybridisation to Ap^R transcripts, as probably observed in Figure 4.5, should not occur using this probe. PYK/LacZ mRNA abundances were derived relative to the internal control mRNAs, actin and 18s rRNA (18s rRNA results not shown). These controls are based on the assumption that

Figure 4.4 **Activation Factors Associated With PYK1 Coding Region Fragments**
Activation factors are derived as follows;

(LacZ cpm/actin cpm) / (LacZ cpm/actin cpm)
[test plasmid] [YCpPKZ1]

cpm are derived from dot blot data as follows; cpm for each duplicate dilution are averaged and plotted against RNA dilution. Providing cpm are directly proportional to concentration of RNA, cpm associated with two RNA dilutions were averaged and used in the above activation factor determination.
An activation factor is an expression of the abundance of PYK1/LacZ mRNA in test transformants relative to that measured in YCpPKZ1 transformants. It is not an expression of percentage of total mRNA.
In (a) the average activation factors associated which each plasmid construct is noted.
In (b) the activation factors calculated for individual transformants is noted. The abundance of PYK1/LacZ mRNA was measured in 3 individual transformants of YCpPKZ1, YCpPKZ3 and in 2 individual transformants of YCpPKZ2 and YCpPKZ4. As the abundance of PYK1/LacZ mRNA in all test transformants is measured relative to YCpPKZ1 appears as a single measurement in the the Figure.
In (c) the constructs are outlined. Location of fragments relative to the PYK1 translational initiation codon are indicated.

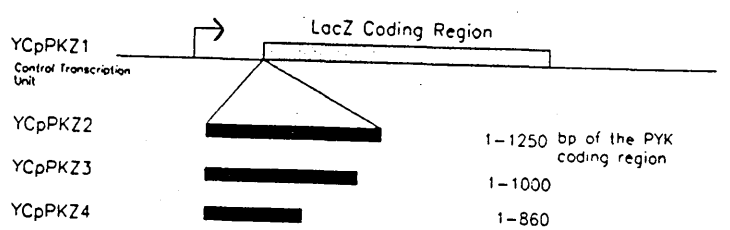
A

CONSTRUCT	PYK1/LacZ mRNA (Activation Factor)
YCpPKZ1	1.0
YCpPKZ2	5.3
YCpPKZ3	7.0
YCpPKZ4	5.7

B

CONSTRUCT	PYK1/LacZ mRNA (Activation Factor)
YCpPKZ1	1.0
YCpPKZ2	6.0 4.5
YCpPKZ3	8.0 7.5 5.5
YCpPKZ4	4.5 6.7

C



actin mRNA and 18s rRNA abundances do not vary between transformants.

Differences in PYK/LacZ mRNA abundances between strains carrying the control plasmid YCpPKZ1, and strains carrying other plasmids, are represented by "Activation Factors". For example, as shown in Figure 4.4(a) the plasmid YCpPKZ3 is associated with an activation factor of 7. This means that there is a seven-times higher steady-state PYK/LacZ mRNA abundance in strains carrying YCpPKZ3 than in strains carrying YCpPKZ1 (Figure 4.4a shows the mean activation factor associated with each construct).

By comparing the activation factors associated with each construct across the full range of transformants, the influence of PYK1 coding region fragments upon the abundance of the hybrid mRNA can be deduced. Constructs containing PYK1 coding region fragments (YCpPKZ2, YCpPKZ3 and YCpPKZ4) are associated with PYK/LacZ abundances significantly higher than that of the control plasmid YCpPKZ1 (up to 7-fold higher). This demonstrates that insertion of PYK1 coding region fragments into the PYK/LacZ transcription unit, downstream of the transcription and translation start sites, results in an increase in PYK/LacZ mRNA abundance. The reduction in mRNA abundance observed by Purvis et al., (1987), was therefore not the result of repression of transcription by sequences contained within the LacZ coding region. Rather it was due to the deletion of sequences contained within the PYK1 coding region which contribute to high mRNA abundance.

This experiment delimits the sequences required for transcriptional activation to within the first 800 bp of the PYK1 coding region.

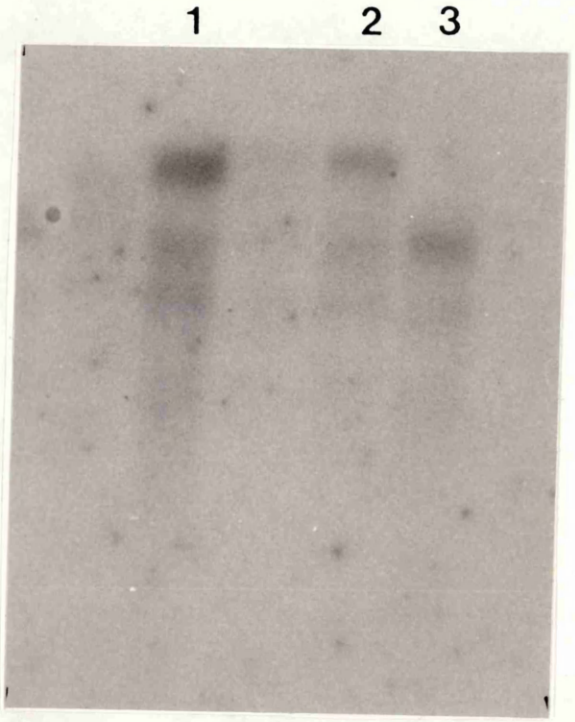
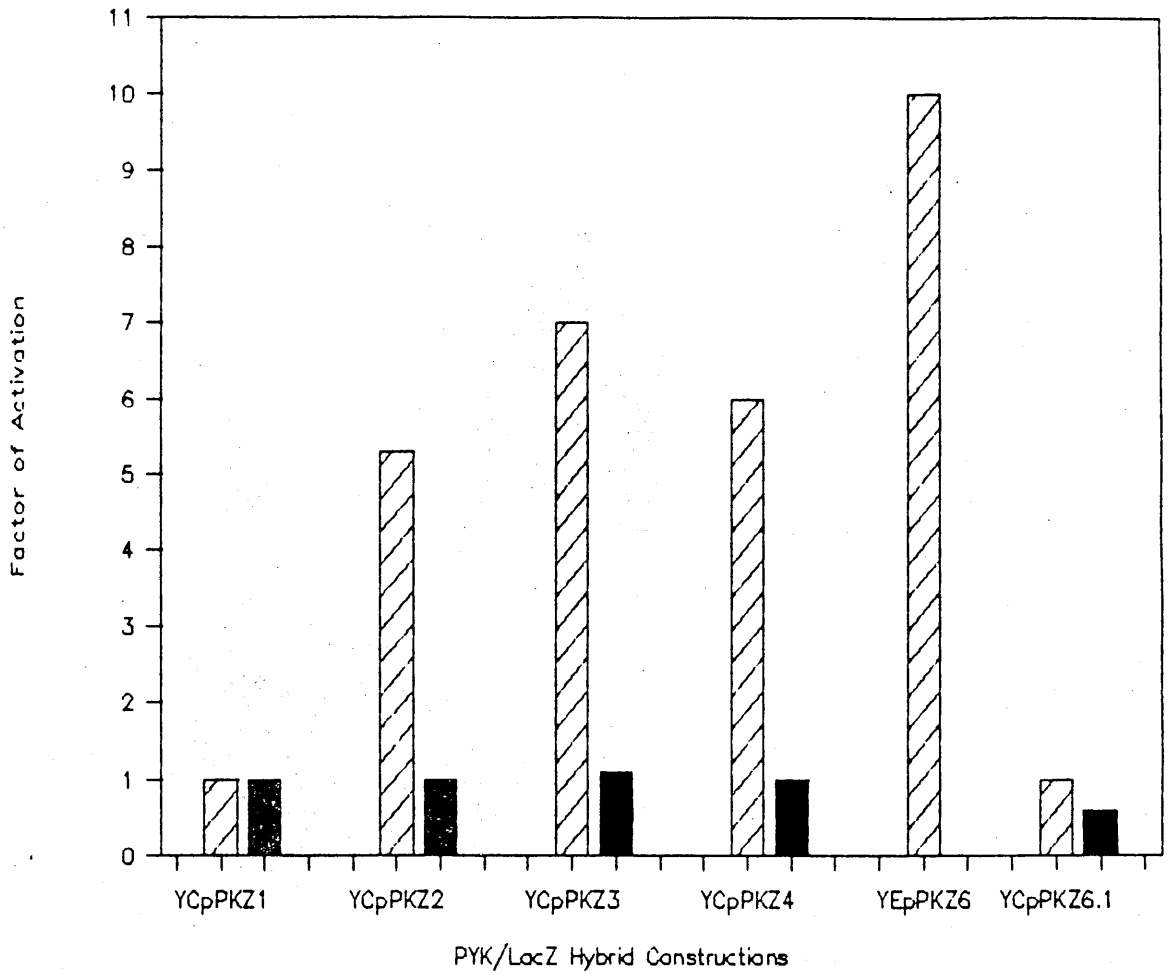


Figure 4.5 Northern Blot of PYK1/LacZ mRNA in Yeast.
RNA from the following transformants was northern blotted, probed with nick-translated pMC1856, washed with 0.5 x SSCP at 55°C and then autoradiographed.
Lane 1: YCpPKZ3. Lane 2: YCpPKZ1. Lane 3: YCpPKZ6.1
On longer exposure of this autoradiograph, bands coresponding to mRNA derived from YCpPKZ2 and YCpPKZ4 transformants can be visualised. This Northern is discussed in 4.3.

CONSTRUCT	PYK1/LacZ mRNA (Activation Factor)	
	[Glucose]	[Lactate]
YCpPKZ1	1.0	1.0
YCpPKZ2	5.3	1.0
YCpPKZ3	7.0	1.1
YCpPKZ4	5.7	1.0

Figure 4.6 Carbon Source Dependence for Activation by the PYK1 Coding Region.
(a) Activation factors represent the abundance of PYK1/LacZ mRNA relative to that measured in YCpPKZ1 transformants (see legend Figure 4.4). Transformants are grown upon 2% (w/v) glucose of 2% (w/v) lactate.



(b) This histogram depicts the activation factors associated with the PYK1/LacZ coding region in various transformants grown on 2% glucose (striped boxes) or 2% lactate (filled boxes).

	Glucose	Lactate
PYK1/LacZ mRNA [YCpPKZ2 Transformants]	0.67	0.55

Figure 4.7 Carbon Source Dependant mRNA Abundance in YCpPKZ2 Transformants
Steady-state mRNA abundance is expressed as a percentage of total mRNA. Actin mRNA is consistently measured as 0.25 x PGK1 mRNA abundance. PGK1 mRNA constitutes 1% of total mRNA (see 4.1). The abundance of PYK1/LacZ mRNA was therefore calculated as follows;

(PYK1/LacZ cpm/actin cpm) x 0.25

Note that this applies only when the DNA probes are of similar length and specific radioactivity.

4.2.3 Elevation of The PYK/LacZ mRNA Abundance by the PYK1 Coding Region is Carbon Source Dependent

The expression of glycolytic enzyme genes may be co-ordinately regulated in response to carbon source (Maitra and Lobo, 1971; Cohen et al., 1986; Nishisawa et al., 1989; See Introduction 1.3). In the case of at least one gene, this regulation is mediated at the level of transcription (Cohen et al., 1987). It is possible that any mechanism which contributes to the expression of a glycolytic enzyme gene may be subject to carbon source regulation. The following experiment was undertaken to assess whether the elevation of hybrid mRNA abundance by PYK1 coding region fragments was dependent upon the nature of the carbon source.

The previous observations (4.2.2) were made on cells grown on the fermentative carbon source, glucose. This experiment was repeated on cells grown to an identical Abs. on 2% lactate, a non-fermentative carbon source (Figure 4.6). mRNA abundances were measured as previously described (Figure 4.4).

No differences were observed in the mRNA abundance associated with each construct, including the control construct YCpPKZ1. This contrasts with the results obtained when these transformants were grown on 2% glucose. The carbon source appears to in some way influence activation by the PYK1 coding region.

An additional observation was made from these data. Figure 4.7 shows the PYK/LacZ mRNA abundances in transformants carrying YCpPKZ2. The abundance of this mRNA relative to actin mRNA is similar in cells grown on lactate and cells grown on glucose. This suggests that transcription of the hybrid PYK/LacZ gene is not induced by growth on glucose. This result is perhaps surprising

when one considers that PYK specific activity is induced 20-fold by glucose (Maita and Lobo, 1971). However, there is no evidence in the literature which demonstrates that the observed induction is due to increased transcription of PYK1.

The observation of constitutively high expression of the PYK1 gene may be artefactual. Actin mRNA abundance may also be induced by glucose or the lactate containing media may be contaminated with a low level of glucose, which is sufficiently high for glucose induction. In addition, there is evidence to suggest that glucose induction of glycolytic enzymes may be highly strain dependent and is perhaps influenced by undefined parameters of the culture conditions (see Introduction 1.3).

Circumstantial evidence supporting the observation that transcription of PYK1 is constitutive is outlined in Chapter 7.

4.3.4 Elevation of PYK/LacZ mRNA abundance by the PYK1 Coding Region is Independent of mRNA Stability

The observed elevation in mRNA abundance (4.3.2) could be due to either an increased rate of transcription, an increase in the chemical stability of the PYK/LacZ transcripts, or a combination of both. In order to differentiate between these possibilities the stabilities of hybrid mRNAs derived from the control plasmid YCpPKZ1 and from YCpPKZ4 were measured. YCpPKZ4 is associated with high PYK/LacZ mRNA abundance.

A number of methods have been used to measure the

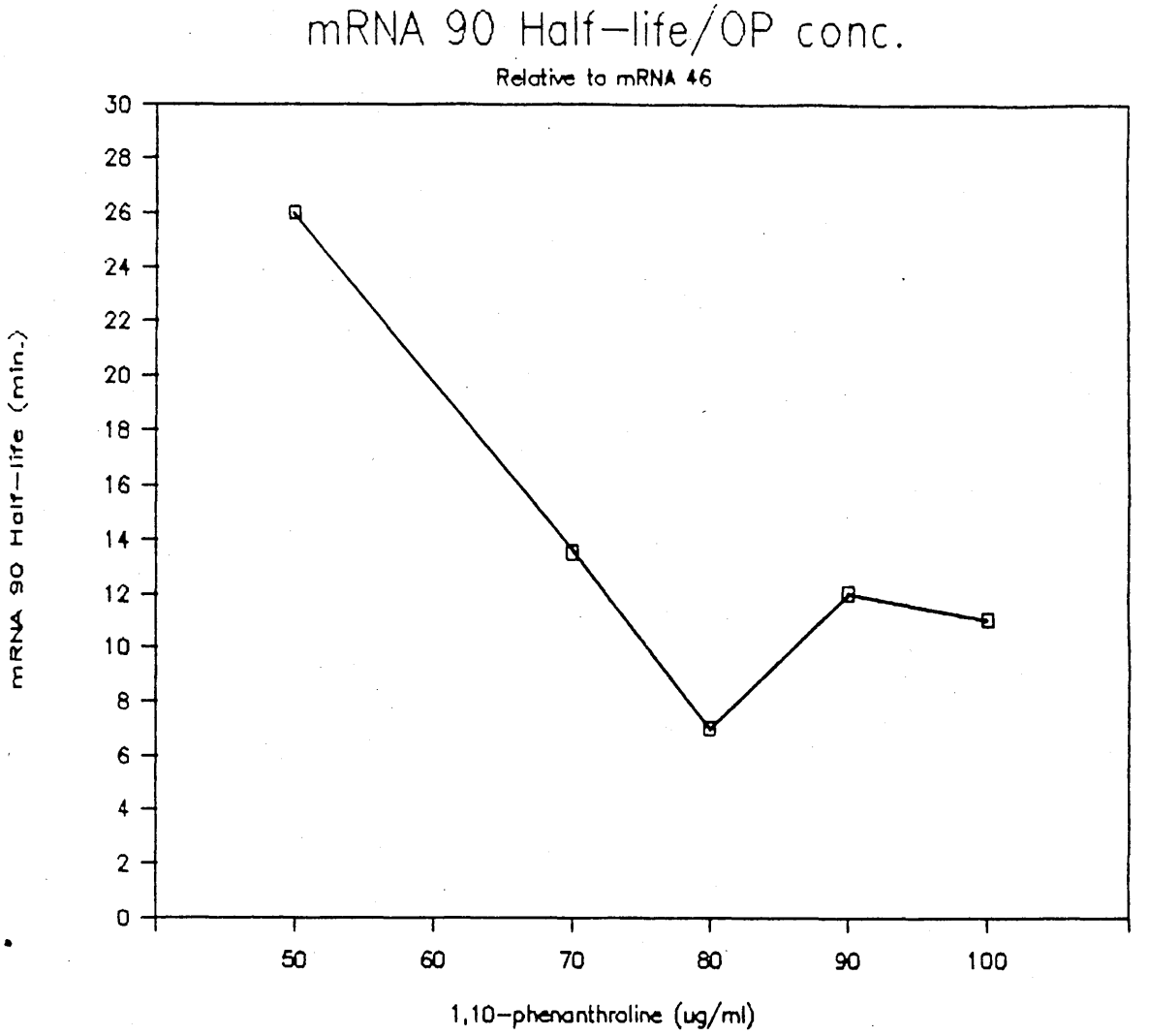


Figure 4.8 Determination of 1,10-Phenanthroline (OP) Concentration for mRNA Stability Measurements
This graph illustrates that the observed half-life of a specific mRNA is dependant upon the concentration of 1,10-phenanthroline (OP) added to the yeast culture. Half-lives were calculated by measuring the abundance of mRNA90 at ten time points after the addition of OP to the culture.

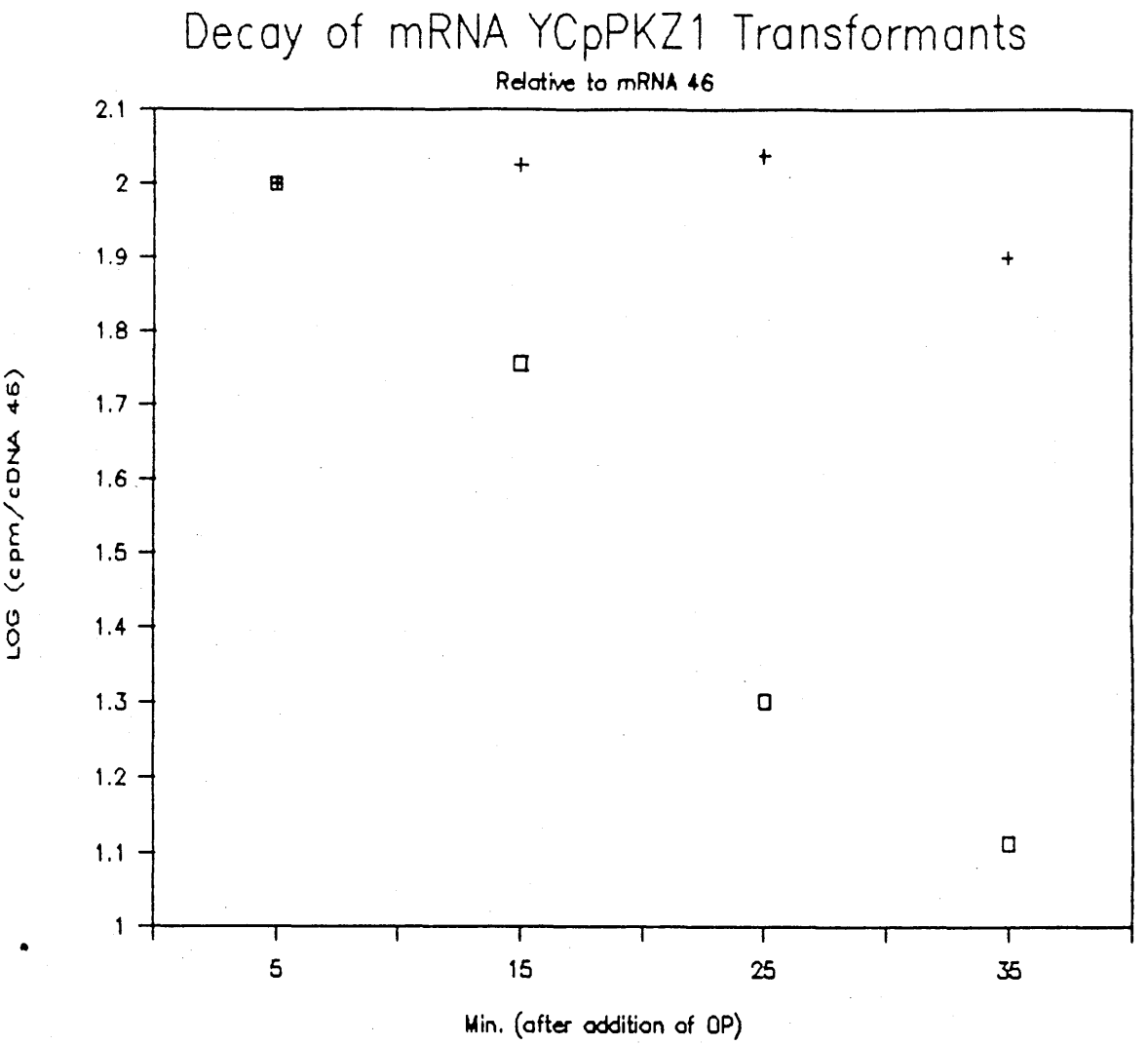


Figure 4.9 (a) Decay of PYK1/LacZ mRNA in YCpPKZ1 Transformants.
Squares represent the abundances of mRNA90 relative to mRNA46, at four time points after the addition of 80 ug/ml OP to an exponentially growing culture of a YCpPKZ1 transformant. The crosses represent the relative abundance of PYK1/LacZ mRNA at the same time points. A comparison should be made between this graph and that shown overleaf in Figure 4.9 (b).

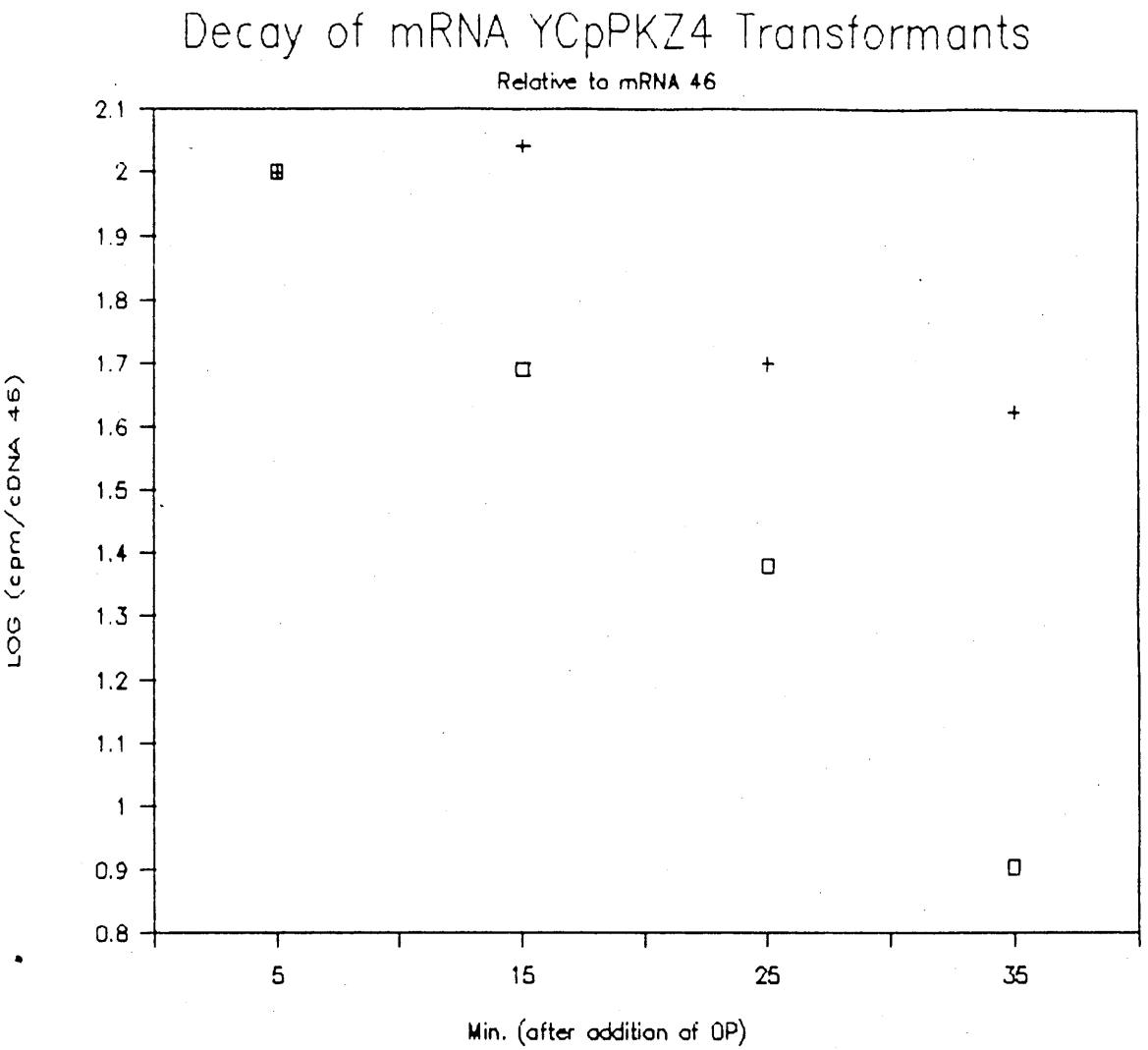


Figure 4.9 (b) Decay of PYK/LacZ mRNA in YCpPKZ4 Transformants.
Squares represent the abundances of mRNA90 relative to mRNA46, at four time points after the addition of 80 ug/ml OP to an exponentially growing culture of a YCpPKZ4 transformant. The crosses represent the relative abundance of PYK1/LacZ mRNA at the same time points. A comparison should be made between this graph and that shown previously in Figure 4.9 (a).

chemical half-lives of yeast mRNAs. The decay kinetics of an mRNA can be measured by in vivo pulse labelling of RNA or by inhibition of transcription followed by mRNA quantitation by hybridisation techniques. The method used in this study follows that of Santiago et al., (1986). Under appropriate conditions, the zinc ion chelator, 1,10-phenanthroline inhibits RNA synthesis (including mRNA), whilst only minimally inhibiting translation. 1,10-phenanthroline was added to an exponentially growing culture in rich medium (YPG) and the rate at which specific mRNA sequences were cleared from the cytoplasm was measured relative to the highly stable 18s ribosomal RNA (Santiago et al., 1986) by hybridisation procedures. In this study, different concentrations of ACS grade 1,10-phenanthroline (Sigma) were assessed for their ability to reproduce half-life measurements obtained by Santiago et al.,(1986). The results of this assessment are shown on Figure 4.8. Previous measurements have shown that the rate decay of mRNA46 (a random yeast RNA the cDNA clone of which was isolated by Santiago et al.) is indistinguishable from that of 18s rRNA, and that the half-life of mRNA90 (which is known to encode the ribosomal protein L3; Joyce Moore, unpublished data) is 6.5min. Figure 4.8 shows that a half-life of 6.5min was obtained in this study for mRNA90, relative to mRNA46, using a 1,10-phenanthroline concentration of 80mg/ml. This concentration of 1,10-phenanthroline was therefore used in all subsequent experiments.

Figure 4.9 (a and b) depicts the abundances of mRNA90 and PYK/LacZ mRNA (relative to mRNA46) at four time points after the addition of 1,10-phenanthroline (80mg/ml) to cultures of strains harbouring YCpPKZ1 or YCpPKZ4. mRNA 90 half-lives were measured to control for any spurious non-specific effects on mRNA stability in these strains. It should be noted that to determine the half-life of an mRNA accurately, its abundance should be measured at 8 or

10 time points during a 35min time period. The above experiment does not generate accurate mRNA half-lives, rather it gives an indication of the gross relative stabilities of the PYK/LacZ mRNAs derived from different transcription units. The results obtained are consistent with previous measurements of both mRNA90 and PYK/LacZ mRNA (YEpPKG441 derived) half-lives (PYK/LacZ : mRNA half-life =27min; mRNA90 : mRNA half-life =6.6min; Santiago et al.,1986; Purvis et al., 1987).
In the present experiment, no gross differences in the half-lives of the hybrid mRNAs were observed (compare Figure 4.9 a and b). This suggests that the observed differences in mRNA abundances (4.2.2) are not due to differences in mRNA stability. These results are consistent with the proposal of Purvis et al., (1987) that mRNA stability differences cannot account for the decrease in mRNA abundance observed when the PYK1 coding region is replaced with the LacZ coding region. It would appear that the PYK1 coding region contains information which activates transcription of the PYK1 gene.

