Transcription Factor Interactions at the *PGK* Promoter in Yeast

Elizabeth A. Packham BSc

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy, February 1996

Statement

This thesis is based on work conducted by the author, in the Department of Genetics, University of Nottingham, during the period between October 1992 and July 1995.

All the work recorded in this thesis is original, unless otherwise acknowledged in the text or by references. None of this work has been submitted for any other degree at this or any other university. Results from Chapters 3, 4 and 5 are accepted for publication (Packham E. A., Graham I. G. and Chambers A. (1996) The multifunctional transcription factors Abf1p, Rap1p and Reb1p are required for full transcriptional activation of the chromosomal *PGK* gene in *Saccharomyces cerevisiae*. *Mol Gen Genet* in press).

Signed 

Date 9.2.96

Acknowledgements

I would like to thank the Medical Research Council for the financial support which allowed me to carry out this work.

I would also like to thank my supervisor, Dr. Alistair Chambers, for his advice and help during the past three years, and I hope that I have not caused him too many problems. Particular thanks go to Dr. Ian Graham, without whose friendship and support at every stage of this PhD the work would not have progressed as smoothly as it did, and to Louise Jenkins, Margaret Duffy and Sue Miles for making the lab a fun place to work. Thankyou also to Alistair, Ian and Louise for letting me discuss unpublished results.

Thanks to members of the Genetics Department, past and present, for a great three years. Especially to Jon for speedy techniques, Sue for everything, Alan for freely given sarcasm, Caroline and Nikki for all those lengths-Tuesdays will never be the same again, the cricket team for an incredible number of ducks, the volleyball team for many bruises, and Stuart for encouragement and not being asleep when it mattered.

Finally, thankyou to Mum and Dad for many years of support and encouragement through all the stages of my education.

Table of Contents

Abstract

			Page
Chapter 1	**Introduction**		**1**
1.1	RNA polymerase II transcription		1
	1.1.1	Yeast RNA polymerase II	1
1.2	Phosphoglycerate kinase: a model system for studying transcription		2
1.3	General transcription factors		5
	1.3.1	TFIID	5
	1.3.2	TFIIA	5
	1.3.3	TFIIB	6
	1.3.4	TFIIF	7
	1.3.5	TFIIE	7
	1.3.6	TFIIH	8
1.4	Binding of basal transcription factors to form an active initiation complex		8
1.5	Activated transcription		9
	1.5.1	Transcriptional activator proteins	10
	1.5.2	DNA binding domains	10
	1.5.3	Activation domains	11
	1.5.4	Interaction between upstream and downstream promoter elements	12
	1.5.5	Targets for transcriptional activators	12
	1.5.6	Coactivators	12
1.6	Activated transcription at *PGK*		13
1.7	Transcription factors which bind to the *PGK* promoter		16
	1.7.1	Rap1p	16
	1.7.2	Gcr1p	19
	1.7.3	Abf1p	20
	1.7.4	Reb1p	22
	1.7.5	Cpf1p	25
	1.7.6	yATF	27
1.8	Aims of thesis		28

2.25 Determination of plasmid copy number 41
2.26 Polymerase chain reaction (PCR) 41
2.26.1 Primers 41
2.26.2 Cloning of PCR products 43
2.27 DNA sequencing 43
2.27.1 Sequencing midiprep or caesium pure DNA 43
2.27.2 Sequencing of miniprep DNA 44
2.28 Pop-In/Pop-Out homologous recombination 44
2.29 recovery of plasmids from yeast 44

Chapter 3 Identification of New Transcription factor Binding Sites in the *PGK* Promoter 46
3.1 Introduction 46
3.2 Results 46
3.2.1 Does Abf1p bind upstream of the UAS ? 46
3.2.2 Does Rap1p bind upstream of the UAS ? 48
3.2.3 Identification of a new transcription factor binding site for Cpf1p 48
3.2.4 Identification of a new transcription factor binding site for Reb1p 49
3.3 Discussion 51

Chapter 4 Construction of Yeast Strains to Investigate the Roles of Transcription Factors at the Chromosomal *PGK* Locus 54
4.1 Introduction 54
4.1.1 The "Pop in-Pop out" method of homologous recombination 55
4.2 Results 56
4.2.1 Constructing chromosomal deletions of *PGK* transcription factor binding sites 56
4.2.2 Rap1p is central to the activation of transcription at *PGK* in the chromosome 60
4.2.3 The role of Cpf1p at the chromosomal *PGK* locus 61
4.2.4 Is the *PGK* promoter bidirectional ? 61
4.2.5 YCR11c probes 62
4.3 Discussion 62

Chapter 2 Materials and Methods **29**

2.1 Strains used in this work 29

2.2 Bacterial growth media 29

2.3 Yeast growth media 29

2.4 Growth conditions 30

2.5 Isolation of DNA from *E. coli* 30

2.5.1 Miniprep 30

2.5.2 Midiprep 30

2.5.3 Large scale preparation of DNA from a caesium chloride gradient 31

2.5.4 "Wizard" DNA preps 31

2.6 Isolation of DNA from *S. cerevisiae* 31

2.6.1 "Ten minute" chromosomal DNA prep 31

2.6.2 Copy number DNA prep 31

2.7 Digesting DNA 32

2.8 Ligations 33

2.9 Addition of phosphate groups to artificially synthesised oligonucleotides 33

2.10 Transformation of *E. coli* 33

2.11 One step transformation of *S. cerevisiae* 34

2.12 Lithium acetate transformation of *S. cerevisiae* 34

2.13 Preparation of total RNA from yeast 34

2.14 Preparing a total protein extract from yeast 34

2.15 Biorad protein assay 35

2.16 β-galactosidase assay 35

2.17 *In vitro* protein production 35

2.18 Partial purification of overexpressed Reb1p 36

2.19 Protein gel 37

2.20 Gel retardation assay 37

2.20.1 Preparation of retardation probes 37

2.20.2 Binding reactions 38

2.21 Preparation of probes for Southern and Northern blots 38

2.21.1 Nick translation 38

2.21.2 Random oligolabelling 38

2.22 Southern blots 39

2.23 Northern blots 39

2.24 Removal of hybridised probe from nitrocellulose filters 40

Chapter 5 The Effect of Transcription Factor Binding Site Deletions on Regulation of a Plasmid-Borne Copy of *PGK*, and the Use of a Minimal Promoter to Determine Roles for Transcription Factors **66**

5.1 Introduction 66

5.2 Results 67

5.2.1 Multicopy plasmid-borne deletions of the Reb1p and Abf1p sites from the PGK UAS 67

5.2.2 The effect of deleting the potential yATF site from the *PGK* promoter 68

5.2.3 A plasmid-borne copy of *PGK* in a *cpf1* background 69

5.2.4 Transcriptional activity of the redefined UAS 70

5.2.5 The *PGK* yATF sequence in a minimal promoter plasmid 71

5.2.6 The CRE consensus is no more active than the yATF sequence 72

5.2.7 An interaction between Rap1p and Gal11p 74

5.3 Discussion 74

Chapter 6 Construction of a Genomic Yeast Library **79**

6.1 Introduction 79

6.2 Results 80

6.2.1 Construction of pAJ800 80

6.2.2 Construction of yeast library 81

6.2.3 Transforming the library into yeast 82

6.3 Discussion 83

Chapter 7 General Discussion **86**

7.1 Multifunctional transcription factors at the *PGK* promoter 86

7.2 yATF does not have a role in the regulation of transcription at the *PGK* locus 89

7.3 A model for the regulation of transcription at the *PGK* promoter 92

7.4 Further work 93

References

Abstract

Two new transcription factor binding sites have been identified within the phosphoglycerate kinase (*PGK*) gene promoter in the yeast *Saccharomyces cerevisiae*. The binding sites are upstream of the previously defined UAS, and are bound *in vitro* by the multifunctional transcription factors Reb1p and Cpf1p. A deletion of the Reb1p binding site was made from a *PGK* gene construct on a multicopy plasmid, and also targeted to the chromosomal copy of *PGK*. Deletions of the Rap1p and Abf1p binding sites in the UAS were also targeted to the chromosome. Analysis of RNA from the chromosomal deletion strains confirmed the central role of Rap1p in the activation of transcription from *PGK*. Reb1p and Abf1p were also found to be important for transcriptional activation. This is in contrast to results from experiments using multicopy plasmids carrying Reb1p or Abf1p binding site deletions from *PGK*. In this situation, neither the Reb1p site, nor the Abf1p site, plays a role in transcriptional activation.

A role for Cpf1p at the *PGK* promoter was examined using a *cpf1* null strain of yeast. Northern blot analysis was used to assay transcription from the chromosomal *PGK* gene in the absence of Cpf1p, and also transcription from a multicopy plasmid carrying the wild type *PGK* gene in the *cpf1*$^-$ background. In both cases, the absence of Cpf1p was found to have very little effect on the level of transcription.

In addition, a role for the potential yATF binding site at the 3' end of the *PGK* UAS was investigated. Oligonucleotides containing this sequence were inserted upstream of a minimal promoter, and levels of a β-galactosidase reporter were assayed. No activation over the basal level was observed. A deletion of the potential yATF binding site from the UAS was made from a multicopy plasmid construct, and also from the chromosomal locus. Transcription from the deleted constructs was found to be no different from transcription from the wild type gene.

Finally, DNA sequences which are able to complement the C-terminus functions of Rap1p were identified. A yeast genomic library was generated downstream of the N-terminus and DNA binding domain of Rap1p. This library was transformed into a *rap1*^{ts} strain of yeast to look for complementation of the ts phenotype. Transformants which grew at the non-permissive temperature were obtained. Results from the analysis of the DNA sequences in these transformants are presented.

Chapter 1

Introduction

1.1 RNA Polymerase II Transcription

Transcription, catalysed by DNA-dependent RNA polymerase, is the process in which an RNA message is made from a gene. Bacteria have one RNA polymerase consisting of three core subunits with which additional polypeptides become associated, for recognising specific promoters and for gene regulation. In eukaryotic cells there are three forms of RNA polymerase which each transcribe a subclass of nuclear genes; RNA polymerase I transcribes rDNA genes encoding large ribosomal RNAs, RNA polymerase II transcribes the protein-coding messenger RNAs, and RNA polymerase III transcribes small RNAs such as tRNA and 5S RNA. In the case of class I and class III genes, the transcript, either rRNA or tRNA, is the final product, but transcription from the protein encoding class II genes produces an mRNA which is translated by the ribosomes to give a protein. As the majority of genes fall into class II, most studies have focused on RNA polymerase II (RNA pol II).

RNA pol II is a multisubunit enzyme which shares some subunits with RNA polymerases I and III. It consists of two large subunits, which form the structural and functional core of the enzyme, and several smaller ones (Sawadogo and Sentenac 1990). The two large subunits show considerable homology to the bacterial RNA polymerase subunits β and β' which indicates a conservation of the mechanism of transcription between prokaryotes and eukaryotes. As the structure and function of RNA pol II subunits has been well conserved between yeast and higher eukaryotes, *S. cerevisiae* is a good model for studying the enzyme (Young 1991).

1.1.1 Yeast RNA Polymerase II

Yeast RNA pol II consists of at least eleven subunits, all of which have been cloned and sequenced (see review Young 1991). The three largest subunits make up the core of the enzyme. The largest is 220 kDa (Young and Davis 1983), and shows considerable homology with the large subunit of RNA pol III and the large prokaryotic subunit, β' (Allison *et al.* 1985), except for the C-terminus which encodes a very unusual structure consisting of 26 heptapeptide repeats with the consensus amino acid sequence Pro Thr Ser Pro Ser Tyr Ser. This C-terminal domain (CTD) is unique to eukaryotic polymerase II enzymes. In yeast, a minimum of 9-11 of the repeat units are necessary for viability (Allison *et al.* 1988), and

replacing the yeast repeats with *Drosophila* CTD repeats gives rise to a recessive lethal (Allison *et al.* 1988). This suggests that the role of the repeat units is species specific.

The state of phosphorylation of the CTD may determine the state of transcriptional activation. Two forms of RNA pol II can be isolated, RNA pol II0 and RNA pol IIA. The II0 form appears to be much more transcriptionally active than the IIA form and is also highly phosphorylated (Bartholomew *et al.* 1986, Cadena and Dahmus 1987). It seems to be the unphosphorylated form, RNA pol IIA, which loads onto the promoter (Chesnut *et al.* 1992). Thus active RNA pol II could be formed by multiple phosphorylations of its largest subunit after it has associated with the promoter.

A yeast CTD kinase has been purified (Lee and Greenleaf 1989), and a protein with 50% homology to a phosphatase has been identified, which allows transcription of *HIS4* in the absence of BAS1, BAS2 and GCN4, which are normally required for basal and activated transcription (Arndt *et al.* 1989). If phosphorylation of the CTD was required for the elongation phase of transcription, then RNA pol II could be "locked" into an initiation complex by maintaining the unphosphorylated form. Other possible roles for the CTD include destabilising histone/DNA interactions during elongation, and interacting with transcription factors during initiation (Lee and Greenleaf 1989).

The second largest subunit of RNA pol II in yeast is 150 kDa in size and homologous to the prokaryotic β subunit, whilst the third core subunit is 45 kDa and a partial homologue of the bacterial α subunit (Sawadogo and Sentenac 1991). Some of the RNA pol II subunits are components of all three RNA polymerases; these are known as the common subunits and in yeast have molecular weights of about 27 kDa, 23 kDa and 14.5 kDa . Finally there are about five small subunits. The resulting RNA polymerase is unable to initiate selectively at promoters without the help of other factors, but can catalyse template dependent synthesis of RNA. The binding of RNA pol II to specific regions of promoters is aided by initiation factors; the TATA associated RNA pol II complex is sufficient for accurate transcription, but the level of mRNA production is regulated by gene specific factors which bind to the upstream of the promoter.

1.2 Phosphoglycerate Kinase: A Model System for Studying Transcription

There has been much interest in the phosphoglycerate kinase gene as it is a highly expressed gene in yeast, responsible for producing 1-5% of the total cellular RNA and protein (Holland and Holland 1978). It catalyses the conversion of 1, 3-bis-

phosphoglycerate to 3-phosphoglycerate during glycolysis, with the concomitant production of ATP. The *PGK* gene is regulated to some extent by carbon source, being three to four times more active in yeast cells grown on glucose than in yeast cells grown on a non-fermentable carbon source (Maitra and Lobo 1971, Chambers *et al.* 1989). Regulation of the gene also occurs in response to a heat shock (Piper *et al.* 1988), mediated by a heat shock element at -372.

The promoter of *PGK* was a good candidate promoter for the production of heterologous proteins in yeast, since when the *PGK* gene is present on a high copy number plasmid, the levels of PGK protein within the cell can be as high as 50% (Mellor *et al.* 1985). Once the components which make up the efficient upstream activation sequence were identified, the *PGK* promoter was found to be a good model system for studying the transcriptional mechanisms necessary to achieve high levels of gene expression.

PGK was cloned (Hitzeman *et al.* 1980) by immunological screening of a yeast library in *E. coli* with a mouse antibody to purified yeast PGK. This identified a 3.1kb *HindIII* fragment containing the entire *PGK* sequence. All known *pgk* mutants mapped to chromosome III (Lam and Marmur 1977), and it was demonstrated that there was only a single copy of the gene in yeast by rescuing *pgk* mutants with the *PGK* clone. Later, Dobson *et al.* (1982) mapped *PGK* to a 2.95 kb *HindIII* fragment from a λ library, and demonstrated that it was "identical" to Hitzeman's 3.1 kb fragment.

The sequence of *PGK* (Dobson *et al.* 1982, Hitzeman *et al.* 1982) revealed an open reading frame of 1248 bp encoding 416 amino acids which gave a protein with a predicted molecular weight of 44 kDa (Hitzeman *et al.* 1982). This was found to have 65% homology with horse and human PGK amino acid sequences. Examination of codon bias showed that out of a possible 61 codons, only 38 were used, and 95% of the codons utilised only 25 of the 61 (Hitzeman *et al.* 1982). This restricted codon usage may be one of the factors which allows PGK levels to be so high. Similar bias is found in other highly expressed yeast genes (Bennetzen and Hall 1982).

Both the chromosomal copy of *PGK,* and the 3.1kb clone on a plasmid, produced a 1.5kb mRNA. This suggested that all of the control sequences necessary for transcription of *PGK* were present in the clone. The *PGK* mRNA has a half life of 70-80 minutes, which is unusually long, and may be another factor contributing to the high expression of this gene (unpublished results, cited in Chen *et al.* 1984). The 5' end of the *PGK* RNA transcript was mapped to -36, relative to the ATG of *PGK*,

by cDNA extension (Hitzeman *et al.* 1982). Polyadenylation was found to occur at sites 86-93bp downstream of the TAA stop codon (Hitzeman *et al.* 1982).

The *PGK* promoter is similar in structure to other yeast RNA pol II promoters (Figure 1.1). Pol II promoters consist of a core promoter, from which a low level of transcription can be initiated, and an upstream activation sequence (UAS) which provides the gene with its regulatory properties, and can activate transcription to as much as five thousand times the basal level. Some promoters also have sequences for the repression of transcription, upstream repression sequence (URS). The core elements of the promoter are the TATA box and RNA start site (RIE on Figure 1.1), and RNA pol II promoters contain either both of these, or just the RNA initiation sequences (Weis and Reinberg 1992). The TATA box (consensus TATAAA) in yeast RNA pol II promoters is found between 40 and 120 base pairs upstream of the RNA start site (Chen and Struhl 1985, Furter-Graves and Hall 1990). This is in contrast to higher eukaryotes where the TATA element is only 25 to 30 base pairs from the start site (Breathnach and Chambon 1981). It acts as a nucleation point for the formation of an initiation complex containing RNA polymerase II and general transcription factors by allowing TFIID to bind. If the promoter does not contain a TATA box, then the initiator provides the site for the formation of the initiation complex, possibly through initiator-TAF interactions which tether TFIID (Concino *et al.* 1984, Roeder 1991, Weis and Reinberg 1992, Struhl 1994, Martinez *et al.* 1994).

The basal promoter of *PGK* contains two potential TATA sequences, TATA1 at -152 and TATA2 at -114, and a CT rich region upstream of the RNA start site at -39. Deletion of TATA2 has no effect on steady state RNA levels or the amount of PGK protein in the cell. Deletion of both potential TATA sites causes a dramatic reduction in protein levels (Rathjen and Mellor 1990). When the site of RNA initiation was investigated, it was found that initiation from TATA1 alone was at the wild type site, but from TATA2 the amount of wild type initiation was reduced, and initiation occurred at sites downstream. If neither TATA site was present then initiation occured from sites in the coding region. Thus, for initiation to occur at -39, TATA1 is sufficient but TATA2 is only partially functional. The RNA initiation site is not determined by distance from the TATA box, or by the CAAG sequence at the site of initiation. Rather, there is a sequence called the determinator, between the CT rich region and CAAG, which determines a single RNA start site. The role of the CT rich region is unclear. Such regions are found in many yeast promoters, but deletions in this region had no effect on levels of *PGK*, or the site of initiation.

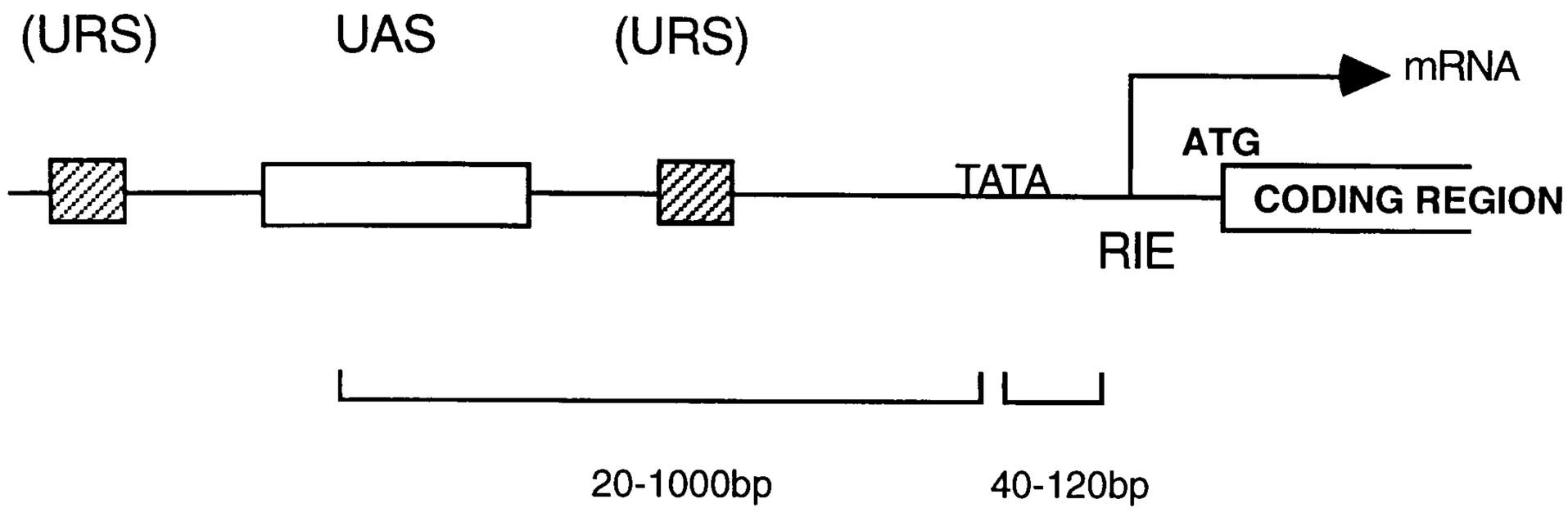


Figure 1.1 The general organisation of a yeast RNA polymerase II promoter. UAS: Upstream Activation Sequence, URS: Upstream Repression Sequence, TATA: TATA box, RIE: RNA Initiation Element.