4.4 Discussion of Results

The results presented above are consistent with the hypothesis that the PYK1 coding region contains an element which contributes to the rate of transcription of the PYK1 gene.
The above data demonstrate that fragments of the PYK1 coding region will, when inserted into the hybrid PYK/LacZ transcription unit, elicit an elevation in mRNA abundance of five to eight fold. As there is no measurable increase in mRNA stability associated with this elevation in abundance it is probable that the PYK1 coding region fragments activate the rate of transcription. The observation of an activation activity of the PYK1 coding region, indicates that the

transcriptional defect associated with YCpPKZ1, noted by Purvis et al., (1987) is unlikely to be due to the presence of a transcriptional repressor or silencer within the LacZ coding region but, rather is due to the absence of sequences required for the normal rate of transcription.

Further to this, it is apparent that activation does not require the entire PYK1 coding region. For example, the 800bp PYK1 insert present in YCpPKZ4 activates transcription, on average, greater than five-fold.

It has also been noted that different sizes of PYK1 insert are associated with different degrees of elevation of mRNA abundance. Perhaps these differences represent minor changes in mRNA stability. The mRNAs synthesised from YCpPKZ2, YCpPKZ3 and YCpPKZ4 vary in structure; primarily they are of very different lengths. This may result in minor differences in mRNA stability which would be reflected in mRNA abundance. For example, the mRNA from the YCpPKZ2 plasmid is longer than those with which it is being compared. As there is an inverse correlation between mRNA length and stability in yeast (Santiago et al., 1986) then we might expect this mRNA to be less abundant than its counterparts, even when it is synthesised at the same rate. These minor differences are unlikely to be observed by the relatively crude measurement of mRNA stability used in this study.

Interestingly, the activation activity of the PYK1 coding region is carbon source dependent. It is a formal possibility that transcription of the PYK1 gene is dependent upon the presence of the PYK1 coding region when cells are grown on glucose but is independent of the coding region activator when cells are grown on lactate. The observed abundances of the PYK1/LacZ mRNAs are similar irrespective of growth on a fermentative or non-fermentative carbon source. This shall be disscused in Chapter 7.

Differences in activation factor were observed in different transformants carrying the same plasmid construct (figure 4.4b). This illustrates the importance of making measurements on more than a single transformant. It should be stressed that, although activation factors vary between transformants, every measurement of PYK/LacZ mRNA from plasmids containing PYK1 coding region fragments is significantly greater than that from the control plasmid YCpPKZ1. Thus, the conclusions outlined above are based on significant trends in mRNA abundances.

A further experiment has been carried out by Purvis to delimit sequences required for elevation of PYK/LacZ mRNA abundance. The first 500 bp of the YCpPKZ4 PYK1 coding region fragment were deleted. The resultant transcription unit was placed on an episomal based vector (for convenience of construction), producing YEpPKZ6. The mRNA abundance associated with the presence of this plasmid (Figure 4.1) was measured and compared to the control plasmid YEpPKZ1 (Figure 4.10). The specific PYK/LacZ mRNA abundances shown in Figure 4.10 are adjusted for the plasmid copy number, as measured by Southern analysis. These data suggest that the presence of the remaining 360 bp PYK1 coding region fragment is sufficient to elevate the abundance of the PYK/LacZ mRNA. This result suggests that this fragment of the PYK1 coding region contains an activator of transcription. This fragment is present in all the constructs decribed in this chapter.

In chapter 7 the results of the experiments described above shall be considered in the context of other published work on the regulation of yeast genes in general, and glycolytic enzyme genes in particular. In the following two chapters, experiments designed to elucidate the mechanism of PYK1 coding region activation are described.

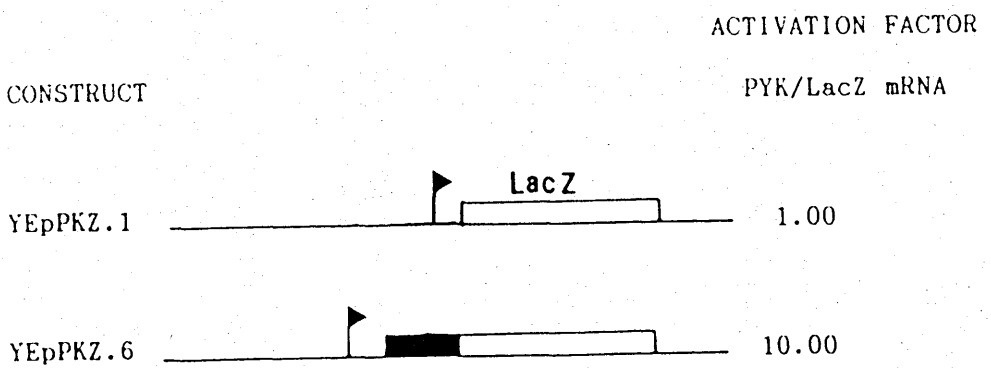


Figure 4.10 Elevation of mRNA Abundance by a 360 bp Fragment of the PYK1 Coding Region.
Activation factors represent the abundance of PYK/LacZ mRNA relative to that measured in the control transformant YEpPKZ.1. The abundances are adjusted for differences in plasmid copy number. YEpPKZ.1 is a multicopy plasmid containing an identical PYK1/LacZ transcription unit to that present on the YCpPKZ1 control vector (Figure 4.3). YEpPKZ1.6 is a modification of YEpPKZ.1 where a 360 bp fragment (500-860 bp) of the PYK1 coding region has been inserted, in frame, between the translational start site and the LacZ coding region (Figure 4.3).

CHAPTER 5

THE PYK1 CODING REGION MAY INTERACT WITH A YEAST PROTEIN

5.1 Introduction

This chapter describes experiments designed to test the hypothesis that the PYK1 coding region contains a cis-acting binding site(s) for a specific regulatory protein(s).

Results outlined in the previous chapter are consistent with the presence within the PYK1 coding region, of a transcriptional activator of the PYK1 gene. This activator may function in cis, (as Upstream Activation Sites or UASs; see 1.2) or in trans. Trans-activation would almost certainly require the information specifying transcriptional activation to reside within a post-transcriptional product of the PYK1 gene. The element may even function by a combination of both cis and trans actions. The issue of trans-activation of transcription by the PYK1 coding region will be addressed in the following Chapter.

On the assumption that the PYK1 coding region contains a cis-acting transcriptional activator, work has been undertaken to elucidate how such an activator may function.

The mechanisms by which UASs, located 5' to the I-site, function by binding specific regulatory proteins have been discussed in 1.2.3. It is conceivable that, in addition to UASs, yeast genes contain "Downstream" Activation Sites (DASs; Mellor et al., 1987). These DASs may function by a mechanism similar to UASs, acting as recognition sites for specific proteins

The mechanism of transcriptional activation by a DAS would have to accommodate transcription by RNA polymerase II through the region of protein/DNA interaction. RNA polymerase III transcription provides a precedent for regulation by a downstream site. Transcription factor IIIA (TFIIIA) binds to an internal promoter in the Xenopus oocyte gene encoding 5S RNA (see Introduction

1.2.3). The binding of TFIIIA to the internal promoter of this gene is compatible with transcription. Taking this into account, there are no mechanistic constraints to prevent a transcribed sequence specifying binding of a regulatory protein. It is thus conceivable that yeast RNA polymerase II transcribes DNA which interacts specifically with a protein factor. Experiments were therefore carried out to determine if the activating 360 bp fragment of the PYK1 coding region interacts specifically with a yeast protein.

5.2 PYK1 Coding Region/Protein Interactions

As discussed previously, protein/DNA interactions can be detected by a number of techniques. These range from simple filter binding assays, where protein/DNA complexes are sequestered by a nitro-cellulose membrane filter, to in vivo photofootprinting, where in vivo interactions can be revealed. In this study, gel retardation or band shift methods were used in the first instance (see 3.3). With cautious interpretation, gel retardation experiments yield much information about the nature of protein/DNA interactions and are technically straightforward. Gel retardation experiments were undertaken in collaboration with Dr. Ian Dawes of the University of New South Wales, Australia. Dr. Dawes prepared total yeast protein extract by the method of Huet et al.(1985) as outlined in Chapter 3 (3.3).

5.2.1 Preparation of PYK1 DNA Fragment for Gel Retardation

As outlined in 4.2.5, a 360 bp fragment of the PYK1 coding region has been shown by Ian Purvis to be sufficient for the activation of transcription from the PYK1 5'-flanking region. This fragment may contain a

binding site for a trans-acting regulatory protein. The 360 bp PYK1 coding region fragment was purified for gel retardation experiments.

100ug of plasmid pSPPYK2 (the construction of this plasmid is summarised in Figure 6.1a) was restricted with XbaI and BamHI and ethanol precipitated. The 360 bp PYK1 fragment was purified by gel electrophoresis and geneclean methods and the yield assessed. 1ug of purified fragment was end-labelled by incorporation of ^{32}P-dCTP and further purified by the geneclean method to remove unincorporated nucleotides and Klenow fragment protein, which otherwise may interfere with the gel retardation assays. The labelled fragment was electrophoresed through a 5% (w/v) non-denaturing polyacrylamide gel to assess it's purity. Occasionally, a second band of higher molecular weight was shown to be present. Further purification by geneclean removed the contamination which led to the formation of this second band. It is likely that the contaminant observed was due to restriction enzyme or Klenow fragment forming a retarded complex with the end-labelled DNA fragment.

4.2.2 Detection of Protein/DNA Interactions

Protein/DNA interactions were detected by a procedure identical to that described in chapter 3. Protein extracts were prepared as in 3.4. Retarded DNA bands were observed in binding assays which contained protein eluted from the heparin-Sepharose column at 0.25M ammonium sulphate. Only two out of fifty five protein fractions (Nos.32 and 33) were capable of eliciting retardation of the PYK1 coding region fragment. Fraction 33 was shown to initiate formation of up to 3 retarded complexes with the PYK1 gene fragment in the presence of a 2000-fold excess of non-specific DNA (phage lambda DNA). Figure 5.1 shows that the most mobile complex (CI) is formed

preferentially at low protein concentrations and the least mobile complex (CIII) is formed preferentially at high protein concentrations. These complexes were clearly absent in the protein-free binding assay control. Therefore, a component of protein fraction 33 retards the electrophoretic mobility of a PYK1 fragment. Retarded DNA is observed in the form of one of three states or complexes. These complexes almost certainly represent different associations of the DNA with protein molecules. For example, CI could represent a complex of a single DNA molecule and a single protein molecule. CII could represent an association of a single DNA molecule with a protein molecule which is distinct from that present in CI or it could represent the protein/protein interactions. As CI forms at low protein concentrations, CI is mainly considered throughout the following discourse.

5.2.3 A Component of Protein Fraction 33 Apparently Interacts Selectively with a Fragment of PYK1 Coding Region DNA

In vitro complex formation is not, in itself, sufficient evidence that an in vivo interaction occurs. If however, the interaction is shown to take place between a specific DNA sequence and a specific protein then it is a reasonable assumption that this may reflect a biologically important relationship. Specificity can be assessed by a number of criteria. Perhaps the most commonly applied criterion is the ability of a protein to bind to a specific DNA fragment with a higher affinity than to non-specific DNA fragments. This ability is simply in binding assays where two types of DNA molecule compete for the binding of a specific protein. When interpreting such competition experiments, one has to bear in mind that, at sufficiently high concentrations,

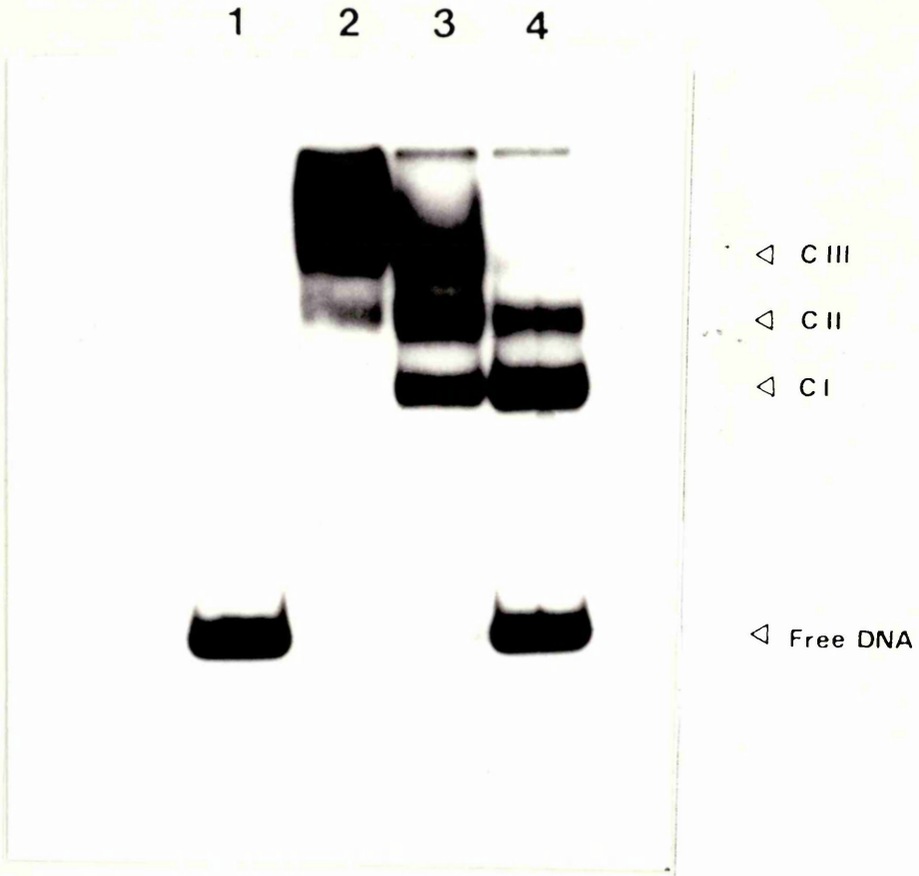


Figure 5.1 Interaction of a Yeast Factor With a Fragment of the PYK1 Coding Region

Unbound (free) PYK1 360 bp fragment DNA and the position of three distinct complexes are indicated (c : complex). The contents of the binding assay mixes which were electrophoresed on this 5% non-denaturing polyacrylamide gel before autoradiography were as follows;

All Lanes : 500pg [^{32}P]-dCTP labelled 360 bp (500-860 bp) PYK1 DNA and 1ug of unlabelled phage lambda DNA.

Lane 1 : No protein added.

Lane 2 : 6ug of protein (heparin-Sepharose fraction 33).

Lane 3 : 3ug of protein (fraction 33).

Lane 4 : 1.5ug of protein (fraction 33).

Binding assays were carried out as described in 2.13.

any protein will retard the migration of DNA through a gel. Additionally non-specific DNA sequences will, if present at high enough concentrations, competitively inhibit formation of specific interactions. The specificity of an interaction defined by the different degrees to which specific and non-specific DNA fragments interact with a protein. In practice, the ability of unlabelled DNA to inhibit the interaction between protein and specific end-labelled DNA is tested.

Figure 5.2 illustrates a standard competition experiment carried out to assess the specificity of the complex described in 5.2.2. The contents of the individual binding assays are given in the figure legend. The ability of unlabelled 360 bp PYK1 DNA fragment to competitively inhibit complex formation was examined in lanes 1 to 5. As the amount of unlabelled PYK1 DNA fragment is reduced from 100ng to 2.5ng, a significant decrease in the amount of uncomplexed (free) DNA was observed. This decrease is concurrent with an increase in complex I (CI) formation in lanes 1 to 3. In lane 1, where the CI formation is most markedly reduced, the unlabelled PYK1 DNA fragment is in a 400-fold molar excess. Lanes 6 to 11 are binding assays in which decreasing amounts (400ng to 0.1ng) of unlabelled plasmid DNA (supercoiled and open circular pSP65) have been added as non-specific competitors. No general trend in the amount of free DNA present in the binding assay was observed across the range of plasmid competitor concentrations. Therefore, under the conditions of this experiment, the 360bp PYK1 DNA fragment is more successful at competitively inhibiting complex (CI) formation than pSP65 plasmid. However, this experiment does not reveal the degree of differential affinity.

In addition, the plasmid competitor assessed in these experiments is circular but the PYK1 DNA fragment is linear. This conformational difference may account for

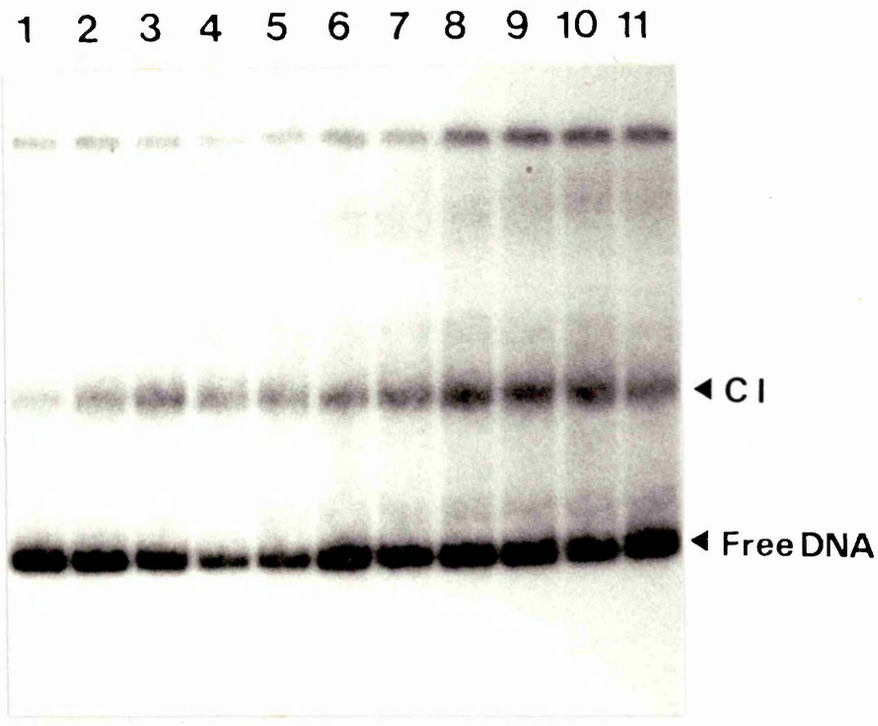


Figure 5.2 Assessment of Specificity of the Interaction Between a Yeast Factor and PYK1 Coding Region DNA; Part 1.
Binding assay mixes contained the following;
All lanes : 250pg radiolabelled 360 bp PYK1 DNA, 500ng of unlabelled phage lambda DNA and 0.7ug of protein (fraction 33).
Lane 1 : 100ng unlabelled 360 bp PYK1 fragment.
2 : 50ng unlabelled 360 bp PYK1 fragment.
3 : 10ng unlabelled 360 bp PYK1 fragment.
4 : 5ng unlabelled 360 bp PYK1 fragment.
5 : 2.5ng unlabelled 360 bp PYK1 fragment.
6 : 400ng unlabelled pSP65 plasmid.
7 : 200ng unlabelled pSP65 plasmid.
8 : 50ng unlabelled pSP65 plasmid.
9 : 10ng unlabelled pSP65 plasmid.
10 : 1ng unlabelled pSP65 plasmid.
11 : 0.1ng unlabelled pSP65 plasmid.
pSP65 DNA is a mixture of supercoiled and open circular plasmid.

the observed differences in competitiveness.

Figure 5.3 represents a similar competition experiment in which twice the amount (1.4ug) of protein fraction is added to the binding assays. At this protein concentration complexes II and III are clearly observed. In this experiment the degree of competition was compared for linear plasmid DNA and supercoiled/open-circle plasmid DNA (lanes 7 and 8) . Unlabelled 360 bp PYK1 DNA fragment was again used as competitor (in decreasing amounts across lanes 2 to 5). Lanes 7 and 8 illustrate that linear plasmid and a mixture of supercoiled/open-circle plasmid prevent complex formation to the same extent. A comparison of lanes 2 and 7 illustrates that 100ng of PYK1 DNA fragment competes with the labelled fragment for binding to a greater extent than 100ng of pSP65.

Figure 5.4 illustrates an experiment which confirmed the observation that equal amounts (expressed in micrograms) of PYK1 DNA fragment and linear pSP65 compete with labelled fragment for complex formation with different efficiencies. A comparison of the amount of free DNA present in lanes 1 and 8 shows that 50ng of PYK1 DNA fragment competes for binding more successfully than 50ng of pSP65. Only when 500ng of pSP65 DNA was added to the binding assay did it compete to the same extent as 50ng of PYK1 DNA (compare the intensity of the free DNA bands in lane 1 and 5). This may be considered evidence of selective binding to the PYK1 DNA. This is discussed further in 5.2.4.

The competition experiments described above suggest that a component of fraction 33 associates preferentially with a PYK1 DNA fragment. The experiment to be described below attempts to confirm this observation and examines the efficacy of PYK1 fragment association with a component of fraction 33 in comparison to a smaller DNA fragment present in the binding assays at a higher molar

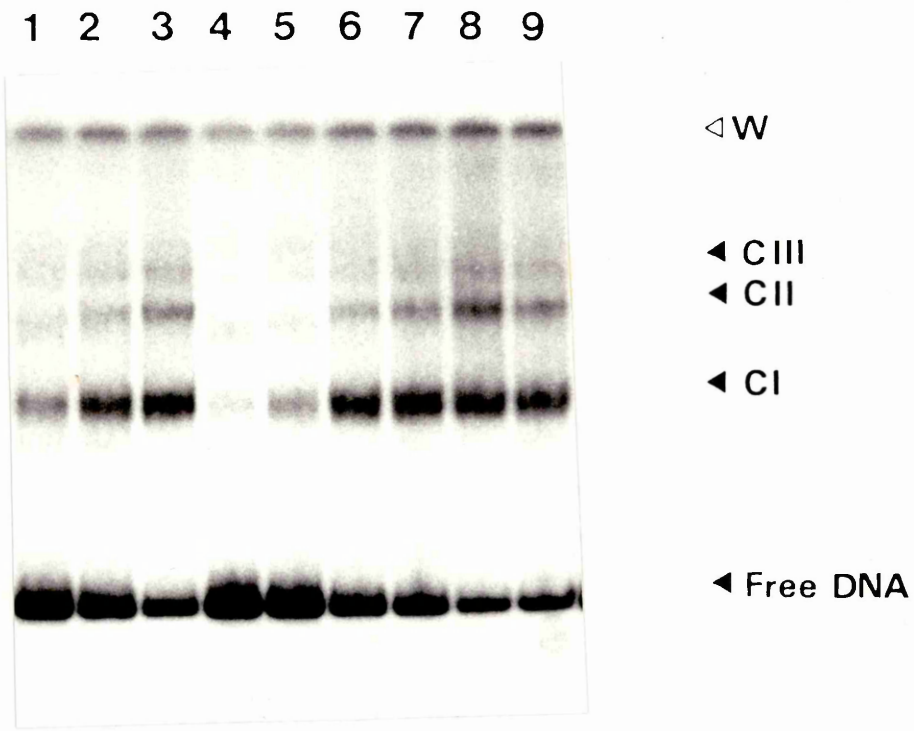


Figure 5.4 Assessment of Specificity of the Interaction Between a Yeast Factor and PYK1 Coding Region DNA; Part 3 .
Binding assay mixes contained the following;
All lanes : 250pg radiolabelled 360 bp PYK1 DNA, 500ng of unlabelled phage lambda DNA and 0.7ug of protein (fraction 33).

Lane 1 : 50ng unlabelled 360 bp PYK1 fragment.
2 : 25ng unlabelled 360 bp PYK1 fragment.
3 : 10ng unlabelled 360 bp PYK1 fragment.
4 : 1ug unlabelled linear pSP65.
5 : 500ng unlabelled linear pSP65.
6 : 100ng unlabelled linear pSP65.
7 : 75ng unlabelled linear pSP65.
8 : 50ng unlabelled linear pSP65.
9 : 10ng unlabelled linear pSP65.

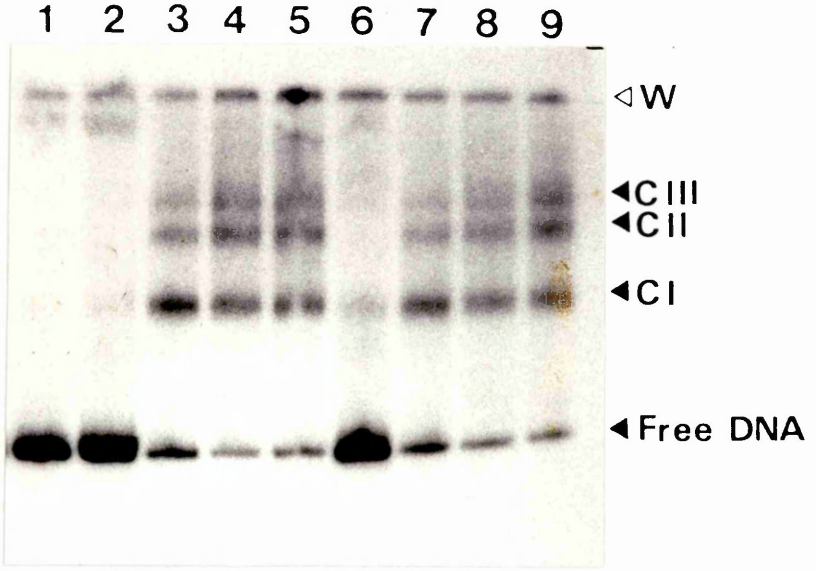


Figure 5.3 Assessment of Specificity of the Interaction Between a Yeast Factor and PYK1 Coding Region DNA; Part 2.
Binding assay mixes contained the following;
All lanes : 250pg radiolabelled 360 bp PYK1 DNA, 500ng of unlabelled phage lambda DNA and with the exception of Lane 1, 1.4ug of protein (fraction 33).
Lane 1 : No protein. No unlabelled PYK1 or pSP65 DNA
2 : 100ng unlabelled 360 bp PYK1 fragment.
3 : 10ng unlabelled 360 bp PYK1 fragment.
4 : 1ng unlabelled 360 bp PYK1 fragment.
5 : 0.1ng unlabelled 360 bp PYK1 fragment.
6 : 1ug unlabelled linear pSP65 plasmid.
7 : 100ng unlabelled linear pSP65 plasmid.
8 : 100ng unlabelled supercoiled/open-circle pSP65 plasmid.
9 : 10ng unlabelled linear pSP65 plasmid.