The initiation of transcription requires many components, not just RNA polymerase (Table 1.1). This became clear when it was found that many components were necessary for reconstituting RNA pol II initiation activity *in vitro*. Purified RNA polymerase II was not sufficient for the initiation of transcription from the Adenovirus Major Late promoter (Matsui *et al.* 1980). However, addition of a fractionated human cell extract allowed initiation to occur. The cell extract fractions were termed transcription factor (TF) IIA-TFIID, depending upon which fraction they appeared in after separation from a crude extract by chromatography on phosphocellulose (Matsui *et al.* 1980). The fractions have since been studied intensively to discover the nature of the general transcription factors. Some of these fractions later turned out to contain more than one component, and the exact number of initiation factors necessary for transcription is not yet known.

1.3 General Transcription Factors

1.3.1 TFIID

This fraction contains the protein which binds to the TATA box; TATA Box Binding Protein (TBP). TBP in yeast is a 27 kDa protein, encoded by a unique gene (Hahn *et al.* 1989), with 180 amino acids at the C-terminus which are 80 to 90% conserved between species. These C-terminal amino acids (63-240) are required both for binding to the minor groove of the TATA box, and also for a basal level of transcription (Horikoshi *et al.* 1990). The N-terminus of yeast TBP shares no sequence similarity with the N-termini of TBPs from other species, and it has been proposed to be the site of interaction with transcription factors to give an activated level of transcription. However, yeast cells expressing TBP lacking the N-terminus grow at least as well as cells with wild type TBP, suggesting that this region is not required for the essential functions of TBP (Cormack *et al.* 1991).

In vitro translated yeast TBP is able to bind to TATA elements, and can also complement a mammalian system lacking TFIID (Buratowski *et al.* 1988). Interestingly, initiation occured at the mammalian distance of 30 bp downstream of the TATA box, suggesting that the distance of initiation from the TATA box is not determined by the TATA binding factor. As TBP is the only one of the basal transcription factors able to bind in a sequence specific manner to the TATA box, it is thought to recruit other general transcription factors, and also RNA pol II to promoters.

1.3.2 TFIIA

Table 1.1 General Initiation Factors Required by Yeast RNA Polymerase II, and their Human Homologues

Yeast	M Wt (kDa)	Human
d	27	TFIID
e	41	TFIIB
b	85, 75, 55, 50, 38	TFIIH
g	105, 54, 30	TFIIF
a	66, 43	TFIIE
TFIIA	32, 13.5	TFIIA

The role of TFIIA in assisting the formation of a stable complex between TFIID and the TATA box is unclear. Buratowski *et al.* (1988) found TFIIA not to be essential for transcriptional initiation but it appeared to have a stimulatory effect; other studies suggest that TFIIA is necessary to achieve a high level of transcription (Ranish and Hahn 1991), or that TFIIA is not required for promoter recognition (Sayre *et al.* 1992a). This confusion may be partly explained by the identification of a second activity (TFIIJ) in the TFIIA fraction (Cortes *et al.* 1992). TFIIA had a stimulatory effect on transcription and TFIIJ was required for initiation.

In yeast, TFIIA consists of two polypeptides (32 and 13.5 kDa) both of which are required for its activity. Yeast TFIIA can functionally replace mammalian TFIIA in basal transcription. It is not a DNA binding protein, and it binds to the TFIID/TATA complex with greater affinity than it binds to TFIID not bound to DNA (Ranish and Hahn 1991).

One possible role for TFIIA is to block the interaction of negative regulators with the basal transcription factors (Auble and Hahn 1993). ADI (ATP dependent inhibitor) prevents TBP from binding to the TATA box in yeast. The interaction of ADI and TBP is through the C-terminus of TBP and this is blocked by TFIIA.

1.3.3 TFIIB

After TFIID has bound to the TATA box, this complex is bound by TFIIB via a protein-protein interaction with TBP. The presence of TFIIB is absolutely required for the initiation of transcription (Reinberg and Roeder 1987), where it is necessary for the recruitment of RNA pol II to the initiation complex (Buratowski *et al.* 1989). RNA pol II and TFIIF are recruited through another protein-protein interaction between the small subunit of TFIIF and TFIIB (Flores *et al.* 1991). The order of the interactions between TFIIB and its targets is not known; TFIIB can bind stably to RNA pol II in solution (Koleske and Young 1994) so the association with TBP does not necessarily precede association with RNA pol II.

TFIIB acts as a bridge between TFIID and RNA pol II which suggests that it contains separate domains for the two protein-protein interactions. The C-terminus has been shown to be both necessary and sufficient for an interaction with the TBP/DNA complex (Hisatake *et al.* 1993), and the N-terminus is important for an interaction with TFIIF (Ha *et al.* 1993). Yeast initiation factor e is the homologue of human TFIIB (Tschochner *et al.* 1992). It is a 41 kDa protein, encoded by *SUA7*, which interacts specifically with RNA polymerase II and, as has been demonstrated with gel retardation assays, also with TFIID and TFIIA which are bound to promoter

DNA. Factor e is also proposed to be responsible for determining the RNA start site (Pinto *et al.* 1992).

1.3.4 TFIIF

This general transcription factor consists of two subunits in humans, RAP30 and RAP74 (Flores *et al.* 1989, Sopta *et al.* 1989), and three subunits in yeast, 105, 54 and 30 kDa (Tan *et al.* 1994). The small subunit alone is sufficient for the recruitment of RNA pol II to the DNA/IID/IIA/IIB complex (Flores *et al.* 1991) but the complex is formed more efficiently if both subunits are present (Tyree *et al.* 1993). The association of TFIIF with RNA pol II can occur in the absence of DNA.

TFIIF is unique amongst the general transcription factors, since it is found associated with the transcription complex after initiation is complete. This means that it can interact with both phosphorylated and unphosphorylated forms of RNA pol II. Whilst TFIIF is absolutely required for the formation of the initiation complex, it also has a role in the elongation of transcription (Flores *et al.* 1989) where it may cooperate with the elongation factor TFIIS. TFIIF is able to increase the rate of elongation about six-fold (Bengal *et al.* 1991), by suppressing pausing at some sites on the template (Tan *et al.* 1994).

1.3.5 TFIIE

TFIIE is a heterotetramer of two 56 kDa subunits and two 34 kDa subunits which enters the pre-initiation complex after RNA pol II (Inostroza *et al.* 1991). It is not required at all promoters, being essential at the adenovirus major late promoter but not at the immunoglobulin heavy chain promoter (Parvin *et al.* 1992). When present it probably interacts through RNA pol II; TFIIE has been shown to bind both RNA pol II and TFIIB in solution (Reinberg and Roeder 1987). Another protein-protein contact may be involved in the recruiting of TFIIH which binds after TFIIE; the interaction between TFIIE and TFIIH may be cooperative (Flores *et al.* 1992). TFIIE could contain an ATPase activity, since after addition of ATP or dATP, TFIIE dissociates from the initiation complex (Buratowski *et al.* 1989).

The yeast equivalent of TFIIE is factor a, but human TFIIE cannot substitute for factor a. Factor a consists of two polypeptides of 66 kDa and 43 kDa both of which are required for transcriptional activation (Sayre *et al.* 1992b). When a multi-component complex containing RNA pol II and the initiation factors e, b and g was isolated from yeast (see Table 1.1), factor a was required, along with TBP, for selected transcription *in vitro* (Koleske and Young 1994).

1.3.6 TFIIH

This complex consists of eight subunits and is the only general transcription factor to have an associated enzyme activity (reviewed in Drapkin *et al.* 1994). TFIIH copurifies with a DNA-dependent ATPase activity, an ATP dependent DNA helicase activity and a CTD kinase activity. TFIIE is not only able to recruit TFIIH to the pre-initiation complex, but also to regulate its enzyme activities. Subunits of TFIIH have been shown to be encoded by genes involved in excision repair, thus providing a link between transcription and DNA repair (Drapkin and Reinberg 1994).

The yeast homologue of TFIIH is factor b; originally thought to be made up of three polypeptides with weights 85 kDa, 75 kDa and 50 kDa (Feaver *et al.* 1991a), it is now known to contain two other subunits of 55 and 38 kDa (Feaver *et al.* 1993). Associated with factor b is a DNA-dependent ATPase activity which is stimulated by DNA fragments containing promoter sequences. There is also a protein kinase activity which will phosphorylate the yeast RNA polymerase II CTD (Feaver *et al.* 1991b). Like human TFIIH, yeast factor b has a role in nucleotide excision repair; the 85 and 50 kDa subunits are encoded by *RAD3* and *SSL1*, and another excision repair protein, RAD25/SSL2, interacts with, but is not a component of, factor b (Feaver *et al.* 1993).

1.4 Binding of basal transcription factors to form an active initiation complex.

An ordered assembly of general transcription factors and RNA pol II onto a promoter has been proposed (Buratowski *et al.* 1989, reviewed in Conaway and Conaway 1993, and Buratowski 1994; see Figure 1.2). In this, TFIID and TFIIA bind to the TATA box of the promoter forming a stable complex which is recognised by TFIIB. TFIIF can interact with both TFIIB and RNA polymerase II, and it is TFIIF that brings RNA pol II to the pre-initiation complex. Two further factors are required before transcription can occur; TFIIE binds and recruits TFIIH forming a complete pre-initiation complex. ATP is then required to form an activated pre-initiation complex, capable of initiating transcription when nucleoside triphosphates are supplied (Goodrich and Tjian 1994). There is evidence (Koleske and Young 1994) that rather than associating with the promoter in a linear fashion, initiation factors and RNA polymerase II could assemble into a multisubunit complex in the absence of DNA and then bind to the promoter.

Once the initiation complex has formed, the addition of nucleotides allows transcription to take place. The general transcription factors do not travel with the

i. TATA + TFIID/TFIIA

ii. + TFIIB

iii. + RNA pol II via TFIIF

iv. +TFIIE

v. + TFIIH

vi. + ATP

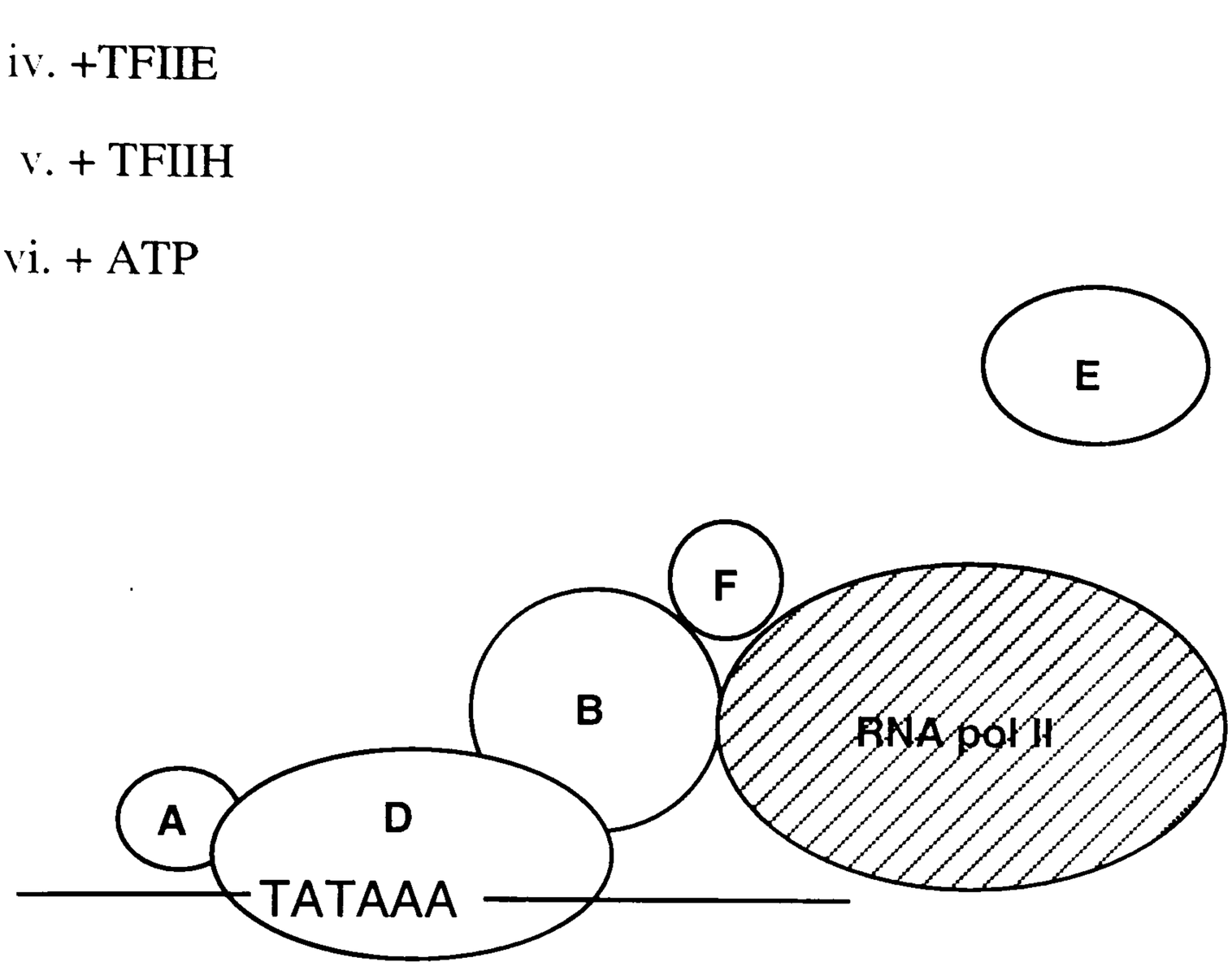


Figure 1.2 The assembly of general transcription factors and RNA polymerase II onto a promoter to form a pre-initiation complex. Rather than assembling in a linear fashion as shown above, it is possible that components of the general transcription factors assemble into a multisubunit complex in the absence of DNA and then bind to the promoter (Koleske and Young 1994).

elongating polymerase, but rather are lost from the initiation complex as it moves into elongation (Zawel *et al.* 1995). TFIID remains bound to the promoter ready to recycle TFIIB as it is released. TFIIE is lost by the time the nascent RNA reaches +10, and TFIIH sometime after elongation reaches +30. The only general trancription factor which has been found associated with RNA polymerase after initiation has occured is TFIIF. This is released from the initiation complex after +10, but has the ability to reassociate with a polymerase if it stalls at a transcriptional block. Once the stalled polymerase starts to elongate TFIIF is released once more. This release of general transcription factors after initiation allows these factors to be recycled, and may increase the numbers of polymerases loading onto a promoter. This may be more important for achieving a high level of transcription than the formation of the first initiation complex (Zawel *et al.* 1995).

1.5 Activated Transcription

Both the activation of transcription, and the regulation of transcription in response to physiological signals, are a result of the transcription factor binding sites in its upstream activating sequence (UAS; see Figure 1.1) The UAS mediates the regulation of gene expression in response to physiological signals, and stimulates the activity of the core promoter over a basal level. It is similar to the enhancer elements of higher eukaryotes in that it can function in both orientations, and at a variable distance from the TATA box. However the distance over which it will function is only up to about a kilobase, not several, and neither will it work if placed downstream of the initiation site (Guarente and Hoar 1984, Struhl 1984).

In some constitutively expressed genes, the UAS consists of a poly (dA-dT) sequence, *eg PET56*, *HIS3* and *DED1* (Struhl 1985). This activation of transcription may be due to the recognition of the poly (dA-dT) sequence by a yeast DNA binding protein, or because the general transcription factors are able to gain increased access to a DNA template with a disrupted chromatin structure, possibly due to the exclusion of nucleosomes. This would explain the constitutive nature of genes with poly (dA-dT) UAS. However, the poly (dA-dT) regions of *PET56*, *HIS3* and *DED1* are all able to form nucleosome cores (Losa *et al.* 1990). A gene encoding a yeast protein that recognises nonalternating oligo(A)·oligo(T) tracts has been cloned (Reardon *et al.* 1993), this may play a role in the promotion of transcription from such promoter elements.

Some yeast genes may rely not only on an upstream activating sequence for maximum levels of transcriptional activation. In certain cases the presence of a downstream activation sequence (DAS) which influences transcription initiation has

been suggested. A transcription factor binding site(s) within the coding region, whose deletion reduces the level of transcription but not by decreasing mRNA stability, was proposed to explain the observation that some highly expressed yeast promoters do not activate heterologous gene constructs to levels which are as high as might have been expected (Chen *et al.* 1984). This has been noted for *PGK*, *PYK* and *SRP1* (Mellor *et al.* 1987, Purvis *et al.* 1987, Fantino *et al.* 1992). Indeed, Abf1p has been shown to bind to the *PGK* coding region at position +79/+91 (Ian Graham, unpublished results).

1.5.1 Transcriptional Activator Proteins

Proteins which activate transcription consist of two domains, a DNA binding domain and an activator domain (Hope and Struhl 1986). The DNA binding domain anchors the transcription factor to the promoter so that the activation domain is able to interact with other factors present at the promoter (Brent and Ptashne 1985). The high specificity of the DNA/protein interaction provides a mechanism by which genes can be differentially expressed. The activation domain from activator protein A can be fused to the DNA binding domain of activator B, and this hybrid will still activate transcription as long as binding sites for B are present in the promoter (Ptashne 1986).

Since the association of TFIID with the TATA box is a slow, rate limiting step *in vitro* (Hoopes *et al.* 1992) various roles for gene specific transcriptional activators can be proposed: They could i) act on previously potentiated promoters to increase the frequency of initiation, ii) act cooperatively with the general transcription factors to help with the assembly of the initiation complex, iii) act independently to help assemble the initiation complex, or iv) activate repressed promoters by disrupting chromatin structure and allowing general transcription factors to bind. The recruitment of TBP to the basal promoter has been shown to be increased by the presence of an activator (Klein and Struhl 1994a).

1.5.2 DNA Binding Domains

Transcriptional activator proteins utilise various structural motifs in their DNA binding domains in order to bind DNA. These include helix-turn-helix (HTH), zinc finger, leucine zipper and helix-loop-helix (HLH). The helix-turn-helix motif consists of two α-helices separated by about four amino acids forming a β-turn. HTH proteins often bind to DNA as dimers, the second helix fits in the major groove whilst the first helix lies across it in contact with the DNA backbone. The recent identification of the DNA binding motif of Rap1p provides an example of an HTH

protein which does not bind as a dimer (Rhodes D, unpublished). The HTH motif is often part of a larger DNA binding domain which is important for recognition of the binding site (Pabo and Sauer 1992).

The zinc finger motif was first noted in TFIIIA (Miller *et al.* 1985), and consists of two cysteines and two histidines positioned to chelate a zinc ion. Once zinc is bound, the "finger" will bind DNA. Yeast zinc fingers are not homologous to the TFIIIA motif, since although they bind zinc, they have four to six cysteines per finger (Pfeifer *et al.* 1989). Yeast activators with Zn fingers, for example, Gal4p, show considerable homology to each other, and this extends either side of the finger. It is possible that the finger interacts non-specifically with the DNA and the DNA binding specificity is determined by sequence carboxy terminal to this (Pfeifer *et al.* 1989).

The leucine zipper (Landshulz *et al.* 1988) has a basic region, which usually forms a helix, for interaction with DNA, and a dimerization region containing 4 or 5 leucines spaced exactly seven amino acids apart. Thus, all the leucines are on the same face when a helix is formed. Initially, the leucines were thought to interdigitate when dimerization occured (Landshulz *et al.* 1988), but it is now known that the two helices form a coiled coil (O'Shea *et al.* 1989). For DNA binding to occur, the coiled coil fits over the centre of the binding site, and the basic helices extend in opposite directions along the major groove.

The helix-loop-helix motif (Murre *et al.* 1989a) is similar to the leucine zipper motif in that it contains a basic region followed by a dimerization domain. In this case the dimerization domain is an α-helix, followed by a six to ten amino acid loop, and a second α-helix. Like leucine zipper and HTH proteins, HLH proteins can form both homodimers and heterodimers (Murre *et al.* 1989b) which allows regulation of the transcription factor itself, and can also generate new DNA binding specificities.

1.5.3 Activation Domains

There are several classes of activator domains, glutamine-rich, for example, Gal11p, proline-rich and the most common class, acidic activators, which seem to be able to activate transcription in all eukaryotes tested. Initial studies of the acidic activation domains of Gal4p and Gcn4p showed them to be small and negatively charged (Hope and Struhl 1986). There was no specific sequence requirement and extensive deletions into the Gcn4p activation domain could be made without affecting its *in vivo* function. This suggested that tertiary structure was not important, although the activation domain was proposed to form α-helices (Hope *et al.* 1988). The idea that

acidic activation domains are amphipathic helices, or unstructured regions with a high concentration of negative charge (Sigler 1988) has been challenged by studies which suggest that the activation domains of Gal4p and Gcn4p form β-sheets (Leuther *et al.* 1993, Van Hoy *et al.* 1993), although the acidic activator VP16 has not been shown to do so. Thus, it is possible that there is more than one mechanism by which activators function, and that the activator domain should have a flexible structure, for interaction with its targets.

1.5.4 Interaction Between Upstream and Downstream Promoter Elements

RNA pol II and the general transcription factors form a complex over the basal promoter which may be situated at some distance from the UAS where the activator binding sites are located. As only two or three proteins can bind adjacent to the initiation complex, proteins upstream must be brought into contact with the basal factors. Various suggestions have been made including twisting, sliding and oozing (reviewed in Ptashne 1986), but the mechanism now generally accepted for eukaryotes is looping (Schleif 1992). In this case two proteins separated by several kilobases of DNA can be brought into contact if the intervening DNA forms a loop.

1.5.5 Targets for Transcriptional Activators

That activators are able to increase the level of transcription from a gene predicts that the activator proteins make contacts with the basal transcription factors. Many of the general transcription factors have been demonstrated to interact with activators. Acidic activators interact with TBP (Lee and Struhl 1995), TFIIB (Lin and Green 1991, Roberts *et al.* 1993) and TFIIH (Xiao *et al.* 1994), whilst the glutamine-rich Sp1 can interact with TFIIE (Peterson *et al.* 1991). Also, the largest subunit of RNA pol II may interact with activators via the CTD (Allison and Ingles 1989). Activators are able to mediate their effects at different stages in the formation of the initiation complex. Thus, some act early to increase recruitment of TBP (Klein and Struhl 1994a) or other general transcription factors, binding sites for basal transcription factors on TFIIB are revealed in the presence of an acidic activator (Roberts and Green 1994). Others interact with the final components of the complex, possibly helping to promote open-complex formation, or chain elongation (Xiao *et al.* 1994).

1.5.6 Coactivators

Further components of the basal transcription machinery were identified when it was demonstrated that TBP could not replace the TFIID fraction in responding to transcriptional activators such as Sp1 and GAL-VP16 (Pugh and Tjian 1990, Berger

et al. 1990), although basal transcription was not affected. The other factors present in the TFIID fraction are termed TBP Associated Factors (TAFs). It was proposed that the TAFs might interact with the N-terminus of TBP, but an N-terminal truncation of human TBP contains all the major TAFs which form a complex with full length TBP. The truncated TBP will support transcriptional activation both from different classes of activation domain i) Sp1-glutamine-rich, ii) Gal4-AH-acidic and iii) Zta-not rich in any particular amino acid, and also from a TATA-less promoter (Zhou 1993). Yeast TBP has also been found to be stably associated with other factors, yeast TAFs (Poon and Weil 1993).

Other protein factors are involved in transmitting the signal from activators to the basal complex, these factors can be titrated by a strong activator causing transcription to be reduced from genes without binding sites for that activator (Berger *et al*. 1990, Kelleher *et al*. 1990). These adaptors, mediators or coactivators are proteins which do not bind to DNA but make protein/protein contacts, and they have been identified in screens where activators function without their activation domains, or where weak activators work as strong ones, *eg* GAL11, SUG1, ADA2 and ADA3 (Himmelfarb *et al*. 1990, Swaffield *et al*. 1992, Berger *et al*. 1992, Pina *et al*. 1993).

Recently a mediator complex was isolated from yeast, which enabled acidic activators to activate transcription of a system reconstituted with essentially homogenous basal factors, and RNA pol II. The mediator was found to be made up of about twenty proteins including three subunits of TFIIF, GAL11, SUG1 and SRB2, 4, 5 and 6. At the same time a holoenzyme was isolated which allowed purified basal factors to respond to activators. The holoenzyme was found to consist of the mediator and RNA pol II (Kim *et al*. 1994). A similar holoenzyme, stimulated by the activator GAL4-VP16, was isolated by a different group (Thompson *et al*. 1993, Koleske and Young 1994). The need for the holoenzyme to allow basal transcription factors to respond to activators suggests that a direct interaction between activators and general transcription factors may not be sufficient for activated transcription.

1.6 Activated Transcription at *PGK*

After the primary structure of *PGK* had been determined, the promoter was studied to find regions which were required for efficient expression of the gene (Ogden *et al*. 1986). A series of unidirectional deletions showed that sequence upstream of -620 could be deleted without any effect on the levels of transcription from *PGK*. However, deletions which removed promoter sequence to -350 caused a dramatic

reduction in the amount of RNA produced. The identification of a region which was important for transcriptional activation allowed fine deletions to focus on the sequence between -620 and -350. Window deletions were made within this region and the boundaries of the upstream activation sequence were defined as -402 to -479 (Ogden *et al.* 1986). This upstream activation sequence (UAS) could be moved closer to the RNA start site without reducing levels of transcription, and insertion of the UAS fragment into a deletion window which reduced *PGK* trancription resulted in a return to wild type levels of *PGK* RNA. In addition, the *PGK* UAS was able to enhance the transcription of a heterologous gene construct consisting of the *TRP1* promoter linked to the human interferon α-2 coding region (Ogden *et al.* 1986).

The *PGK* UAS was divided into two fragments Y (-461 to -531) and Z (-402 to -460) which were incubated in gel retardation assays with protein extracts from cells grown in glucose or acetate (Stanway *et al.* 1987). Different retardation complexes were formed depending upon whether the carbon source was fermentable or non-fermentable. DNaseI footprinting revealed an area of protection over the region -523 to -496 in cells grown on glucose (termed Yfp, Y footprint) which was absent when the cells had been grown on acetate. The UAS appeared to contain two domains, one for controlling transcription in response to carbon source and one for activating transcription. Also present were three repeats of a sequence, CTTCC, thought to be of potential functional significance (Stanway *et al.* 1987).

Window deletions of small regions of the *PGK* UAS were made to look at their effects on transcriptional activation (Chambers *et al.* 1988). These showed that, whilst important for transcriptional activation, the CTTCC blocks do not all activate to the same extent. Removal of CT block 1 caused a 50% decrease in the level of transcription, removal of block 2 reduced transcription by 75%, and deletion of all the UAS sequence upstream of block 1 results in about a 90% decrease in activation. A window deletion which removed the region protected in footprinting (Yfp) had no effect on the level of transcription, but a deletion of sequence 3' to this protected region, not including the CTTCC blocks, caused an 80% drop in levels of RNA. Thus another functional element of the *PGK* UAS, termed the activator core (AC), was identified (Chambers *et al.* 1988). Fragment Z (-402 to -460) was extended to include the new sequence (Z^+). When Z^+ was used in a gel retardation assay, a specific protein interaction was found which was not due to the protein which bound the protected region on fragment Y. DNaseI footprinting showed protection of the AC region, and also some protection of the CT blocks. This protection was different on the coding and non-coding strands suggesting asymetric protein binding.

The *PGK* UAS contains three elements; three CTTCC blocks, an activator core and a region of strong protein interaction, the Yfp. These elements were examined for transcriptional activity in the context of a minimal promoter plasmid (Stanway *et al.* 1989). The minimal promoter was constructed by taking TATA and RNA start site sequences from the *PGK* promoter and linking them to the IFN coding region. Subfragments of the *PGK* UAS were cloned upstream of the minimal promoter, and the levels of interferon RNA used to determine the relative activation potential of each fragment. The whole UAS could activate transcription to a high level and the Yfp was able to activate weakly, but the three CTTCC blocks were inactive, as was the activator core in combination with one CT block. However, the activator core in conjunction with either all three CT blocks, or the Yfp, was able to activate to a moderate level.

The activator core was shown to bind the multifunctional transcription factor Rap1p (Chambers *et al.* 1989) and this site was also important for the carbon source regulation of *PGK*. The binding of Rap1p to the *PGK* promoter was investigated using nuclear protein extracts from cells which had been grown in glucose or pyruvate. Rap1p binding to the Z^+ fragment was seen with the glucose extract but not the pyruvate extract. As levels of Rap1p mRNA are not affected by carbon source this suggests that regulation of *PGK* transcription in response to carbon source is mediated at the level of binding Rap1p to the *PGK* UAS.

Such regulation could be achieved by post translational modification of Rap1p. When a nuclear protein extract from cells grown in glucose was treated with phosphatase, binding to the Z^+ fragment was abolished (Tsang *et al.* 1990). It could be restored by including a phosphatase inhibitor, ammonium molylbdate, or protein kinase in the phosphatase reaction. Treatment with phosphatase caused binding of Rap1p to the Z^+ fragment to decrease, but when sequences from the *PGK* promoter normally found 5' to Z^+ were included in retardation reactions, binding of phosphatased Rap1p increased. This may mean that the 5' end of the Rap1p binding site is involved in the stability of Rap1p binding. Similar responses were found when just the DNA binding domain of Rap1p was treated with phosphatase.

Rap1p appears to play a central role at the *PGK* promoter. It mediates regulation of transcription in response to carbon source (Chambers *et al.* 1989), and although the Rap1p binding site alone was unable to activate a minimal promoter (Stanway *et al.* 1989), in conjunction with the CTTCC boxes it is important for an activated level of transcription (Chambers *et al.* 1988, Stanway *et al.* 1989, Henry *et al.* 1994).

The 5' region of the *PGK* UAS, termed the Yfp, contained a perfect match to the Abf1p consensus sequence. A protein which bound to this region was purified from yeast and shown to have identical properties to Abf1p (Chambers *et al.* 1990). Also, *in vitro* translated Abf1p bound to the Yfp (Chambers *et al.* 1990). The role of Abf1p in the *PGK* promoter is unclear, although deletion of its binding site did not affect transcription from a multicopy plasmid borne copy of *PGK* (Chambers *et al.* 1988), the Abf1p site was able to activate a minimal promoter weakly (Stanway *et al.* 1989). Neither did a deletion of the Abf1p site affect carbon source regulation (Chambers *et al.* 1989), although the binding of protein to this site appeared to be affected by carbon source (Stanway *et al.* 1987, Chambers *et al.* 1989).

It was thought that a protein bound to the CTTCC blocks in the *PGK* UAS, since *in vitro* footprinting had revealed some protection over these regions (Chambers *et al.* 1988), and binding of Gcr1p to the CTTCC motif had been demonstrated *in vitro* (Baker 1991). However, *in vivo* binding was harder to demonstrate, but eventually Gcr1p was found to bind to just two of the three CTTCC blocks in the *PGK* promoter by *in vivo* footprinting in *GCR1* and *gcr1*$^-$ strains (Henry *et al.* 1994). Protection of -454 and -453 in CTTCC block 1 and -429 and -428 in block 3 was seen but there was no evidence of protection of block 2. The binding of Rap1p to its binding site was not affected by the presence or absence of Gcr1p. Gcr1p was shown to positively influence transcription from *PGK* but it requires Rap1p to be bound to the promoter to do so (Drazinic and Baker, unpublished).

The yeast co-activator Gal11p has been shown to have a positive effect on transcription from *PGK* in both fermentable and non-fermentable carbon sources (Stanway 1994). When levels of *PGK* were examined in *GAL11* and *gal11*$^-$ strains Gal11p was found to stimulate a two-fold increase of *PGK* transcription. This effect is only seen if the Rap1p site is present; if the Rap1p site is deleted from a copy of *PGK* on a high copy number plasmid there is no decrease in the activity of the construct in a *gal11*$^-$ strain. When the *PGK* UAS was footprinted in both *GAL11* and *gal11*$^-$ strains no difference in protection was seen, suggesting that binding of transcription factors is not affected by Gal11p.

1.7 Transcription Factors which Bind to the *PGK* Promoter

1.7.1 Rap1p

Repressor/Activator protein, or Rap1p, is an essential (Shore and Nasmyth 1987), abundant transcription factor in the yeast cell. There are estimated to be about 6000 molecules of Rap1p per haploid nucleus (Verdier *et al.* 1990) most of these have

been found, by immunolocalisation, to be at the ends of paired bivalent meiotic chromosomes (Klein *et al.* 1992). The many roles of Rap1p within the cell are reflected in the number of names under which the protein was studied: TUF, SBF-E, TBA and GRF1 (Huet *et al.* 1985, Shore and Nasmyth 1987, Longtine *et al.* 1989, Buchman *et al.* 1988a) before it was cloned (Shore and Nasmyth 1987). Rap1p was purified by affinity chromatography and the gene was then isolated from a genomic library using an antibody generated against the purified protein. The DNA sequence of *RAP1* contained one ORF encoding a protein with a predicted molecular weight of 92.5kDa, although the apparent weight on SDS-PAGE was 120kDa.

Rap1p is encoded by an essential gene (Shore and Nasmyth 1987) which contains within its promoter four Rap1p binding sites. These are not essential for efficient activation of Rap1p (Graham and Chambers 1994a), but Rap1p is thought to have a role in negative autoregulation (Graham and Chambers, unpublished results). The consensus DNA sequence to which Rap1p binds was derived by Buchman *et al.* (1988a) and has recently been extended at both the 5' and 3' ends (Graham and Chambers 1994b) to give 5' RTRCACCCANNCMCC 3'. This consensus binding site has a conserved core, with 5' and 3' flanking regions which may be required to stabilise the interaction of Rap1p with DNA.

The DNA binding domain (DBD) of Rap1p (Henry *et al.* 1990) is in a central region of the protein (amino acids 361-596), as demonstrated by testing N- and C-terminal truncations of the protein in gel retardation assays. The Rap1p DBD does not contain any obvious DNA binding motifs, but at a recent meeting the crystal structure was shown to be an HTH motif with homology to c-myc (Rhodes D., unpublished). The 236 amino acid DBD is large when compared with the DNA binding domains of other yeast transcription factors such as GAL4, GCN4 and HAP1 whose DBDs range from 60 to 148 amino acids. However, another multifunctional transcription factor, Abf1p, also has a large DNA binding domain, but it does contain a potential Zn finger motif at its N-terminus.

Rap1p binding sites had been found at silencers, in UASs and at telomeres suggesting roles in both activation and silencing (Buchman *et al.* 1988a, Buchman *et al.* 1988b). This has been demonstrated using yeast strains containing temperature sensitive Rap1p (Kurtz and Shore 1991). *MATα*, which has a Rap1p binding site in its UAS, shows decreased levels of transcription in *rap1*^{ts} strains, whilst in the same strains, partial derepression of *HMR* silencing occurs (Kurtz and Shore 1991). Individual Rap1p binding sites from UASs, silencers and telomeres were all shown to activate transcription of a β-galactosidase reporter when placed in a UAS-less promoter (Buchman *et al.* 1988b). Since a mutated Rap1p binding site at a silencer

can be functionally replaced with a true binding site from either a UAS, or a silencer (Shore and Nasmyth 1987), the context of the DNA binding site appears to determine the role of Rap1p. In other words, the sequences which flank Rap1p binding sites at the silencers cause Rap1p to silence transcription, whereas those surrounding Rap1p sites in UASs cause it to activate transcription.

The C-terminus of Rap1p contains regions which are important for activation of transcription and also transcriptional silencing. Hardy *et al.* (1992a) fused regions of the Rap1p C-terminus to the $GAL4_{DBD}$ and found that amino acids 630-695 were important for transcriptional activation, whilst derepression required amino acids 678-827. These two regions overlap but the fact that some of the hybrids acted only to activate, or to derepress suggests that there may be two separate domains for these functions. The *rap1*^{s} mutants of Sussel and Shore (1991), which are defective in the silencing function of Rap1p, all map to the C-terminal region of Rap1p necessary for silencing, but not activation. A further domain has since been defined in the C-terminus, responsible for telomeric silencing. Mutations in the last twenty eight amino acids of the C-terminus have been demonstrated to be essential for both telomeric and HML silencing, and also to play a role in the regulation of telomere length (Liu *et al.*1994).

Most studies of Rap1p have focused on the C-terminus since a large deletion of the N-terminus is still viable (cited in Hardy *et al.* 1992a). However, *in vitro* studies suggested that the binding of Rap1p causes the DNA to bend, and that this requires the N-terminus. This bending does not ocur at the Rap1p recognition sequence but at a site 5' to it (Vignais and Sentenac 1989, Gilson *et al.* 1993). Vignais and Sentenac proposed that this bending by Rap1p required two domains, one to bind the DNA and a second to affect the bending, since mutations which affect the strength of Rap1p binding do not affect DNA bending. In support of this, DNaseI and chemical footprinting of the Rap1p/DNA complex *in vitro* revealed that full length Rap1p or just the Rap1p DNA binding domain caused a distortion within the consensus recognition sequence, but only the full length protein was able to induce a bend in the DNA (Gilson *et al.* 1993). The DNA bend has been shown to be greater than 50^{o}, but it is reduced by removing 230 N-terminal amino acids (44-274), and increased by removing C-terminal amino acids (Muller *et al.* 1994).

The many roles of Rap1p within the yeast cell have been suggested to be a result of the context of the Rap1p binding site. This hypothesis predicts that there will be other proteins in yeast which will interact with Rap1p at these loci in order for the Rap1p binding site to differentiate between activation or silencing. Such proteins have been identified: Silent Information Regulator (SIR1-4) proteins interact with

Rap1p at silencers and telomeres, as does Rap1p-Interacting Factor (RIF1). Mutations in *RIF1* result in defective silencing, and also in the lengthening of telomeres, but a deletion of *RIF1* does not affect the essential activation functions of Rap1p (Hardy *et al.* 1992b). The two hybrid system has been used to show that the interaction between Rap1p and Rif1p involves the C-terminus of Rif1p and amino acids 667-827 of Rap1p. This region of Rap1p is also needed for its role in silencing which suggests that Rap1p/Rif1p interactions may be restricted to silencers and telomeres. An interaction between Rap1p and the SIR proteins is suggested by evidence that the sub-nuclear localisation of Sir3p and Sir4p is similar to that of Rap1p (Palladino *et al.* 1993), and that the silencing defect of rap1^{s} can be supressed by overexpression of Sir1p or Sir4p (Sussel and Shore 1991). A direct interaction between the C-terminus of Rap1p and Sir3p and Sir4p was demonstrated using the two hybrid system (Moretti *et al.* 1994). The SIR proteins do not bind DNA and thus Rap1p may be involved in their recruitment to *HM* loci and telomeres for the establishment of silencing. A protein which works in concert with Rap1p in the activation of transcription at glycolytic loci, is Gcr1p (GlyColysis Regulator).

1.7.2 Gcr1p

Gcr1p was first identified as a mutation in yeast which affected several of the glycolytic enzymes (Clifton *et al.* 1978). Whilst the *gcr1* mutation decreases the levels of most glycolytic enzymes, more so when cells are growing on sugars rather than without sugars, it is not in an actual glycolytic gene (Clifton and Fraenkel 1981). The *GCR1* gene was cloned by complementation, and sequencing showed it to be a 94kDa protein (Baker 1986) which was not essential to the cell. There is some evidence that Gcr1p can interact with Rap1p without contacting DNA, and Rap1p and Gcr1p have been shown to coimmunoprecipitate (Tornow *et al.* 1993), suggesting that they can form a complex *in vivo*. However, Gcr1p is able to bind DNA independently of Rap1p (Baker 1991), but it is possible that the interaction of Gcr1p with DNA is stabilised by the presence of Rap1p. Binding sites for Gcr1p are found in the promoters of many glycolytic genes adjacent to Rap1p binding sites (Reviewed in Chambers *et al.* 1995).