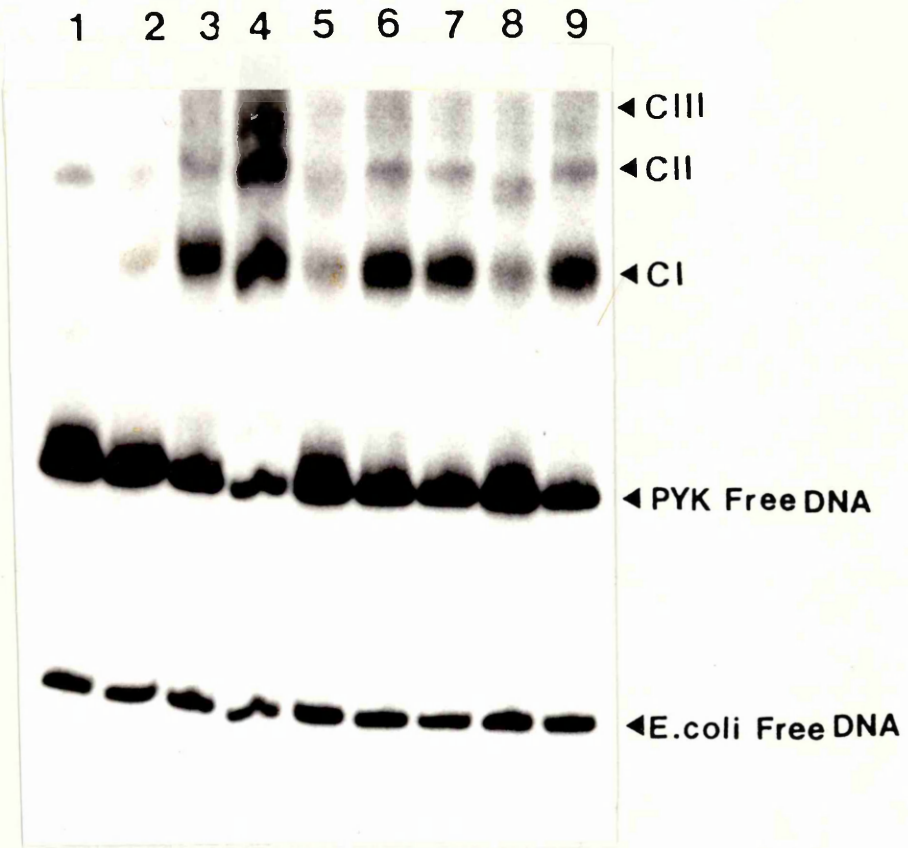


Figure 5.5 Assessment of Specificity of the Interaction Between a Yeast Factor and PYK1 Coding Region DNA; Part 4.
Binding assay mixes contained the following;
All Lanes : 250pg of radiolabelled 360 bp PYK1 DNA, 500ng of unlabelled phage lambda DNA, 250pg of radiolabelled 120 bp E.coli DNA fragment.
Lane 1 : No protein.
Lane 2 : 0.35ug protein (fraction 33).
Lane 3 : 0.7ug protein (fraction 33).
Lane 4 : 1.4ug protein (fraction 33).
Lane 5 : 0.7ug protein (fraction 33) and 50ng 360 bp unlabelled PYK1 DNA fragment.
Lane 6 : 0.7ug protein (fraction 33) and 25ng 360 bp unlabelled PYK1 DNA fragment.
Lane 7 : 0.7ug protein (fraction 33) and 10ng 360 bp unlabelled PYK1 DNA fragment.
Lane 8 : 0.7ug protein (fraction 33) and 500ng linear pSP65.
Lans 9 : 0.7ug protein (fraction 33) and 100ng linear pSP65.

concentration. Figure 5.5 illustrates an experiment which compares the protein binding affinity of two labelled DNA fragments. Each binding assay contains equal microgram quantities of end-labelled PYK1 DNA fragment and end-labelled E.coli ARG repressor binding site DNA. The specific radioactivity of the E.coli DNA fragment is less than that of the PYK1 fragment but the E.coli fragment is present in a three-fold molar excess. By examining the free DNA band in lanes 3, 4, 7 and 9, it is clear that PYK1 DNA is sequestered into complexes more readily than E.coli DNA. This may also be considered evidence of a selective interaction between a specific yeast protein fraction and a fragment of the PYK1 coding region.

5.2.4 Conclusions from Gel Retardation Experiments

A component of a heparin-Sepharose protein fraction (33) has been shown to interact with a fragment of PYK1 coding region thought to be responsible for elevation in mRNA abundance (4.4). The specificity of this interaction has been examined by binding competition experiments. A 360 bp PYK1 DNA fragment was shown to compete for protein binding ten-fold more favourably than equal amounts of various forms of pSP65 plasmid. This suggests that protein contained in Fraction 33 interacts selectively with the PYK1 fragment. However, one should be aware that equal molar concentrations of the PYK1 360 bp fragment and pSP65 plasmid, compete for binding to an equal extent (Figure 5.4 lanes 1 and 5). Therefore, if the protein responsible for eliciting band retardation binds to DNA termini, or if the binding of one protein molecule to the plasmid excludes the possibility of any further protein binding, then no selective binding to the PYK1 fragment has occurred. If one of these conditions arise, or if the non-specific competitor (pSP65) contained a sequence identical to the specific yeast binding site, then the

relative molarities of the non-specific and specific competitor DNA would be relevant to assessment of binding selectivity. However, Figure 5.5 shows clear evidence for selective binding to the 360 bp PYK1 fragment. In this experiment, despite the presence of a non-PYK1 control fragment in molar excess, binding remains confined to the 360 bp PYK1 DNA fragment.

The selective binding properties of protein Fraction 33 cannot be assumed to be physiologically relevant. The extent to which Fraction 33 protein selectively interacts with PYK1 DNA compares unfavourably with previously observed specific interactions, in which binding proteins display up to 1000 fold greater affinity for a specific DNA binding site.

Low selectivity of binding could be due to a number of factors. For example, binding may be highly dependent on the conditions of incubation of protein and DNA. Although the conditions used in this preliminary study have been used previously to identify a specific yeast DNA binding protein (Huet et al., 1985) they may not be optimal for the PYK1 DNA/protein interaction. Alternatively, the binding protein may require a second protein factor to stabilise binding. Co-operative binding is displayed by some yeast regulatory proteins (see Introduction 1.2.3.2; Forsburg and Guarente 1989; Oleson et al., 1987). The putative stabilising factor may not be present in Fraction 33 or may be unstable upon extraction from yeast.

Binding to the 360 bp fragment may be unstable or relatively non-specific due to the absence of a second DNA binding site normally present within the intact PYK1 gene. Such a site may be present within the 5'-flanking region of PYK1 or within the coding region, outwith the 360 bp fragment.

5.3 Footprinting Analysis of Complexes

5.3.1 Introduction

Footprinting was carried out to further examine the interaction, between a component of protein Fraction 33 and the 360 bp PYK1 DNA fragment. Footprinting can reveal whether a protein binds to a specific site on the DNA molecule. On binding a specific site, a protein may protect the DNA with which it comes in contact from an endonuclease or base modifying activity. Binding of a protein to a specific site within a DNA molecule indicates a particular affinity for that site compared to other areas of the DNA duplex. This may be representative of an interaction which occurs *in vivo*. As CI (4.2.2) is an interaction of low specificity (under the conditions of this study), footprinting was carried out to determine whether a specific interaction occurs at a site (or sites) on the PYK1 fragment.

DNase I has commonly been used as an agent of nuclease activity but more recently, hydroxyl radical cleavage has been shown to have distinct advantages over enzymatic cleavage. The accessibility of the DNA sugar-phosphate backbone to small hydroxyl radical reagents, renders all unprotected nucleotides susceptible to cleavage by such reagents (Tullis and Dombroski, 1986). This is in contrast to DNaseI activity which can be sterically hindered by the bound, protecting protein. As a result, patterns of protection obtained from DNaseI cleavage experiments tend to overestimate the extent of direct protein/DNA contact.

The DNA cleavage activity of the 1,10-phenanthroline-cuprous ion [$(OP)_2Cu^+$ or OP Cu] (Sigman, 1986) has been utilised previously to generate footprints of *E.coli* RNA polymerase-lac promoter complexes (Kuwabara and Sigman,

1987). The authors have shown that DNA fragments can be cleaved within a polyacrylamide gel matrix by bathing the gel fragments in cleaving reagents. Therefore, non-complexed DNA can be separated from complexed DNA before footprinting and different complexes formed in a single binding assay can be examined. $(OP)_2Cu^+$ is a co-reactant with H_2O_2 in the cleavage reaction which can be outlined as follows -

$$(OP)_2Cu^+ + DNA \longrightarrow (OP)_2Cu^+\text{-}DNA \xrightarrow{H2O2} (OP)_2Cu^{2+}.OH^- \text{-}DNA \longrightarrow \text{oligonucleotides.}$$

in this reaction $(OP)_2Cu^+$ binds directly to the DNA (Sigman, 1986).

5.3.2 Preparation of labelled DNA for Footprinting experiments

Fragments used for footprinting were labelled at a single end. Two methods were used to achieve this. The first method utilised the property of Klenow fragment of incorporating labelled nucleotides into single-stranded DNA with a 3'-OH primer but not with a 5'-phosphate. The 360 bp PYK1 coding region DNA fragment was purified from pSPPYK2 (plasmid construction reviewed in Figure 6.1) following digestion of the plasmid with BamHI and PstI, or alternatively with XbaI and EcoRV, and end-labelled with radiolabelled dCTP and dATP. To ensure that only the BamHI generated end incorporated label, the labelled 360 bp BamHI/PstI fragment was restricted with Sau3A (which specifically cleaves the fragment 165 bp from the 5'-overhang). This restriction digest was electrophoresed through a 5% (w/v) polyacrylamide gel and on autoradiography a single band, running below the xylene cyanol dye front (which has mobility equivalent to a 260 bp DNA fragment) was observed. The absence of any other labelled band indicated that end-labelling had taken place at only the BamHI generated end of the DNA molecule

and therefore the end-labelled BamHI/Pst1 fragment could be used directly in footprinting experiments.
The second method used to obtain a fragment labelled at one end was as follows. A plasmid was linearised by restriction digestion close to the PYK1 insert and end-labelled by Klenow fragment activity. One terminus of the molecule was then removed by restriction digestion of the DNA at a convenient site. For example, pSPPYK1 (Figure 6.1) was digested with Xba1 and end-labelled. After purifying the DNA by the geneclean method, it was restricted with Pst1 which removed a very small end-labelled fragment. As this fragment is only 10 bp, it does not interfere with the footprinting procedures and hence no further purification is required.

5.3.3 Preparative Gel Retardation

To obtain sufficient amounts of DNA/protein complexes for footprinting analysis, binding assays were scaled up at least 20 fold. Binding assay mixes containing large amounts of protein and DNA were electrophoresed through a non-denaturing polyacrylamide gel. Complexes I, II, and III were clearly seen, with CI forming preferentially at the lower protein concentration. Each band representing free or complexed DNA was excised from the gel and the DNA was cleaved in situ with the OP-Cu reagent following the method of Kuwabara and Sigman (1987). This method is outlined in 2.13

5.3.4 Protein binding to the PYK1 coding region fragment does not protect a specific site from cleavage.

Lane 1 of Figure 5.7 illustrates the cleavage pattern obtained when PYK1 DNA is treated with cleavage reagents in situ. The cleavage pattern is not entirely even throughout the length of the phosphate-ribose backbone.

As this pattern is repeatable, it probably reflects properties of the DNA conformation under the experimental conditions. It is not possible however, to predict the function of specific structural features as revealed by this type of cleavage because it is not known whether there is a strong structure-function relationship between *in vitro* DNA conformation and *in vivo* transcription.

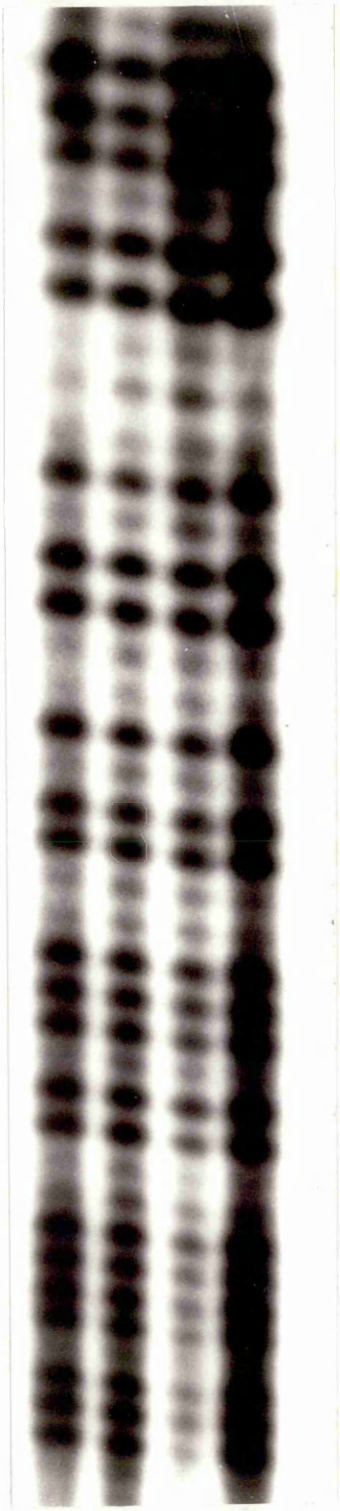
Lanes 2, 3 and 4 of Figure 5.6a show cleavage patterns obtained with DNA fragments which were retarded within complexes on non-denaturing gels. No areas of protection were observed in these tracks when compared with the control lane 1. That is, the cleavage patterns obtained from complexed and non-complexed DNA are identical. This suggests that specific protein/DNA interaction does not take place within the area of the *PYK1* fragment represented by this gel.

It should be noted that areas of protection were revealed when this method of footprinting was applied to a specific control complex; that of a specific yeast protein fraction and a DNA fragment from the 5'-region of the lipoamide dehydrogenase gene (Joe Ross, Gordon Lithgow, Ian Dawes; Unpublished data). We can conclude that the methods used will reveal areas protected from cleavage by a specific interaction, further indicating that there is an absence of specific protein interaction with the *PYK1* fragment.

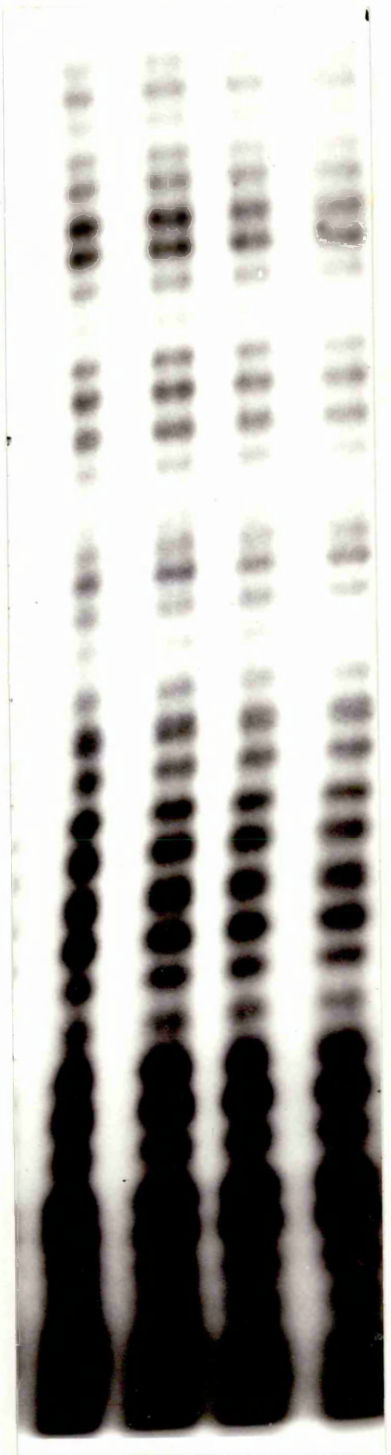
As the *PYK1* fragment of interest is 360 bp long it was difficult to resolve all bands corresponding to cleavage products. Thus, if a protein bound close to the non-*Bam*HI generated terminus, then the protection pattern may be unobservable. To overcome this problem, labelling was carried out at the *Xba*I generated site of the fragment (that is the opposite end from *Bam*HI). Figure 5.6b shows cleavage of the *Xba*1/*Eco*RV *PYK/LacZ* fragment in solution. Cleavage was carried out without prior purification of complexes by gel retardation. A protein concentration

Figure 5.6 Search For Footprint on 360 bp PYK1 DNA Fragment
Cleavage of DNA by the activity of 1,10-phenanthroline-copper ion is discussed in 2.14. Individual complexes and free DNA (Figure 5.1) were subjected to cleavage.
(a) All complexes examined in this footprinting gel contain the 360 bp PYK1 DNA fragment end-labelled at a single termini. DNA equvailent to 10^6 cpm was loaded onto each lane.
Cleavage patterns were obtained with DNA from the following sources;
Lane 1 : Uncomplexed (free) DNA.
Lane 2 : Complex I (see Figure 5.1)
Lane 3 : Complex II
Lane 4 : Complex III
(b) All complexes examined in this footprinting gel contain XbaI/EcoRV PYK1/LacZ DNA fragment from pSPPYK1 (This plasmid is shown in figure 6.1). DNA equvailent to 10^6 cpm was loaded in each lane. Cleavage was carried out in solution.
Lane 1 : No protein, uncomplexed DNA.
Lane 2 : 4 ug protein (fraction 33).
Lane 3 : 8 ug protein (fraction 33).
Lane 4 : 16 ug protein (fraction 33).
The ratio of DNA to protein in the binding assay was optimal for complex formation.

a



b



which maximises complex formation, as shown by retardation gels, was used in solution cleavage. Fragments incubated with yeast protein fractions prior to the cleavage reaction, exhibit an identical cleavage pattern to the control (see Figure 5.6b). The same result was obtained when the large Xba1/Pst1 fragment of pSPPYK2 was used in these assays.
These results suggest that no area of protection occurs throughout the entire length of the PYK1 360bp fragment. This indicates that CI formation is not due to a specific protein/DNA interaction.

5.4 Conclusions

The experiments outlined in this chapter were carried out to assess whether a 360 bp fragment of the PYK1 coding region, which apparently activates transcription in vivo, binds a specific yeast protein in vitro. Section 5.2.4 outlines the evidence for a specific protein/PYK1 DNA interaction based on gel retardation experiments. In order to obtain further evidence of specificity, footprinting experiments were carried out. These experiments did not produce any indications that binding occurs to a specific region of the PYK1 DNA fragment.
There is, therefore, an apparent contradiction between the results of the two types of experiments. The results of competition experiments suggest that a component of protein Fraction 33 binds a 360 bp PYK1 DNA fragment with perhaps a ten-fold greater specificity than it does non-specific DNA. One would perhaps expect that this selectivity reflects preferential binding to a specific sequence within the PYK1 fragment. If this is the case, then one might expect to observe a ten-fold preference for binding at a particular site on the fragment and that this preference might be detected by a decrease in susceptibility to cleavage at this specific site. Such an

area of protection is not observed. There are a number of possible explanations for this anomaly. It may be that the interaction between protein Fraction 33 and PYK1 DNA is non-specific and that selective binding observed in gel retardation experiments was artifactual. Alternatively, as the degree of selectivity is low, and hence the protein/DNA interaction is possibly weak, the footprinting method may not be sensitive enough to detect binding. Low selectivity of binding may reflect low affinity of the protein for the putative binding site, which in turn may lead to cleavage sensitivity. However, binding of protein to the 360 bp PYK1 fragment results in repeatable band retardation which suggests that the protein/DNA complex is stable under the conditions used. A second alternative is that the footprinting procedures (outlined in 2.13) disrupt the protein/DNA complex before or during the cleavage reactions. This could be investigated by examining the mobility of the cleaved DNA on a non-denaturing polyacryamide gel. If the complex has remained intact during the footprinting procedures then the DNA should be retarded relative to uncomplexed, free DNA. The DNA duplex would remain intact after the cleavage reaction as cleavage occurs on average only once per single strand.

In summary, an interaction between a component of a yeast protein fraction and a 360 bp DNA fragment from the PYK1 coding region has been investigated. The protein has been shown to bind to the PYK1 DNA fragment with a higher affinity than to non-PYK1 DNA. It has not been demonstrated that the protein binds to a specific site on the PYK1 fragment. For this reason the biological significance of this interaction remains unclear.

In an attempt to establish the validity of the interaction described above a number of experiments

should be carried out;

(i) The binding of a component of fraction 33 to PYK1 coding region DNA should be assayed at different salt concentrations, to optimise conditions for specific binding to the fragment.

ii) Binding of pooled heparin-Sepharose fractions should be assayed, as co-operative binding may occur.

(iii) Binding to the PYK1 coding region could be examined by in vivo footprinting following the method of Selleck and Majors (1987a and 1987b).

(iv) Yeast extract from cells grown on different carbon sources should be assayed for binding activity as a regulatory protein may only exhibit binding competence under specific culture conditions.

CHAPTER 6

THE *PYK1* CODING REGION MAY ACTIVATE TRANSCRIPTION IN *TRANS*

6.1 Introduction

Evidence was presented in Chapter 4, which indicates that the PYK1 coding region contains a transcriptional activator. It has been shown that 360 bp of this region is sufficient for transcriptional activation (4.4, Ian J. Purvis, unpublished data). The previous chapter details experiments undertaken to establish whether the PYK1 coding region contains a binding site for a specific yeast protein. The presence of such a binding site might suggest that the PYK1 element responsible for transcription activation acts in cis. In contrast, the experiments described in this Chapter indicate that transcriptional activation by the PYK1 coding region is dependent upon a post-translational product of the PYK1 gene. In addition, over-expression of the PYK1 coding region or PYK1 coding region fragments appeared to elicit an increase in the abundance not only of PYK1/LacZ mRNA (as described in Chapter 4) but also of at least one other glycolytic enzmye mRNA.

In this Chapter, consideration will be given to the results of this study. How these results relate to the large body of published work on yeast gene expression and co-ordinate regulation will be discussed later (Chapter 7).

The presence of a transcriptional activator within the PYK1 coding region is intriguing as there is no reason a priori why information pertinent to transcriptional control should arise within a protein coding region.

Yeast protein coding genes are, in the main, controlled by elements which reside in the 5'-flanking regions.

There seems to be very little limitation on the number and diversity of control elements, located within 5'-flanking regions which can influence transcription (1.2.2). Given this, it is unclear why a transcriptional activator should occur within the coding region of PYK1. Perhaps PYK1 is relatively unique in there having been selective pressure for the evolution of a coding region located transcriptional activator. Alternatively, many yeast genes may contain transcriptional activators within their coding regions and there has been general selective pressure for such elements within the coding regions of many genes. It is clear that the evolutionary constraints on such a region will differ from those acting upon an element in the 5'-flanking region. For example, PYK1 has an extremely high codon bias. One might suppose that the drift towards high codon bias may be contrary to the preservation of a sequence specific binding site.

The possibility that transcriptional activation is not dependent on binding of a trans-acting regulatory protein to a specific DNA sequence within the PYK1 coding region shall now be considered. The information specifying transcriptional activation may not reside intrinsically within the DNA sequence but is present within a post-transcriptional product of this gene. Transcriptional activation may be dependent upon sequences within coding region of the PYK1 mRNA or a structural component of the pyruvate kinase protein.

If transcriptional activation was dependent upon a post-transcriptional product (PYK1 protein) then one may expect activation of transcription from the PYK1 initiation site, to be independent of the relative location of the expressed PYK1 coding region. Let us consider this possibility since it is pertinent to the interpretation of the experiments that follow.

The above scenario predicts that the PYK1 coding region

would activate transcription in trans and that there is no absolute requirement for the PYK1 coding region in cis for high transcription. However, we may already have evidence that the PYK1 coding region does not activate transcription in trans. The PYK1/LacZ hybrid transcription unit on the plasmid YCpPKZ1 (Figure 4.4) does not contain PYK1 coding region sequences. Insertion, in cis, of PYK1 coding region fragments into this transcription unit results in an increase in transcription of the hybrid gene (4.3.2, Figure 4.5). Transformants containing the control plasmid in this experiment, YCpPKZ1, contain a fully functional PYK1 gene at the chromosomal locus. Therefore, the presence in trans of the PYK gene does not alleviate the transcriptional defect affecting the hybrid transcription unit. One may conclude from this result that the PYK1 coding region does not activate transcription in trans. But there may be another explanation. Perhaps there is a defect in hybrid gene transcription because there is a deficiency of a trans-acting factor. The information specifying high transcription, derived from the chromosomal PYK1 gene, may be limiting. In this situation the control hybrid transcription unit would appear to require a cis acting element from the PYK1 coding region for elevated transcription, when in fact over-expression of the PYK1 coding region in trans would be sufficient.

In summary, although it may seem unlikely that transcriptional activation by the PYK1 coding region occurs entirely in trans, it remains a formal possibility.

It is also possible that transcriptional activation by the PYK1 coding region requires both a post-transcriptional product and cis-located PYK1 coding region sequences. This seems to be an overly cumbersome hypothesis which does not in any way help to explain how a transcription activation site evolves within a coding

region.
The hypothesis of trans-activation of PYK1 expression by PYK1 expression opens a new possibility; that of the trans-activation of other genes. This issue will discussed later.
Having considered the theoretical plausibility of trans-activation by the PYK1 coding region, experimental means of establishing it's validity will now be discussed.
Circumstantial evidence for trans-activation was obtained by investigating whether transcriptional activation requires translation of PYK1 coding region sequences. This is based on the supposition that, if the information specifying elevated transcription rate was located within the pyruvate kinase enzyme, activation would be dependent upon translation of the PYK1 coding region sequences.

6.2 Activation of Transcription by the PYK1 Coding Region May Depend Upon Translation

The requirement of translation for transcriptional activation by the PYK1 coding region, was investigated initially by the construction and analysis of YCpPKZ6.1. This plasmid is CEN based and contains a modified version of the transcription unit present in YEpPKZ6 (4.4). A frameshift mutation has been introduced within PYK1 coding region fragment previously shown to activate transcription of the fusion PYK1/LacZ gene (see 4.4).
As a result of this frameshift mutation, translation terminates 27 codons downstream of the ATG initiation codon.
If the PYK1 fragment of this plasmid exhibited transcriptional activation properties then it is unlikely that activation is dependent upon a post-translational product. Alternatively, if the fragment did not exhibit transcriptional activation in this plasmid, then either

activation requires translation or activation is dependent upon the integrity of a sequence at the site of the frameshift mutation.

6.2.1 Construction of YCpPKZ6.1

An outline of the construction strategy is shown in figure 6.1 and is summarised as follows. The BglII-Sst1 fragment of plasmid YEpPKZ6 was inserted into a BamHI-SstI restricted pSP64. The product was called pSPPYK1. Subsequently the XbaI-EcoRV fragment from pSPPYK1 was ligated to the large fragment generated by a XbaI-EcoRV restriction digest of YCpPKZ1. The product, YCpPKZ6.1, contains a 360 bp PYK1 coding region fragment fused to LacZ coding region flanked by the PYK1 flanking regions (see Figure 6.1) The PYK1 coding region fragments of the hybrid transcription unit on YCpPKZ6.1 differs from that on YEpPKZ6 as follows;

YEpPKZ6	(Fragment X)	ATG TCT AGG ATC TGT.....etc
		START
YCpPKZ6.1	(Fragment Y)	ATG TCT AGA GGA TCT GT.....etc

The PYK1 coding region fragments present in YEpPKZ6 and YCpPKZ6.1 will be termed X and Y respectively.

The hybrid transcription units are very similar but fragment Y contains an AG insertion (underlined above) just downstream of the transcriptional start of the 360 bp PYK1 insert. This insertion abolishes translation at a site 27 codons downstream.

As predicted YCpPKZ6.1 does not confer B-galactosidase activity as shown by plates containing X-gal.

Figure 4.5 (chapter 4) illustrates an autoradiograph of a Northern blot containing RNA extracted from a YCpPKZ6.1 transformant and probed with a nick-translated plasmid containing LacZ (pMC1856). An RNA species of approximately the predicted length is detected (3.46

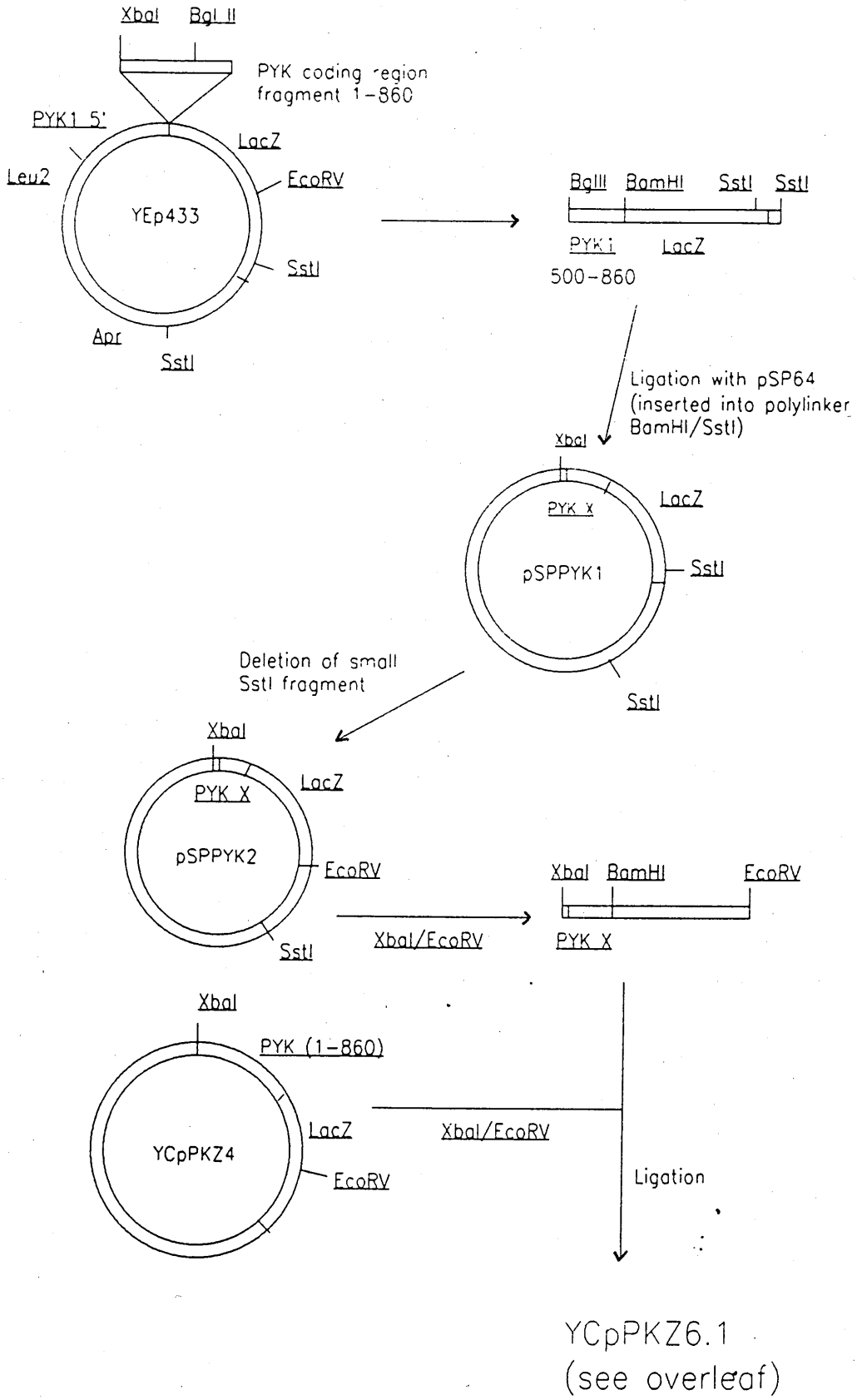


Figure 6.1 (a) Construction of pSPPYK1, pSPPYK2 and YCpPKZ6.1
The progenitor plasmid YEp433 was constructed by IJP Purvis. YCpPKZ4 is discussed in Chapter 4.1. The designations fragment X and Y are discussed in 6.2.1. The numbering of the PYK1 fragment refers to the number of bp from the first nucleotide of the translational initiation codon of the PYK1 gene.