The DNA binding domain has been mapped to the C-terminal 154 amino acids of Gcr1p (Huie *et al.* 1992), and a consensus recognition sequence derived with CTTCC at its core (Baker 1991). Gcr1p also contains an activation domain at its N-terminus (Tornow *et al.* 1993). This is essential for the function of Gcr1p *in vivo*, as demonstrated by complementation in *gcr1*$^{-}$ cells (Tornow *et al.* 1993). Recently Gcr1p has been demonstrated to contain a leucine zipper structure necessary and

sufficient for dimerization *in vitro* and *in vivo*, although dimerization does not seem to be essential for functional Gcr1p (Deminoff *et al.* 1995)

Gcr1p at glycolytic gene promoters does not only interact with Rap1p. A screen looking for other activities which affect the expression of glycolytic genes has identified two further Gcr genes; *GCR2* and *GCR3* (Uemura and Fraenkel 1990, Uemura and Jigami 1992a). Mutations in both of these genes have similar phenotypes to *gcr1* mutations, *viz* defective growth on fermentable carbon sources, and near normal growth on non-fermentable carbon sources. The amino acid sequence of Gcr2p shows a region of similarity with Gcr1p, and the two hybrid system has been used to demonstrate an interaction between these proteins (Uemura and Jigami 1992b). In this system a $GAL4_{AD}$/Gcr1p hybrid was able to complement a *gcr2*$^-$ strain suggesting that Gcr2p provides Gcr1p with an extra activation domain. In agreement with this a Rap1p/Gcr2p fusion is able to complement *gcr1*$^-$ mutants. Thus Gcr1p and Gcr2p may form a complex in which Gcr1p binds DNA and Gcr2p donates an activation domain. The role of Gcr3p remains unclear.

1.7.3 Abf1p

Abf1p (ARS Binding Factor), like Rap1p, is a multifunctional yeast protein. Also known as SBF-B, BAF1, REB2, GF1, TAF, SUF, OBF1 and Y protein (Shore *et al.* 1987, Halfter *et al.* 1989a, Morrow *et al.* 1989, Dorsman *et al.* 1990, Hamil *et al.* 1988, Dorsman *et al.* 1989, Francesconi *et al.* 1989, Stanway *et al.* 1987), Abf1p binding sites are found at silencers, ARSs and in the UASs of genes with a wide range of functions such as ribosomal protein genes, glycolytic genes, *COX6*, *CAR1* and *ILV1* (Buchman *et al.* 1988a, Dorsman *et al.* 1989, Della Seta *et al.* 1990, Trawick *et al.* 1992, Kovari and Cooper 1991, Remacle and Holmberg 1992). This wide range of roles for the protein has been demonstrated using a series of temperature sensitive lethal mutations in *ABF1* (Rhode *et al.* 1992). At the semi-permissive temperature, CEN-ARS plasmids have an ARS-specific instability, the cell cycle G_1 - S phase transition is not efficient, no activation is seen by an Abf1p UAS, there is reduced RNA synthesis, and DNA synthesis is reduced to 25% of wild type.

The purified protein (Sweder *et al.* 1988, Diffley and Stillman 1988) was found to have a molecular mass of 135kDa, and antibodies raised against purified Abf1p were used to facilitate its cloning (Rhode *et al.* 1989, Halfter *et al.* 1989b). The single copy of *ABF1* is located on chromosome V. It is an essential gene encoding a 731 amino acid protein with a predicted molecular weight of 81.6kDa (Rhode *et al.* 1989). The difference between the predicated weight and the weight as seen in SDS-

PAGE is probably due to post-translational modification, or to the conformation of Abf1p affecting its migration through the gel. The protein has a basic N-terminus and an acidic C-terminus and contains four asparagine clusters along its length. The DNA binding activity of Abf1p requires the N-terminal two thirds of the protein (Halfter *et al.* 1989b). This part of the protein contains a consensus, $CysX_7HisX_3HisX_4CysX_4Cys$, for an atypical metal binding "finger" which seems to be involved in the binding of Abf1p to DNA, since substitution of residues His57 or Cys71 abolishes specific DNA binding.

A consensus binding site for Abf1p has been determined, R**T**C**R**YNNNNNN**ACG** (Dorsman *et al.* 1989), containing two conserved elements, shown in bold type, separated by a variable sequence of a given length which is equivalent to one turn of the DNA helix. An alteration of the spacing between the two conserved elements abolishes DNA binding, as does the substitution of the final G for a C. Binding is reduced by point mutations of the first T and C of the consensus (Dorsman *et al.* 1989. Della Seta *et al.* 1990) and this reduced ability to bind DNA leads to a reduction in the level of transcription from the mutated Abf1p binding site (Della Seta *et al.* 1990). Methylation interference, missing contact analysis and potassium permanganate footprinting demonstrate that the interaction between DNA and Abf1p involves both the coding and non-coding strands (Della Seta *et al.* 1990, McBroom and Sadowski 1994a). Contact between Abf1p and its recognition site in the DNA is not limited to the consensus sequence; ethylation studies (McBroom and Sadowski 1994a) have indicated that phosphates important for binding extend some four or five base pairs on either side of the consensus. Thus not only is Abf1p in close contact with its consensus recognition sequence, but also with regions of flanking DNA.

These extensive interactions may result in the ability of Abf1p to bend DNA. The bend is towards the minor groove at an angle of approximately 120^{o} with its centre about 7bp 5' to the Abf1p consensus sequence (McBroom and Sadowski 1994b). When DNA bends induced by proteolytic fragments of Abf1p are analysed the centres, angles and planes of the bends are different from those induced by the full length protein. McBroom and Sadowski suggest a model of DNA bending which requires bends induced by three regions of Abf1p to combine, making up the full 120^{o}.

The region of ARS1 at which Abf1p binds is associated with bent DNA (Snyder *et al.* 1986, Diffley *et al.* 1994), and removal of this region affects the *in vivo* function of the ARS. The role of Abf1p here may be to prevent transcription from upstream genes affecting ARS function, or to maintain the region in a nucleosome free state

(Snyder *et al.* 1986). Deletions made in ARS1 which move the ARS consensus sequence (ACS) such that it is incorporated into nucleosome DNA reduce the efficiency of the ARS as measured by plasmid stability (Simpson 1990). Five matches to the consensus Abf1p binding site are found in the promoter of *ABF1* (Halfter *et al.* 1989b) which suggests that, like Rap1p, Abf1p may be involved in some kind of autoregulation.

The role of Abf1p in transcription seems to be as a weak activator; oligonucleotides which bind Abf1p have been shown to activate transcription weakly in a *CYC1* reporter plasmid (Brand *et al.* 1987, Halfter *et al.* 1989a, Buchman and Kornberg 1990). The level of activation of transcription was increased if the oligonucleotides were assayed in combination with the T-rich region from the *DED1* promoter (Buchman and Kornberg 1990). This synergism between an Abf1p binding site and a T-rich region has also been noted for ribosomal protein genes (Goncalves *et al.* 1995). The synergistic effect is not unique to T-rich sequences in combination with Abf1p binding sites. Similar effects are seen if the Abf1p site is replaced with one for either Rap1p or Reb1p (Buchman and Kornberg 1990, Goncalves *et al.* 1995), although not by the remaining multifunctional transcription factor, Cpf1p.

Rap1p and Abf1p are both members of the family of multifunctional transcription factors in yeast. Two other members of this family, Reb1p and Cpf1p, are discussed below as their potential interactions with the promoter of the phosphoglycerate kinase gene are investigated in this thesis.

1.7.4 Reb1p

A third multifunctional transcription factor is Reb1p, named after its ribosomal enhancer binding ability (Morrow *et al.* 1989). Reb1p has also been studied as RBP1, factor Y, GRF2 and QBP (Kulkens *et al.* 1989, Fedor *et al.* 1988, Chasman *et al.* 1990, Brandl and Struhl 1990). Reb1p was purified from yeast as a protein with a molecular weight of 125kDa (Morrow *et al.* 1990, Chasman *et al.* 1990), and a consensus binding site was determined, YNNYYACCCG (Chasman *et al.* 1990). A search of yeast DNA sequences with this consensus revealed binding sites for Reb1p in upstream activation sequences, at centromere CEN4 and subtelomeric regions X and Y, as well as the 35S rRNA enhancer. The sequences which flank the Reb1p consensus are important for binding; sequences identical at all positions of the consensus bind Reb1p with different affinities. Two binding sites for Reb1p are found in the promoter of the *REB1* gene (Morrow *et al.* 1990, Morrow *et al.* 1993a), and deletion of these almost completely abolishes transcription (unpublished results, cited in Lang *et al.* 1994).

When the gene for Reb1p was cloned, it was found to be an essential, single copy gene on chromosome II (Ju *et al.* 1990). The open reading frame predicted a molecular weight of 92kDa which differs from the weight of the purified protein. This difference is probably due to the phosphorylation of serine residues after translation (Morrow *et al.* 1990). There were no common DNA binding motifs in the Reb1p sequence, although homology with the oncoprotein Myb, whose DNA binding region includes periodic tryptophan repeats, was detected (Ju *et al.* 1990). The DNA binding activity of Reb1p was localised to the C-terminal half of the protein (Morrow *et al.* 1993a) by N- and C-terminal deletions. When *S. cerevisiae* *REB1* sequence was compared with the homologous sequence from *K. lactis* two regions of homology were found in the Reb1p C-terminus (Morrow *et al.* 1993a). These were separated by 140 amino acids in *S. cerevisiae* and by 40 amino acids in *K. lactis*. This suggested that Reb1p might have a bipartite DNA binding domain. Indeed, internal deletions of the Reb1p C-terminal sequence can be made which do not affect its ability to bind to DNA (Morrow *et al.* 1993a). Since Reb1p binds to DNA as a monomer, and protects a footprint of about 20-25bp, the two DNA binding domains must be brought together when the protein folds.

Whilst the N-terminus of Reb1p can be deleted without affecting DNA binding, this region of the protein appears to have an important role in the biological function of Reb1p since an N-terminal truncation, missing the first 201 amino acids, is not able to support growth of yeast cells whose genomic *REB1* gene has been disrupted (unpublished results, cited in Morrow *et al.* 1993a).

Roles for Reb1p in both transcriptional activation and repression of RNA polymerase II genes, and also as a terminator for RNA polymerase I transcription have been demonstrated. A Reb1p binding site acts as a weak activator of transcription when present in a reporter plasmid, but a synergistic effect is seen when the Reb1p binding site is assayed in conjunction with a T-rich sequence of DNA (Buchman and Kornberg 1990, Chasman *et al.* 1990, Graham and Chambers 1994a). Whilst a Reb1p binding site alone can activate transcription from a minimal promoter, the same site, if placed between a UAS and TATA box, can reduce the level of activated transcription from the UAS (Wang *et al.* 1990). Thus the position of the Reb1p binding site in the promoter could determine its role in transcription. Reb1p appears to play a role in both activation and repression of transcription in the glycolytic gene *ENO1*, which has an upstream repression sequence (URS) in its promoter as well as a UAS. Both the URS and the UAS bind Reb1p (Carmen and Holland 1994).

Whilst Reb1p was first identified as a protein which could bind to the ribosomal enhancer (Morrow *et al.* 1989, Kulkens *et al.* 1989), it is possible that the Reb1p binding site is not an essential part of the enhancer. Although mutations in the Reb1p binding site in the enhancer reduce transcription from an adjacent operon, this is not as great an effect as deleting either the entire enhancer region, or the Reb1p site in the ribosomal promoter (Kulkens *et al.* 1992). The enhancer appears to contain redundant elements; a 45bp region at the 3' end of the 190bp enhancer is sufficient to supply enhancer function, but the Reb1p site, Abf1p site and T-rich region are all able to confer some enhancer function (Morrow *et al.* 1993b).

This minor role for Reb1p at the ribosomal enhancer may be because this site is essential for RNA polymerase I termination (Lang and Reeder 1993, Shultz *et al.* 1993). Thus, deletion of the enhancer Reb1p binding site may reduce transcription from adjacent operons as a consequence of upstream transcription running on into the enhancer and affecting its function. Transcriptionally active ribosomal genes are followed by nucleosome free enhancers (Dammann *et al.* 1995), which could be due to the presence of Reb1p at the terminator.

A Reb1p binding site has been shown to be an essential part of the RNA pol I terminator (Lang and Reeder 1993), where it acts to pause RNA pol I whilst its transcript is released (Lang *et al.* 1994). The Reb1p binding site seems to act as a non-specific pause element, since it is able to stop RNA polymerases I, II and III. However it does not seem to act simply as a block, since the presence of a physical block on the template, such as a bead or *lac* repressor, does not allow transcript release (Lang *et al.* 1994). Whilst the Reb1p DNA binding domain appears to be sufficient for its pausing function in termination, if the orientation of the binding site is reversed (which does not affect Reb1p binding) then termination is abolished (Lang *et al.* 1994). This orientation dependence is a property of terminators in higher eukaryotes.

One role proposed for Reb1p is that of altering chromatin structure to facilitate access of other transcription factors. Reb1p has been implicated in the positioning of nucleosomes in the *GAL1-GAL10* intergenic region (UAS_G) to generate a 160bp nucleosome free region (Fedor *et al.* 1988). The nucleosomes across the *GAL1-GAL10* intergenic region are in positions dependent on binding of Reb1p to UAS_G, even when the promoter is repressed during growth on glucose. On induction of transcription the arrangement of nucleosomes does not appear to alter, although the DNA becomes more accessible to micrococcal nuclease and methidiumpropyl-EDTA iron(II) cleavage (Fedor and Kornberg 1989). However, it has been suggested

that the activation domain of Gal4p is able to displace a nucleosome when transcription from UAS_G is induced (Axelrod *et al.* 1993).

Reb1p has been shown to be necessary for transcriptional activation of a hybrid (*gal-his3*) TATA independent hybrid promoter, consisting of UAS_G upstream of a Gcn4p binding site (Brandl and Struhl 1990). Deletion analysis showed that the Reb1p binding site was necessary for activation, and that it behaved in a distance dependent manner. This requirement for Reb1p suggests a possible role for nucleosome positioning allowing GCN4 to activate a TATA-less promoter.

1.7.5 Cpf1p

Centromere Promoter Factor 1 was first purified as a 57-64kDa protein which bound specifically to the CDE1 sequence in centromeres (Bram and Kornberg 1987). A CDE1 consensus sequence (Hieter *et al.* 1985) was identified in the promoter of the *GAL2* gene, and if the CDE1 site was placed between the *GAL1* UAS and TATA sequence in a *GAL1::his3* fusion, transcription was repressed (Bram and Kornberg 1987). Thus Cpf1p was the first of the group of proteins referred to as multifunctional transcription factors to be identified. Other groups purified specific CDE1 binding proteins with a surprising range of molecular weights, 16kDa, 58kDa, 37 and 64kDa (Cai and Davis 1989, Baker *et al.* 1989, Jiang and Philippsen 1989). However these were found to be degradation products of the same protein when the gene was cloned (Baker and Masison 1990, Cai and Davis 1990, Mellor *et al.* 1990). The gene for Cpf1p, which maps to chromosome X, has an open reading frame of 351 amino acids and encodes a 39.4kDa protein. Unlike the other multifunctional transcription factors, Rap1p, Abf1p and Reb1p, Cpf1p is not encoded by an essential gene. A strain of yeast in which *CPF1* has been disrupted shows phenotypes of slow growth, chromosome loss and methionine auxotrophy. This suggests that Cpf1p has roles in both chromosome maintenance and in transcriptional regulation. However, the CDE1 consensus sequence shows no transcriptional activation of a minimal promoter reporter plasmid and does not show a synergistic effect when assayed in conjunction with a T-rich DNA sequence (Mellor *et al.* 1990, Buchman and Kornberg 1990).

Cpf1p has been shown to bind to its consensus binding site as a dimer (Mellor *et al.* 1990), for which the C-terminus of the protein is necessary. The C-terminus of Cpf1p contains two potential amphipathic helices preceded by a string of basic amino acids, a bHLH motif (Cai and Davis 1990, Mellor *et al.* 1990), and also a dimerisation domain situated C-terminal to the bHLH motif (Dowell *et al.* 1992). The C-terminal 85 amino acids are sufficient for dimerisation and can be replaced

with the leucine zipper dimerisation domain from USF, a human homologue of Cpf1p. The N-terminal 209 amino acids appear to be dispensible for Cpf1p function; although this region of the protein contains two clusters of acidic amino acids, (A1 and A2 in Figure 1.3) these can be deleted without losing methionine prototrophy, and with only a small increase in plasmid loss (Mellor *et al.* 1990).

Binding sites for Cpf1p have been identified at centromeres and in promoters, suggesting that the protein has at least two roles in the yeast cell. Attempts have been made to separate the transcriptional and chromosomal functions of Cpf1p. Mellor *et al.* (1991) found that mutations reducing or abolishing the ability of Cpf1p to bind to DNA did not affect methionine prototrophy but did cause increased chromosome loss, whilst PCR random mutagenesis of *CPF1* (Foreman and Davis 1993) produced two classes of mutations, those which were more compromised for transcriptional activation and those more compromised for chromosome loss. However, Masison *et al.* (1993) found that there was a correlation between the ability of mutant Cpf1p to support methionine independent growth and rescue chromosome loss and its affinity for binding the CDE1 consensus.

That a *cpf1* strain should be a methionine auxotroph suggests a role for Cpf1p in the regulation of methionine biosynthetic genes. Two CDE1 sites were identified in the promoter of *MET25* and shown to be required for a high level of activated transcription (Thomas *et al.* 1989), and Cpf1p is also required for the transcription of *MET16* (Thomas *et al.* 1992). In these cases Cpf1p does not appear to behave as a conventional transcriptional activator, since it seems that the CDE1 site is required rather than of the protein itself (Kent *et al.* 1994). Whilst Thomas *et al.* (1992) showed by Northern blotting in CPF1 and *cpf1* strains that neither *MET16* nor *MET25* was transcribed in the absence of Cpf1p, Kent *et al.* (1994) used an RNase protection assay to show that both *MET16* and *MET25* are expressed in a *cpf1* strain. This anomaly can been explained by the fact that the conditions of growth used by the two groups were different, and expression of *MET16* and *MET25* is more complex than just requiring Cpf1p for activation. Both Met4p and Gcn4p play roles in the induction of these genes in reponse to methionine starvation and to general amino acid starvation (Thomas *et al.* 1992, O'Connell *et al.* 1995).

One way in which Cpf1p might affect transcription is through the modulation of chromatin proteins, such as Spt21p which regulates histone biosynthesis (McKenzie *et al.* 1993). A screen to find high copy number suppressors of methionine auxotrophy in *cpf1* yeast isolated a DNA sequence containing the 3'd element of Ty 1-17. This sequence can be used as a promoter by genes such as *SPT21* and *RPD1*. When double disruptions were made *eg cpf1/spt21,* the double mutants were found

to grow without methionine. That the strains could not complement the centromere defects of *cpf1* strains again suggests distinct functions for regions of Cpf1p.

The family of multifunctional transcription factors in yeast contains Rap1p, Reb1p, Abf1p and Cpf1p (Figure 1.3). These proteins share a number of features which separate them from other yeast transcription factors. They are all abundant within the cell, as might be expected from the fact that they each have many binding sites. All four proteins have roles in the regulation of transcription as well as roles influencing the structure of chromatin, and the importance of these proteins to *Saccharomyces cerevisiae* is demonstrated by the fact that all except for Cpf1p are encoded by essential genes.

The homologues of Abf1p, Reb1p and Cpf1p have been cloned from *Kluyveromyces lactis* by complementation (Goncalves *et al.* 1992, Morrow *et al.* 1993a, Mulder *et al.* 1994) and Rap1p has been cloned in the same organism by homology to a subfragment of *S. cerevisiae RAP1* (Larson *et al.* 1994). *K. lactis* diverged from *S. cerevisiae* about 10^8 years ago, and comparisons of the protein sequences from both species can help to identify regions of functional importance. Abf1p, Reb1p and Cpf1p from *K. lactis* can functionally complement *S. cerevisiae,* but *K. lactis* Rap1p is unable to complement *S. cerevisiae rap1^{ts}* strains. This may be because the N-terminal truncation of *K. lactis* Rap1p has resulted in the loss of domains required for interacting with *S. cerevisiae* proteins, or because the *K. lactis* protein does not have an activation domain. Interestingly, *K. lactis CPF1* is an essential gene. It is possible that in *S. cerevisiae*, other proteins have evolved to interact with Cpf1p and overcome the lethal effects of the gene disruption. Thus studying these proteins in different yeast species may provide insights into domains to which no function has yet been assigned.