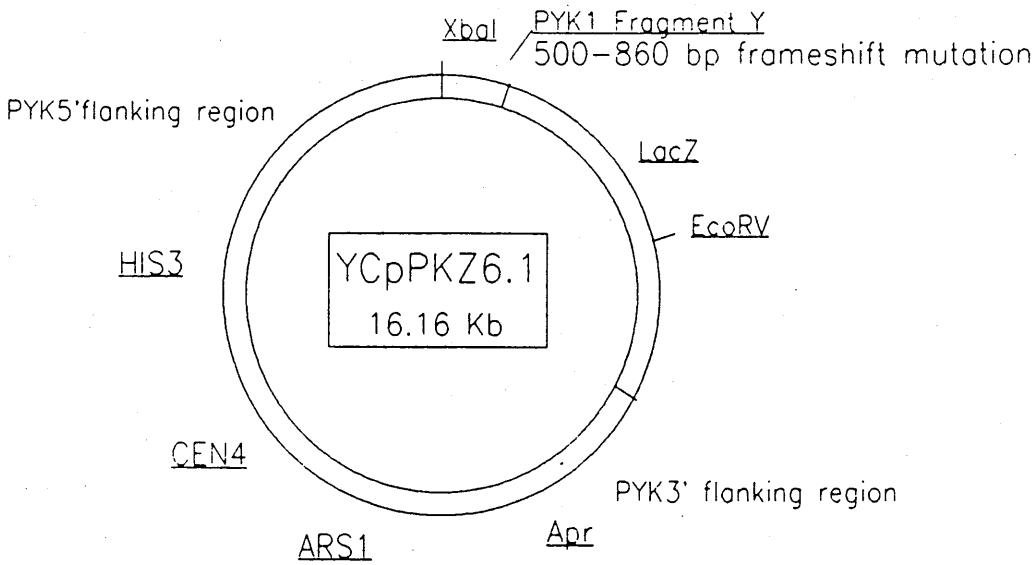


Figure 6.1 (b) Structure of YCpPKZ6.1
This plasmid is similar to YCpPKZ1 (Figure 4.1) however in this construct a PYK1 coding region fragment has been inserted between the translational initiation codon and the LacZ coding region. This insertion results in a frameshift and subsequent termination of translation of the PYK1 coding region fragment and the LacZ gene.

Kb). Thus the PYK/LacZ gene in YCpPKZ6.1 is transcribed.

6.2.2 PYK1 Coding Region Fragment Y does not Activate Transcription in YCpPKZ6.1

To establish whether fragment Y activates transcription in YCpPKZ6.1, transcription from this plasmid was compared to transcription from the control plasmid YCpPKZ1 (see Figure 6.1 and 6.2). The control plasmid is identical to YCpPKZ6.1 except the hybrid gene does not contain the PYK1 coding region fragment.

The abundance of PYK1/LacZ mRNA relative to actin mRNA was compared in cells containing either YCpPKZ6.1 or YCpPKZ1 (see Figure 6.2). In each case, measurements were made on three individual transformants to control for mutation which may arise during the process of transformation. These data indicate that the presence of PYK1 fragment Y has no affect on the abundance of hybrid mRNA.

This result contrasts with that obtained from a YEpPKZ6/YEpPKZ1 comparison (4.4) where the presence of fragment X is associated with a ten-fold elevation of hybrid mRNA abundance. The insertion of two base pairs into the 5'-end of fragment X (producing fragment Y and abolishing translation) suppresses elevation of mRNA abundance normally associated with these PYK1 coding region sequences.

There is more than one possible explanation for the above result. Either fragment Y failed to activate transcription or, the AG insertion decreases the mRNA stability. The latter explanation may seem likely for the following reason. mRNA stability may be dependent upon the degree to which the mRNA is being translated (Brown, 1989). Fragment Y contains a frameshift mutation which results in the loss in B-galactosidase activity, therefore it is likely that the ribosome loading on this

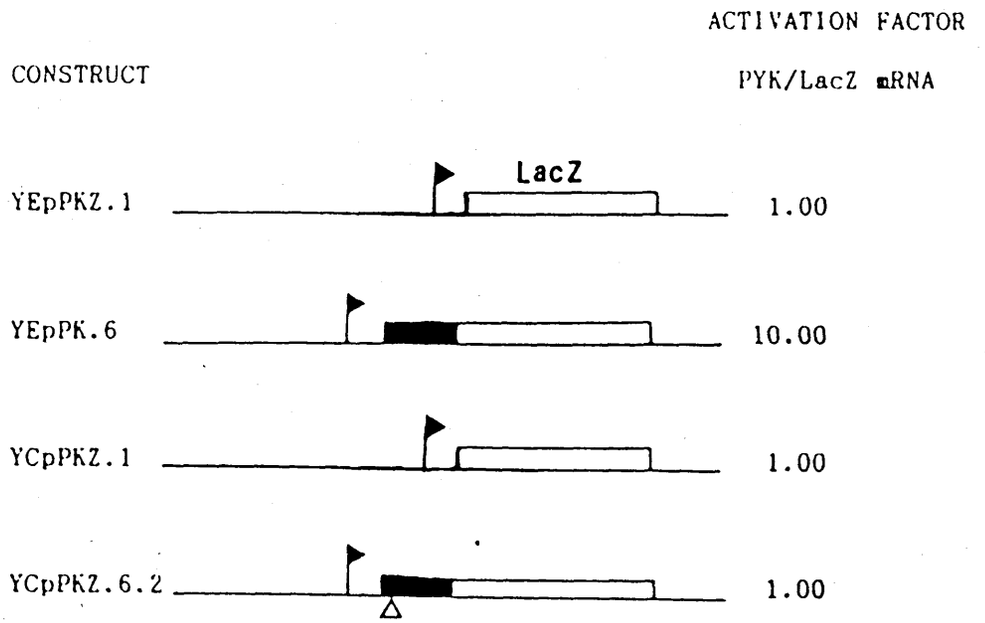


Figure 6.2 Activation Factors Associated with PYK1 Coding Region Fragment X and Fragment Y.
Activation factors are calculated relative to the constructs such that YEpPKG6 is compared to YEpPKZ1 and YCpPKZ6.1 is compared to YCpPKZ1. The activation factors were calculated as outlined previously (see legend to Figure 4.4) The PYK1/LacZ gene is translated in YEpPKZ6 transformants but not in YCpPKZ6.1 transformants due to the presence of a frameshift mutation and subsequent stop codon at the 5'-end of the PYK1 insert.

hybrid mRNA is low. As this may leave the hybrid mRNA more accessible to endonucleolytic activities, the mRNA stability may be much reduced. It follows that even if fragment Y activates transcription, the PYK1/LacZ mRNA abundance may appear low because the mRNA synthesised is unstable. In fact, the abundance may appear similar to that observed in strains which carry the control plasmid YCpPKZ1, where transcription is not activated, but the hybrid mRNA synthesised is stable (4.3.2 and 4.3.4). To firmly establish whether fragment Y (containing the frameshift mutation) activates transcription, the stabilities of the hybrid mRNAs originating from YCpPKZ1 and YCpPKZ6.1 were measured.

This experiment could have yielded one of two possible results. If mRNA originating from YCpPKZ6.1 (containing fragment Y) was found to be eight to ten times less stable than the mRNA originating from YCpPKZ1 then this would imply that fragment Y activates transcription. In contrast, if the stabilities of these mRNAs were found to be similar then it is unlikely that fragment Y activates transcription.

The stabilities of the hybrid mRNAs were measured in an identical fashion to that outlined in 4.2.4. mRNA 90 stabilities were also measured in each strain as an internal control to reveal any non-specific effects on mRNA stability. The half-lives of mRNA 90 in DBY746 is 6.6 +/-0.67 min, as measured previously (Santiago et al., 1986).

The decay of the hybrid mRNAs and mRNA 90 is depicted in Figure 6.3 (a) and (b). The half-life of mRNA 90 is similar in both strains. These results are in good agreement with previously published data (Santiago et al., 1986). Similarly, the stabilities of the hybrid mRNAs derived from YCpPKZ1 (control plasmid) and YCpPKZ6.1 (containing fragment Y) are comparable.

As the abundances and stabilities of these hybrid mRNAs

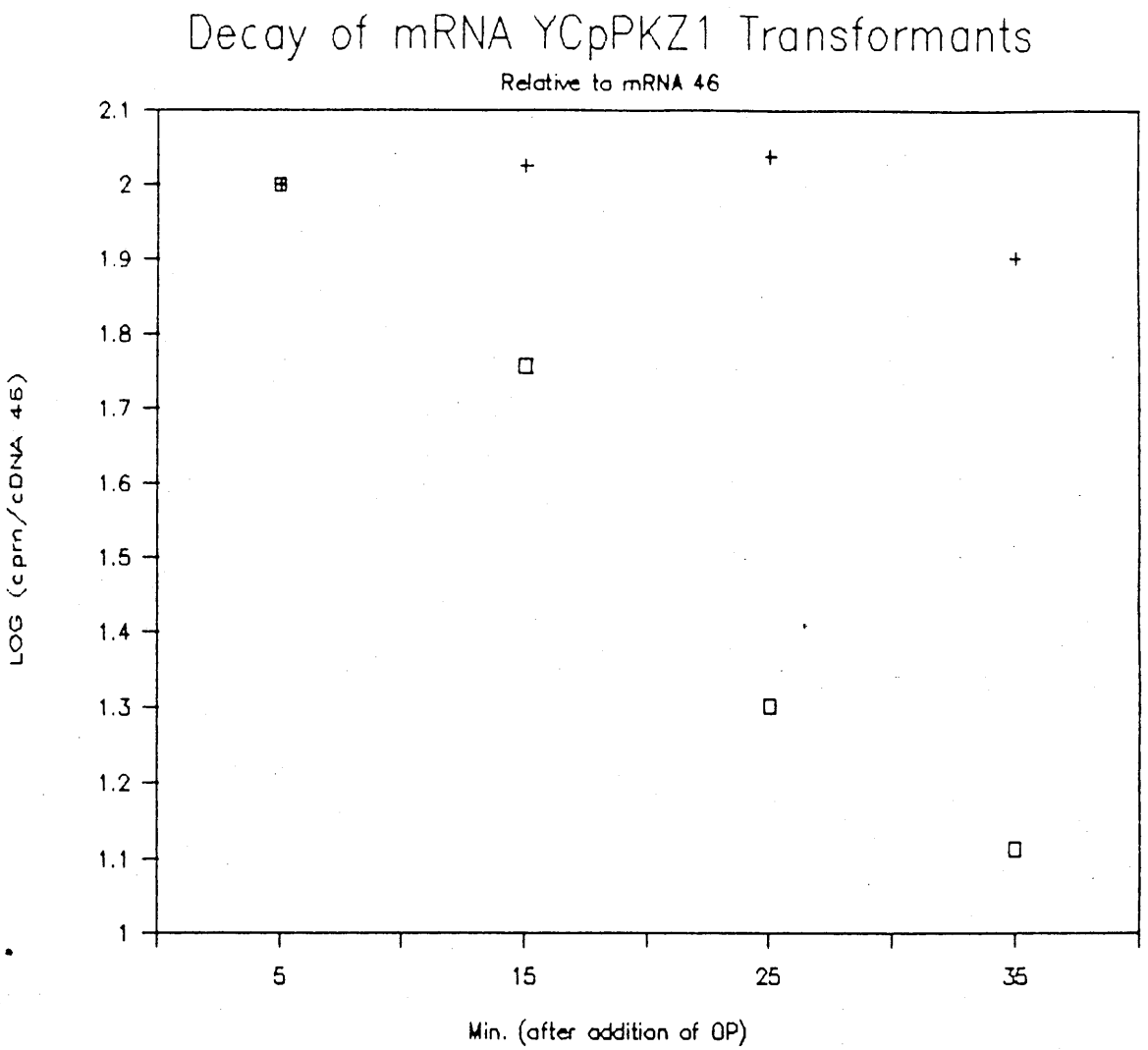


Figure 6.3 (a) Decay of PYK1/LacZ mRNA in YCpPKZ1 Transformants.
Squares represent the abundances of mRNA90 relative to mRNA46, at four time points after the addition of 80 ug/ml OP to an exponentially growing culture of a YCpPKZ1 transformant. The crosses represent the relative abundance of PYK1/LacZ mRNA at the same time points. A comparison should be made between this graph and that shown overleaf in Figure 6.3 (b).

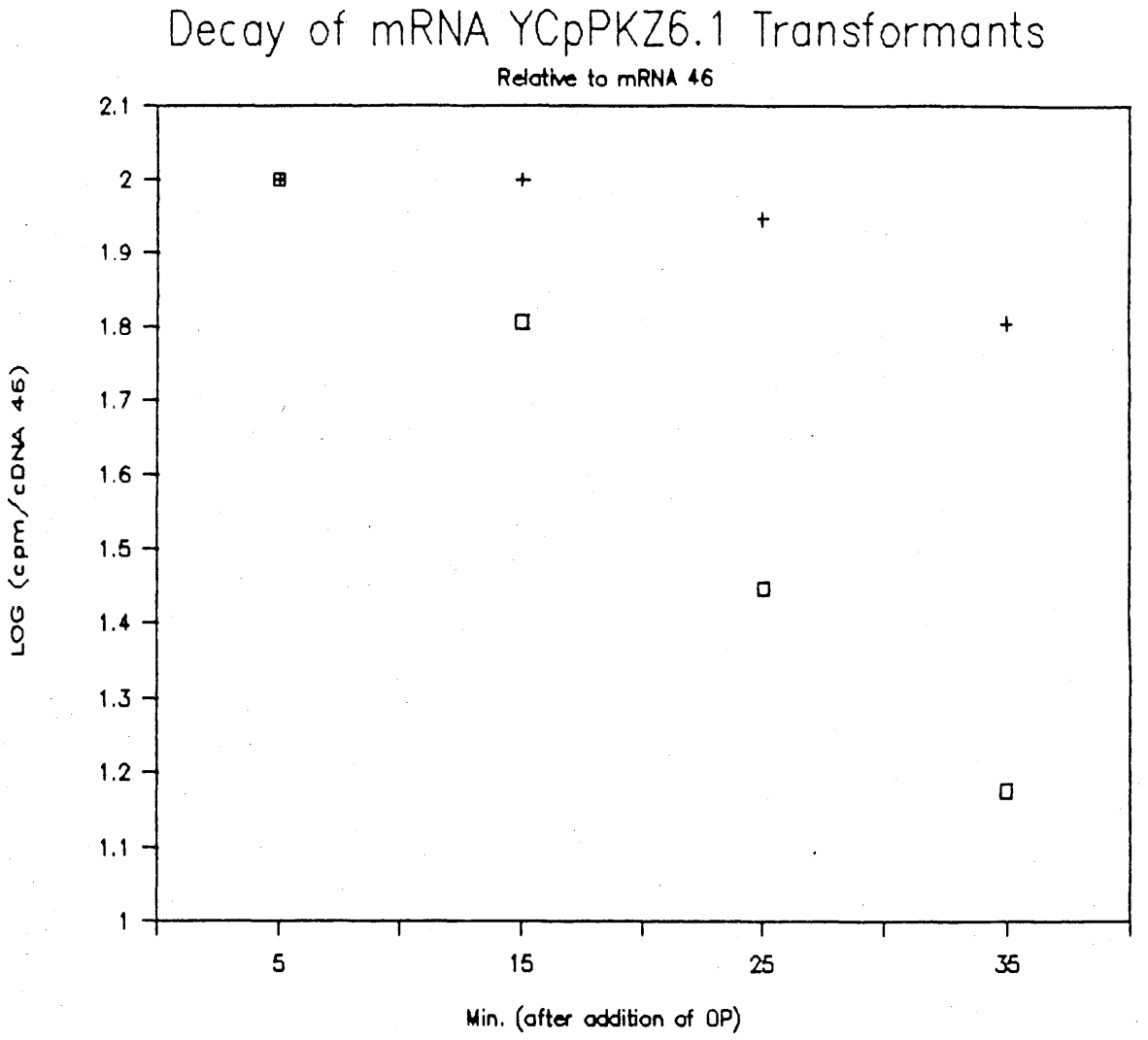


Figure 6.3 (b) Decay of PYK/LacZ mRNA in YCpPKZ6.1 Transformants.
Squares represent the abundances of mRNA90 relative to mRNA46, at four time points after the addition of 80 ug/ml OP to an exponentially growing culture of a YCpPKZ6.1 transformant. The crosses represent the relative abundance of PYK1/LacZ mRNA at the same time points. A comparison should be made between this graph and that shown previously in Figure 6.3 (a).

have been shown to be similar, it is concluded that the rates of mRNA synthesis are also similar. This implies that the presence of PYK1 fragment Y does not have an effect on the transcription of the hybrid gene.
Fragment X is associated with transcriptional activation whereas Fragment Y is not. Fragment Y contains an additional two base pairs which constitute a frameshift mutation.
The observed difference in transcriptional activation activity may be the result of an undefined mutation or it may be due to the aforesaid frameshift mutation. It is not possible, without further experimentation to rule out the possibility that a non-defined mutation could have influenced the observed mRNA abundances. But, as abundance was measured in 3 yeast transformants, such a mutation would have to have occurred within the plasmid prior to yeast transformation.
There are a number of ways in which the predicted frameshift mutation could have influenced the transcriptional activation function of the PYK1 coding region. It is possible that the AG insertion may have disrupted a cis-acting sequence motif which is essential for transcriptional activation. It is impossible without further study to rule out this possibility, but there is no evidence that such a sequence motif occurs exactly at this position. Alternatively, transcriptional activation may have been abolished by disrupting translation of the PYK1/LacZ mRNA by the insertion of a frameshift mutation at the 5'-end of the fusion coding region. The second explanation is consistent with the hypothesis that the PYK1 coding region activates transcription in trans via a post-translational product. In this case, the property of transcriptional activation is assigned to the pyruvate kinase/B-galactosidase fusion protein.
In summary, there is some evidence to support the notion of translation dependent-transcriptional activation by

the PYK1 coding region. Alternative explanations of the observations made have been outlined above.

6.3 Cells Expressing PYK1/LacZ or PYK1 Genes Have Elevated Phosphoglycerate Kinase (PGK1) mRNA Abundance

The data described above suggests that transcriptional activation by the PYK1 coding region may be dependent upon translation.

In the introduction to this Chapter the theoretical possibility of trans-activation by the PYK1 coding region was considered. It was proposed that, even in the presence of a chromosomally located PYK1 gene, experiments could be carried out to reveal trans-activation. The argument being that, when a reporter gene is introduced into the cell, it fails to be trans-activated by the chromosomal PYK1 gene because the trans-activator (a product of PYK1) is limiting. Only when an additional copy of the PYK1 coding region is introduced into the cell (in cis or in trans with respect to the reporter gene), can the reporter gene be transcribed at high levels. As yet, the only evidence in support of this is that activation by the PYK1 coding region, may be dependent upon translation. The theoretical consideration of trans-activation should now be extended.

For the sake of argument, let us accept that PYK1 trans-activates its own transcription. That is to say, transcription from the PYK1 "promoter" can be activated by information within the PYK1 coding region but that there is no requirement for this information in cis. What would be the consequences of such a mechanism? The most obvious consequence is that there would be a positive feedback loop operating, where increased expression of PYK1 would lead to further increases in expression. Therefore, there would appear to be a requirement for down regulation. This point will be discussed further in

the final assessment given in Chapter 7. One other conceivable consequence is that other genes may be subject to trans-activation by PYK1. It is this possibility that is investigated next.

It is quite likely that the expression of the glycolytic enzyme genes is co-ordinately regulated in response to carbon source (Maitra and Lobo, 1971; Cohen, 1986). At least part of this regulation takes place at the level of transcription (1.3). To facilitate such co-ordinate transcriptional regulation glycolytic genes share sequence elements within the upstream flanking regions. The corollary is then that the commonality of regulation of glycolytic genes extends to trans-activation by PYK1. That is, as this subset of genes are regulated in similar ways, it is possible that PYK1 can trans-activate other glycolytic enzyme genes.

It is likely that trans-activation by the PYK1 coding region requires a cis-acting site within the PYK1 promoter. Other genes containing similar cis-acting sites would also be subject to activation by the PYK1 coding region. The alternative is that PYK1 is regulated by an exceptional mechanism.

A candidate for trans-activation by PYK1, is PGK1. This gene shares many common features within the 5'-flanking region with PYK1 and apparently also requires coding region sequences for optimal transcription (Mellor et al., 1985 and 1987; 4.1). The relationship of the PGK1 "Downstream Activation Site" and the PYK1 coding region activator is unknown but it may be that a common regulatory mechanism is operating for both genes. PGK1 may contain cis-acting sequences (which presumably also occur in PYK1) which facilitate trans-activation by PYK1. Experiments were carried out to determine whether the expression of PYK1 coding region or coding region fragments had any influence on the expression of PGK1.

6.3.1 PGK1 mRNA Abundance is Elevated in Cells Expressing PYK1/LacZ Fusion Genes

To investigate the possibility that the PYK1 coding region containing hybrid PYK1/LacZ gene can trans-activate the transcription of other genes, the steady state abundance of PGK1 mRNA was measured in transformants containing plasmids YCpPKZ1 to 5 and YCpPKZ6.1. These transformants were described earlier (4.3 and 6.2).

mRNA abundance measurements were made on two or three individual transformants for each plasmid, to control for any non-specific mutation which may have occurred during the transformation process. RNA from exponentially growing glucose cultures of each transformant was serially diluted, immobilised on nitro-cellulose membrane as previously described and probed with a nick-translated plasmid containing the PGK1 coding region and bacterial plasmid sequences including the AP^R gene (as described in m&M). This probe has previously been shown to hybridise, predominantly, to PGK1 mRNA (I.J. Purvis, unpublished data) such that on a normal 16 hour exposure of a probed Northern blot only a single band is visible. The results of these measurements are shown in Figure 6.4 (a), (b) and (c). The activation factors are derived by firstly calculating PYK1/LacZ mRNA abundances relative to actin, then normalising the abundances relative to those obtained in YCpPKZ1 transformants. For example, the PGK1 mRNA abundance in cells containing YCpPKZ4 is 3.0 times greater than in cells containing the control plasmid, YCpPKZ1. This derivation is identical to that used in 4.3 when calculating the effect of PYK1 coding region fragments on the transcription of the PYK1/LacZ hybrid gene.

The results (Figure 6.4) show that in all transformants containing YCpPKZ3 and YCpPKZ4, the abundance of the PGK1

Figure 6.4 Trans-Activation Associated With PYK1 Coding Region Fragments
Activation factors are derived as follows;

(PGK1 cpm/actin cpm) / (PGK1 cpm/actin cpm)
[test plasmid] [YCpPKZ1]

cpm are derived from dot blot data as follows; cpm for each duplicate dilution are averaged and plotted against RNA dilution. Providing cpm are directly proportional to concentration of RNA, cpm associated with two RNA dilutions are averaged and used in the above activation factor determination.
An activation factor is an expression of the abundance of PYK1/LacZ or PGK1 mRNA in test transformants relative to that measured in YCpPKZ1 transformants. It is not an expression of percentage of total mRNA.
In (a) the average activation factor associated with each plasmid construct is noted.
In (b) the activation factors calculated for individual transformants is noted. The abundance of PYK1/LacZ and PGK1 mRNA was measured in 3 individual transformants of YCpPKZ1, YCpPKZ3, YCpPKZ6.1 and in 2 individual transformants of YCpPKZ2 and YCpPKZ4. As the abundance of PYK1/LacZ and PGK1 mRNA in all test transformants is measured relative to YCpPKZ1 a single measurement in the Figure for transformants of these plasmids.
In (c) the constructs are outlined. Numbers refer to the number of nuleotides from the PYK1 translational initiation codon.
The data noted here pertaining to the abundance of PYK/LacZ mRNA was previously shown in figure 4.4.

Activation Factors

A

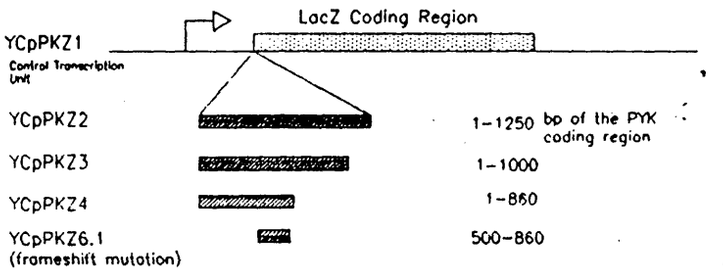
CONSTRUCT	PYK/LacZ	PGK
YCpPKZ1	1.0	1.0
YCpPKZ2	5.5	1.5
YCpPKZ3	7.0	3.5
YCpPKZ4	5.5	3.0
YCpPKZ6.1	1.0	1.2

Activation Factor

B

CONSTRUCT	PYK/LacZ	PGK
YCpPKZ1	1.0	1.0
YCpPKZ2	6.0	2.0
	4.5	1.0
YCpPKZ3	8.0	4.0
	7.5	2.5
	5.5	4.0
YCpPKZ4	4.5	1.5
	6.7	4.0
YCpPKZ6.1	0.5	1.4
	1.5.	1.0
	1.0	1.2

C



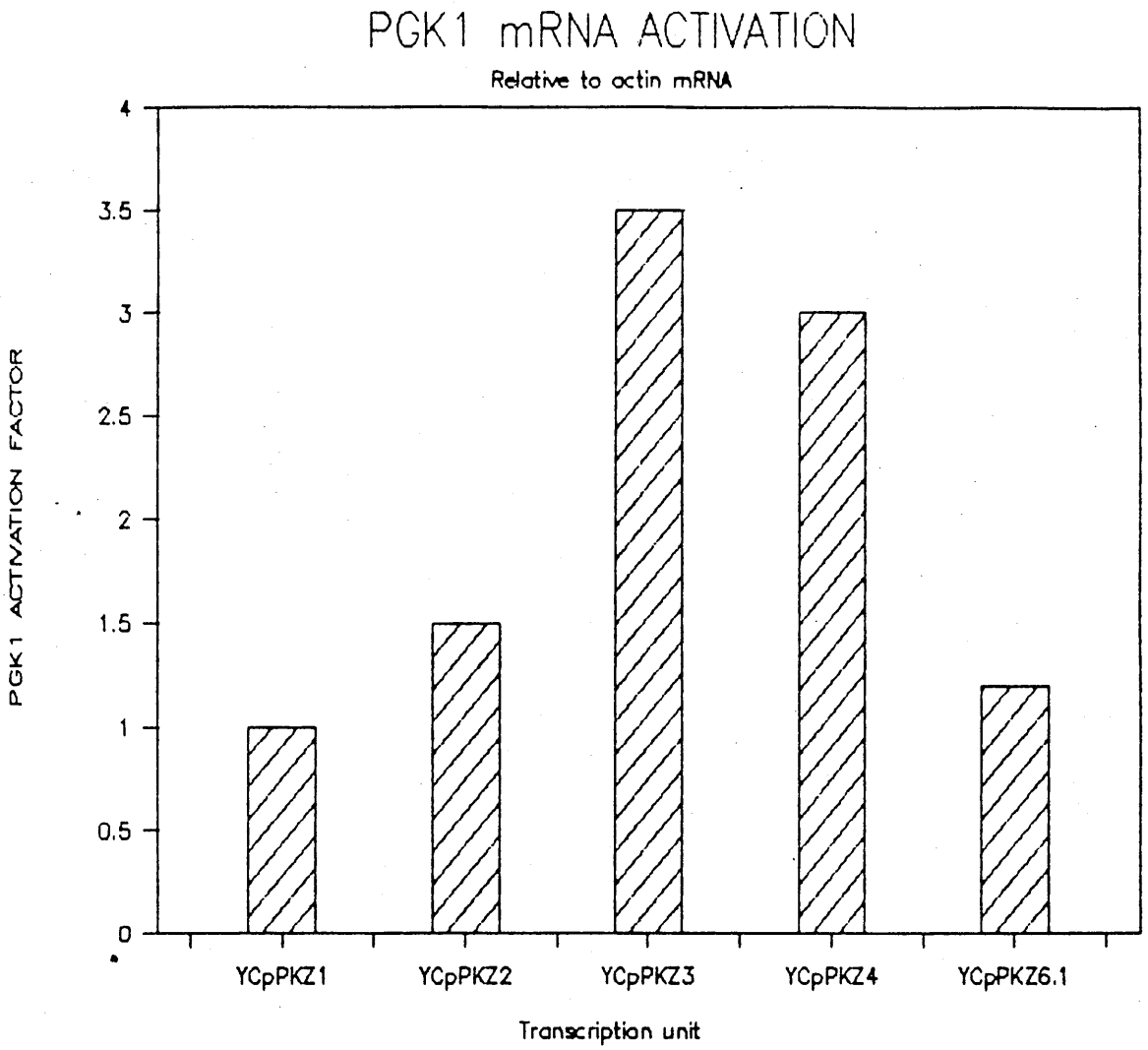


Figure 6.4 (d) Trans-Activation Associated with PYK1 Coding Region Fragments.
This histogram depicts data outlined on the previous page [Figure 6.4 (a)].

mRNA is elevated at least three-fold over the PGK1 mRNA abundance measured in transformants containing the control plasmid YCpPKZ1. PGK1 mRNA is normally very abundant in yeast, comprising 1% of the total mRNA (Mellor et al., 1985). In transformants containing YCpPKZ3 or YCpPKZ4, the abundance is increased to 3.5% of total mRNA. These results are shown graphically in Figure 6.4 (d).

In the above experiment, a yeast strain has been transformed with plasmids which differ only in the composition of the PYK1/LacZ gene in that some of these plasmids contain fragments of the PYK1 coding region. The abundance of PGK1 mRNA is apparently unequal in these transformants. All six transformants examined carrying YCpPKZ3 and 4, which contain PYK1 coding region fragments, have higher PGK1 mRNA abundances than transformants carrying YCpPKZ1, which does not carry any of the PYK1 coding region. Therefore, manipulation of the PYK1/LacZ hybrid gene apparently modifies the expression of the unlinked PGK1.

These data are consistent with trans-activation of PGK1 mRNA abundance by the expression of PYK1 coding region fragments.

6.3.2 Trans-Activation of PGK1 by the PYK1 Coding Region is Dependent Upon Carbon Source

It has been previously stated that it is likely that under some conditions, expression of the glycolytic enzyme genes is co-ordinately regulated in response to carbon source (4.2.3). The basis of this regulation is not yet fully defined. Therefore, any mechanism which contributes to the transcription of a glycolytic gene may be subject to carbon source control. If expression of the PYK1 coding region positively affects the transcription of PGK1, then this mechanism itself may be carbon source

Activation Factors

CONSTRUCT	PGK1	PGK1
YCpPKZ1	1.00	1.0
YCpPKZ2	1.50	nd
YCpPKZ3	3.50	1.4
YCpPKZ4	3.00	0.7
	[Glucose]	[Lactate]

Figure 6.5 Carbon Source Dependence of Trans-Activation by the PYK1 Coding Region Fragments
Activation factors represent the abundance of PYK1/LacZ and PGK1 mRNA relative to that measured in YCpPKZ1 transformants (see legend Figure 4.4). Transformants are grown upon 2% (w/v) glucose of 2% (w/v) lactate.

dependent.
Specific activities of glycolytic enzymes are depressed on non-fermentative gluconeogenic carbon sources, and are high on fermentative carbon sources (Maitra and Lobo, 1979). This prompted the investigation of the affect of PYK1 expression upon PGK1 expression in cells grown on the non-fermentative carbon source, lactate. Previously, transcriptional activation of PYK1/LacZ by PYK1 coding region fragments was observed with cells grown on glucose but not with those grown on lactose (4.2.3).
PGK1 mRNA abundances were measured in transformants grown on 2% lactate medium (YPL). Activation factors relative to the control plasmid YCpPKZ1 were calculated as described above. These abundance measurements are shown in Figure 6.5.
There is no detectable difference in the abundance of PGK1 mRNA between lactate-grown cells carrying test and control plasmids. There is therefore no evidence of trans-activation of PGK1 transcription by the PYK1 coding region when cells are grown on lactate. The phenomenon of trans-activation of PGK1 by PYK1 is carbon source dependent.
It is formally possible that trans-activation by PYK1 mediates carbon source regulation of the PGK1 gene.

6.3.3 Cells Containing Multiple Copies of the PYK1 Gene Exhibit Elevated PGK1 mRNA Abundance

The results outlined above indicate that expression of hybrid PYK1/LacZ genes containing PYK1 coding region fragments influence PGK1 mRNA abundance. In an attempt to confirm that this observation is due to a PYK1 coding region specific phenomenon, experiments were carried out in which the entire PYK1 gene was expressed in yeast from a multicopy plasmid, and the PGK1 mRNA abundance measured. The plasmids used in this particular experiment

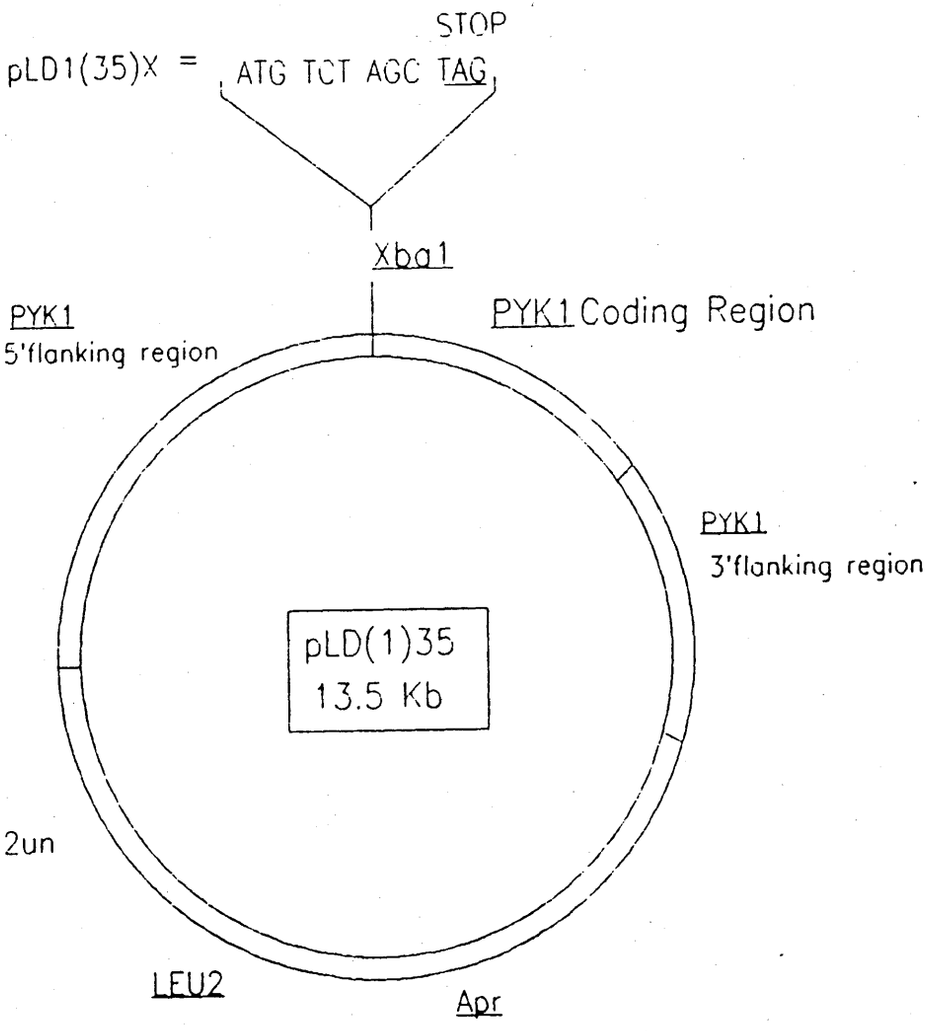


Figure 6.6 Structure of pLD1(35) and pLD1(35)X
The stop codon was introduced into pLD1(35) to produce pLD(35)X by Paul Moore (unpublished data).

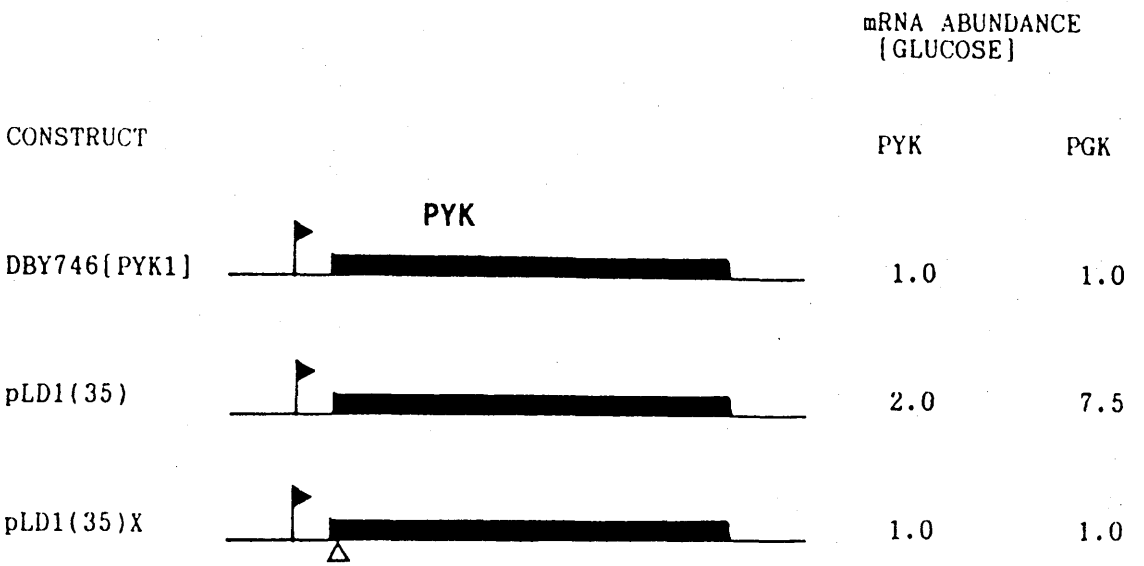


Figure 6.7 Trans-Activation of PGK1 mRNA Abundance by the Intact PYK1 Gene.
(A) Activation factors are calculated relative to the abundance of PGK1 mRNA in untransformed DBY746 cells. Activation factors are calculated as described previously (see Figure 6.4).
An autoradiograph of a dot blot from which these data are derived is illustrated overleaf (Figure 6.7b)

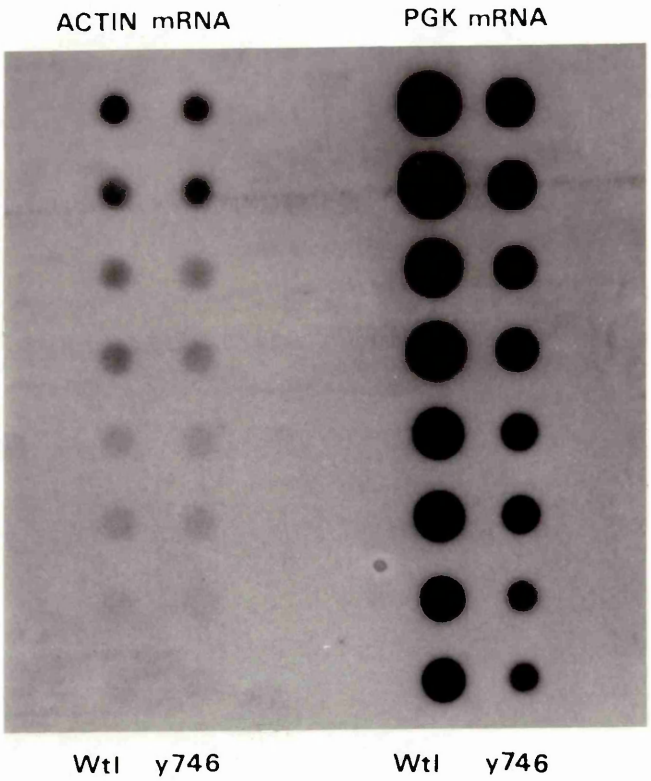


Figure 6.7 Trans-Activation of PGK1 mRNA Abundance by the Intact PYK1 Gene.
(b) This autoradiograph illustrates a dilution series (low dilution/top with duplicate loadings of each dilution) of RNA extracted from untransformed DBY746 (y746) and Wt1. Wt1 strain is a pLD1(35) transformant expressing the entire PYK1 gene. Identical filters were incubated with probes specific for either actin or PGK1 mRNA (as indicated above). The hybridisation of the actin specific probe indicates that both "lanes" of each filter contain equal amounts of RNA. The PGK1 specific probe indicates that the abundance of PGK1 mRNA is elivated in Wt1 cells. The relative hybridisation has been quantified and noted in Figure 6.7 (a).

differ from those previously described in that they do not contain any LacZ sequences nor do thy contain the sequences associated with centromeric based plasmids.
The abundance of PYK1 and PGK1 mRNA was measured in cells containing multiple copies of the PYK1 gene. The PYK1 gene was introduced on a multicopy 2um based vector, pLD1(35) (Figure 6.6) which also contains the LEU2 gene for selection in yeast. PYK1 and PGK1 mRNA abundances were also measured in cells containing pLD1(35)X (Figure 6.6). This plasmid contains an entire PYK1 gene into which a stop codon has been introduced after the third amino acid of the PYK1 coding region, thus abolishing translation of PYK1 mRNA. One would perhaps assume that the previously observed trans-activation of PGK1 transcription (6.3.1) was dependent upon a post-translational product of the PYK1/LacZ transcription unit. Measuring PGK1 mRNA abundance in cells carrying pLD1(35)X may have revealed whether trans-activation by PYK1 is dependent upon a post-translational product.
Before considering the consequences of PYK1 expression on PGK1, observations on the expression of PYK1 itself in cells containing these plasmids shall be discussed.
In this laboratory, Paul Moore has measured the abundance of PYK1 mRNA and plasmid copy number in a number of transformants containing pLD1(35) and pLD1(35)X. It was shown that the presence of an increased PYK1 copy number does not lead to a proportional increase in PYK1 mRNA abundance (Moore et al., submitted for publication). That is, as extra copies of the PYK1 gene are introduced into the cell, there is a decrease in the abundance of the PYK1 mRNA per gene copy. For example, when PYK1 is present at 5-6 copies, only a 2-fold increase in PYK1 mRNA abundance is observed. Such a "dosage compensation" was not observed when the PYK1 coding region was replaced with the LacZ coding region (Purvis et al., 1987). One interpretation of this observation is the following;

limitation of mRNA abundance at high copy number may be mediated, at least partly, by the PYK1 coding region.
Transformants containing pLD1(35)X do not exhibit elevated PYK1 mRNA abundance. This suggests that the insertion of a stop codon after the ATG of the PYK1 coding region abolishes the elevation in PYK1 mRNA abundance associated with the presence of the PYK1 transcription unit on a high copy number plasmid.
There are a number of possible explanations of this observation.
The introduction of this stop codon may have disrupted a sequence essential for transcription. However, this explanation seems highly unlikely as hybrid mRNA abundance in cells transformed with YEpPKZ6, which does not contain the first 500 bp of the PYK1 coding region, is high (4.2.2). Alternatively, the abolition of translation may destabilise the otherwise stable PYK1 mRNA. This explanation also appears unlikely, as the stability of PYK1 mRNA is independent of translation (Purvis et al., 1987b). Formally, it is possible that abolition of translation prevents elevation of PYK1 mRNA abundance because the normal transcription rate is dependent upon translation of the PYK1 coding region. This explanation is consistent with the hypothesis that the PYK1 coding region activates its own transcription via a post-transcriptional product.
We will now consider the consequences of expression of plasmid borne PYK1, on PGK1 mRNA abundance. PGK1 mRNA abundances were measured relative to actin mRNA, in wild type host cells, and cells transformed with PLD1(35) or pLD1(35)X. The abundances were then normalised to the PGK1 mRNA abundance of the host strain DBY746. These relative abundances are shown in Figure 6.7. In a pLD1(35) transformant (wt1), PGK1 mRNA abundance was apparently elevated over seven-fold when compared with the host strain DBY746. In contrast, a strain containing

pLD1(35)X (mut7) exhibits a PGK1 mRNA abundance comparable to the host strain DBY746.
It would appear from these data that the presence of a multicopy plasmid carrying the PYK1 gene causes an increase in the abundance of the PGK1 mRNA, but the presence of a stop codon at the beginning of the PYK1 coding region prevents mRNA abundance elevation.
These results are consistent with two previous observations; expression of PYK1 trans-activates the expression of PGK1, and transcriptional activation activity of the PYK1 coding region is dependent upon translation (previously this observation was made on the transcription of the hybrid PYK1/LacZ transcription unit).
It should be noted that these measurements have been made on single transformants and therefore the data may be attributable to non-defined mutations in the transformed strains. However, these data are fully consistent with the hypothesis that the PYK1 coding region activates transcription in trans, and with previous observations.

6.4 Expression of the PYK1 Coding Region May Affect the Abundance of a Number of Glycolytic Enzyme mRNAs

The data previously described, strongly suggests that expression of PYK1 coding region sequences may affect the abundance of PGK1 mRNA. As glycolytic genes are likely to be co-ordinately expressed it is possible that trans-activation by PYK1 is not restricted to PYK1 and PGK1. The abundances of a number of other glycolytic enzyme mRNAs were measured to investigate whether expression of PYK1 influences the expression of other genes.
It is not possible to predict which glycolytic genes may be trans-activated by PYK1 and therefore genes were chosen for investigation essentially at random, although an attempt was made to include genes coding for enzymes

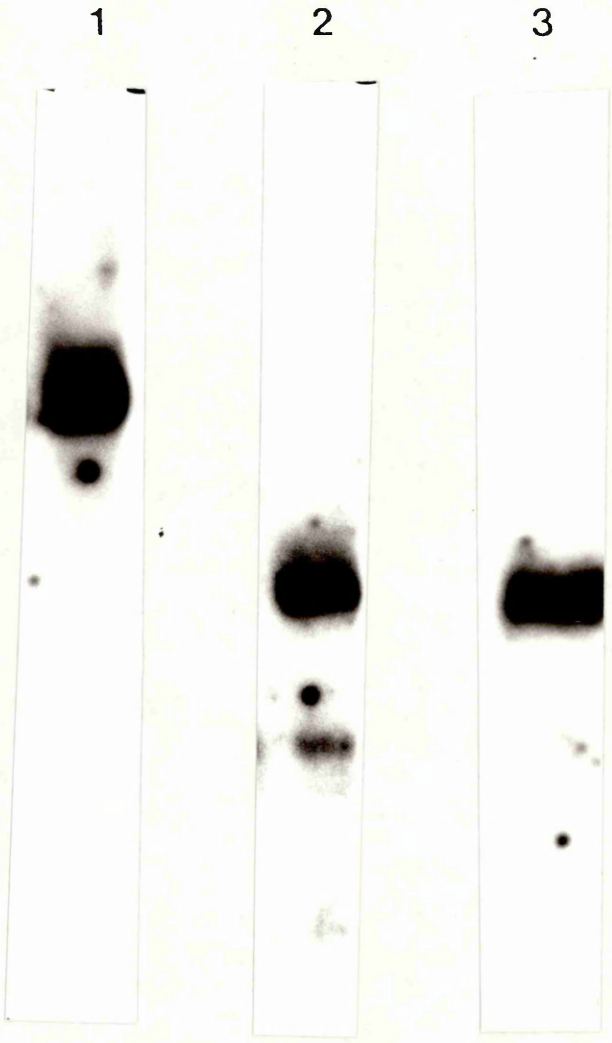


Figure 6.9 Northern Blot of Glycolytic Enzyme mRNAs
RNA from untransformed DBY746 was blotted and individual lanes probed with glycolytic enzyme specific oligonucleotide probes. The hybridisation conditions are noted in Figure 2.1. Three autoradiograph examples are illustrated.
Lane 1 : oligo ENO (enolase mRNA specific)
Lane 2 : oligo TPI (triose phosphate isomerase mRNA specific)
Lane 3 : oligo PGM (phosphoglycerate mutase mRNA specific)

A

Activation Factors

CONSTRUCT	PYK/LacZ	PGK	HXK	TPI	PFK1	PFK2	PFK1,2	ENO	PGM
YCpPKZ1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
YCpPKZ2	5.5	1.5	nd	nd	nd	nd	1.2	nd	nd
YCpPKZ3	7.0	3.5	2.3	1.2	2.7	1.6	2.3	1.2	0.9
YCpPKZ4	5.5	3.0	nd	nd	nd	nd	2.5	nd	nd
YCpPKZ6.1	1.0	1.2	1.0	1.0	1.8	0.8	1.0	0.9	0.6

B

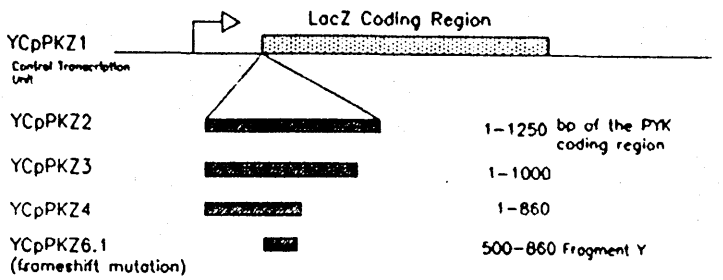


Figure 6.10 Effect of PYK1 Coding Region Fragment Expression on Glycolytic Enzyme mRNA Abundance

(a) Activation factors are calculated relative to the abundance of the appropriate mRNA measured in the control YCpPKZ1 transformants by the method previously outlined (see legend to Figure 6.4). The data corresponding to PYK1/LacZ and PGK1 mRNA abundances has been shown previously (Figure 6.4).

nd = Not Determined.

(b) Sketch of contructs. These are shown in 4.3.

from different parts of the glycolytic pathway.
Total RNA from strains containing YCpPKZ1, YCpPKZ4 or YCpPKZ6.1 was probed with nick-translated plasmid, or radiolabelled oligonuleotides, complementary to specific glycolytic genes.
Oligonucleotide probes were designed such that they hybridised to regions of sequence common to all known members of the enzyme family. For example, two genes have been sequenced which code for yeast hexokinase (HEXA and HEXB). An oligonucleotide was synthesised which is complementary to 33 bp of sequence which is shared by both genes.
RNA was probed with a nick translated plasmid containing the phosphofructokinase 1 gene (PFK1) which codes for phosphofructokinase subunit A and hybridises to sequences of both PFK1 and PFK2 under the conditions used (P. Moore, unpublished data). The RNA was also probed with oligonucleotides complementary to the mRNAs of hexokinase (HEXA and HEXB), triose phosphate isomerase (TPI), enolase (PENO8 and PENO46), phosphoglycerate mutase (PGM), phosphofructokinase 1 (PFK1) and a separate oligonucleotide which was complementary to PFK2. These oligonucleotides are listed in figure 2.1. Each oligonucleotide was shown to hybridise to a single band on a Northern blot. The hybridisation conditions were established empirically by varying the concentration of formamide in the hybridisation solution and adjusting stringency of washing. Figure 6.9 illustrates a Northern blot (prepared by Lynn Taylor) which has been probed with oligonucleotides specific to the enolase, triose phosphate isomerase and phosphoglycerate mutase mRNAs. (Hybridisation conditions for oligonucleotides complementary to hexokinase and phosphofructokinase mRNAs were established by Lynn Taylor, unpublished data)
The abundance of the various specific mRNAs was measured relative to actin and then normalised to the mRNA

abundance observed in transformants containing the control plasmid, YCpPKZ1.

Figure 6.10 shows the measured abundance of a number of mRNAs in cells containing YCpPKZ4 and YCpPKZ6.1 relative to the control plasmid YCpPKZ1. The previously determined relative abundances of PYK1/LacZ and PGK1 mRNA are also noted.

Differences in the abundance of some mRNAs can be observed between the various transformants. In particular, PFK1 and HEX mRNA abundance is at least two-fold higher in strains carrying YCpPKZ4 than in strains carrying YCpPKZ1, in which PYK/LacZ and PGK1 mRNA abundance is not elevated. The increase in PFK1 mRNA abundance is reflected in the total abundance of both PFK mRNAs as measured by the nick-translated plasmid probe. However, no other mRNA abundance appears to be elevated to the extent of PYK\LacZ or PGK1 mRNA.

It is difficult to understand the significance of these results. Each of the measurements discussed above should be made on a number of transformants containing each plasmid to allow an accurate assessment of the effects of expression of the PYK\LacZ hybrid gene on glycolytic enzyme mRNAs. Nevertheless, this data does not contradict the view that PYK1 coding region expression affects the abundance of more than one glycolytic mRNA and the possible significance of PYK1 trans-activation of PGK1, HEX and PFK will be discussed in Chapter 7.

6.5 Cells With Elevated PGK1 mRNA Abundance Are Not Undergoing a "Heat Shock" Stress Response

In the early part of this Chapter transcriptional activation by a post-translational product of PYK1 was considered. The correlation of PGK1 mRNA abundance elevation and translation of the PYK1 gene may be

considered strong evidence of the existence of such a mechanism. Trans-activation of the expression of one metabolic enzyme gene by the expression of another is a mechanism without precedent. Thus, one must carefully consider whether the observations made are artifactual or whether they allude to an unusual regulatory mechanism. For example, over-expression of PYK1 or PYK1/LacZ genes may have an indirect, non-specific affect on transcription patterns.

In an effort to explain why strains expressing PYK1 coding region have elevated PGK1 mRNA abundance, consideration has been given to circumstances in which PGK1 mRNA synthesis is modulated. On heat shock for example, the abundance of PGK1 mRNA is raised six-fold. This abundance increase is mediated by the presence of the Heat Shock Element (HSE) approximatly 300 bp upstream of the transcriptional start site (Piper et al., 1986 and 1988).

Heat Shock Response can be stimulated by a number of environmental or intra-cellular conditions, perhaps even by the presence, within the cell, of denatured proteins (Pelham, 1988). One could argue that the apparent "trans-activation" of PGK1 mRNA abundance by the expression of the PYK1 coding region, is in fact, due to a "heat shock" type stress response incited by the synthesis of PYK/B-galactosidase protein or over-production of the PYK protein. The general stress response in turn may then activate the transcription of the PGK1 gene.

The heat shock response is a transient event (usually subsiding after 60 min.), so one would envisage that the stress response would be activated only a short time before RNA was prepared for quantitation in the above experiments.

In order to establish the validity of the above explanation, cells with high PGK1 mRNA levels were examined for symptoms of stress. There are a number of

characteristic changes in cell metabolism during stress response and this is reflected by comprehensive changes in the transcription of a number of genes.
The stress state of PGK1 overexpressing transformants was assessed by measuring growth rate and the abundance of a heat shock and a non-heat shock specific mRNA.

6.5.1 Growth Rates of YCpPKZ1, YCpPKZ3 and YCpPKZ4 Transformants Are Identical

Figure 6.11 shows the growth rates of strains containing YCpPKZ1, YCpPKZ4, and YCpPKZ6.1. One may expect that if a strain over-expressing PGK1 (YCpPKZ4 transformants) was stressed, then it may exhibit a lower growth rate than the control strain in which PGK1 mRNA abundance is normal. This however, is not observed. Transformants containing YCpPKZ3 and YCpPKZ4 have slightly higher growth rates than transformants containing the control plasmid YCpPKZ1.