1.7.6 yATF

The mammalian ATF/CREB family of proteins (Hai *et al.* 1988) bind an optimal binding site TGACGTCA, known as the CRE (cAMP Response Element; Lin and Green 1988). Each of the two halves of this palindromic sequence are likely to be bound by a protein monomer (Sellers *et al.* 1990). A yeast protein with the same binding specifivity as ATF was identified in yeast (Lin and Green 1989, Jones and Jones 1989), and its binding site was shown to have the ability to activate transcription. A potential yeast Activating Transcription Factor (yATF) binding site has been identified in the *PGK* promoter (Lin and Green 1989). This potential yATF binding site is a poor match to the CRE consensus as it contains only one half of the palindrome. However a member of the yeast ATF/CREB family has been shown to

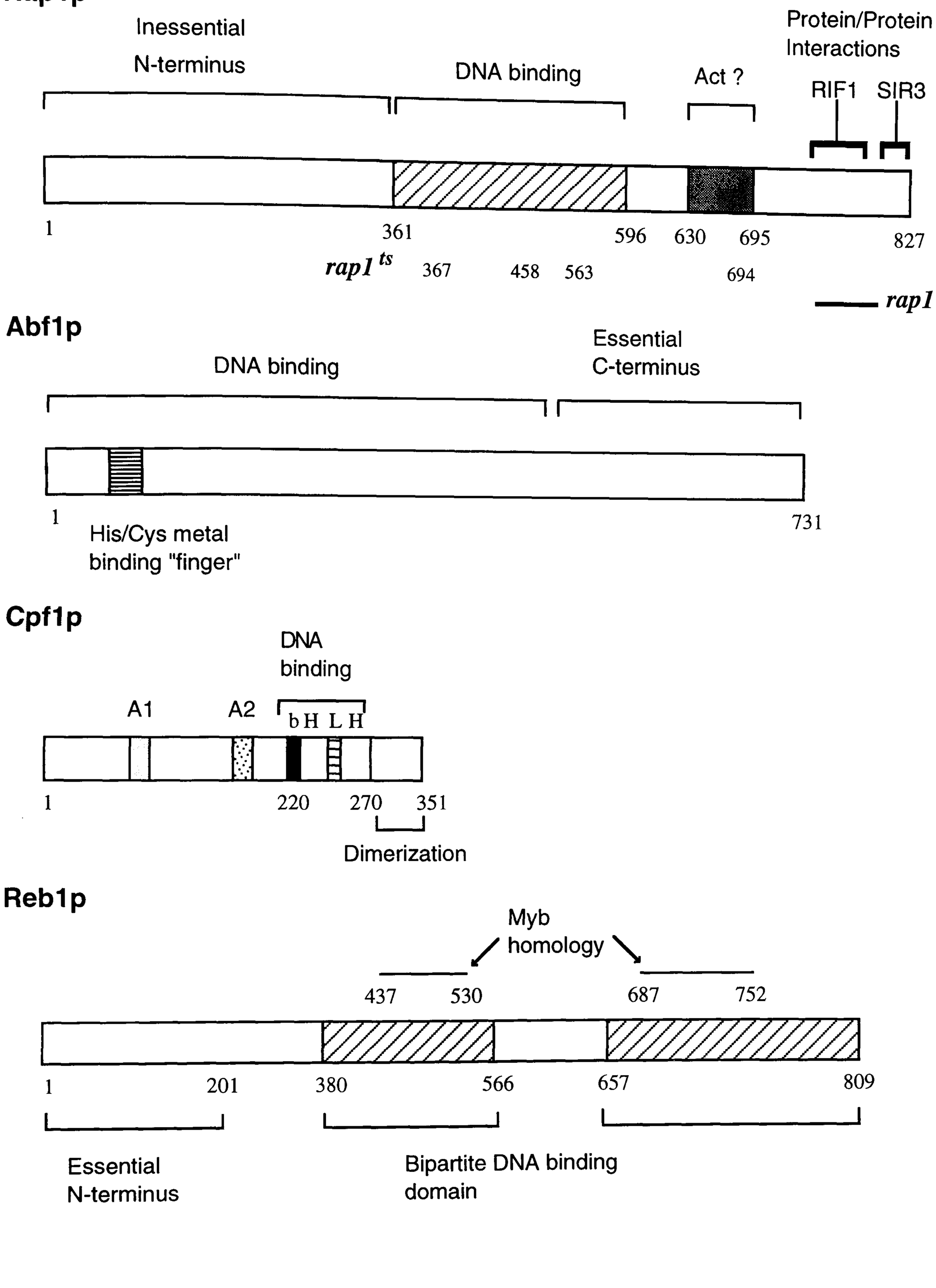


Figure 1.3 A summary of the structures of the family of yeast multifunctional transcription factors. The proteins are described in more detail in the text. Act ?: the position of the potential activation domain in Rap1p; A1 and A2: clusters of acidic amino acids.

bind to such a sequence, containing only one half-site (Nehlin *et al.* 1992), and so further investigation of the potential *PGK* site was undertaken in this work.

1.8 Aims of thesis

At the start of this work the *PGK* promoter was known to consist of a basal promoter, and UAS which bound the transcription factors Rap1p, Abf1p and Gcr1p. A potential yATF binding site had also been identified at the 3' end of the UAS. Deletions of transcription factor binding sites had been made from *PGK* constructs on multicopy plasmids in order to investigate roles for the transcription factors known to bind to the promoter. Since these transcription factors were identified after deletion analysis of the promoter to find regions important for the activation of transcription, one aim of this work was to carry out a systematic search of the promoter with transcription factor consensus binding sites to see whether any other factors interacted at this locus. If so, *in vitro* binding to the new sites would be investigated.

The potential yATF binding site would be inserted into a minimal promoter plasmid to look for the ability to activate transcription, and a deletion of the potential binding site would be made from a multicopy plasmid. Finally, the deletions of transcription factor binding sites which had been made from multicopy plasmid contructs of *PGK* would be introduced into the chromosomal *PGK* locus by homologous recombination. This would allow the roles of the transcription factors to be examined in the absence of artifacts such as high plasmid copy-number, interference from plasmid sequences and plasmid chromatin structure.

Chapter 2

Materials and Methods

Techniques employed were those commonly used in molecular biology and were used, with occasional alterations, as described in *Molecular cloning: a laboratory manual* (Sambrook J., Fritsch E. F. and Maniatis T. (1989 Second ed.) Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY.).

2.1 Strains used in this work

E. coli

MC1061: *F⁻ araD139 Δ(ara-leu)7696 Δ(lac)X74 galU galK hsdR2 ($r_K^- m_K^+$) mcrB1 rpsL (Str^r)*

BL21(DE3): *F⁻ ompT hsdS_B (r_B^- m_B^-)*, with a λ prophage carrying the T7 RNA polymerase gene (Studier *et al*. 1990)

Saccharomyces cerevisiae

DBY745: *α ade1-100 leu2-3 leu2-112 ura3-52*

W303-1A: *a ade2-1 trp1-1 can1-100 leu2-3, 112 his3-11, 15 ura3-52*

R884-1C: *a ade2-1 trp1-1 can1-100 leu2-3, 112 his3-11, 15 ura3-52 gal11-313*

YAG93: *α ade1-100 leu2-3 leu2-112 ura3-52 cpf1 Δ10-351*

YLP1: *α ade1-100 leu2-3 leu2-112 ura3-52 PGK Δ-463/-475*

YLP2: *α ade1-100 leu2-3 leu2-112 ura3-52 PGK Δ-503/-516*

YLP3: *α ade1-100 leu2-3 leu2-112 ura3-52 PGK Δ-552/-562*

YLP4: *α ade1-100 leu2-3 leu2-112 ura3-52 PGK Δ-427/-415*

YDS413: *α ade 2-1 can1-100 his3-11 his3-15 leu2-3 leu2-112 trp1-1 ura3-1 rap1-4*

YDS410: *α ade 2-1 can1-100 his3-11 his3-15 leu2-3 leu2-112 trp1-1 ura3-1 rap1-5*

YDS409: *α ade 2-1 can1-100 his3-11 his3-15 leu2-3 leu2-112 trp1-1 ura3-1 rap1-2*

2.2 Bacterial Growth Media

LB:1% bacto-tryptone, 0.5% yeast extract, 1% NaCl (+1.5-2% bacto-agar for plates)

LBAp: LB with 50μg/ml ampicillin

2.3 Yeast Growth Media

YPD (complete): 2% bactopeptone, 1% yeast extract, 2% glucose (+2% agar for plates)

SC (minimal): 0.67% yeast nitrogen base with no amino acids, 1% glucose (+2% agar for plates). After autoclaving the SC media was supplemented with a rich amino acid mix (100X stock: 0.2% adenine, 0.2% tryptophan, 0.2% histidine, 0.2% arginine, 0.2% methionine, 0.3% isoleucine, 0.3% lysine, 0.5% phenylalanine, 1% glutamate, 1% aspartate, 1.5% valine, 2% threonine, 4% serine) excluding either leucine or uracil to maintain selective pressure for plasmids.

Buffered SC agar: SC-agar was made up to half of the required volume with water, and then 0.2M sodium phosphate buffer pH7 was added to give the final volume. After autoclaving amino acids were added as required, and finally X-gal was added to give a final concentration of 40µg/ml.

5-FOA SC-agar: After autoclaving, SC-agar was allowed to cool to about 40°C before 0.003% leucine, 0.002% uracil, 0.002% adenine and 1mg/ml 5-FOA were added.

2.4 Growth Conditions

Unless otherwise stated *E. coli* were grown at 37°C in LB or LBAp, and yeast strains were grown at 30°C in either YPD or SC medium.

2.5 Isolation of DNA from *E.coli*

2.5.1 Miniprep: 2ml LBAp was inoculated with a single bacterial colony and grown for 6-7 hours or overnight. 1.5ml cells was harvested at high speed for 2 minutes in a MicroCentaur and resuspended in 100µl GTE (50mM glucose, 25mM TrisHCl, 10mM EDTA). 200µl NaOH/SDS (0.2M NaOH, 1% SDS) was added to lyse the cells, followed by 150µl potassium acetate (3M K^+, 5M acetate). Precipitated proteins were removed with a 5 minute high speed spin and 400µl supernatant transferred to a fresh Eppendorf. [If the initial culture had been grown overnight the supernatant was subjected to a phenol/chloroform extraction at this stage.] An equal volume of ethanol was added and after a brief vortex, tubes were spun at high speed for 5 minutes. The pellet of nucleic acid was resuspended in 20µl water.

2.5.2 Midiprep: 40ml LBAp was inoculated and grown overnight to stationary phase. Cells were harvested (4000rpm, 5 minutes) in a bench top centrifuge and resuspended in 2ml GTE. 4ml NaOH/SDS was added and then 3ml potassium acetate. After removing cell debris and precipitated proteins (4000rpm, 5 minutes), 2.5 volumes of ethanol were added to the supernatant to precipitate the nucleic acid

(4000rpm, 10 minutes). The pellet was resuspended in 400µl TE (10mM Tris, 1mM EDTA) to which 5µl RNase (DNase free, 500µg/ml, Boehringer Mannheim) was added, and left at 37°C for 15 minutes. Tubes were then heated to 70°C for 10 minutes with 20µl 10% SDS before two extractions with phenol/chloroform. DNA was precipitated with 1ml ethanol containing 0.3M sodium acetate (EtOH/NaOAc), washed with 70% ethanol, and resuspended at a concentration of 0.5-1µg/µl.

2.5.3 Large Scale Preparation of DNA from a Caesium Chloride Gradient: A 40 ml culture of *E. coli* was treated as described above until the first precipitation of nucleic acid. The pellet was then resuspended in 4.6 ml STE (10mM NaCl, 10mM TrisHCl pH 7.4, 1mM EDTA). Gradients were made by adding 4.5g CsCl and 180µl of 10mg/ml ethidium bromide. Small gradient tubes were filled, balanced in pairs, and sealed before spinning at 40000rpm, 20°C for 20-24 hours in a Centrikon T-2055 ultracentrifuge. After harvesting the supercoiled plasmid band from the gradient, the ethidium bromide was removed by successive extractions with butanol equilibrated with NaCl-saturated STE. The excess salt was diluted out with 4ml water and the DNA was then precipitated with 10ml ethanol (-70°C, 30 minutes). The pellet was washed with 70% ethanol and resuspended at 1µg/µl.

2.5.4 "Wizard" DNA Preps

Miniprep DNA and midiprep DNA were also be prepared using a "Wizard" kit (Promega) following manufacturer's instructions.

2.6 Isolation of DNA from *Saccharomyces cerevisiae*

2.6.1 "Ten minute" chromosomal DNA prep. (Hoffman and Winston 1987)

Yeast cells grown to saturation overnight in 10ml YPD medium were harvested at 4000rpm for 10 minutes and then washed with 500 µl water. The cell pellet was resuspended in 200µl of a solution comprising 2% Triton X-100, 1% SDS, 100mM NaCl, 10mM TrisHCl pH8, 1mM EDTA, and 100µl each of phenol and chloroform was added. After the addition of glass beads to just below the meniscus, cells were vortexed for 2 minutes. Then 200µl TE was added and tubes were spun at high speed for 5 minutes. The aqueous layer was precipitated with 1ml ethanol and the pellet dissolved in 400µl TE to which 3µl RNase (500µg/ml, DNase free, Boehringer)was added. This was left at 37°C for 15 minutes. After an ammonium acetate precipitation (10µl 4M NH_4OAc, 1ml ethanol) the DNA was extracted once with phenol/chloroform and reprecipitated using EtOH containing 0.3M NaOAc.

2.6.2 Copy Number DNA Prep

Used for the preparation of high quality yeast chromosomal DNA, or to prepare both chromosomal DNA and plasmid DNA from yeast for determining plasmid copy number.

DAY 1: Mid-log yeast cells (4-6 x 10^6 cells per ml) from 50ml of culture were resuspended in 700µl solution 1 (0.9M sorbitol, 20mM EDTA, 14mM β-mercaptoethanol) in a 25ml Corex tube. To form spheroplasts, 20µl Lyticase (Sigma, 8000U/ml in 50mM sodium phosphate buffer (pH7)) was added and the cells were left at 37°C for 2-3 hours. To check that spheroplasts had formed, 10µl cells in 100µl of water were examined under a light microscope and their appearance compared with that of 10µl cells in 1M sorbitol; spheroplasts lyse in water. Spheroplasts were spun out at 3000rpm for 5 minutes in a Sorvall centrifuge and resuspended in 700µl TE. To this were added 70µl 500mM EDTA pH8, 30µl 1M TrisHCl pH7.4, 100µl 10% SDS and 2µl diethyl pyrocarbonate (DEPC, Sigma) and the Corex tubes were left in a 65°C waterbath for 30 minutes without their metal caps. After adding 160µl 5M potassium acetate, they were left on ice for 20 minutes and spun at 4000rpm, 5 minutes in a benchtop centrifuge. The supernatant was transferred to a 50ml Falcon tube to which 2ml ethanol was added, and mixed gently. This was spun at 3500rpm for 5 minutes and the pellet was left to resuspend in 600µl TE overnight at room temperature.

DAY 2: 5µl RNase was added to the resuspended nucleic acid and left at 37°C for 30 minutes. The DNA was then precipitated with 60µl 3M NaOAc and 375µl propan-2-ol. The pellet was then washed with 900µl propan-2-ol and left to resupend in 200µl TE overnight at room temperature.

DAY 3: The DNA was extracted twice with phenol/chloroform and once with chloroform (mixing well each time) before being precipitated with 500µl EtOH/NaOAc. The pellet was washed twice with 70% ethanol and resuspended in 100-200µl water for at least 4 hours.

2.7 Digesting DNA

Colonies of transformed *E. coli* were screened by digesting miniprep DNA with the relevant test restriction enzyme(s). The digests were routinely performed in a 10µl volume containing 3-5µl miniprep DNA, 1µl 10X restriction buffer (supplied with the enzyme by the manufacturer), 0.5-1µl restriction enzyme and 0.5µl RNase. The digests were incubated for about 1 hour at the correct temperature for the enzyme(s). 1µl of loading buffer (50% sucrose, 50mM EDTA, 0.1% Bromophenol blue, 0.01%

xylene cyanol, 0.01% SDS) was added and the fragments generated by the digest were resolved on a 1% agarose gel in TBE (90mM Tris-borate, 2mM EDTA) buffer. The gel was stained with ethidium bromide (5µg/ml) for a few minutes before the DNA fragments were viewed on an ultraviolet transilluminator.

Midiprep DNA and caesium pure DNA was generally digested in a 20µl reaction volume containing 1-3µg DNA, restriction enzyme and its buffer, but digests of total yeast chromosomal DNA were in much larger volumes, 50-100µl, and were generally left to digest overnight.

2.8 Ligations

Vector DNA was prepared by linearising 2µg parental plasmid and then removing 5' phosphate groups by treating the digested plasmid with 1-2µl Calf Intestinal Alkaline Phosphatase (Pharmacia) at 37°C for 30 minutes. After a phenol/chloroform extraction the vector was precipitated and resuspended in about 30-50µl water. Insert DNA was isolated from agarose or polyacrylamide gels.

The ligation reactions were set up with vector:insert ratios of 1:1 and 1:5. Thus, 50ng vector with 50ng insert and 50ng vector with 0.5µg insert. A 20µl reaction also contained 4µl 5X ligase buffer (GIBCO-BRL) and 0.5µl T4 DNA ligase (GIBCO-BRL). The ligations were incubated overnight at 14°C, or for 4-6 hours at room temperature if the insert and vector had sticky ends, before being transformed into *E. coli*.

2.9 Addition of Phosphate Groups to Artificially Synthesised Oligonucleotides

Artificially synthesised oligonucleotides (Table 2.1C) do not have 5' phosphate groups, so before cloning 5' phosphate groups were added in a reaction catalysed by T4 polynucleotide kinase (GIBCO-BRL) before pairs of oligonucleotides were annealed. 1.25µg of each pair of oligonucleotides was put into a 25µl reaction with 2.5µl 10X C buffer (500mM Tris pH7.5, 100mM $MgCl_2$, 10mM DTT), 2.5µl 10mM rATP and 10U T4 kinase. After an hour at 37°C the kinase reaction was heated to 70°C for 10 minutes, and then left to cool to room temperature which allowed the oligonucleotides to anneal together. They could then be used directly in a ligation reaction.

2.10 Transformation of *E. coli*

MC1061 was grown to OD_{600} of 0.6 in 20ml LB medium at 37°C. Cells were harvested at 4000rpm for 5 minutes in a bench top centrifuge. After resuspension in 10ml 50mM $CaCl_2$ on ice for 20 minutes cells were harvested as before and resuspended in 1ml 50mM $CaCl_2$. 100µl of competent cells were added to 20µl of DNA solution and left on ice for 30 minutes before being given a 30 second heat shock at 55°C and being returned to ice for a few minutes. Transformations were spread on LB plates containing ampicillin (50µg/ml) and grown overnight at 37°C.

2.11 One Step Transformation of *Saccharomyces cerevisiae* (Chen *et al.* 1992)

1.5 ml of mid-log phase DBY745 were harvested and resuspended in 100µl One Step Buffer (0.2M lithium acetate, 40% PEG, 100mM DTT). Approximately 1µg of plasmid DNA was added to the resuspended cells which were then incubated at 45°C for 30 minutes. During this incubation the cells were mixed gently every 10 minutes to ensure that they stayed in suspension. Transformants were selected at 30°C on SC with no leucine (SC-leu), or SC with no uracil (SC-ura).

2.12 Lithium Acetate Transformation of *S. cerevisiae* (Becker and Guarente 1991)

100 ml of mid-log phase cells were harvested, at 3000rpm for 5 minutes, and resuspended in 5ml TE. The spin was repeated twice, with the cells being resuspended first in 5ml TE/100mM lithium acetate, then in 1ml TE/lithium acetate. The cells were then shaken at 30°C for an hour, before 400µg sonicated salmon sperm carrier DNA was added. 1µg of transforming DNA was added to 100µl cells, and the reaction was incubated at 30°C for 30 minutes. 700µl of a solution containing 35% (w/v) PEG 4000, 100mM lithium acetate/TE, pH 7.5, was added, before a further 50 minutes incubation at 30°C. The cells were then heat shocked for 5 minutes at 42°C, washed with 500µl TE, and resuspended in 100µl TE for plating on selective media.

2.13 Preparation of Total RNA from Yeast

Mid-log cells from a 50ml overnight culture were harvested and washed with 300µl LET buffer (100mM TrisHCl pH7.4, 100mM LiCl, 0.1mM EDTA). The cell pellet was resuspended in 100µl LET buffer and glass beads added to below the meniscus. Then 100µl phenol equilibrated with TNE (10mM TrisHCl pH7.4, 140mM NaCl, 1mM EDTA) was added and cells were lysed by vortexing for 20 seconds. 40µl TNES (TNE+0.1%SDS), 100µl water and 100µl chloroform were added. After

shaking for 45 seconds and spinning at high speed in a MicroCentaur for 5 minutes the aqueous layer was removed; this was re-extracted 3 times with phenol/TNE and chloroform before the RNA was precipitated with 20µl 3Mpotassium acetate and 700µl ethanol at -80°C for 30 minutes. The pellet of RNA was washed with 80% ethanol and resuspended at 0.5-1µg/µl in water. RNA samples were used as quickly as possible after they had been prepared, but if necessary they were stored at -80°C

2.14 Preparing a total protein extract from yeast

Cells from a 50ml overnight culture were harvested (3000rpm, 10 minutes) and washed twice with 2ml 25mM sodium phosphate buffer pH7 (7mM Na_2HPO_4, 5mM NaH_2PO_4; for 500ml 200mM stock 57.7ml 1M Na_2HPO_4 was added to 42.3ml 1M NaH_2PO_4). After resuspending the washed pellet in 300µl 25mM $NaPO_4$ buffer with 1mM PMSF (Phenylmethylsulfonyl Fluoride) glass beads were added to just below the meniscus and tubes were vortexed for 1-2 minutes in 30 second bursts. In between vortexing the samples were kept on ice. After a 30 second spin in a MicroCentaur the supernatant was transfered to a fresh Eppendorf and the beads and cell debris were washed with a further 200µl sodium phosphate buffer/PMSF. Supernatants were pooled and spun at high speed for 10 minutes at 4°C, after which the protein concentration was determined by Biorad Assay and extracts were stored at -20°C.