6.5.2 Heat Shock Protein 90 and Ribosomal Protein L3 mRNA Abundances Indicate PGK1 Over-Expressing Cells Are Not Undergoing a Stress Response

On heat shock, yeast initiates the synthesis of a number of specific heat shock proteins and temporarily reduces, or ceases the synthesis of non-heat shock proteins. An example of this second class is the ribosomal proteins. On heat shock, the abundance of ribosomal protein mRNAs undergoes a sudden decline The observed 87% reduction in abundance is a result of a decrease in mRNA synthesis and stability (Herruer et al., 1988). Therefore, an increase in the synthesis of an Heat Shock Protein (HSP) mRNA and/or a significant reduction in the abundance of ribosomal protein mRNAs is indicative of a Heat Shock response.

Transformant	Growth Rate [Doubling Time]
YCpPKZ1	140
YCpPKZ3	120
YCpPKZ4	125

Figure 6.11 Growth Rates of Transformants Expressing PYK1/LacZ Genes
Doubling Times were determined on cultures growing exponentially on YPG broth by measuring optical density at A_{600}. The significance of these measurements is discussed in 6.5.1.

To investigate the stress status of transformants used in this study, the abundance of a Heat Shock protein mRNA (HSP90) and of ribosomal protein L3 mRNA were measured.
An oligonucleotide complementary to part of the Heat Shock Protein 90 coding region (Figure 6.8) was synthesised for the detection of the HSP90 mRNA. This oligonucleotide may however hybridise to at least one other yeast gene; Heat Shock Cognate 90 (HSC90). HSP90 and HSC90 are approximately 90% similar at the nucleotide level (Farrelly et al., 1982), however the sequence of HSC90 remains unpublished. In contrast to HSP90, the level of HSC90 expression is only mildly altered on heat shock. Thus, if the HSP90 oligonucleotide hybridised strongly to the HSC90 mRNA, it may be impossible to differentiate between stressed and unstressed cells using the above probe.
To determine whether probing RNA with this oligonucleotide could differentiate between a stressed and an unstressed culture, yeast was grown on GYNB and heat shocked (23^{o}C to 38^{o}C transition). RNA was then prepared at 0, 15, 30, and 60mins after heat shock. The heat shock status of the culture was confirmed by examining the protein synthesis profile (Jane MacKenzie, unpublished data). The HSP90 oligonucleotide was shown to hybridise to a single band on a Northern blot of RNA prepared at 15 min after heat shock (Figure 6.12). This result is consistent with the work of Finkelstein et al., (1982) who demonstrated that induced synthesis of HSP90 was regulated at the level of mRNA abundance. One can assume therefore, that the presence of HSP90 mRNA as detected by the HSP90 oligonucleotide, is diagnostic of cells undergoing a heat shock or stress response.
Equal quantities of RNA from host strain DBY746 and wt1 [pLD1(35)] transformants were dotted onto nitrocellulose. To ensure that low levels of HSP90 mRNA could be detected, ten times the standard amount of RNA was

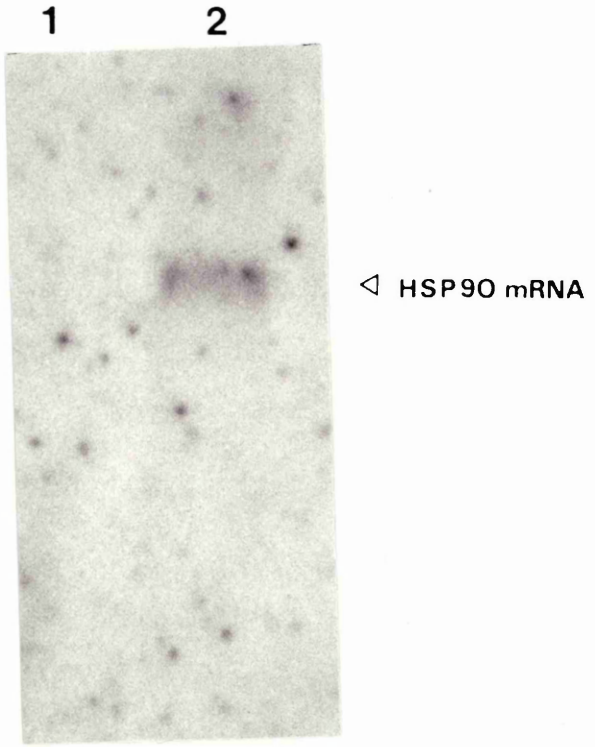


Figure 6.12 Heat Shock RNA Contains HSP90 mRNA
RNA was extracted from heat shocked and non-heat shocked cells (see 6.5.2, Jane McKenzie, unpublished data) and northern blotted. This filter was then probed with an oligonuceotide probe specific for the HSP90 mRNA.
Lane 1 : Non-heat shock RNA
Lane 2 : Heat shock RNA (as defined by protein synthesis profile; Jane McKenzie, unpublished data).

probed with end-labelled HSP90 oligonucleotide. The amount of probe hybridised was compared and found to be very similar for each RNA sample.

There are two possible interpretations of the above result. As the HSP90 probe hybridises to RNA from both strains with equal intensity then its possible that both strains, DBY746 and wt1, are stressed to the same extent. If this is so, then stress response does not explain why PGK1 mRNA abundance is elevated in the pLD1(35) transformant. Alternatively, perhaps the signal represents hybridisation to HSC90 mRNA. If so, then HSP90 mRNA is undetectable in PGK1 over-expressing cells. Again, PGK1 over-expression cannot be attributable to the stress response.

RNA from cells carrying YCpPKZ1, YCpPKZ4 and YCpPKZ6.1 were also probed for the presence of HSP90 mRNA but no detectable hybridisation was obtained. Therefore, HSP90 mRNA could not be detected in cells in which the expression of PGK1 is apparently trans-activated. The absence of probe hybridisation in this experiment and the presence of hybridisation in previous experiments may be the result of differences in the total amounts of RNA loaded onto each filter.

An oligonucleotide complementary to the ribosomal Protein L3 mRNA (RP1; Figure 2.1) was end-labelled and used to probe RNA from transformants containing plasmids YCpPKZ1, YCpPKZ2, YCpPKZ3, YCpPKZ4, and YCpPKZ6.1. The abundance of RP1 mRNA in these transformants was calculated relative to actin mRNA and is shown in Figure 6.13. These data demonstrate that the abundance of ribosomal protein L3 mRNA does not vary significantly between strains.

It would appear that the over-expression of PGK1 does not correlate with depressed RP1 mRNA abundance. That is to say, there is no evidence of stress, as defined by RP1 mRNA abundance, in cells in which the abundance of PGK1 is elevated.

CONSTRUCT	RP1 cpm/Actin cpm
YCpPKZ1	0.05
YCpPKZ2	0.04
YCpPKZ3	0.06
YCpPKZ4	0.06
YCpPKZ6.1	0.05

Figure 6.13 Relative Abundances of RP1 mRNA
These figures represent the relative abundance of ribosomal protein 1 mRNA between transformants carrying YCpPKZ plasmids. It should be noted that these figures do not represent the percentage of total mRNA nor do they reflect the relative abundance of actin and RP1 mRNAs as the length and specific activities of the actin and RP1 probes vary. Rather, these figures indicate that the abundance of RP1 mRNA, relative to actin, does not vary in a fashion consistent with the presence of PYK1 coding region fragments in the plasmid borne PYK1/LacZ transcription unit.

This result suggests that PGK1 over-expression observed in this study is not due to cell stress as defined by the heat shock response.
The conclusions derived from the above three experiment are made on the assumption that the abundance of actin mRNA does not radically alter on stress.

6.6 Summary of Results

At the beginning of this Chapter the question of why a sequence specific transcriptional control element should arise within an area of sequence constraint, the coding region, was considered. Transcriptional activation by a post-translational product of the PYK1 gene was then introduced as a formal possibility. Such a mechanism contrasts with that by which UASs influence transcription (reviewed 1.2).
The experiments outlined in this Chapter were carried out to investigate the hypothesis that transcriptional activation by the PYK1 coding region is dependent upon a post-transcriptional product.
Section 6.2 presents mRNA abundance and mRNA stability results which indicate that the insertion of a frameshift mutation close to the start of the 360bp PYK1 coding region fragment abolishes the transcriptional activation previously associated with this PYK1 fragment. The frameshift mutation may inhibit transcriptional activation by disrupting an essential sequence or by abolishing translation of the hybrid mRNA.
The results of these experiments are consistent with the hypothesis that translation is required for transcriptional activation by the PYK1 coding region. If transcriptional activation was determined exclusively by a specific DNA sequence, then it is unlikely that translation of these sequences would have any bearing upon transcription rate.

Clear evidence for trans-activation of mRNA abundance by the PYK1 coding region comes from examining PGK1 mRNA in cells over-expressing PYK1 coding region. Over-expression of the entire PYK1 coding region, or fragments of the PYK1 coding region fused to LacZ, apparently can cause an elevation in PGK1 mRNA abundance (6.3). This trans-activation of PGK1 gene expression is dependent upon carbon source. Trans-activation is observed when cells are grown on the fermentative carbon source, glucose, but not on the non-fermentative carbon source, lactate.

Activation of expression of one metabolic enzyme gene by the expression of another is an unprecedented observation. It is prudent therefore to consider carefully whether this observation is artifactual or whether it reflects an unusual regulatory mechanism. One source of artifactual results was thought to be transcriptional activation via the stress response, however no symptoms of cell stress were observed in yeast in which the expression of PGK1 was trans-activated.

If indeed a PYK1 gene product participates in an unusual transcriptional trans-activation mechanism, then there are a number of possible consequences. These will be discussed in the following Chapter. In addition, consideration will be given to ways in which PYK1 expression could affect PGK1 expression and whether or not there is evidence in the literature to support the observations made in this Chapter.

CHAPTER 7

TRANSCRIPTION OF THE PYK1 GENE IN YEAST
A DISCUSSION

7.1 Transcriptional Activation by the PYK1 Coding Region

In this section the results presented in the previous three Chapters, pertaining to the "transcriptional activator" located within the PYK1 coding region, will be considered. Models will be presented which will describe how this particular activator may function. Evidence from the literature, both in support of, and contradictory to, each model, will be presented.

As part of the assessment of these models, the possible consequences of each will be speculated upon freely. By speculating extensively, it is not intended to inflate the importance of the results previously presented or to solicit unjustified support for a particular model. Rather, it is to reveal possible experimental approaches by which particular models may be tested.

Results presented in Chapter 4 indicate that the PYK1 coding region contains information, which, if present within a transcriptional unit, will lead to an increase in the rate at which that unit is transcribed. The chromosomally located PYK1 coding region does not compensate for the absence of the PYK1 coding region from a plasmid borne gene. This strongly suggests that the PYK1 coding region contains a cis-acting transcriptional activator.

On the whole, this interpretation of the results compares favourably with that of Mellor et al., (1987) who propose the existence of a Downstream Activation Site (DAS) within the PGK1 coding region. Both PGK1 and PYK1 apparently contain an element, at this unusual location, which contributes to transcription. It is a formal possibility that the PGK1 "DAS" is involved in the transport of the PGK1 mRNA from the nucleus to the cytoplasm. PGK1 RNA lacking the DAS may be unstable in the nucleus (Jane Mellor, personal communication), however there is no direct evidence to support this. As

both genes code for glycolytic enzymes, one may expect their expression to be coordinately regulated. Comparison of the 5'-flanking regions of the two genes reveals similarities. Both contain TATA-like homologies and CT blocks (3.1) but more importantly, both appear to contain sequences which act as a binding site for the RAP1/TUF1 protein. Indeed, DNA from the PYK1 5'-flanking region has been shown to inhibit the interaction of a PGK1 UAS fragment with a specific yeast protein in vitro (Chambers et al., 1989) suggesting that the two genes are under the influence of at least one common trans-acting factor. It is therefore conceivable that a trans-acting factor binds to a common site within the coding region of both genes. Based on this, one may propose the following model: "The PYK1 coding region contains a cis-acting site(s) which activates transcription by acting as a binding site(s) for a trans-acting regulatory protein". One may also add that, "this protein also binds to the PGK1 DAS, activating transcription and thus co-ordinately regulating the expression of both genes". We will call this the "binding site" model (Figure 7.1).

Some evidence, which in principle supports the binding site model, has come recently from the unpublished work of Joy Rathjen and Jane Mellor. They have shown that, when the PGK1 coding region is replaced with that of the PYK1 coding region, large amounts of PYK protein are synthesised (personal communication). This is circumstantial evidence that the transcriptional defect normally associated with the deletion of the PGK1 coding region DAS is compensated for by the insertion of the PYK1 coding region suggesting that coding region transcriptional activators are at least, complementary.

One would predict from the binding site model that the PYK1 and PGK1 coding regions would contain a common sequence which specifically binds a trans-acting protein. A comparison of sequences do show that some sequence

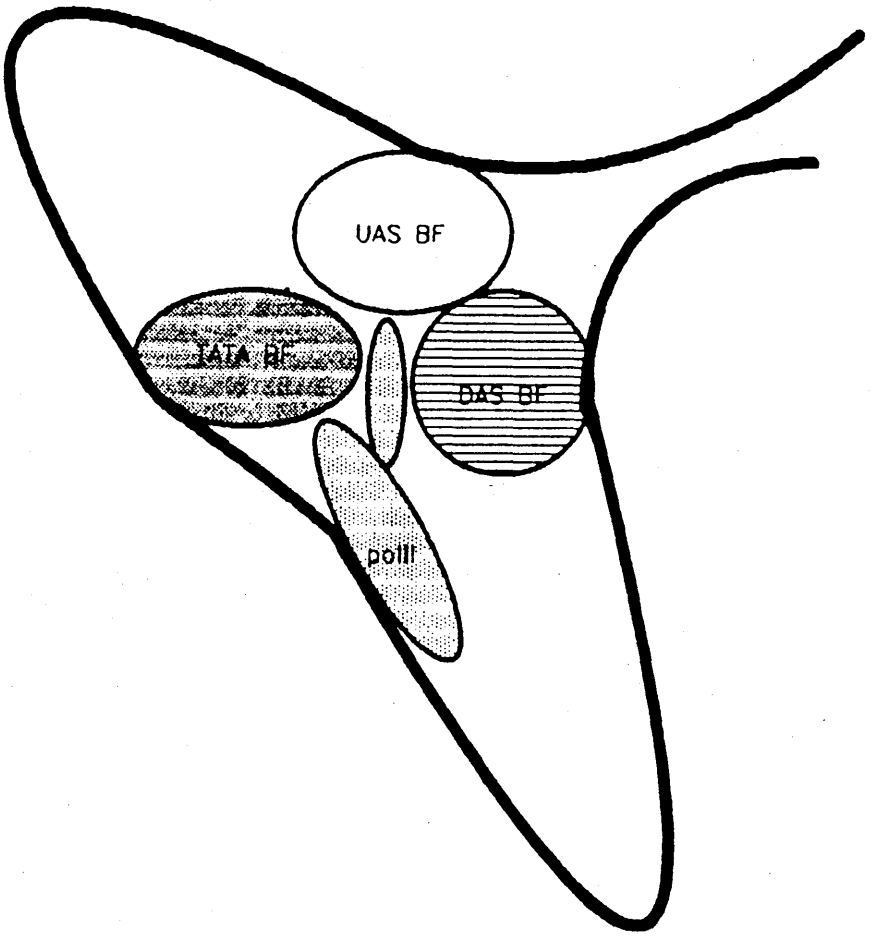


Figure 7.1 Binding Site Model.
This figure represents a possible arrangement of factors during the initiation of transcription of PYK1. Black line: PYK1 DNA. UASBF: UAS binding factor. TATABF: TATA element binding factor. polII: RNA polymerase II (with "tail" represented by a connected ellipse). DASBF: Downstream Activation Site binding factor. There is no evidence for the existence of DASBF other than that described in Chapter 4. This is one model which explains the requirement for the PYK1 coding region for maximal expression of the PYK1 transcription unit. An alternative model is outlined in Figure 7.2.

motifs are shared. For example, a 14 bp consensus sequence (with a 2 to 6 bp gap) is found in the coding regions of not just PYK1 and PGK1 but also in ENOA, TRP1, HEXA, HEXB and GLD (Alistair Brown and Andy Bettany, personal communication). In all cases, translation of these sequences yields an identical set of amino acids. Therefore, the occurence of this motif could be attributed to a combination of high codon bias and primary amino acid sequence, and not to the conservation of a specific protein binding site. A sequence motif, 5'-TGGTGGTGG-3', occurs in the sense-strand of PGK1 and in the antisense-strand of the coding region of PYK1 (Peter Moir, personal communication). However, this sequence lies outwith the 360 bp region which has been shown by Purvis to activate transcription (4.4). The 360 bp PYK1 fragment has been compared to the PGK1 sequence using the GAP programme of the MGC software package. This revealed a 9 bp match between a sequence, 5'-AAGGACAAG-3,' within the PYK1 activator fragment and the PGK1 coding region. However, as with the 14 bp consensus described above, this sequence also codes for identical amino acids in both genes.

In the light of the inability to identify a significant common sequence motif, a modification of the above model could be made such that "the PYK1 and PGK1 coding regions bind an identical trans-acting factor, but at dissimilar DNA sequences" or that "the PYK1 and PGK1 coding region do not bind a common trans-acting factor".

With this in mind, we will now consider other sequences within the PYK1 coding region which may act as potential binding sites. A search for sequences (utilising the programme BESTFIT) which have been previously characterised as transcriptional elements, reveals no significant similarities. Particular attention was paid to a search for the sequence, 5'-CACGTGA-3' which may be involved in the transcriptional regulation of the LPD1

gene from a downstream site. Yeast LPD1 (which codes for lipoamide dehydrogenase) contains two copies of the sequence 5'-CACGTGA-3'; one is within the 5'-flanking region and the second is located 75 bp downstream of the translational start site. A very similar sequence also has a role in initiating transcription of the TRP1 gene (Mellor et al, 1988). In addition, the sequence forms the core of a consensus, found within all yeast centromeres, representing the binding site for centromere binding protein (CP1). The function of this sequence in LPD1 is not yet defined but a DNA fragment containing the downstream site has been shown to interact with a specific yeast protein in vitro (Joe Ross, PhD thesis; Edinburgh, 1988). This protein may be CP1. In addition, preliminary evidence tentatively suggests that expression of a heterologous coding region from the LPD1 promoter is enhanced by the presence, in cis, of the LPD1 coding region (Zaff Zaman, personal communication). The LPD1 gene, therefore, may contain a DAS. However, the PYK1 coding region does not contain the 5'-CACGTGA-3' sequence.

Another "downstream" transcriptional element is the 5S ribosomal RNA gene internal control region (1.2.2). RNA polymerase III transcription is dependent upon the integrity of the binding site for transcription factor IIIA. A search has been made for RNA polymerase III promoter sequences within the PYK1 coding region, but no significant similarities were discovered.

To test the binding site models proposed above, experiments were carried out to search for protein/DNA interactions in vitro formed by the 360 bp fragment of PYK1 coding region which was previously shown to activate transcription. These experiments were, to a degree, encouraging in that a protein was shown to preferentially bind to a PYK1 coding region fragment, but the results were by no means conclusive (protein exhibited 10-fold

affinity preference for the PYK1 fragment; results reviewed in Chapter 5). Further work is required to confirm that a specific interaction takes place in vivo. Specific experiments which should now be carried out toward this end, have been stated elsewhere (5.4).

To complete the examination of the "binding site" model we should consider how a protein bound downstream of the I-site could contribute to transcription.

By analogy with transcriptional factors which bind to UAS's, a DAS binding protein may activate transcription by protein/protein interactions (1.2.3). This interaction could be with another transcription factor, such as TATA binding factor, or with RNA polymerase itself, perhaps stabilising an otherwise unstable (pre-)initiation complex. As outlined in the Introduction (1.2.1), yeast RNA polymerase II has a repeated heptapeptide "tail" which is apparently required for transcription (Allison et al., 1985). It has been hypothesised that this tail contacts activator domains of regulatory proteins (Allison et al., 1988). There are no obvious reasons why the polymerase tail could not make contact with a protein bound to a downstream site (Figure 7.1). However, it would seem unlikely that the tail could make identical contacts with both a protein bound upstream of the I-site and a protein bound downstream. Therefore, perhaps the downstream binding factor would make distinctly different contacts with polymerase, than the UAS bound protein.

Alternatively, a protein bound within the coding region may influence transcription rate by altering the conformation of the DNA and associated nucleosome positioning around the I-site. As outlined in the introduction, there are a number of changes in DNA conformation associated with transcriptional initiation (1.2.3). The removal of nucleosomes from promoter sequences has been shown to accompany transcriptional induction (Almer et al., 1986).

The proposed binding of CP1 to sites within and upstream of the LPD1 coding region may result in nucleosome "clearing" from a large region around the transcription initiation region allowing access for other transcription factors (Ian Dawes, personal communication).

In summary, the "binding site" model is consistent with a number of results presented in this study and with data collected from the study of PGK1 and LPD1. Also, there are a number of potential mechanisms by which a protein bound at a downstream site could influence the rate of transcription. However, we will now discuss some inconsistencies between predictions prompted by the binding site model and experimental observations.

For example, one may predict that activation would be independent of orientation of the PYK1 coding region element and occur when the element is placed at variable distances downstream from the I-site. Likewise, the activator should function when located outwith the coding region or within a region which is untranslated. It follows that the nuclear function of the activator should be independent of translation in the cytoplasm. The experiments described in the early part of Chapter 6 (6.2) test whether translation of the 360 bp PYK1 activator fragment (fragment X) is required for transcriptional activation. The results obtained are consistent with the prediction that activation is dependent on translation, although it is possible that absence of activation was due to disruption of a cis-acting site. If translation is a prerequisite of transcriptional activation then the "binding site" model, in itself, is not a sufficiently accurate description of the mechanism by which PYK1 transcription rate is determined. The requirement for a post-translational product would suggest that a cis(coding region)-located site is not required for transcriptional activation. In the introduction of Chapter 6 , consideration was given

to the probability of a sequence specific element evolving within a yeast gene coding region. In general, it was considered that such an element would be more likely to evolve within a flanking region. This, in addition to the experimental observations described above might suggest that the PYK1 transcriptional activator operates in trans. However, as noted above the PYK1 chromosomal locus does not appear to compensate for the deletion of the PYK1 coding region from a plasmid borne PYK1 gene; when the coding region is restored, transcription is elevated. This strongly suggests that a cis-acting site within the coding region is required. However, as outlined in the introduction to Chapter 6, it is formally possible that activation is due to a post-translational product, but that the level of this product derived from the chromosomal locus is insufficient for activation. The presence of an additional PYK1 promoter, in the absence of additional post-translational product, then results in an observed reduction in transcription from the plasmid-borne PYK1 promoter.

A purely trans-mediated transcriptional activation mechanism perhaps seems rather unlikely but remains a formal possibility.

The binding site model appears increasingly inadequate, when one considers the measurements made on the abundance of PGK1 mRNA in transformants overexpressing the entire PYK1 gene or PYK1 coding region fragments fused to LacZ (6.3). These results strongly suggest that the PGK1 mRNA abundance is influenced by the expression of the PYK1 coding region. This process is also dependent upon translation of the PYK1 coding region and confirms that the PYK1 coding region can activate transcription in trans. Of course, these results may be artifactual and this issue was examined, with particular reference to the "stress" response in 6.5.

For the sake of argument, consider now a model based on

these observations, and on the formal possibility that the PYK1 coding region is not required in cis for transcriptional activation. The model is that, "PYK protein positively regulates the transcription of the PYK1 gene and that of other glycolytic enzyme genes, either directly or indirectly" (Figure 7.2). This shall be called the "trans feedback" model in order to differentiate it from the "protein feedback" model proposed by Chen et al., (1987) which describes the requirement of the PGK1 coding region for high transcription (4.1).

Perhaps the experimental evidence for this model is best summarised by the graph illustrated in Figure 7.3. The Figure legend contains a description of how this graph has been constructed. Basically, each point on the graph represents a transformant and its position is an indication of the PYK1/LacZ and PGK1 mRNA abundance. The transformants vary in the plasmids which they carry and the carbon source on which they were grown. The interesting feature of this graph is that it reveals a qualitative correlation between the abundance of PYK1/LacZ mRNA and the abundance of PGK1 mRNA.

In considering the "trans feedback" model, we are immediately confronted with a positive feedback loop; as PYK protein is synthesised, PYK1 transcription is activated leading to an increase in the abundance of the PYK1 mRNA, which in turn is translated to produce more protein. For this system to function efficiently, negative regulation is required. This may be achieved simply by limiting a specific factor required for PYK1 expression, either at the level of transcription or translation. Research undertaken in this laboratory, but independent of this study, has yielded strong evidence to suggest that the synthesis of PYK protein is limited at the level of translation (Moore et al., manuscript submitted to gene). Pyruvate kinase activities and the

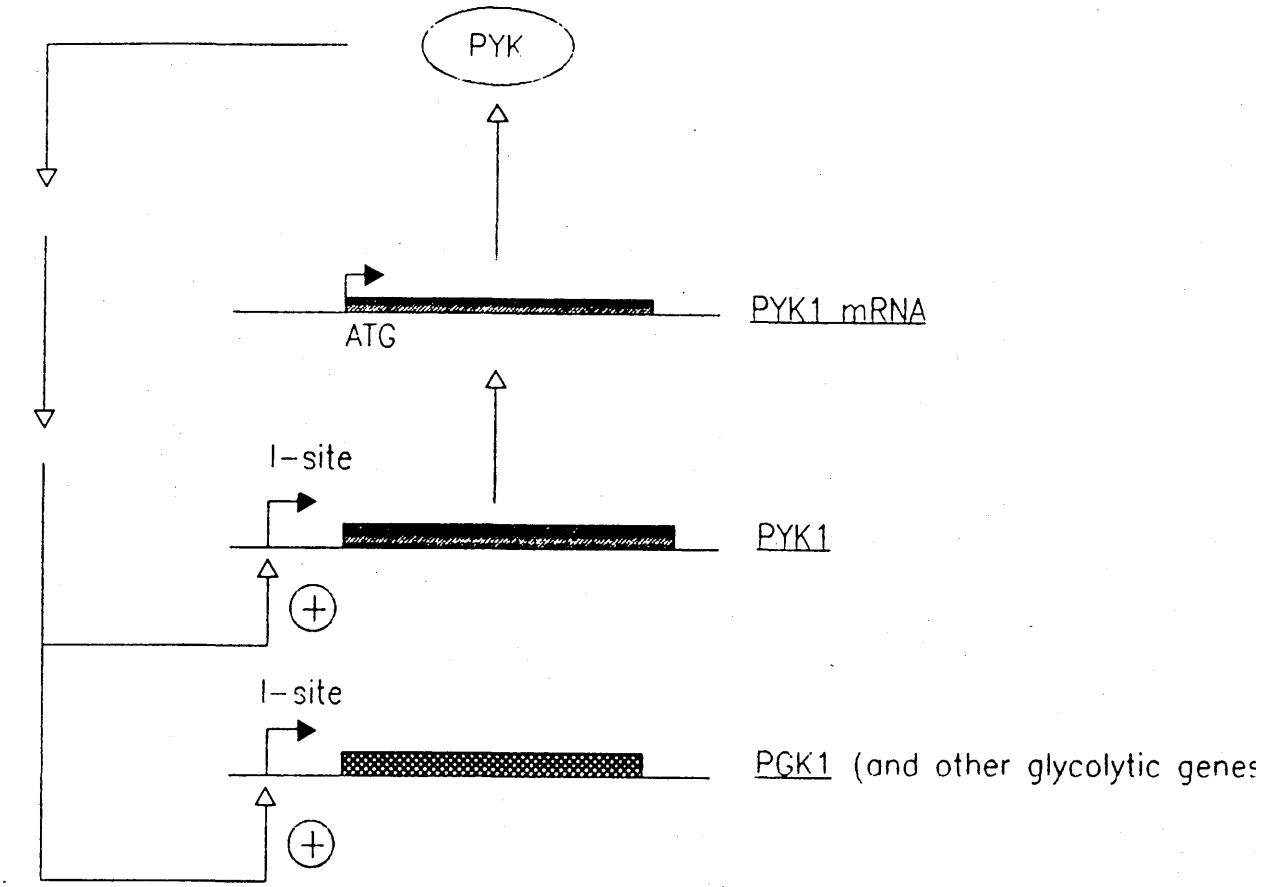


Figure 7.2 Trans-Feedback Model
This figure depicts the possible consequences of the synthesis of PYK protein in yeast. The filled arrows indicate initiation sites of transcription (for *PYK1* and *PGK1*) and translation (*PYK1* mRNA). The clear arrows outline the flow of information which specifies transcriptional activation. The model implies that the synthesis of PYK protein in yeast indirectly induces the transcription of *PYK1*, *PGK1* and perhaps other glycolytic enzyme genes. The evidence for and against this model is discussed in Chapter 7. This is one model which explains the requirement for the *PYK1* coding region for maximal expression of the *PYK1* transcription unit. An alternative model is illustrated in Figure 7.1.