2.15 Biorad Protein Assay (Bradford 1976)

Standard solutions of Bovine Serum Albumin (BSA) containing 1-30µg of protein were prepared in 800µl water and 200µl of Biorad reagent was added. After shaking, and allowing the colour to develop, the OD_{595} was read in a Unicam SP1800 Spectrophotometer. A standard curve of µg protein against OD_{595} was drawn from which the concentration of unknown protein samples could be determined. 5µl of each yeast total protein extract was diluted with 795µl water and 200µl Biorad reagent was added. Once the colour had developed the OD_{595} was read and the concentration of the original samples was calculated.

2.16 β-Galactosidase Assay (Miller 1972)

5µg of total protein extract was mixed with 150µl 2X β-galactosidase buffer (120mM Na_2HPO_4, 80mM NaH_2PO_4, 2mM $MgCl_2$, 100mM β-mercaptoethanol, 1.33mg/ml o-nitrophenyl β-D-galactopyranoside (ONPG)), in a 300µl volume. After a brief vortex the reactions were incubated at 37°C for 10-30 minutes whilst a yellow colour developed. The reaction was stopped by vortexing with 500µl 1M

Na_2CO_3. The OD_{420} was read and the β-galactosidase activity was calculated and expressed as units per mg protein per minute. [1 OD unit min^{-1} mg^{-1} = 222 nmoles substrate min^{-1} mg $protein^{-1}$]

2.17 *In vitro* protein production

The same reaction conditions were used with SP6, T3 or T7 RNA polymerases. First an RNA transcript was produced from a DNA template by incubating linearized plasmid template with a mix containing rNTPs and RNA polymerase. The transcription reaction was set up at room temperature, using reagents from a kit from Promega, in the following order: 4µl water, 4µl 5X transcription buffer (provided with the enzyme), 2µl DTT, 0.5µl RNasin, 1µl 10mM ATP, 1µl 10mM CTP, 1µl 10mM UTP, 0.5µl 1mM GTP, 2µl 5mM cap analogue (m^7G5'ppp5'G), 2µl (1µg) template DNA, 2µl RNA polymerase. This reaction was incubated at 37°C for 10 minutes before adding 1µl 10mM GTP and continuing the incubation for a further 80 minutes.

The DNA template was removed by the addition of 1µl undiluted Promega RQ1 DNase at 37°C for 15 minutes and after a phenol/chloroform extraction the RNA was precipitated with 10µl 3M sodium acetate and 330µl ethanol. The pellet of RNA was washed with 80% ethanol and resuspended in 12µl water. 2µl of this was run on a 1.5% agarose gel along with some of the DNA template to check that the RNA transcript was not degraded. 1µg of RNA was then used as a template for the *in vitro* translation reaction.

1µl amino acid mix, 3µl ^{35}S methionine, 1µl RNasin, 2µl (1µg) RNA and 8µl water were added to 35µl rabbit reticulocyte lysate (Promega) and incubated at 30°C for 1 hour. Mock translation reactions were prepared at the same time, the components were identical but the reactions were not primed with RNA. The translation reactions were stored at -20°C.

2.18 Partial Purification of Overexpressed Reb1p

Overexpressed Reb1p was partially purifed using a modified version of the protocol described in Morrow *et al.* (1993a). 100ml cultures of the *E. coli* strain which overexpresses Reb1p, and a control strain containing parental plasmid with no Reb1p insert, were grown to mid-log phase (OD_{550} ~0.5) before they were induced with 4mM IPTG. Induced cells were grown for a further 3 hours before they were harvested at 4500rpm for 10 minutes and resuspended in 5ml TE (pH8). After a rapid freeze/thaw the cells were sonicated (3x 15 seconds) on ice. Cell debris was removed with a 15000rpm spin for 15 minutes and resuspended in 500µl TE (pH8),

whilst three volumes (15ml) 40% saturated $(NH_4)_2SO_4$ was added to the supernatant. After a 30 minute incubation on ice the supernatant was spun at 15000rpm for 15 minutes. The 40% pellet was resuspended in 1ml TE pH7.2, and the supernatant was made up to 70% saturated with $(NH_4)_2SO_4$ and left on ice for a further 30 minutes before being spun once more at 15000rpm for 15 minutes. The 70% pellet was resuspended in 500µl TE pH7.2 and the supernatant was discarded.

The high concentration of ammonium sulphate was reduced in the 40% and 70% pellets by spin dialysis through a Microsep (Amicon). After spinning at 6000rpm for 1 hour, a further 500µl TE pH7.2 was added, and the spin continued for another hour. The concentration of the protein fractions was then determined with a Biorad protein assay and 5µg was incubated with labelled DNA fragments in a gel retardation assay.

2.19 Protein Gel

A 7.5% SDS polyacrylamide gel was made up in the Biorad Mini-Protean Apparatus for resolving proteins. The protein gel (2ml 30:0.8 acrylamide:bis-acrylamide, 2ml 1.5M Tris pH8.8, 80µl 10% SDS, 3.92ml water, 70µl 10% APS, 7µl TEMED) was poured to below the depth of the comb and overlayed with propan-2-ol. After the protein gel had set, the propan-2-ol was poured off and a stacking gel (500µl 30:0.8 acrylamide:bis-acrylamide, 750µl 0.5M Tris pH6.8, 30µl 10% SDS, 1.7ml water, 30µl 10% APS, 3µl TEMED) was poured on top. The comb was inserted and the gel left to set.

Protein samples were mixed with 5µl 4X loading buffer (200mM Tris-Cl pH 6.8, 400mM dithiothreitol, 8% SDS, 0.4% bromophenol blue, 40% glycerol) in a 20µl volume and heated to 95°C for 5 minutes. They were then loaded immediately and the gel was run in 1X protein gel running buffer (5X protein gel running buffer: Tris 15g/l, glycine 72g/l, SDS 5g/l). The gel was run at 150V until the bromophenol blue had moved through the stacking gel when it was turned up to 200V and left until the dye had reached the bottom of the gel.

If the gel had been used for the electrophoresis of radioactive samples it was dried onto Whatman paper and exposed to Fuji X-ray film. Otherwise protein bands were visualised by simultaneously fixing the gel in methanol: glacial acetic acid, and staining it with Coomassie Brilliant Blue. The gel was immersed in 100ml Coomassie stain (0.25g Coomassie Brilliant Blue R250 (Sigma) dissolved in 90ml Methanol: water (1:1 v/v), 10ml glacial acetic acid) and left at room temperature for

30 minutes. The stain was removed by soaking the gel in destain (30% methanol, 10% acetic acid) at room temperature for 30 minutes to overnight.

2.20 Gel Retardation Assay

2.20.1 Preparation of Retardation Probes

Probes were end-labelled with [γ-^{32}P] in the presence of T4 polynucleotide kinase. After digesting DNA to release the probe fragment, digest reactions were treated with Calf Intestinal Alkaline Phosphatase (Pharmacia) to remove 5' phosphate groups. The required fragment was then purified from a 1% agarose gel using a Gene Clean kit (Stratech) and the DNA left in 16µl water. To this were added 2µl 10XC buffer (500mM TrisHCl, 100mM $MgCl_2$, 10mM DTT), 1µl T4 kinase and 1µl [γ–^{32}P] ATP (185TBq mmol^{-1}, Amersham). The labelling reaction was then left at 37°C for 1-2 hours.

Before use the labelled probe was separated from the unincorporated [γ–^{32}P] ATP by running the labelling reaction through a Nick Column (Pharmacia) or through a 1ml spun column made from G-50 Sephadex (Pharmacia).

Spun column: A 1ml syringe was plugged with glass wool and filled with Sephadex G-50 (Pharmacia) which had been pre-swollen in TE. It was spun at 3000rpm for 3 minutes to compact the Sephadex. The syringe was topped up with Sephadex and respun as often as necessary to get a bed volume of 1ml. The column was equilibrated with two 100µl aliquots of TE, spinning for 3 minutes each time. Then the reaction volume was made up to 100µl and it was loaded onto the column. The eluate, approximately 100µl, was collected and precipitated before the labelled probe was resuspended in water at a concentration of about 100 counts per 5µl.

2.20.2 Binding Reactions

20µl reactions were set up containing 2µl 10X binding buffer (50% glycerol, 10mM EDTA, 100mM β mercaptoethanol, 250mM TrisHCl pH7.5, 250mM NaCl, 200mM KCl), 2µl 0.5µg/µl poly[dI.dC] (Pharmacia), 5µl (100 counts) labelled probe, protein (total protein extract or *in vitro* translated protein) and water. These were incubated at room temperature for 45 minutes before being resolved on a 5% polyacrylamide gel in 0.5X TBE buffer (45mM Tris borate, 1mM EDTA). After running for 1.5-2 hours at 150V gels were dried and exposed to X-ray film.

2.21 Preparation of probes for Southern and Northern blots

DNA fragments were isolated from 1% agarose gels and labelled either by nick translation or by random oligolabelling depending upon the length of the probe and the sensitivity required. Random oligolabelling is a very efficient method of labelling a piece of DNA but is not suitable for fragments shorter than 0.5kb.

2.21.1 Nick translation (Rigby *et al.* 1977)

Reagents for nick translation were from a kit supplied by GIBCO-BRL. To 14µl DNA were added 2µl solution A2 (0.2mM each of dATP, dGTP, dTTP), 2µl solution C (0.4U/µl DNA polymeraseI, 40pg/µl DNase I) and 1µl [α-^{32}P] dCTP (Amersham) Reactions were left at 15°C for 1-5 hours since incorporation of labelled nucleotide increases linearly with time. Nick translated probes were separated from unincorporated nucleotide using a Nick column (Pharmacia). The eluate from the Nick column was used directly to probe filters.

2.21.2 Random oligolabelling (Feinberg and Vogelstein 1983)

Probe DNA in 21µl water was boiled for 3 minutes and cooled on ice for 2 minutes. Then 6µl 5X oligo labelling buffer (25mM $MgCl_2$, 120mM Tris pH8, 500mM PIPES pH6.5, 10mM DTT, 0.25mM dATP, 0.25mM dTTP, 0.25mM dGTP, 25U/ml random hexanucleotide primers) was added with 1µl 10mg/ml BSA, 1µl Klenow (BRL) and 1µl [α–^{32}P] dCTP. The labelling reaction was left at 37°C for 1-2 hours. The labelled probe was precipitated with 100µl EtOH/NaOAc at -80°C for 30 minutes before being resuspended in 400µl TE.

2.22 Southern Blots

DNA samples were run through 1% agarose gels made with TAE buffer (40mM Tris-acetate, 1mM EDTA) either for 5-6 hours at 80V or overnight at 30-40V. The gel was then acid depurinated by washing in 0.25M HCl for 15 minutes. The DNA on the gel was denatured in 0.2M NaOH and 1.5M NaCl for 20 minutes before the gel was washed twice for 15 minutes with neutralization solution (1.5M NaCl, 0.5M TrisHCl, 1mM EDTA). The washed gel was placed upside down on a platform covered with 3MM Whatman paper with a wick into 20X SSC (3M NaCl, 0.3M sodium citrate) and surrounded with parafilm. DNA transfer was onto a nitrocellulose filter with 0.45µm pores (Sigma) which was covered with 2 pieces of 3MM Whatman and then a stack of paper towels. A weight of 0.5-1kg was placed on top of the towels and the blot was left for 6 hours or overnight for transfer of the

DNA to occur. The DNA was fixed by baking the filter in a 80°C vacuum oven for 2 hours. The filter was then stored at room temperature until required.

Before probing a Southern filter it was washed in pre-hybridization buffer (6X SSC, 5X Denhardt's solution (50ml of 100X Denhardt's contains 1g Ficoll, 1g polyvinylpyrrolidine and 1g BSA), 0.5% SDS) to which denatured salmon sperm DNA had been added, to give a concentration of 50µg/ml, at 65°C for at least an hour. The washing was carried out in hybridization bottles in a Hybaid hybridization oven.

After the probe had been denatured (100°C, 5 minutes) it was added to the pre-hybridization buffer making sure that the concentrated probe did not touch the filter. Hybridization took place at 65°C overnight. Non-specific binding of the probe to the filter was reduced by washing the filter at 65°C in 2X SSC, 0.1% SDS twice for 15 minutes and then increasing the stringency of the washes depending upon how homologous the probe was to target DNA. The filter was then covered in Saran wrap and exposed either to X-ray film or to a screen for the Molecular Dynamics Phosphorimager 425.

2.23 Northern Blots

For a 100ml gel, 1.25g agarose was melted in 87ml water and 10ml 10X MOPS/EDTA (200mM MOPS, 50mM NaOAc, 10mM EDTA) was added. When this had cooled to 50°C 5.1ml 37% formldehyde was added and the gel was poured immediately. It was left to set for at least 1 hour before use.

For each lane 10-20µg total RNA was made up to about 20µl with water and 15µl Northern sample buffer (750µl formamide, 150µl 10X MOPS/EDTA, 240µl 37% formaldehyde, 100µl water, 80µl 1%w/v bromophenol blue) was added. The samples were incubated at 65°C for 15 minutes and then loaded onto the gel. Gels were run in 1X MOPS/EDTA buffer at 30V overnight.

Unused pieces of gel were cut off and the gel soaked in 0.05M NaOH, 10X SSC for 15 minutes, and then twice in 10X SSC for 15 minutes. The washed gel was placed upside down on a platform covered with 3MM Whatman paper with a wick into 10X SSC. The gel was surrounded with Parafilm and a piece of 0.2µm nitrocellulose filter (Schleicher & Schuell), cut to size, was placed on top. This was overlayed with 2 pieces of 3MM Whatman, a stack of paper towels and a weight of 0.5-1kg. The transfer was left overnight. After dismantling the blot, RNA was fixed to the

nitrocellulose filter by baking in an 80°C vacuum oven for 2 hours. The filter could then be stored at room temperature until required.

Northern filters were pre-hybridized at 42°C for at least an hour in 5X SSPE (20X SSPE: 3M NaCl, 0.2M NaH_2PO_4, 20mM EDTA), 50% formamide, 5X Denhardt's solution and 0.5% SDS to which denatured salmon sperm DNA had been added to a final concentration of 100µg/ml. The pre-hybridisation was carried out in hybridisation bottles in a Hybaid oven.

The denatured probe was added to the pre-hybridization buffer and left to hybridize overnight at 42°C. Filters were washed at 42°C, twice in 2X SSPE, 0.1% SDS for 15 minutes and then in 1X SSPE, 0.1% SDS for 30 minutes. They were then covered with Saran wrap and exposed either to X-ray film or to screens for the Molecular Dynamics Phosphorimager 425.

2.24 Removal of Hybridised Probe From Nitrocellulose Filters

Probes which had been hybridised to nitrocellulose filters were removed so that the filter could be hybridised to a second probe. The first probe was stripped by boiling the filter in 0.1% SDS, 1mM EDTA for twenty minutes. To ensure that the filter was clean it was exposed to X-ray film or to phosphorimager screens. A stripped filter was washed in 2X SSC or 2X SSPE before the pre-hybridisation procedure was repeated.

2.25 Determination of Plasmid Copy Number

5-10µg of copy number DNA was digested with about 40U *EcoR1* in a 100µl reaction volume overnight at 37°C to ensure that the chromosomal DNA was cut to completion. Digests were then precipitated with EtOH/NaOAc and resuspended in 20µl water before being electrophoresed on a 1% agarose/TAE gel and transfered to nitrocellulose. The filter was probed firstly with a plasmid specific probe, and secondly with a ribosomal probe (Petes *et al.* 1978) to recognise chromosomal DNA. The ratio of plasmid to chromosomal DNA as indicated by these probes gave an indication of the plasmid copy number. In some cases the autorads were examined by eye to look for gross differences in ratio, but after the arrival of the Molecular Dynamics PhosphorImager 425, a more accurate estimation could be achieved by scanning the filters.

2.26 Polymerase Chain Reaction (PCR)

2.26.1 Primers

Oligonucleotide primers (see Table 2.1A) were synthesised by the Biopolymer Synthesis and Analysis Unit (University of Nottingham). 250µl of aqueous oligonucleotide solution was precipitated with 700µl EtOH/NaOAc and washed with 75% ethanol before being resuspended in 50µl water. A 1 in 400 dilution of this oligonucleotide solution was made and the OD_{260} of this was determined. The concentration of oligonucleotide was calculated, using A_{260} of 1 as equivalent to a concentration of 30µg oligonucleotide per ml, and it was then diluted as required to give a 10µM stock solution..

In general a 25µl PCR reaction was set up containing 100ng DNA template, 2.5µl 10µM primer A, 2.5µl 10µM primer B, 2.5µl 10X PCR buffer (500mM KCl, 100mM Tris pH8.8, 25mM $MgCl_2$, 2mM dNTPs), and 0.25µl *Taq* polymerase (Boehringer). The reaction was overlayed with mineral oil and subjected to about 30 repeated cycles of DNA denaturation, primer annealing and product extension in a Techne PHC-3 thermal cycler.

Standard programme:	Denature	94°C 40 seconds
	Anneal	55°C 1 minute
	Extend	72°C 2 minutes 30 seconds

The exact conditions for the PCR programme depended on variables such as the melting temperature of the primers, and the predicted size of the amplified product. The theoretical melting temperature of a primer was calculated by allowing 4°C for each G/C pair and 2°C for each A/T pair. If both primers had high melting temperatures then the annealing temperature in the PCR reaction was high, but if there were any mismatches between the primer sequence and the target DNA sequence then this annealing temperature was lowered. The extension time was determined by assuming that the polymerase synthesized DNA at a rate of about 1kb per minute.

In some instances when the primers were not totally homologous to the DNA template, spurious products were amplified due to the annealing of the primers to non specific sites on the template. To reduce the amount of non-specific product a "touchdown" PCR programme was used (Don *et al.* 1991). This starts with an initial annealing temperature which is higher than the expected melting temperature of the primers and decreases by 2°C every second cycle until the desired annealing

temperature is reached. Since the amount of product in a PCR reaction increases exponentially, *ie* after every cycle there is twice as much product as was present in the previous cycle, a difference in melting temperature between the correct and incorrect annealings will give a two-fold advantage per cycle. Thus the desired product can be amplified preferentially.

A touchdown PCR programme

94°C 40 seconds	2 cycles annealing at 70°C		72°C 2 minutes 30 seconds
" "	" "	68°C	" "
" "	" "	66°C	" "
" "	" "	64°C	" "
" "	" "	62°C	" "
" "	" "	60°C	" "
" "	" "	58°C	" "
" "	" "	56°C	" "
94°C 40 seconds	20 cycles annealing at 55°C		72°C 2 minutes 30 seconds

After a PCR programme had been completed, 5µl of each reaction was run on a 1% agarose gel to check that a product of the right size had been amplified. PCR products could generally be digested with no cleaning of the reaction, but for cloning into T-vector (Promega) the reaction was extracted once with phenol/chloroform and then precipitated.

2.26.2 Cloning of PCR Products

If the PCR primers had suitable restriction enzyme sites within them the amplified product was cleaved and cloned directly into a vector with compatible ends. However, restriction enzyme sites within a few bases of the end of a DNA fragment are not always recognised efficiently, often requiring long incubation times for more than 50% of template to be cut. In such cases it is helpful to utilise the template independent addition of a single adenosine residue to the 3' end of a PCR product by some polymerases.

pGEM-T vector from Promega has been cut within its polylinker with *EcoRV*, and a 3' terminal thymidine added to both ends making it suitable for the direct cloning of PCR products. This can overcome the difficulty of incorporating restriction enzyme sites into PCR primers to facilitate later cloning of the product. Such restriction sites

within the primers can still be utilised to excise the PCR product from the T-vector, or alternatively restriction sites within the T-vector polylinker can be used.

2.27 DNA Sequencing

2.27.1 Sequencing Midiprep or Caesium Pure DNA

Chain termination DNA sequencing based on the method of Sanger *et al.* (1977) was carried out using a Sequenase kit (United States Biochemical). Double stranded DNA was denatured using alkaline denaturation, the sequencing primer (Table 2.1B) was annealed and extension in the presence of [α-^{35}S] dATP and dideoxynucleotides allowed to occur.