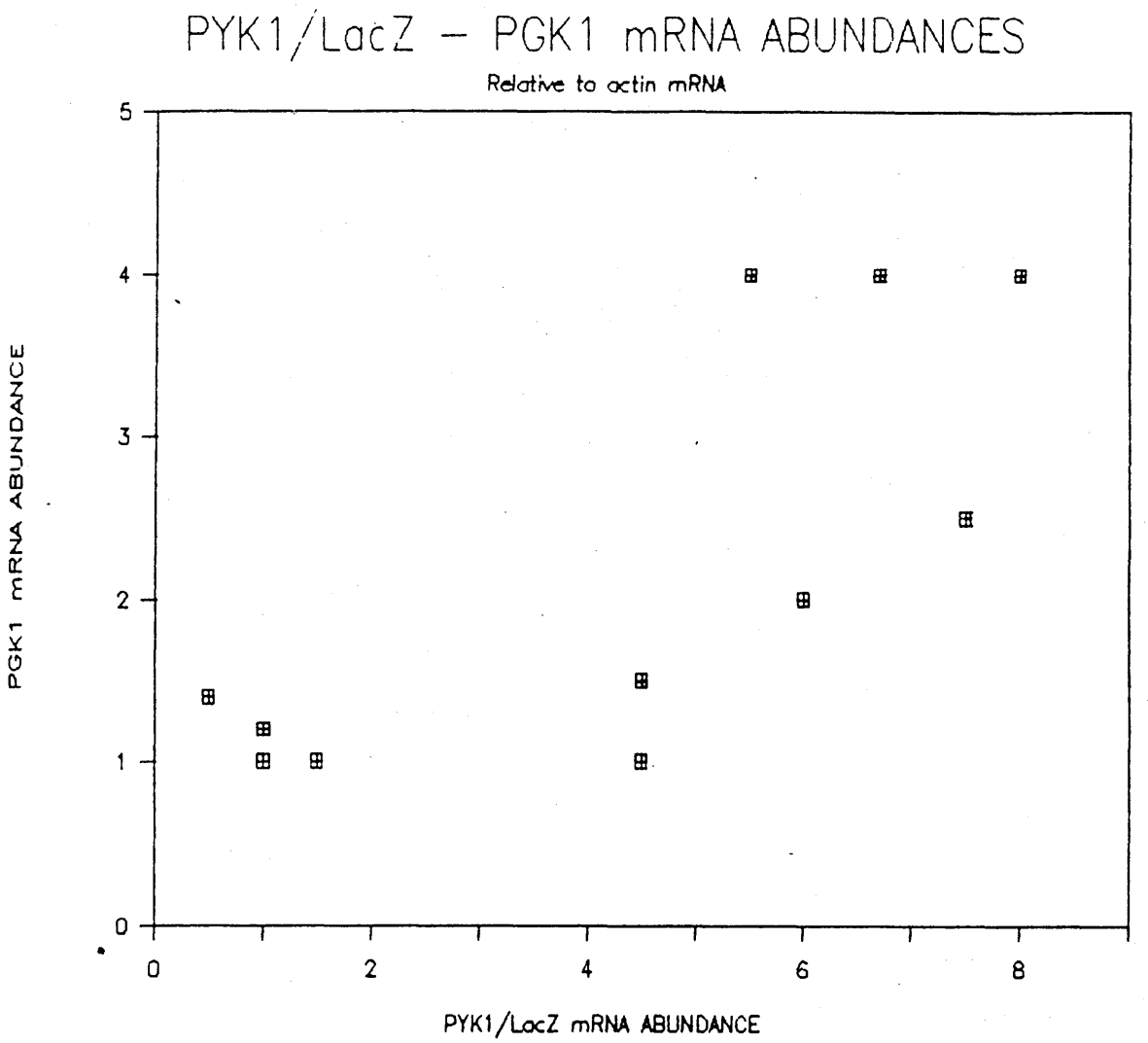


Figure 7.3 Relationship Between PYK1/LacZ and PGK1 mRNA Abundance.
Activation factors derived from the abundances of PYK/LacZ and PGK mRNA have been calculated (see Figure 6.4 and 6.5) For individual transformants, the activation factor associated with PYK1/LacZ has been plotted against the activation factor associated with PGK1. As the activation factor is a function of mRNA abundance relative to actin mRNA, the graph axis have been labelled simply "mRNA abundance". The numbers do not correspond to percentage of total mRNA

ribosome loading of PYK1 mRNA were assessed in transformants containing different PYK1 mRNA abundances. As mRNA abundance increases, the ratio of enzyme activity:mRNA decreases, as does the ribosome loading of the PYK1 mRNA. The reduction in ribosome loading was shown to be specific to PYK1 mRNA. These data suggest either that the PYK protein specifically inhibits the translation of PYK1 mRNA or that a PYK1 specific translation factor becomes limiting when PYK1 mRNA abundance is elevated. Theoretically, this mechanism provides a point at which the positive feedback loop, inherent in the trans feedback model, could be constrained.

Perhaps the most striking feature of the trans feedback model, is that it implies that a metabolic enzyme has influence over the transcription of its own gene and the genes of other metabolic enzymes. There is no precedent for such a regulatory feature. However, at least one intriguing publication indicates that some metabolic enzymes may be related to the transcriptional regulation of other metabolic enzyme genes. Anton et al., (1987) have shown that a eukaryotic transcriptional repressor protein, the product of the Neurospora crassa qa-IS gene, is homologous to part of the arom multifunctional enzyme complex. The qa gene cluster encodes the three enzymes which are required for the conversion of quinate to protocatechuic acid, thus allowing growth on quinate as a sole carbon source. The second enzyme activity of this catabolic pathway is identical to the third step of the biosynthetic shikimate pathway, which is catalysed by the arom enzyme complex. Thus, not surprisingly, the qa-3 gene is homologous to that part of the arom complex which is associated with the same enzymatic activity (Anton et al., 1987). The qa gene cluster is activated by the product of the qa-IF gene and repressed by the product of the qa-IS gene. Quinate, the substrate, induces

transcription of the qa genes. It is thought that qa-IS protein represses transcription by interacting with the qa-IF activator protein (Huiet and Giles, 1986). The remarkable homology of the qa-IS repressor to the arom complex, and to qa-3, is the only known example of unequivocal homology between a transcriptional repressor protein and a metabolic enzyme. It is possible that the common ancestor of the transcriptional repressor protein and the arom complex had both regulatory and metabolic capabilities. The trans feedback model requires that pyruvate kinase has a similar dual role.

We will now deliberate on how a post-translational product of the PYK1 gene could influence the abundance of its own mRNA and other mRNA species. This could occur by either a direct or an indirect mechanism.

Possibly the most direct mechanism by which mRNA abundance may be influenced is activation of transcription by direct contact with the transcriptional initiation complex. This could be achieved by interaction of the PYK protein with a binding site close to the PYK1 gene. Prompted by this, a search was made for potential DNA binding characteristics within the primary structure of the PYK1 protein. The primary amino acid sequence was examined for the occurrence of cysteines and histidines at intervals which are consistent with the formation of "zinc finger" DNA binding domains (1.2.3.1). No such suitable spacing was observed. Likewise, the distribution of leucines are inconsistent with the formation of a "leucine zipper" (1.2.3).

Direct DNA binding would also require that the PYK protein was present within the nucleus. However, there is no amino acid sequence motif which resembles the SV40 nuclear location signal or the general nuclear localisation signal proposed by Newport and Forbes (1987) or indeed any of the sequences required by yeast ribosomal proteins for nuclear localisation (Moreland et

al., 1984).

It would seem that PYK protein does not exhibit any characteristics of a DNA binding protein. However, for completeness the binding of purified PYK to PYK1 sequences could be assayed.

Perhaps it is more realistic to imagine that PYK influences the transcription of genes by a more indirect mechanism. In principle, PYK could regulate any point in the glycolytic transcriptional regulatory circuit. This may occur by sequestration of factors which inhibit the transcription of glycolytic genes or by the activation of factors which are required for transcription. PYK protein domains or surfaces which mediate this regulation may be distinct from those which are involved in catabolic functions.

By considering some of the metabolic consequences of the trans feedback model we may be able to correlate some of the predictions to previously published observations.

The trans feedback model predicts that the expression of PYK1 may lead to an increase in the abundance of a number of other glycolytic enzyme mRNA's. This elevation of mRNA abundances may, in turn, lead to an increase in glycolytic enzyme specific activities. The over-expression of a single metabolic enzyme rarely leads to an increase in metabolic flux. However, if the synthesis of all the enzymes of the glycolytic pathway was co-ordinately increased, then one may expect to see changes in the flux of glucose to ethanol and biomass unless strong product feedback inhibition was operating. In principle, glycolytic flux could be controlled indirectly by the expression of a single gene, PYK1. One would predict from this that, if no other factor was limiting, over-expression of the PYK1 gene would lead to an increase in glycolytic flux. Schaaff et al. 1989, have recently demonstrated that this is clearly not the case. Yeast was transformed with a multicopy plasmid which

carried the PYK1 gene. Ethanol production and biomass were measured in these transformants relative to a control strain transformed with a non-PYK1 containing plasmid. Although PYK specific activity was shown to increase (8.6-fold), no increase in metabolic flux, as defined by ethanol and biomass production, was observed. We can conclude that there is no clear relationship between the expression of PYK1 and glycolytic flux. This result may reflect strong product inhibition of glycolytic enzymes. However, it is feasible that the expression of PYK1 induces not all, but a subset of glycolytic enzyme genes. This would imply that the expression of glycolytic enzyme genes is, to a degree, non-co-ordinated.

The induction of glycolytic enzymes by growth on glucose is discussed in the Introduction (1.3). The results of Maitra and Lobo (1971) show that the specific activities of all glycolytic enzymes increase upon addition of glucose or galactose to the culture broth. However, a non-co-ordinate relationship is evident when the early kinetics of induction is examined. Many enzymes, including phosphoglycerate kinase and enolase, are induced with a greater delay than pyruvate kinase. In fact, pyruvate kinase is induced at a rate higher than any other enzyme. A perfectly logical explanation of this results is offered (Maitra and Lobo, 1971): Enzymes are induced by different metabolites, therefore, the relative delay between the induction of enzymes will be a function of the rate at which the inducers are produced from a common precursor. However, as suggested by Maitra and Lobo, some other rate limiting step could be responsible. This rate limiting step may be the induced synthesis of PYK protein. By applying the trans feedback model to these results, one might suggest that the synthesis of the PYK enzyme is a pre-requisite for the induction of a subset of glycolytic enzymes. Experiments outlined in 6.4

illustrate that this is a possibility. Expression of PYK1 coding region, may elevate the abundance of not just PGK1 mRNA, but also HEX and PFK1 mRNA's. The elevation in the abundances of HEX and PFK1 is small and has only been measured in single transformants. Therefore, this should be considered preliminary data. The significance of trans-activation of these two particular enzymes is not evident. over-production of HEX or PFK does not, in itself, lead to an increase in glycolytic flux (Schaaff et al., 1989) but it would be interesting to determine the affect of highly over-expressing the three "trans-activated" genes, PGK1, HEX and PFK1.

In summary, despite clear experimental indications in this study that expression of PYK1 influences the expression of other genes, there is almost no physiological experimental data to support this conclusion. Therefore, it is unlikely that the "trans feedback" model completely describes the regulatory properties of the PYK1 coding region.

As yet none of the models constructed to describe the activation of transcription by the PYK1 coding region are wholly consistent with all observations and published data. To progress towards a more accurate model a number of experimental approaches should be undertaken. These will be discussed later.

7.2 Carbon Source Regulation of PYK1

The observations made on the carbon source regulation of the PYK1 gene shall now be considered. Maitra and Lobo (1971) clearly show an approximate 20-fold induction of PYK specific activity by glucose. However, as outlined in 1.3, this induction may be strain or culture condition dependent. Genetic evidence for specific regulation of glycolytic gene expression, has led to the general concept that induction is the consequence of an increase

in the rate of transcription of glycolytic genes. Clearly, the transcription of some glycolytic genes is induced by glucose (ENO2; Cohen et al., 1987, see 1.3). However, results presented in Chapter 4 (4.3.3) are inconsistent with transcriptional induction greater than 2-fold of PYK1. These results conflict markedly with the conclusions of, but not the data presented by, Nishizawa et al. (1989). As described in the Introduction (1.3), Nishizawa et al. have carried out deletion analysis of the PYK1 promoter regions fused to a reporter catalase coding region and showed evidence for the existence of two UAS's and an URS (Upstream Repression Sequence). They investigated whether sequences within the 5'-region mediate induction of PYK by glucose. Catalase specific activity was clearly shown to be 3-fold higher when cells were grown on glucose as opposed to a non-fermentable carbon source. The authors suggested from the protein level data that "transcription" from the PYK1 promoter was induced on the fermentable carbon source. mRNA abundance measurements were not presented to substantiate this suggestion. The results presented in 4.3.3 of this study are consistent with constitutively high PYK1 mRNA abundance with respect to carbon source.

Cohen et al., (1987) have previously shown that at least one glycolytic enzyme gene is constitutively expressed. This gene is ENO1. When cells are grown on a non-fermentable carbon source, expression of the ENO1 is quantitatively similar to that of the ENO2 gene. ENO2 is induced 20-fold by growth on glucose, whereas ENO1 expression is not induced. Crucial to the constitutive expression of the ENO1 gene, is a negative regulatory element situated between -143 and -181 bp from the I-site. Deletion of this URS region permitted glucose-dependent induction of ENO1 which was quantitatively similar to that of the ENO2 gene. The presence of this URS confers constitutive expression of this gene.

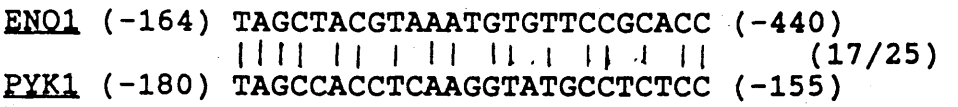


Figure 7.4 Sequence Similarity Between Two Glycolytic Upstream Repression Sites (URS)
GCG BESTFIT search facility was used to compare sequence from the URS of the constitutively expressed ENO1 gene (Cohen et al., 1987) with the PYK1 sequence. The similarity shown above occurs between the ENO1 URS and a PYK1 sequence previously shown to repress expression (Nishizawa et al., 1989).

The sequence of the ENO1 URS was compared, using BESTFIT, to the PYK1 5'-flanking regions. This comparison reveals a region of similarity between the URS and the PYK1 sequences (17/25 identity Figure 7.4). Interestingly, this similarity occurs within a region of the PYK1 gene previously shown to be a URS, repressing expression on both glucose and non-fermentable carbon sources (Nishizawa et al., 1989). Unfortunately, the exact consequences for transcription of deleting this sequence from the PYK1 promoter cannot be determined from the results of Nishizawa et al., as only protein levels were determined. The occurrence of homology between a URS, known to confer constitutive transcription on a glycolytic enzyme gene, and the PYK1 URS may be considered circumstantial evidence which supports the conclusion that PYK1 is constitutively transcribed.

If indeed the PYK1 gene is constitutively transcribed, then the observed induction of PYK by glucose is likely to be due to a translational control mechanism. Experiments carried out in this laboratory, discussed above (7.1), indicate that translation of the PYK1 mRNA may be subject to control by negative feedback by the PYK protein or by a PYK1 specific translation factor. These results may also be considered as circumstantial evidence that the induction of PYK by glucose occurs, not by increased transcription but by increased translation. However, it should be stressed that the supportive experimental evidence, presented in this study, is limited. For example, the experiments described in Chapter 4 were carried out to examine transcriptional regulation properties of the PYK1 coding region, therefore, no specific enzyme activities were measured. Ideally, experiments should be carried out in which measurements are made of both PYK1 mRNA abundances and PYK specific activities, both before and after induction by glucose.

7.3 Proposals for Further Study

Proposals for future research include experiments to confirm previous observations and experiments which further test specific models.

I Confirmation of Observations

(i) Although the techniques employed in this study are considered suitable for the measurement of mRNA abundance and stability, confirmation of many of the observations by different techniques would be desirable. For example, transcription rate has been inferred throughout this study by assessing mRNA abundance and mRNA stability; direct measurement of transcription rate by incorporation kinetics, would justify this inference.

(ii) The trans-activation of PGK1 should be further confirmed by measuring the abundance of PGK1 mRNA in a number of transformants, each expressing PYK1 at different levels. Convincing visual evidence of trans-activation could be obtained by probing northern blots for both PYK1 and PGK1 mRNAs. In addition the abundance of PYK1 mRNA derived from the chromosomally located gene should be determined in cells over-expressing fragments of the PYK1 coding region. Oligonucleotide probes specific for the chromosomal derived PYK1 mRNA should be utilised.

(iii) The degree to which PYK1 expression influences the expression of other glycolytic enzyme genes should be assessed by repeating measurements on the abundances of mRNAs in cells over-expressing PYK1.

(iv) To confirm that trans-activation of PGK1 by PYK1 is independent of heat shock response it should be demonstrated that the Heat Shock Element, within the PGK1 promoter, does not mediate this process. This could be

done by assessing the trans-activation of PGK1 promoter deletion mutants.
(v) Confirmation of the constitutive expression of PYK1 on different carbon sources should be carried out by measuring PYK1 mRNA abundances and PYK specific activity before and after glucose induction.

II Testing Models and Further Defining Regulatory Features of PYK1

(i) The minimum sequences required for transcriptional activation by the PYK1 coding region should be determined by further deletion analysis. To this end the 360 bp PYK1 activator fragment (fragment X) has been cloned into pSP64 to facilitate uni-directional deletions.
(ii) The position dependence of the PYK1 coding region activator should be assessed. These experiments would directly challenge the validity of the "binding site" and "trans feedback" models. The function of the activator should be assessed when it is placed at different sites, relative to the I-site and the translated region, including sites in trans to the reporter transcription unit.
(iii) The stability of PGK1 mRNA should be measured in untransformed cells and transformants over-expressing PYK1. This would reveal whether the increase in PGK1 mRNA in cells over-expressing PYK1 is due to an increase in transcription or a stabilisation of the PGK1 mRNA.
(iv) A search should be made for sequences located within PGK1 which are required for the trans-activation by PYK1.
(v) A search should be initiated for sequences within the PYK1 5'-flanking regions which are required for activation of transcription by the PYK1 coding region.
(vi) It is unknown whether the trans-activation of the PGK1 gene causes an induction of PGK1 protein levels.

Therefore, protein extracts of cell over-expressing PYK1 should be examined for increased PGK specific activity and increased PGK protein by SDS-PAGE.

(vii) The putative translational response to glucose induction should be studied by examining the ribosome loading of PYK1 mRNA before and after addition of glucose to a culture.

(viii) In the longer term, if it was shown that PYK indirectly induces the expression of other glycolytic genes then mutants defective in this regulatory circuit should be isolated. However, this may prove a difficult task, as a screen for mutants would be dependent upon differentiating between five to ten-fold expression levels. Perhaps a plate assay based upon B-galactosidase levels could be utilised for this purpose.

Abel, T. and Maniatis, T. (1989).
Action of leucine zippers - review
Nature 341: 24-25.

Allison, L.A. and Ingles, C.J. (1988).
Mutations in an RNA polymerase II polypeptide that enhance or suppress mutations in GAL4
Yeast Special issue 4: 483.

Allison, L.A., Moyle, M., Shales, M. and Ingles, C.J. (1985).
Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases
Cell 42: 599-610.

Almer, A., Rudolph, H., Hinnen, A. and Horz, W. (1986).
Removal of poitioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA elements
EMBO J. 5: 2689-2696.

Anton, I.A., Duncan, K. and Coggins, J.R. (1987).
A eukaryotic repressor protein, the qa-IS gene product of Neurospora crassa, is homologous to part of the arom multifunctional enzyme
J. Mol. Biol. 197: 367-371.

Baker, H.V. (1986).
Glycolytic gene expression in Saccharomyces cerevisiae: nucleotide sequence of GCR1, null mutants, and evidence for expression
Mol. Cell. Biol. 6: 3774-3784.

Beach, D.H. (1983).
Cell Type switching by DNA transposition in fission yeast
Nature 305: 682-688.

Bender, A. and Sprague, G.F. (1987).
MAT 1 protein, a yeast transcription activator binds synergistically with a second protein to a set of cell-type specific gene
Cell 50: 681-691.

Birnboim, H.C. and Doly, J. (1979).
A rapid alkaline extraction procedure for screening recombinant plasmid DNA
Nucleic Acids Res. 7: 1513-1523.

Bram, R.J. and Kornberg, R.D. (1985).
Specific proteins binding to far upstream activating sequences in polymerase II promoters
Proc. Natl. Acad. Sci. USA 82: 43-47.

Brand, A.H., Micklem, G. and Nasmyth, K. (1987).
A yeast silencer contains sequences that can promote autonomous replication and transcriptional activation

Cell 51: 709-719.

Brent, R. (1985).
Repression of Transcription in Yeast.
Cell 42: 3-4.

Brent, R. and Ptashne, M. (1985).
A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor
Cell 43: 729-736.

Brock, M.L. and Shapiro, D.J. (1983).
Oestrogen stabilises vitellogenin mRNA against cytoplasmic degradation
Cell 34: 207-214.

Brown, A.J.P. (1989).
Messenger RNA stability in yeast
Yeast 5: 239-257.

Buratowski, S., Hahn, S., Sharp, P.A. and Guarente, L. (1988).
Function of a yeast TATA element-binding protein
Nature 334: 37-36.

Burtis, K.C. and Baker, B.S. (1989).
Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides
Cell 56: 997-1010.

Casadaban, M.J., Martinez-Arias, A., Shapira, S.K. and Chou, J. (1983).
B-galactosidase gene fusions for analysing gene expression in E. coli and yeast
Methods in enzymology. Ed. Wu R Grossman L Moldave K 100: 293-308.

Cavallini, B., Huet, J., Plassat, J-L., Sentenac, A., Egly, J-M, and Chambon, P. (1988).
A yeast activity can substitute for the HeLa cell TATA box factor
Cell 334: 77-80.

Chambers, A., Stanway, C., Sang, J., Uemura, H., Kingsman, A.J. and Kingsman, S.M. (1989).
RAP1 is a component of a common transcriptional control circuit of yeast glycolytic genes
J. Cell. Biochem. supplement 13E: N305.

Chen, C.Y. and Hitzeman, R.A. (1987).
Human, yeast and hybrid 3-phosphoglycerate kinase gene expression in yeast
Nucleic Acids Res. 15: 643-660.

Chen, C.Y., Opperman, O. and Hitzeman, R.A. (1984).
Homologous versus heterologous gene expression in the yeast, Saccharomyces cerevisiae
Nucleic 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.

Chen, W. and Struhl, K. (1988).
Saturation mutagenisis of a yeast HIS3 "TATA element" : Genetic evidence for a specific TATA-binding protein.
Proc. Natl. Acad. Sci. USA 85: 2691-2695.

Chen, w., Tabor, S. and Struhl, K. (1987).
Distinguishing between mechanisms of eukaryotic transcriptional activation with bacteriophage T7 RNA polymerase
Cell 50: 1047-1055.

Cherry, J.R., Johnston, T.R., Dollard, C., Shuster, J.R. and Denis, C.L. (1989).
Cyclic AMP-dependant protein kinase phosphorylates and inactivates the yeast transcriptional activator ADR1
Cell 56: 409-419.

Cigan, A.M. and Donahue, T.F. (1987).
Sequence and structural features associated with translational initiation regions in yeast-a review
Gene 59: 1-18.

Ciriacy, D. and Breitenbach, I. (1979).
Physiological effects of seven different blocks in glycolysis in Saccharomyces cerevisiae
J. Bact. 139: 152-160.

Clark, L. and Carbon, J. (1980).
Isolation of a yeast centromere and construction of functional small circular chromosomes
Nature 287: 504-509.

Clark, L. and Carbon, J. (1985).
The structure and function of yeast centromeres
Annu. Rev. Genetics. 19: 29-57.

Clifton, D. and Fraenkel, D.G. (1981).
The gcr (glycolysis regulation) mutation of Saccharomyces cerevisiae
J. Biol. Chem. 256: 13074-13080.

Clifton, D., Weinstock, S.B. and Fraenkel, D.G. (1978).
Glycolysis mutants in Saccharomyces cerevisiae
Genetics 88: 1-11.

Cohen, R., Holland, J.P., Yokoi, T. and Holland, M.J. (1986).
Identification of a regulatory region that mediates glucose-dependant induction of the Saccharomyces cerevisiae enolase gene
Mol. Cell. Biol. 6: 2287-2297.

Cohen, R., Yokoi, T., Holland, J.P., Pepper, A.E. and Holland, M.J. (1987).
Transcription of the constitutively expressed yeast enolase gene ENO1 is meadiated by positive and negative cis-acting regulatory sequences
Mol. Cell. Biol. 7: 2753-2761.

Company, M. and Errede, B. (1987).
Cell-type-dependant activation by yeast transposon Ty1 involves multiple regulatory determinants
Mol. Cell. Biol. 7: 3205-3211.

Constanzo, M.C. and Fox, T.D. (1986).
Product of Saccharomyces cerevisiae nuclear gene required for translation of a specific mitochondrial mRNA
Mol. Cell. Biol. 6: 3694-3703.

Corden, J.L., Cadena, D.L., Ahearn, J.M. and Dahmus, M.E. (1985).
A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II
Proc. Natl. Acad. Sci. USA 82: 7934-7938.

Corton, J.C. and Johnston, S.A. (1989).
Altering DNA-binding specificity of GAL4 requires sequences adjacent to the zinc finger
Nature 340: 724-726.

Derynck, R., Singh, A. and Goeddel, D.V. (1983).
Expression of the human-interferon gamma cDNA in yeast
Nucleic Acids Res. 11: 1819-1837.

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
Nucleic Acids Res. 10: 2625-2637.

Donahue, T.F. and Cigan, A.M. (1988).
Genetic selection for mutations that reduce or abolish ribosomal recognition of the HIS4 translational initiator region
Mol. Cell. Biol. 8: 2955-2963.

Dynan, W.S. (1986).
Promoters for housekeeping genes - review
Trends in Genetics August: 196-197.

Elion, E.A. and Warner, J.R. (1986).
An RNA polymerase I enhancer in Saccharomyces cerevisiae
Mol. Cell. Biol. 6: 2089-2097.

Engelke, D.R., Ng, S-Y., Shastry, B.S. and Roeder, R.G. (1980).
Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes
Cell 19: 717-728.

Fairall, L., Rhodes, D. and Klug, A. (1986).
Mapping of sites of protection on a 5S RNA gene by the Xenopus transcriptional factor IIIA
J. Mol. Biol. 192: 577-591.

Farrelly, F.W., Davis, A.H. and Finkelstein, D.B. (1982).
The yeast HSP90 gene family
J Cell Biol 97: abstract 496.

Feagin, J.E., Jasmer, D.P. and Stuart, K. (1987).
Developmentally regulated addition of nucleotides within the apocytochrome b transcripts in Trypanosoma brucei
Cell 49: 337-345.

Ficca, A.G. and Hollenberg, C.P. (1989).
Functional relationship among TATA sequences. Gene induction and transcription initiation in the beta-galactosidase, Lac4, gene from K.lactis.
Current Genetics 15(4): 261-269.

Finkelstein, D.B., Strausberg, S. and McAlister, L. (1982).
Alterations of transcription during heat shock of Saccharomyces cerevisiae
J. Biol. Chem. 257: 8405-8411.

Fisher, J.A., Giniger, E., Maniatis, T. and Ptashne, M. (1988).
GAL4 activates transcription in Drosophila
Nature 332: 853-856.

Folk, W.R. (1988).
Changing direction in Pol III transcription
Genes and Development 2: 373-375.

Forsburg, S.L. and Guarente, L. (1989).
Identification and characterization of HAP4: a third componant of the CCAAT-bound HAP2/HAP3 heterodimer.
Genes & Development 3: 1166-1178.