5µl 10X denaturing solution (2M NaOH, 2mM EDTA) was added to 5µg of DNA in a total volume of 50µl water and heated at 37°C for 15 minutes. The denatured DNA was precipitated with 5µl 3M NaOAc pH4.5-5.5 and 150µl ethanol at -70°C for 30 minutes. The pellet was washed with 70% ethanol and resuspended in 2µl sequenase buffer, 2µl 10µM sequencing primer and 6µl water. The primer was annealed by heating the reaction to 65°C for 2 minutes and then cooling to below 35°C before chilling on ice. 1µl 0.1M DTT, 2µl 1:5 diluted labelling mix, 2µl 1:8 diluted sequenase and 0.5µl [^{35}S] dATP were added to the annealed DNA mix and this was left at room temperature for 5 minutes to allow incorporation of the dATP label into the extended product. The reactions were terminated by transferring 3.5µl of the labelling reaction to each of 2.5µl ddGTP, ddATP, ddTTP and ddCTP and incubated at 37°C for a further 5 minutes. This reaction was stopped by the addition of 4µl Stop Solution and samples were heated to 80°C immediately before being loaded on an 8% sequencing gel.

For an 8% sequencing gel 25.2g urea was dissolved in 12ml 38:2 acrylamide:bisacrylamide and 6ml 10X TBE and the volume made up to 60ml. The acrylamide was polymerised with 500µl 10% APS and 50µl TEMED. Sequencing gels were run at a constant 30-35W, then dried down and exposed to X-ray film.

2.27.2 Sequencing of Miniprep DNA

Miniprep DNA from cultures which had been grown for 6-7 hours was sequenced directly by adding 1µl 2M NaOH to 9µl miniprep DNA and denaturing at 37°C for 15 minutes. Then 1.5µl 10µM primer, 3µl 3M KOAc and 75µl cold ethanol were added and the DNA/primer mix left to precipitate at -70°C for 30 minutes. The pellet was washed with 75% ethanol, resuspended in 10µl 1X sequenase buffer and left at

37°C for 20 minutes to anneal the primer to the DNA template. The sequencing reactions were then treated as described above.

2.28 "Pop-In/Pop-Out" Homologous Recombination (Scherer and Davis 1979)

Integrating plasmids, pAJ105, pAJ107 and pAJ108, containing *PGK* sequence with deletions in the promoter region, were transformed into DBY745 using the One Step method of transformation. After screening with PCR, to ensure that the inserted gene had not undergone gene conversion, *URA*$^+$ transformants were selected and grown at 30°C overnight in 10ml YPD before the loss of plasmid sequences was selected for by growing the *URA*$^+$ transformants on SC medium containing 1mg/ml 5-fluoroorotic acid (5-FOA). The final strains were confimed by a second PCR screen and Southern blotting.

2.29 Recovery of Plasmids from Yeast (Strathern and Higgins 1991)

1ml of cells from an overnight yeast culture were harvested and resuspended in 200µl TE containing 100mM NaCl, 0.1% SDS. Sterile glass beads were added to just below the meniscus, and the Eppendorf vortexed for 1 minute. The vortexing was repeated, after addition of 200µl phenol. Tubes were then spun in a microcentrifuge for 2 minutes, and the aqueous layer was extracted with a further 200µl phenol. The aqueous layer was then treated with Glassmilk (Geneclean, Stratech), according to the kit instructions, before the DNA was precipitated with EtOH/NaOAc for 30 minutes at -80°C. The DNA pellet was washed with 80% ethanol, and resuspended in 50µl TE. 2.5µl of this was used for each subsequent *E.coli* transformation.

Table 2.1 The sequences of oligonucleotides used in this work. The positions of PCR primers and sequencing primers are indicated relative to the +1 nucleotide at the start of the coding region. The Trp+ primer is from a region upstream of the *TRP1* promoter. The oligonucleotides in Table 2.1C have the transcription factor binding site underlined.

A: PCR Primers

Primer	Sequence (5' to 3')	Gene	Position
521A	GTGTGACGGATCCGGAAGCTGC	*PGK*	-415 / -436
521B	GCTTTCTAACAGATCTATCC	*PGK*	-1483 / -1464
533A	CCAATTTCGGGATCCAACAAGGTCC	*PGK*	-429 / -405
533B	CCGCATTAAAGCTGATCAGAAACGCAG	*PGK*	+1545 / +1519
PGK	CCGTCCAATGGGACGTTG	*PGK*	+92 / +75
PGKC2	TTGATGTTGGATCCATAAAGCACG	*PGK*	-565 / -542
PGKD	CTTTATGAGGGGATCCTCAATTCAAG	*PGK*	-546 / -571
REB1A	TATAGGTCGACCAATATGCC	*REB1*	-15 / +5
REB1B	TTTTCCGGATCCAATTTTCTG	*REB1*	+1440 / +1420
YCR11cA	GGATCGTCCCATAAGAGCAC	YCR11c	+407 / +426
YCR11cB	GGTACAACACGTAGCGGG	YCR11c	+2943 / +2926
U/S1	AGTACGAGATCTAGAAGGGGCAATATG	*RAP1*	-435 / -409
U/S2	ACTATAAGATCTTGCCCGGGTGGCGGCAGA	*RAP1*	+1869 / +1840

Table 2.1 continued

B: Sequencing Primers.

Primer	Sequence (5' to 3')	Gene	Position
2501⊕	TAGACCCAAGAGGCCTG	*RAP1*	+1797 to +1813
Trp+	TTACGCCGGAGCTCCT		
T ' R	AAGCTCGAGCATTGACC	*PGK*	-116 to -132

C: Transcription factor binding site oligonucleotides.

Oligonucleotide	Sequence (5' to 3')
yATF1	GATCCATTTCGTCACACTGCA
yATF2	GTAAAGCAGTGTGACGTCTAG
CRE1	GATCCTGACGTCACTGCA
CRE2	GATCTGCAGTGACGTCAG

Chapter 3

Identification of New Transcription Factor Binding Sites in the *PGK* Promoter

3.1 Introduction

The phosphoglycerate kinase promoter contains binding sites for Rap1p, Gcr1p and Abf1p (Chambers *et al.* 1989, Chambers *et al.* 1990, Henry *et al.* 1994). These were identified in experiments which deleted regions of the promoter to see which ones were necessary for a high level of transcriptional activation. In order to see whether these were the only transcription factors involved in the regulation of transcription of *PGK*, or whether other sequences were necessary for functions not directly related to activation, the sequence of the promoter (see Figure 3.1) was subjected to a search using the sequences for consensus binding sites of various yeast transcription factors. The consensus sequences chosen included those of transcription factors known to bind to other yeast glycolytic promoters *eg* Reb1p; multifunctional yeast transcription factors *eg* Mcm1p and Cpf1p; transcription factors involved in glucose metabolism *eg* Mig1p; and yATF for which a potential transcription factor binding site has already been identified at the 3' of the *PGK* UAS (Lin and Green 1989). Also included were transcription factors whose binding specificities are related to transcription factors mentioned above *eg* Pho4p and yAP1. Finally the *PGK* sequence was compared with the consensus binding sites of Rap1p, Abf1p and Gcr1p to see whether there were other possible matches for these transcription factors.

3.2 Results

The results of the search of the *PGK* promoter sequence, with transcription factor binding site consensus sequences, are shown in Table 3.1. Further possible matches to the Rap1p and Abf1p consensus sequences were identified, as were potential binding sites for yATF, Cpf1p, Pho4p and Reb1p. With the exception of the potential yATF binding site, all of the new matches were found upstream of the previously defined UAS.

3.2.1 Does Abf1p bind upstream of the UAS ?

The two new potential transcription factor binding sites for Abf1p in the *PGK* promoter are at -704 to -692, and -554 to -542. In order to test whether either of these sites could bind Abf1p *in vitro*, two fragments were isolated from pKV501, a

TTAATTTTTTTTTCTTTCCTCTTTTTATTAACCTTAATTTTTATTTTAGATTC

CTGACTTCAACTCAAGACGCACAGATATTATAACATCTGCATAATAGGCA

TTTGCAAGAATTA**CTCGTGAGTAAGG**AAAGAGTGAGGAACTATCGCAT

ACCTGCATTTAAAGATGCCGATTTGGGCGCGAATCCTTTATTTTG**GCTTC**

ACCCTCATACTATTATCAGGGCCAGAAAAAGGAAGTGTTTCCCTCCTTC

TTGAATTG**ATGTTACCCTCATAAAGCACGTGG**CCTCTTATCGAGAA

AGAAATTA**CCGTCGCTCGTGAT**TTGTTTGCAAAAAGAACAA

AACTGAAA**AAACCCAGACACG**CTCGA**CTTCC**TGTCTTCCTAT

TGATTGCAG**CTTCC**AAT**TTCGTCA**CACAACAAGGTCCTAGCG

ACGGCTCACAGGTTTTGTAACAAGCAATCGAAGGTTCTGGAATGGCGGG

AAAGGGTTTAGTACCACATGCTATGATGCCCACTGTGATCTCCAGAGCAA

AGTTCGTTCGATCGTACTGTTACTCTCTCTCTTTCAAACAGAATTGTCCGA

ATCGTGTGACAACAACAGCCTGTTCTCACACACTCTTTTCTTCTAACCAA

GGGGGTGGTTTAGTTTAGTAGAACCTCGTGAAACTTACATTTACATATAT

ATAAACTTGCATAAATTGGTCAATGCAAGAAATACATATTTGGTCTTTTC

TAATTCGTAGTTTTTCAAGTTATTAGATGCTTTCTTTTTCTCTTTTTTTACAG

ATCATCAAGGAAGTAATTATCTACTTTTTACAACAAATATAAAAACAATG

Figure 3.1 The sequence of the phosphoglycerate kinase promoter from -820 to the methionine start codon, the upstream activation sequence (UAS) is in large type. Indicated, in bold type, are regions containing potential binding sites as identified in a search of the promoter with consensus binding sites for different transcription factors. Also shown are the previously identified binding sites for Rap1p, Abf1p and Gcr1p (underlined, bold type), and elements of the basal promoter (underlined).

Table 3.1 The results of a search to identify matches between transcription factor consensus binding sites and the phosphoglycerate kinase promoter. Co-ordinates are numbered relative to the *PGK* ATG.

Transcription Factor	Consensus Sequence[a]	Ref[b]	Match to *PGK* ?
ABF1	RTCRYNNNNNACG	1	-704/-692, -554/-542, -516/-503
ACR1/SKO1	ATGACGTCAT	2	✖
yAP1	STGACTMA	3	✖
yATF1	GTGACGTMR	4	✖
	KWCGTCA	5	-424/-418
CPF1	RTCACRTG	6	-547/-540
GCR1	CTTCC	7	-457/-453, -449/-445, -432/-428
MIG1	WWWWSYGGGG	8	✖
MCM1	DCCYWWWNNRG	9	✖
PHO4	CACGTG	10	-545/-540
RAP1	RTRCACCCANNCMCC	11	-622/-608, -475/-463
	RMACCCANNCAYY	12	-622/-608, -475/-463
REB1	YNNYYACCCG	13	-623/-614, -561/-552

[a] N: A, T, C or G
R: A or G
Y: T or C
S: C or G
K: G or T
W: A or T
M: A orC
D: A, T or G

[b] References:

1 Dorsman *et al.* 1989
2 Vincent and Struhl 1992
3 Harshman *et al.* 1988
4 Lin and Green 1989
5 Jones and Jones 1989
6 Hieter *et al.* 1985
7 Huie *et al.* 1992
8 Nehlin and Ronne 1990
9 Wynne and Treisman 1992
10 Fisher *et al.* 1991
11 Graham and Chambers 1994b
12 Buchman *et al.* 1988a
13 Chasman *et al.* 1990

clone of *PGK* with a *BamH1* site in a deletion window upstream of the known Abf1p binding site at -516 (Chambers *et al.* 1988).

A *BamH1/Cla1* fragment was isolated from pKV501 and cut with *HinF1* to give a 100bp fragment (Fragment 1, see Figure 3.2) containing one potential Abf1p binding site (-554 to -542), and a 130bp fragment (Fragment 2) containing the second potential Abf1p binding site (-704 to -692). These small fragments were isolated from a polyacrylamide gel and end labelled using [γ–^{32}P] ATP and polynucleotide kinase. The labelled fragments were then incubated with Abf1p in a gel retardation assay. The previously identified Abf1p binding site from the *PGK* UAS (-516 to -503) had been cloned into the polylinker of pSP46, so it was possible to release it as a *BglII/EcoR1* fragment. This fragment, containing a known Abf1p binding site, acted as a positive control and was also isolated from polyacrylamide and end labelled.

The protein sources for the gel retardation assays were a yeast total protein extract (TPE), as described in materials and methods (2.14), and *in vitro* translated (IVT) Abf1 protein. pT3ABF1, a clone of the Abf1p coding region (a gift from J. Diffley), was linearized using a *HindIII* site downstream of the coding region and run-off transcripts were produced from the T3 promoter. The RNA was translated in the presence of ^{35}S methionine so that the quality of the protein could be checked by electrophoresis of the translation reaction through a protein gel, and detection of the radioactive Abf1p on X-ray film (Figure 3.3, lanes 1 and 2). A mock translation reaction, not primed with Abf1p RNA, was set up in parallel. This mock lysate acts as a control for proteins which are already present in the rabbit reticulocyte lysate and could interact with DNA fragments non-specifically during a gel retardation assay, it does not contains a protein corresponding to the radioactive Abf1p (Figure 3.3, lanes 3 and 4).

Binding reactions, between *PGK* promoter fragments 1 and 2 and IVT Abf1p, were allowed to incubate at room temperature for approximately 45 minutes as described in materials and methods, section 2.20.2, before protein/DNA complexes were resolved by electrophoresis on 5% polyacrylamide gels. The gels were then dried down and exposed to X-ray film at -80^{o}C.

These gel retardation assays showed that although the *in vitro* translated protein would recognise the known Abf1p binding site, neither fragment 1 nor fragment 2 from the *PGK* promoter was able form a complex (Figure 3.4, compare Abf1p lysate incubated with the known Abf1p site (+) with the Abf1p lysate incubated with

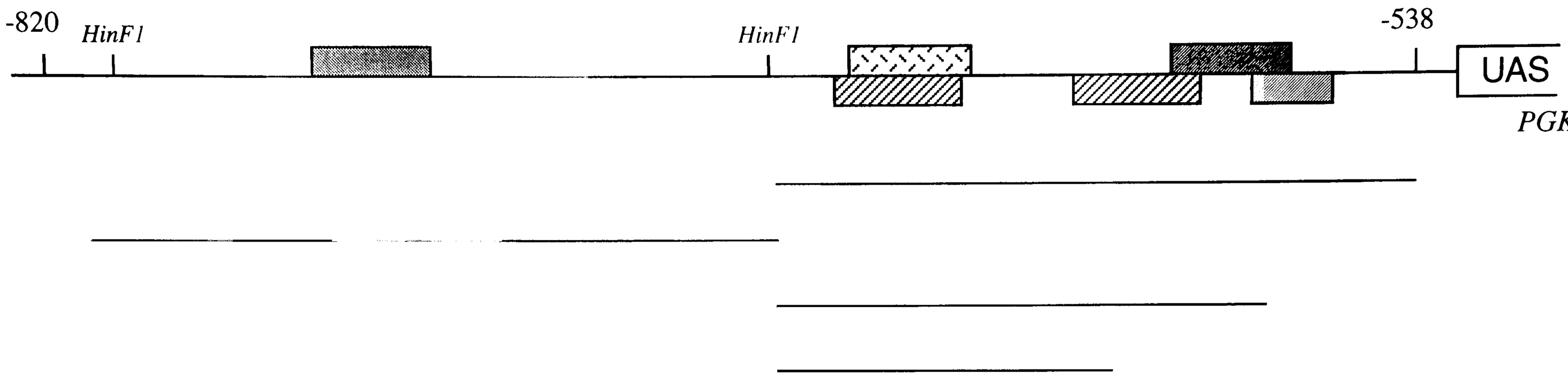


Figure 3.2 Matches found to consensus transcription factor binding sites in the *PGK* promoter upstream of the previously defined UAS. DNA fragments 1, 2, 1A and 1B were isolated from the promoter by restriction enzyme digest and used in gel retardation assays to identify novel interactions between the *PGK* promoter and *S. cerevisiae* transcription factors.

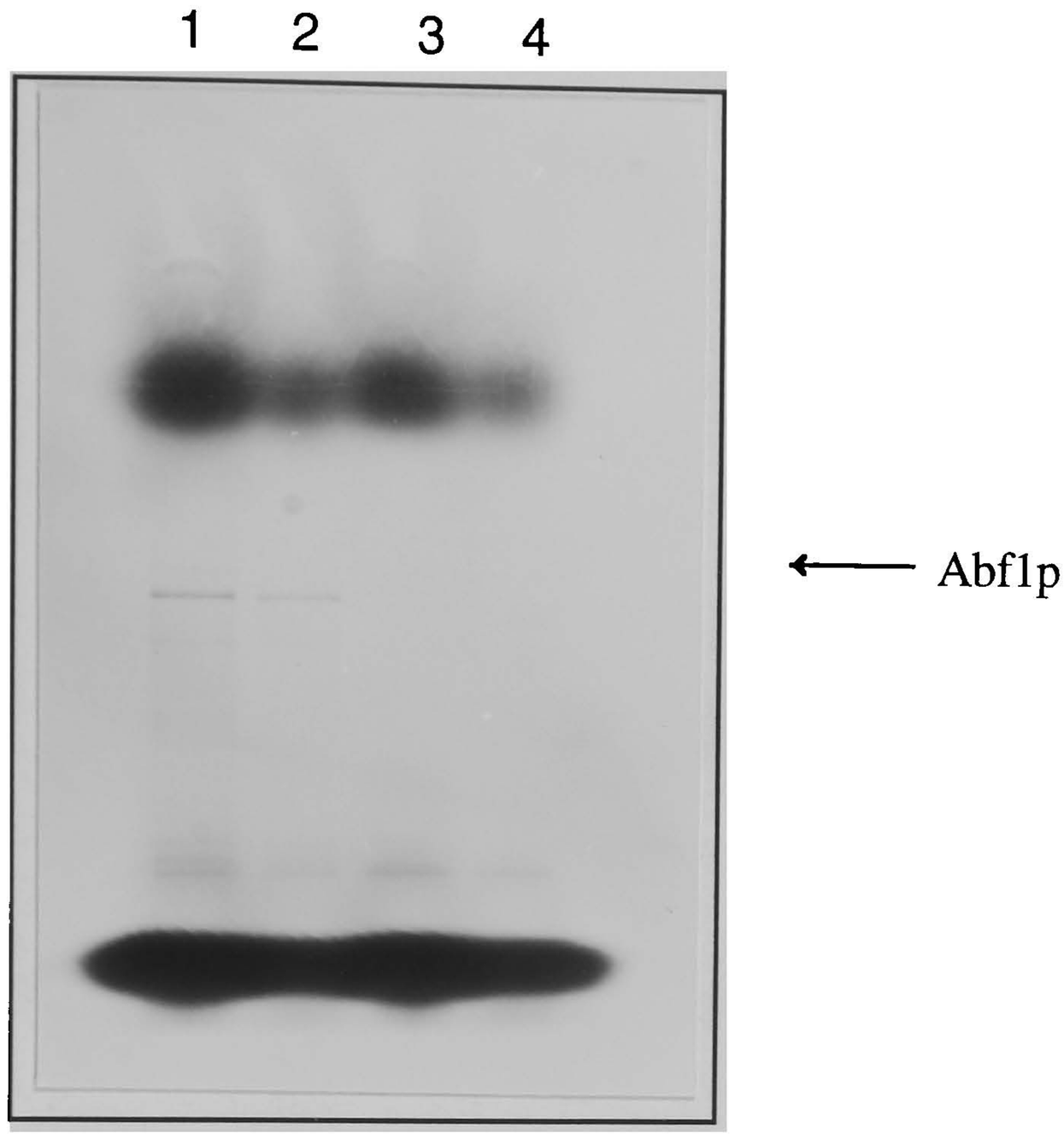


Figure 3.3 A protein gel showing Abf1p translated *in vitro* in the presence of ^{35}S methionine. Lanes 1 and 2 contain two different volumes of Abf1p lysate, whilst lanes 3 and 4 contain similar volumes of a mock lysate which was not primed with Abf1p mRNA.