Fraenkel, D.G. (1982).
Carbohydrate metabolism
The molecular biology of the yeast Saccharomyces: Metabolism and gene expression (Eds; Strathern, J.N., Jones, E.W. and Broach, J.R.)
CSH: 1-37.

Fulton, A.M., Rathjen, P.D., Kingsman, S.M. and Kingsman, A.J. (1988).
Upstream and downstream transcriptional control signals in the yeast retrotransposon, Ty

Nucleic Acids Res. 16: 5439-5458.

Futcher, A.B. and Cox, B.S. (1984).
Copy number and the stability of 2-um circle based artificial plasmids of Saccharomyces cerevisiae
J. Bacteriology 157: 283-290.

Gallwitz, D., Halfter, H. and Mertins, P. (1987).
Splicing of mRNA precursors in yeast
SGM special publication. Gene structure in eukaryotic microbes IRL press, Ed. Kinghorn, J. volume 22: 27-40.

Geiduschek, E.P. and Tocchini-Valentini, G.P. (1988).
Transcription by RNA polymerase III
Ann. Rev. Biochem 57: 873-914.

Gill, G. and Ptashne, M. (1987).
Mutants of GAL4 protein altered in an activation function
Cell 51: 121-126.

Giniger, E. and Ptashne, M. (1987).
Transcription in yeast activated by a putative amphipathic alpha-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.

Guarente, L. and Hoar, E. (1984a).
Upstream sites of the CYC1 gene are active when inverted but not when placed downstream of the TATA box
Proc. Natl. Acad. Sci. USA 81: 7860-7864.

Guarente, L. and Mason, T. (1983).
Heme regulates transcription of the CYC1 gene of Saccharomyces cerevisiae via an upstream activation site
Cell 32: 1279-1286.

Guarente, L., Yocum, R. and Gifford, P. (1982).
A GAL10-CYC1 hybrid promoter identifies the GAL4 regulatory region as an upstream site
Proc. Natl. Acad. Sci. USA 79: 7410-7414.

Guarente, L., Lalonde, B., Gifford, P. and Alani, E. (1984b).
Distinctly regulated tandem upstream activation sites mediate catabolite repression of the CYC1 gene of Saccharomyces cerevisiae
Cell 36: 503-511.

Gurr, S.J., Unkles, S.E. and Kinghorn, J.R. (1987).
The structure and organisation of nuclear genes of filamentous fungi
SGM special publication. Gene structure of eukaryotic

microbes. Ed. Kinghorn J.R. vol 22: 93-139.

Hahn, S. and Guarente, L. (1988).
Yeast HAP2 and HAP3: Transcriptional activators in a heteromeric complex
Science 240: 317-321.

Hahn, S., Hoar, E.T. and Guarente, L. (1985).
Each of three "TATA elememts" specifies a subset of the transcription initiation sites at the CYC1 promoter of Saccharomyces cerevisiae
Proc. Natl. Acad. Sci. USA 82: 8562-8566.

Hall, M.N. and Johnson, A.D. (1987).
Homeo domain of the yeast repressor 2 is a sequence-specific DNA-binding domain but is not sufficient for repression
Science 237: 1007-1012.

Hartshorne, T.A., Blumberg, H. and Young, E.T. (1986).
Sequence homology of the yeast regulatory protein ADR1 with xenopus transcription factor TFIIIA
Nature 321: 283-287.

Hawley, D.K. and McClure, W.R. (1983).
Compilation and analysis of E. coli promoter DNA sequences
Nucleic Acids Res. 11: 2237-2255.

Hereford, L., Bromley, S. and Osley, M.A. (1982).
Period transcription of yeast histone genes
Cell 30: 305-310.

Hereford, L. and Rosbash, M. (1977).
Number and distribution of polyadenylated RNA sequences in yeast
Cell 10: 453-462.

Hereford, L., Osley, M.A., Lugwig, J.R. and McLaughlin, C.S. (1981).
Cell-cycle regulation of yeast histone mRNA
Cell 24: 367-375.

Herruer, M.H., Mager, W.H., Raue, H.A., Vreken, P., Wilms, E. and Planta, R.J. (1988).
Mild temperature shock affects transcription of yeast ribosomal protein genes as well as the stability of their mRNA's
Nucleic Acids Res. 16: 7919-7929.

Hinnebusch, A.G. (1985).
A hierarchy of trans-acting factors modulates translation of an activator of amino acid biosynthetic genes in Saccharomyces cerevisiae
Mol. Cell. Biol. 5: 2349-6360.

Hinnebusch, A.G. (1988).
Mechanisms of gene regulation in the general control of amino acid biosynthesis of Saccharomyces cerevisiae.
Microbial Reviews 52: 248-273.

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
Nucleic Acids Res. 10: 7791-7808.

Hitzeman, R.A., Hagie, F.E., Levine, H.L., Goeddel, d. v., Ammerer, G and Hall, B.D. (1981).
Expression of a human gene for interferon in yeast
Nature 293: 717-722.

Hofbauer, R., Mullner, E., Seiser, C. and Wintersberger, E. (1987).
Cell cycle regulated synthesis of stable mouse thymidine kinase mRNA is mediated by a sequence within the cDNA
Nucleic Acids Res. 15: 741-752.

Holland, J.P., Brindle, P., Willett, C.E. and Holland, M.J. (1989).
Coordinate regulation of glycolytic gene expression
J. Cell. Biol. Supplement 13E UCLA symposia on Molecular and Cellular Biology : N015.

Holland, M.J., Yokoi, T., Holland, J.P., Ambo, K.M. and Innis, M.A. (1987).
The GCR1 gene encodes a positive transcriptional regulator of the enolase and glyceraldehyde-3-phosphate dehydrogenase gene families in Saccharomyces cerevisiae
Mol. Cell. Biol. 7: 813-820.

Holmes, D.S. and Quigly, M. (1981).
A rapid boiling method for the preparation of bacterial plasmids
Analytical Biochem 114: 193-197.

Hope, I.A. and Struhl, K. (1986).
Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast
Cell 46: 885-894.

Horikoshi, M., Carey, M.F., Kakidani, H. and Roeder, R.G. (1988).
Mechanism of action of a yeast activator: direct effect of GAL4 derivatives on mammalian TFIID promoter interactions
Cell 54: 665-669.

Horikoshi, M., Wan, G.C.K., Fujii, H., Cromlish, J.A., Weil, P.A. and Roeder, R.G. (1989).
Purification of a yeast TATA box-binding protein that exhibits human transcription factor-IID activity.
Proc. Natl. Acad. Sci. USA 86(13): 4843-4847.

Hope, I.A., Mahadevan, S. and Struhl, K. (1988).
Structural and functional characterisation of the short acidic transcriptional activation region of yeast GCN4 protein
Nature 333: 635-640.

Huet, J. and Sentenac, A. (1987).
TUF, the yeast binding factor specific for UASrpg upstream activating sequences: identification of the protein and its DNA binding domain
PNAS 84: 3648-3652.

Huet, J., Cottrelle, P., Cool, M., Vignais, M.-.L., Thiele, D., Marck, C., Buhler, J.-.M., Sentanac, A. and Fromageot, P. (1985).
A general upstream binding factor for the genes of the yeast translational apparatus.
EMBO J. 4: 3539-3547.

Huiet, L. and Giles, N.H. (1986).
The qa repressor gene of Neurospora crassa: wild type and mutant nucleotide sequence
Proc. Natl. Acad. Sci. USA 83: 3381-3385.

Irani, M., Tayor, W.E. and Young, E.T. (1987).
Transcription of the ADH2 gene in Saccharomyces cerevisiae is limited by positive factors that bind competitively to its intact promoter region on multicopy plasmids
Mol. Cell. Biol. 7: 1233-1241.

Jinno, Y., Merlino, G.T. and Pastan, I. (1988).
A novel effect of EGF on mRNA stability
Nucleic Acids Res. 16: 4957-4966.

Johnston, M. (1987a).
A model fungal gene regulatory mechanism : the GAL genes of Saccharomyces cerevisiae.
Microbiological Reviews 51(4): 458-476.

Johnston, M. (1987b).
Genetic evidence that zinc is an essentail co-factor in the DNA binding domain of GAL4 protein
Nature 328: 352-355.

Johnston, M. and Dover, J. (1987).
Mutations that inactivate a yeast transcriptional regulatory protein cluster in an evolutionarily conserved DNA binding domain
Proc. Natl. Acad. Sci. USA 84: 2401-2405.

Johnston, P.F. and McKnight, S.L. (1989).
Eukaryotic transcriptional regulatory proteins
Ann. Rev. Biochem. 58: 799-830.

Johnston, S.A., Salmeron, J.M., Jr, and Dincher, S.S. (1989).
Interaction of positive and negative regulatory proteins in the galactose regulon of yeast.

Cell 50: 143-146.

Kaback, D.B., Angerer, L.M. and Davidson, N. (1979).
Improved methods for the formation and stabilisation of R loops
Nucleic Acids Res. 6: 2499-2517.

Kakidani, H. and Ptashne, M. (1987).
GAL4 activates gene expression in mammalian cells
Cell 52: 161-167.

Keegan, L., Gill, G. and Ptashne, M. (1986).
Separation of DNA binding from the transcriptional-activating function of a eukaryotic regulatory protein
Science 231: 699-704.

Keleher, C.A., Goutte, C. and Johnston, A.D. (1988).
The yeast cell-type-specific repressor alpha2 acts cooperatively with a non-cell-type-specific protein
Cell 53: 927-936.

Kim, S., Mellor, J., Kingsman, A.J. and Kingsman, S.M. (1986).
Multiple control elements in the TRP1 promoter of Saccharomyces cerevisiae
Mol. Cell. Biol. 6: 4251-4258.

Klekamp, M.S. and Weil, P.A. (1982).
Specific transcription of homologous class III genes in yeast soluble cell-extracts
J Biol Chem 257: 8432-8441.

Koltunow, A.M., Gregg, K. and Rodgers, G.E. (1987).
Promoter efficiency depends upon intragenic sequences
Nucleic Acids Res. 15: 7795-7807.

Kosche, K.A., Dobkin, C. and Bank, A. (1985).
DNA sequences regulating human beta-globin gene expression
Nucleic Acids Res. 13: 7781-7793.

Kouzarides, T. and Ziff, E. (1989).
Leucine zippers of FOS, JUN and GCN4 dictate dimerization specificity and thereby control DNA binding.
Nature 340: 568-571.

Kuwabara, M.D. and Sigman, D.S. (1987).
Footprinting DNA-protein complexes in situ following gel retardation assays using 1,10-phenanthroline-copper ion: E. coli RNA polymerase-lac promoter complexes
Biochemistry 26: 7234-7238.

Landschulz, W.H., Johnston, P.F. and McKnight, (1988).
The leucine zipper: A hypothetical structure common to a new class of DNA binding protein
Science 240: 1759-1764.

Laughton, A. and Scott, M.P. (1984).
Sequences of a Drosophila segmentation gene: protein structure homology with DNA-binding proteins
Nature 310: 25-31.

Lech, K., Anderson, K. and Brent, R. (1988).
DNA bound Fos proteins activate transcription in yeast
Cell 52: 179-184.

Linquist, S. (1981).
Regulation of protein synthesis during heat shock
Nature 293: 311-314.

Lohr, D. and Hopper, J.E. (1985).
The relationship of regulatory proteins and DNaseI hypersensitivity sites in the yeast GAL1-10 genes
Nucleic Acids Res. 13: 8409-8423.

Lue, N.F., Chasman, D.I., Buchman, A.R. and Kornberg, R.D. (1987).
Interaction of GAL4 and GAL80 gene regulatory proteins in vitro
Mol. Cell. Biol. 7: 4446-3451.

Lycan, D.E., Osley, M.A. and Hereford, L.M. (1987).
Role of transcriptional and post-transcriptional regulation in expression of histone genes in Saccharomyces cerevisiae
Mol. Cell. Biol. 7: 614-621.

Maitra, P.K. and Lobo, Z. (1971).
A kinetic study of glycolytic enzyme synthesis in yeast
J. Biol. Chem. 246: 475-488.

Ma, J. and Ptashne, M. (1987a).
Deletion analysis of GAL4 defines two transcriptional activating segments
Cell 48: 847-853.

Ma, J. and Ptashne, M. (1987b).
A new class of yeast transcriptional activators
Cell 51: 113-119.

Ma, J. and Ptashne, M. (1987c).
The carboxy-terminal 30 amino acids of GAL4 are recognised by GAL80
Cell 50: 137-142.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982).
Molecular Cloning; A Laboratory Manual
Published by Cold Spring Harbour Laboratories : .

Marie, J., Simon, M.-.P., Dreyfus, J.-.C. and Kahn, A. (1981).
One gene, but two messenger RNA's encode liver L and red cellL' pyruvate kinase
Nature 292: 70-72.

Marykwas, D.L. and Fox, T.D. (1989).
Control of the Saccharomyces cerevisiae regulatory gene PET494: Translational repression by glucose and translational induction by oxygen
Mol. Cell. Biol. 9: 484-491.

McAlister, L. and Holland, M.J. (1982).
Targeted deletion of a yeast enolase structural gene
J. Biol. Chem. 257: 7181-7188.

McNeil, J.B. (1988).
Functional charaaterisation of a pyrimidine-rich element in the 5'-noncoding region of the yeast iso-1-cytochrome c gene
Mol. Cell. Biol. 8: 1045-1054.

McNeil, J. and Smith, M. (1986).
Transcription initiation of the Saccharomyces cerevisiae iso-1-cytochrome c gene: Multiple, independant TATA sequences
J. Mol. Biol. 187: 363-378.

Mellor, J. (1989).
The activation and initiation of transcription by the promoters of S.cerevisiae
In Molecular & Cell Biology of Yeasts, Walton, E.F. and Yarranton, G.T. Glasgow & London / New York: Blackie / Van Nostrand Reinhold.

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
Nucleic 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 yeast
Gene 33: 215-226.

Mellor, J., Dobson, M.J., Roberts, N.A., Tuite, M.F., Emtage, J.S., White, S., Lowe, P.A., Patel, T., Kingsman, A.J. and Kingsman, S.M. (1983).
Efficient synthesis of enzymatically active calf chymosin in Saccharomyces cerevisiae
Gene 24: 1-14.

Mellor, J., Dowell, S.J. and Rathjen, J. (1988).
mRNA start site determination in the TRP1 promoter requires an interaction between CP1 and a short repeated sequence
Yeast Special Issue 4: 411.

Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984).
Efficient in vitro synthesis of biologically active RNA and RNA hybridisation probes from plasmids containing a bacteriophage

SP6 promoter
Nucleic Acids Res. 12: 7035-7056.

Merrill, G.F., Hauschka, S.D. and McKnight, S.L. (1984).
tk enzyme expression in differentiating muscle cells is regulated through an internal segment of the cellular tk gene
Mol. Cell. Biol. 4: 1777-1784.

Messenguy, F., Feller, A., Crabeel, M. and Pierard, A. (1983).
Control mechanisms acting at the transcriptional and post-transcriptional levels are involved in the synthesis of the arginine pathway carbamoylphosphate synthase of yeast.
EMBO J. 2: 1249-1254.

Metzger, D., White, J.H. and Chambon, P. (1988).
The human oestrogen receptor functions in yeast
Nature 334: 31-36.

Miller, J.H., MacKay, V.L. and Nasmyth, K.A. (1985).
Identification and comparison of two sequence elements that confer cell type transcription in yeast
Nature 314: 598-601.

Moreland, R.B., Nam, H.G., Hereford, L.M. and Fried, H.M. (1984).
Identification of a nuclear localization signal of a yeast ribosomal protein
Proc. Natl. Acad. Sci. USA 82: 6561-6565.

Moye-Rowley, W.S., Harshman, K.D. and Parker, C.S. (1989).
Yeast YAP1 encodes a novel form of the jun family of transcriptional activator proteins.
Genes and Development 3: 283-292.

Mylin, L.M., Bhat, J.P. and Hopper, J.E. (1989).
Regulated phosphorylation and dephosphorylation of GAL4, a transcriptional factor.
Genes & Development 3: 1157-1165.

Nagashima, K., Kasai, M., Nagata, S. and Kaziro, Y. (1986).
Structure of two genes coding for chain elongation factor EF-alpha from Saccharomyces cerevisiae
Gene 45: 265-273 .

Nagawa, F. and Fink, G.R. (1985).
The relationship between the "TATA" sequence and the transcriptional initiation sites at the HIS4 gene of Saccharomyces cerevisiae
Proc. Natl. Acad. Sci. USA 82: 8557-8561.

Nagoshi, R.N., McKeown, M., Burtis, K.C., Belote, J.M. and Baker, B.S. (1988).
The control of alternative splicing at genes regulating sexual differentiation in D. melanogaster
cell 53: 229-236.

Newport, J.W. and Forbes, D.J. (1987).
The nucleus: structure, function, and dynamics
Annu. Rev. Biochem. 56: 535-565.

Nishizawa, M., Araki, R. and Teranishi, Y. (1989).
Identification of an upstream activating sequence and an upstream repressable sequence of the pyruvate kinase gene of the yeast Saccharomyces cerevisiae
Mol. Cell. Biol. 9: 442-451.

Nonet, M., Sweetser, D. and Young, R.A. (1987).
Functional redundancy and structural polymorphism in the largest subunit of RNA polymerse II
Cell 50: 909-915.

Ogden, J.E., Stanway, C., Kim, S., Mellor, J., Kingsman, A.J. and Kingsman, S.M. (1986).
Efficient expression of the Saccharomyces cerevisiae PGK1 gene depends on an upstream activation sequence but does not require TATA sequence
Mol. Cell. Biol. 6: 4335-4343.

Oleson, J., Hahn, S. and Guarente, L. (1987).
Yeast HAP2 and HAP3 activators both bind to the CYC1 upstream activation site, UAS2, in a interdependant manner
Cell 51: 953-961.

Pabo, C.O. and Sauer, R.T. (1984).
Protein-DNA recognition
Annu. Rev. Biochem. 53: 293-321.

Parker, C.S. and Topol, J. (1984).
A Drosophila RNA polymerase II transcription factor contains a promoter region-specific DNA binding activity
Cell 36: 357-369.

Pelham, H. (1988).
Heat Shock Proteins; Coming in from the cold
Nature 28: 776-777.

Pfeifer, K., Arcangioli, B. and Guarente, L. (1987).
Yeast HAP1 activation competes with a factor RC2 for binding to the upstream activation site UAS1 of the CYC1 gene
Cell 49: 9-18.

Piper, P.W., Curran, B., Davis, M.W., Lockheart, A. and Reid, G. (1986).
Transcription of the phosphoglycerate kinase gene of Saccharomyces cerevisiae increases when fermentative cultures are stressed
Eur. J. Biochem. 161: 525-531.

Piper, P.W., Curren, B., Davis, M.W., Hirst, K., Lockheart, A., Ogden, J.E., Stanway, C.A., Kingsman, A.J. and Kingsman, S.M. (1988).

A heat shock element in the phosphoglycerate kinase gene promoter of yeast
Nucleic Acids Res. 16: 1333-1348.

Poutre, C.G. and Fox, T.D. (1987).
PET3, a S.cerevisiae nuclear gene required for the translation of the mitochondrial mRNA encoding cytochrome c oxidase subunit II
Genetics 115: 637-647.

Ptashne, M. (1986).
Gene regulation by proteins acting nearby and at a distance
Nature 322: 697-701.

Purvis, I.J., Bettany, A.J.E., Loughlin, L., and Brown, A.J.P. (1987a).
The effects of alterations within the 3' untranslated region of the pyruvate kinase messenger RNA upon its stability and translation in Saccharomyces cerevisiae
Nucleic Acids Res. 15: 7951-7962.

Purvis, I.J., Loughlin, L., Bettany, J.E. and Brown, A.J.P. (1987b).
Translation and stability of an Escherichia coli B-galactosidase mRNA expressed under the control of pyruvate kinase sequences in Saccharomyces cerevisiae
Nucleic Acids Res. 15: 7963-7974.

Queen, C. and Baltimore, D. (1983).
Immumoglobulin gene transcription is activated by downstream sequence elements
Cell 33: 741-748.

Rathjen, P.D., Kingsman, A.J. and Kingsman, S.M. (1987).
The yeast ROAM mutation: identification of the sequences mediating host gene activation and cell-type control in the yeast retrotransposon Ty
Nucleic Acids Res. 15: 7309-7324.

Rhodes, D. and Klug, A. (1986).
An underlying repeat in some transcriptional control sequences corresponing to half a double helical turn of DNA
Cell 46: 123-132.

Rudolf, H. and Hinnen, A. (1987).
The yeast PHO5 promoter: phosphate-controlled element and sequences mediating mRNA start-site selection
Proc. Natl. Acad. Sci. USA 84: 1340-1344.

Santiago, T.C., Purvis, I.J., Bettany, A.J.E. and Brown, A.J.P. (1986).
The relationship between mRNA stability and length in Saccharomyces cerevisiae
Nucleic Acids Res. 14: 8347-8360.

Sasson-Corsi, P. and Bornelli, E. (1986).
Transcriptional regulation by trans-acting factors - review
Trends in Genetics August: 215-219.

Sawadogo, M. and Roeder, R.G. (1985a).
Interaction of a gene-specific transcription factor with the Adenovirus Major Late promoter upstream of the TATA box region
Cell 43: 165-175.

Sawadogo, M. and Roeder, R.G. (1985b).
An RNA polymerase II transcription factor binds to an upstream element in the Adenovirus major late promoter
Cell 43: 439-448.

Schaaff, I., Heinisch, J. and Zimmerman, F.K. (1989).
Overproduction of glycolytic enzymes in yeast
Yeast 5: 285-290.

Schmitz, A. and Galas, D. (1979).
The interaction of RNA polymerase and the lac repressor with the lac control region
Nucleic Acids Res. 6: 111-137.

Selleck, S.B. and Majors, J.M. (1987a).
In vivo DNA-binding properties of a yeast transcription activator protein
Mol. Cell. Biol. 7: 3260-3267.

Selleck, S.B. and Majors, J.M. (1987b).
Photo footprinting in vivo detects transcription-dependant changes in yeast TATA boxes
Nature 325: 173-177.

Sentenac, A. and Hall, B. (1982).
Yeast nuclear RNA polymerases and their role in transcription
The molecular biology of the yeast Saccharomyces cerevisiae. Metabolism and gene expression. Ed. Strathern, J.N. Jones, E.W. Broach, J.R. CSH : 561-606.

Shapiro, D.J., Blume, J.E. and Nielson, D.A. (1987).
Regulation of mRNA stability in eukaryotic cells
Bioessays 6: 221-226.

Shepherd, , J.C.W., , McGinnis, W., Carrasco, A.E., DeRobertis, E.M. and Gehring, W.J. (1984).
Fly and frog homoeo domains show homologies with yeast mating type regulatory proteins
Nature 310: 70-71.

Shore, D. and Nasmyth, K. (1987).
Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements
Cell 51: 721-732.

Sigman, D.S. (1986).
The nuclease activity of 1,10-phenanthroline-copper ion
Accounts of Chemical Res. 19: 10-186.

Siliciano, P.G. and Tatchell, K. (1984).
Transcription and regulatory signals at the mating type locus in yeast
Cell 37: 969-978.

Sorger, P.K. and Pelham, H.R.B. (1988).
Yeast heat shock factor is an essentail DNA-binding protein that exhibits temperature-dependant phosphorylation
Cell 54: 855-864.

Sosnowski, B.A., Belote, J.M. and McKeown, M. (1989).
Sex-specific alternative slicing of RNA from the transformer gene results from sequence-dependent splice site blockage
Cell 58: 449-459.

Steitz, T.A., Ohlendorf, D.H., McKay, D.B., Anderson, W.F. and Matthews, B.W. (1982).
Structural similarity in the DNA-binding domains of catabolite gene activator and cro repressor proteins
Proc. Natl. Acad. Sci. USA 79: 3097-3100.

Stillman, D.J., Bankier, A.T., Seddon, A., Goenhout, E.G. and Nasmyth, K.A. (1988).
Characterisation of a transcription factor involved in mother cell specific transcription of the yeast HO gene
EMBO J. 7: 485-494.

St John, T.P. and Davis, R.W. (1979).
Isolation of galactose sequences from Saccharomyces cerevisiae by differential plaque filter hybridisation
Cell 16: 443-452.

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.

Strathern, J.N., Klar, , AJS, , Hicks, J.B., Abraham, J.A., Ivy, J.M., Naysmith, K.A. and McGill, C. (1982).
Homothallic switching of yeast mating type cassettes is initiated by a double-stranded cut in the MAT locus
Cell 31: 183-192.

Struhl, K. (1982).
Regulatory sites for HIS3 expression in yeast
Nature 300: 284-287.

Struhl, K. (1984).
Genetic properties and chromatin structure of the yeast GAL regulatory region: an enhancer like sequence

Proc. Natl. Acad. Sci. USA 81: 7865-7869.

Struhl, K. (1985).
Negative control at a distance mediates catabolite repression in yeast
Nature 317: 822-824.

Struhl, K. and Davis, R.W. (1981).
Transcription of the HIS3 gene region in Saccharomyces cerevisiae
J. Mol. Biol. 152: 535-552.

Struhl, K. (1987).
Promoters, activator proteins, and the mechanism of transcriptional initiation in Yeast
Cell 49: 295-297.

Tekamp, P.A., Valenzuela, P., Maynard, T., Bell, G.I. and Rutter, W.J. (1979).
Specific gene transcription in yeast nuclei and chromatin by added homologous RNA polymerase I and III
J Biol Chem 254: 955-963.

Torchia, T.E.R., Hamilton, R.W., Cano, C.L. and Hopper, J.E. (1984).
Disruption of regulatory gene GAL80 in Saccharomyces cerevisiae: effects on carbon-controlled regulation of the galactose/melibiose pathway genes
Mol. Cell. Biol. 4: 1521-1527.

Tuite, M.F., Dobson, M.J., Roberts, N.A., King, R.M., Burke, D.C., Kingsman, S.M. and Kingsman, A.J. (1983).
Regulated high expression of human interferon-alpha in Saccharomyces cerevisiae
EMBO J. 1: 603-608.

Tullius, T.D. and Dombroski, B.A. (1986).
Hydroxyl radical "footprinting": High resolution information about DNA-protein contacts and application to lambda-repressor and cro proteins
Proc. Natl. Acad. Sci. USA 83: 5469-5473.

Verdiere, J., Gaisne, M., Guiard, B. and Defranoux, N. (1988).
A single mutation in the CYP1 (HAP1) regulatory gene switches the expression of two structural genes encoding isocytochromes c
Yeast special edition 4: 425.

Verdiere, J., Greusot, F., Guarente, L. and Slomimski, P. (1986).
The overproducing CYP1 and the underproducing hap1 mutations are alleles of the same gene which regulates in trans the expression of the structural genes encoding the iso-1-cytochrome c.
Curr. Genet. 10: 339-342.

Warner, J.R. (1989).
Synthesis of ribosomes in Saccharomyces cerevisiae
Microbiological reviews June: 256-271.

Warner, J.R., Mitra, W.F., Schwindinger, M., Studeny, M. and Fried, H.M. (1985).
Saccharomyces cerevisiae coordinates the accumulation of yeast ribosomal proteins by modulating mRNA splicing, translation initiation and protein turnover.
Mol. Cell. Biol. 5: 1512-1521.

Webster, N., Jin, J.R., Green, S., Hollis, M. and Chambon, P. (1988).
The yeast UASg is a transcriptional enhancer in human HELA cells in the presence of GAL4 trans-activator
Cell 52: 169-178.

West, R.W., Yocum, R.R. and Ptashne, M. (1984).
Saccharomyces cerevisiae GAL1-GAL10 divergent promoter region: location and function of the upstream activating sequence UASg
Mol. Cell Biol. 4: 2467-2478.

Yocum, R.R. and Johnston, M. (1984).
Molecular cloning of the GAL80 gene from Saccharomyces cerevisiae and characterisation of a gal80 deletion
Gene 32: 75-82.

Zehring, W.A., Lee, J.M., Weeks, J.R., Jokerst, R.S. and Greenleaf, A.L. (1989).
The C-terminal repeat domain of RNA polymerase II largest subunit is essential in vivo but is not required for accurate transcription initiation in vitro.
Proc. Natl. Acad. Sci. USA 85: 3698-3702.

Zinn, K.D., Dimiao, D. and Maniatis, T. (1983).
Identification of two distinct regulatory regions adjacent to the human B-interferon gene
Cell 34: 865-879.