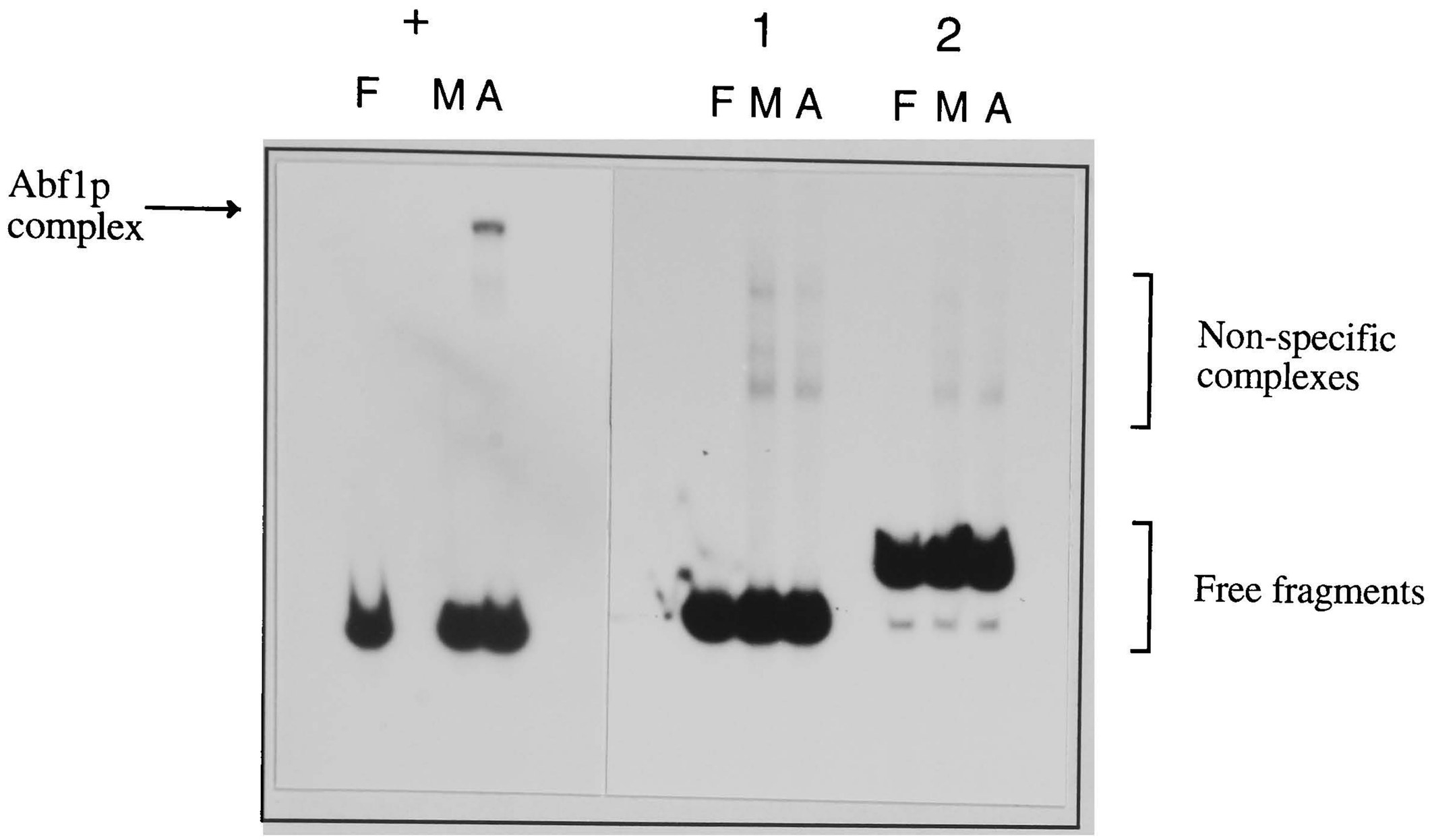


Figure 3.4 Neither fragment 1 nor fragment 2 from the *PGK* promoter will bind *in vitro* translated Abf1p. The gel retardation assays show that although IVT Abf1p will not recognise the *PGK* promoter fragments 1 and 2, it will form a complex with the positive control fragment (+) containing the known Abf1p binding site from the *PGK* promoter. F: fragment alone; M: fragment + mock lysate; A: fragment + Abf1p lysate.

fragments 1 and 2). This shows that neither of the potential Abf1p binding sites upstream of the *PGK* UAS actually binds Abf1p *in vitro*.

3.2.2 Does Rap1p bind upstream of the UAS ?

In vitro translated Rap1p was produced from pPE711 (Chambers *et al.* 1989). The plasmid was linearized with *Xba1* and RNA was transcribed from the SP6 promoter before being translated in a rabbit reticulocyte lysate. The IVT Rap1p was incubated in a gel retardation assay with labelled fragment 1, which contained the potential Rap1p site at -622 to -608, and also with fragment 2 which does not contain a match to the Rap1p consensus sequence. No complex formation was seen between the IVT Rap1p and either of the *PGK* fragments (Figure 3.5), although a complex did form between IVT Rap1p and a DNA fragment containing a known Rap1p binding site, from the *TEF2* promoter, which shows that the *in vitro* produced protein was capable of *in vitro* binding under these assay conditions. Thus there is no new Rap1p binding site upstream of the *PGK* UAS.

3.2.3 Identification of a new transcription factor binding site for Cpf1p

When fragment 1 was incubated with a yeast total protein extract a complex was formed between this promoter fragment and a protein present in the extract (Figure 3.6). This suggests that there is a sequence on the 100bp fragment that binds a protein, or proteins, present in yeast. Fragment 1 contains the potential match to the Cpf1p binding site at -547 to -540.

A clone of the Cpf1p coding region downstream of the T7 promoter was a gift from Jane Mellor (Oxford University) and this allowed Cpf1p to be produced *in vitro*. pSP73-22 (Mellor *et al.* 1990) was linearized with *BamH1* and RNA transcripts were produced from the T7 promoter. After translation the IVT Cpf1p was incubated with promoter fragment 1 which contains at its 3' end a match to the Cpf1p consensus binding site (-547 to -540), and also with a truncated version of this fragment (fragment 1A, Figure 3.2) which had 10bp removed from the 3' end to delete the potential Cpf1p binding site .

When the DNA/protein complexes had been resolved in a retardation assay, binding of IVT Cpf1p was observed with the fragment 1 (Figure 3.7A). However this binding was abolished when IVT Cpf1p was incubated with fragment 1A (Figure 3.7B, lanes M and C). Thus the site of Cpf1p binding was localized to the 3' end of promoter fragment 1, to the region containing the Cpf1p consensus. Fragment 1A was also incubated with the yeast total protein extract to see whether there were any

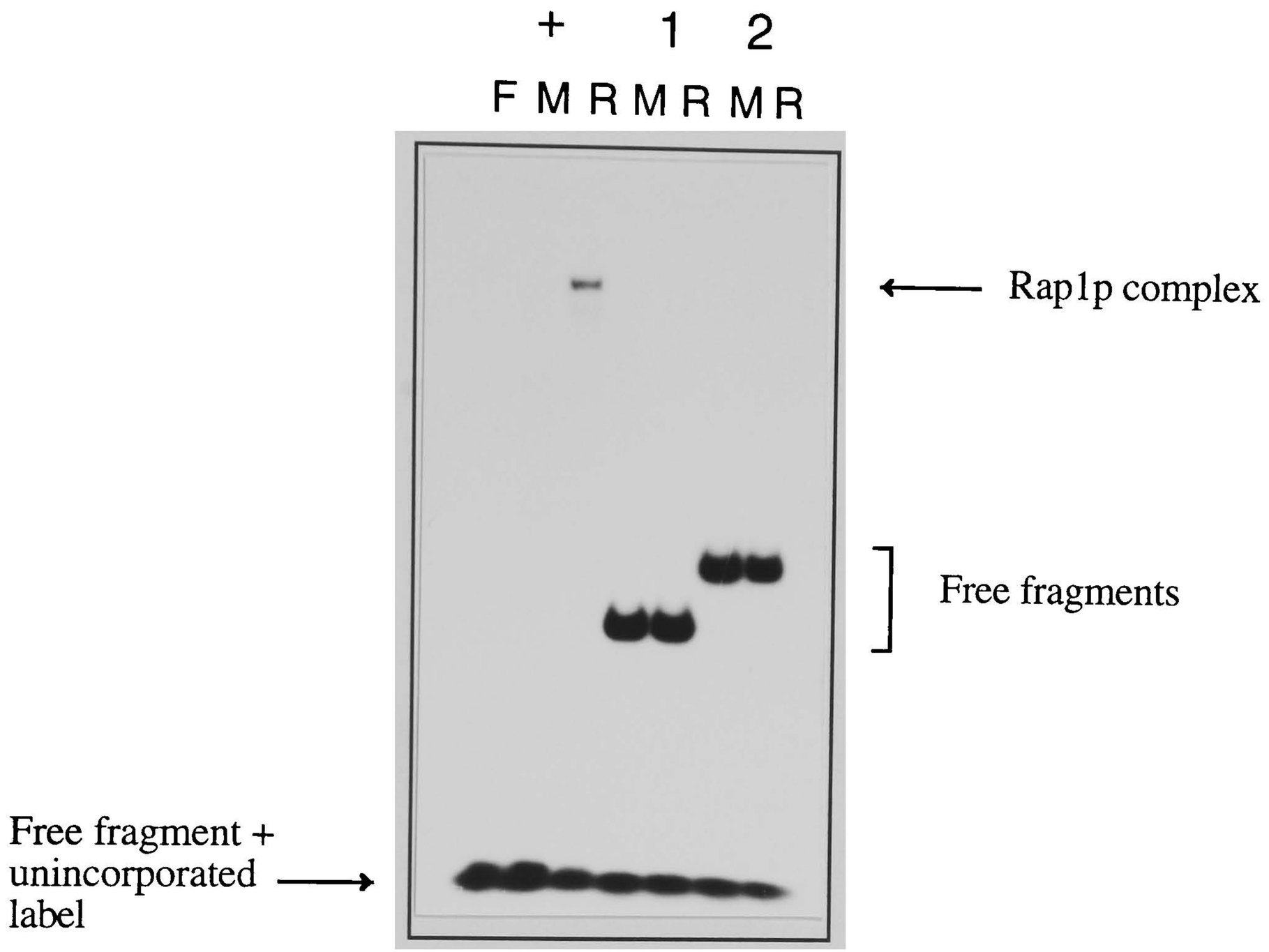


Figure 3.5 Neither fragment 1 nor fragment 2 from the *PGK* promoter will bind *in vitro* translated Rap1p. The gel retardation assay shows that although IVT Rap1p will not recognise the *PGK* promoter fragments 1 and 2, it will form a complex with the positive control fragment (+) containing a known Rap1p binding site. F: fragment alone; M: fragment + mock lysate; R: fragment + Rap1p lysate.

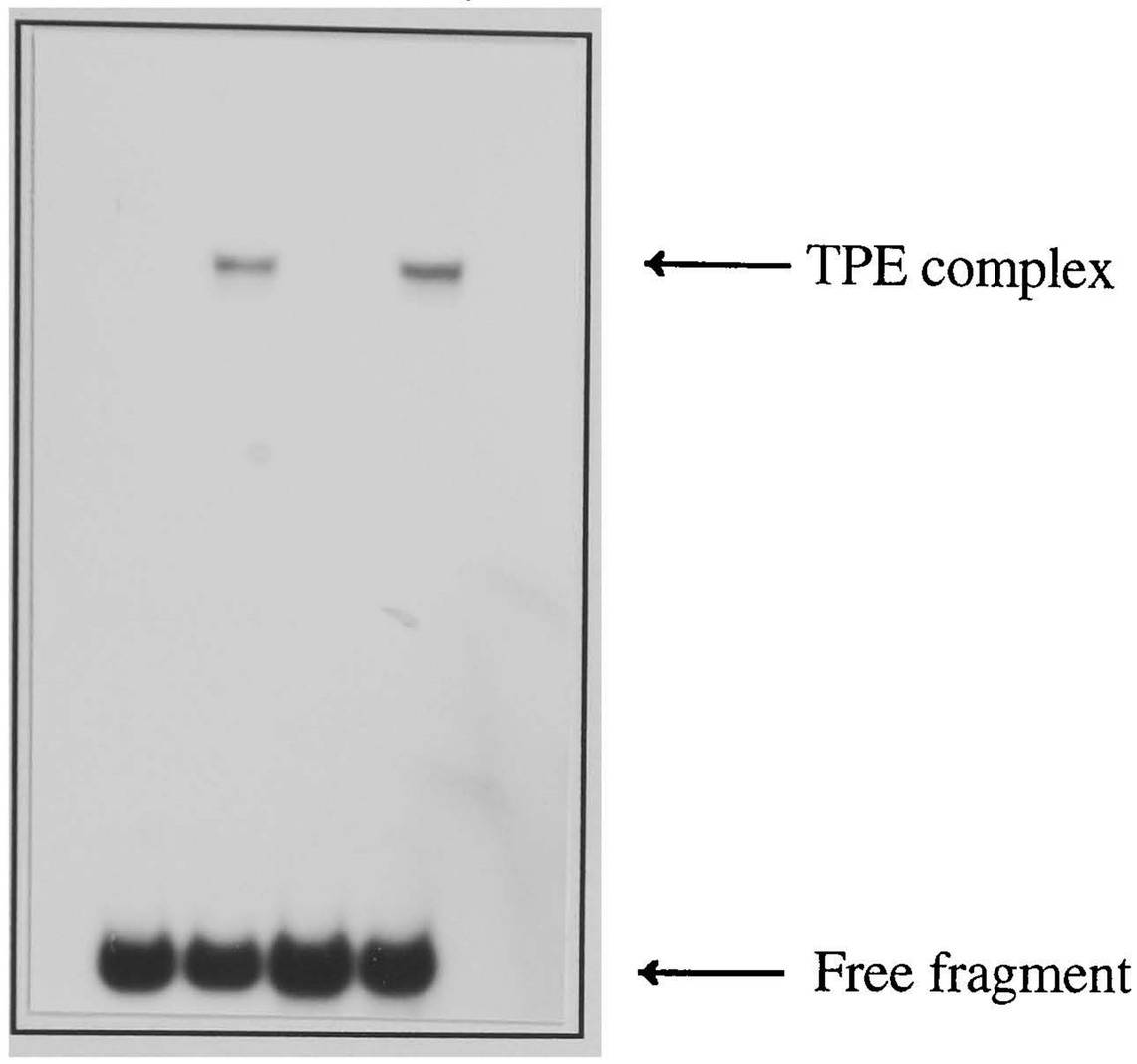


Figure 3.6 A gel retardation assay to show that a protein(s) present in a yeast total protein extract forms a complex with fragment 1 from the *PGK* promoter. Two preparations of fragment 1 were used, the second of which was labelled to a higher specific activity. F: fragment 1 alone; T: fragment 1 + yeast total protein extract.

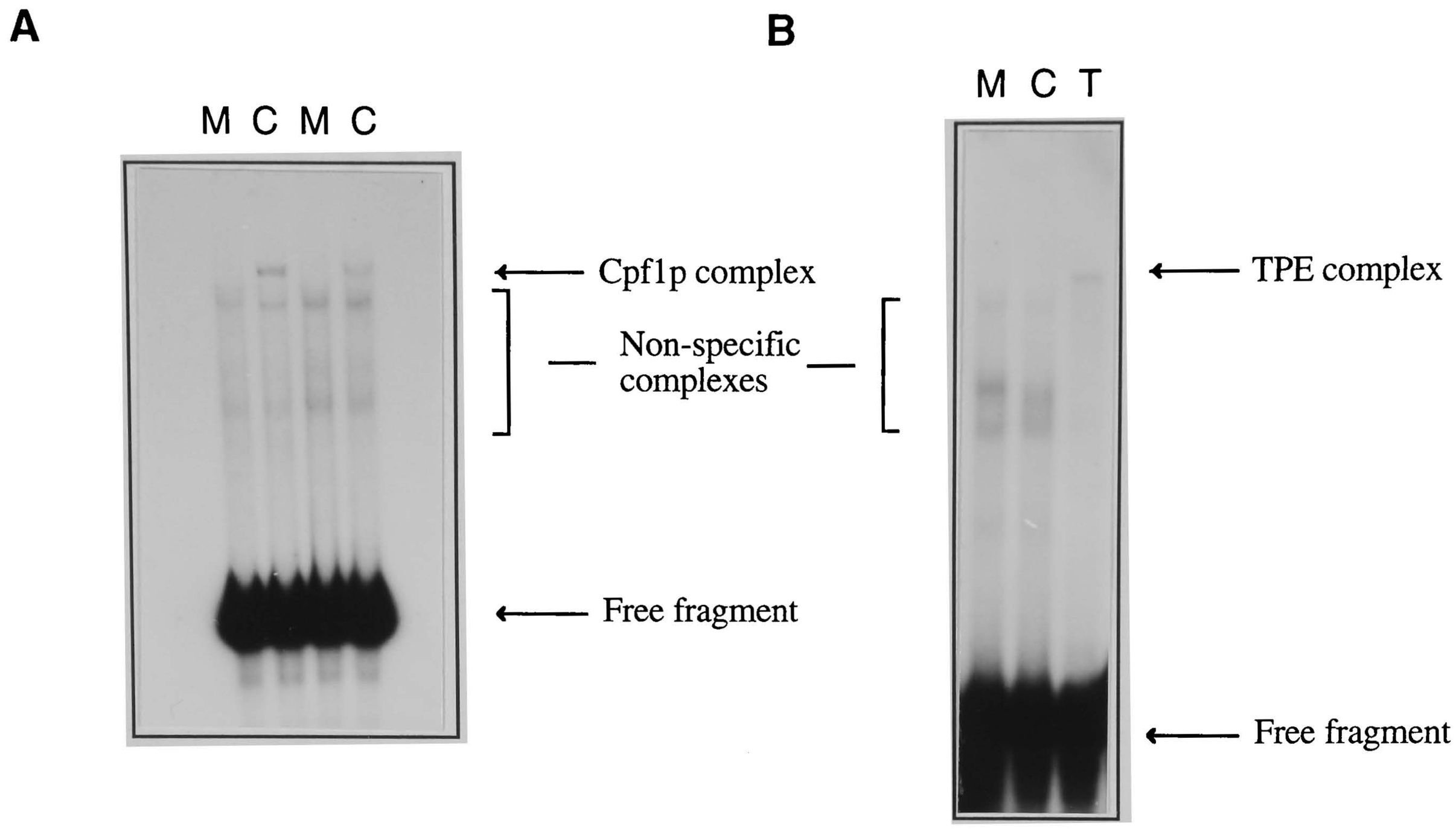


Figure 3.7 Cpf1p will bind *PGK* promoter fragment 1, but not fragment 1A, *in vitro*. **A**: A gel retardation assay showing the formation of a complex between IVT Cpf1p and fragment 1. **B**: Fragment 1A, a truncated version of fragment 1, no longer interacts with IVT Cpf1p but will still form a complex with a yeast total protein extract. M: fragment + mock lysate; C: fragment + Cpf1p lysate; T: fragment + yeast total protein extract.

other proteins which could bind to it. A complex was formed between the TPE and fragment 1A (Figure 3.7B, lane T) suggesting that there was a binding site for another yeast protein on the fragment. Since fragment 1A contains two potential Reb1p binding sites (-623 to -614, -561 to -552) the interaction of Reb1p and this region of the promoter was investigated.

3.2.4 Identification of a new transcription factor binding site for Reb1p

A strain of *E. coli* from which Reb1p could be overexpressed by induction with IPTG (Morrow *et al.* 1993a) was a gift from Jonathan Warner (Albert Einstein College of Medicine, New York). *E.coli*, BL21(DE3), had either been been transformed with pET11a-PCR8, which overexpresses an N-terminal truncated version of Reb1p, or with pET11a, the vector with no Reb1p insert. The pET series of vectors are for cloning and expressing DNA fragments under the control of a T7 promoter (Studier *et al.* 1990). They are transformed into an *E. coli* strain such as BL21(DE3) which has expression of T7 RNA polymerase under the control of the *lac*UV5 promoter which is inducible by IPTG. Thus, addition of IPTG allows T7 RNA polymerase to be produced, and the cloned DNA fragments are expressed (Studier *et al.* 1990). The N-terminal truncation of Reb1p is missing the first 201 amino acids, but this does not affect its binding to DNA since the DNA binding domain is in the C-terminus. The truncated version of the protein undergoes less degradation than the full length protein when expressed in *E. coli* (J. Warner, pers comm).

Two 100ml LBAp cultures were inoculated from 2ml overnights of the overexpressor strain and a strain transformed with just the parental plasmid. The cultures were grown to mid-log phase before the cells were induced with IPTG and grown at 37°C for a further 3 hours. The overexpressed protein was partially purified, using a modified version of Warner's protocol (Morrow *et al.* 1993a), as described in Materials and Methods, section 2.18.

Two protein fractions were obtained from the partial purification proceedure, a 40% and a 70% fraction. The high concentration of ammonium sulphate in these fractions was reduced by spin dialysis since the presence of high salt concentration affected migration of protein/DNA complexes in gel retardation assays. The concentration of the protein fractions was then determined with a Biorad protein assay, and 5µg was incubated with a labelled DNA fragment containing a known Reb1p site from the *RAP1* promoter. The Reb1p binding activity was found in the 70% protein fraction after both 40% and 70% fractions were tested in a gel retardation assay (Figure 3.8).

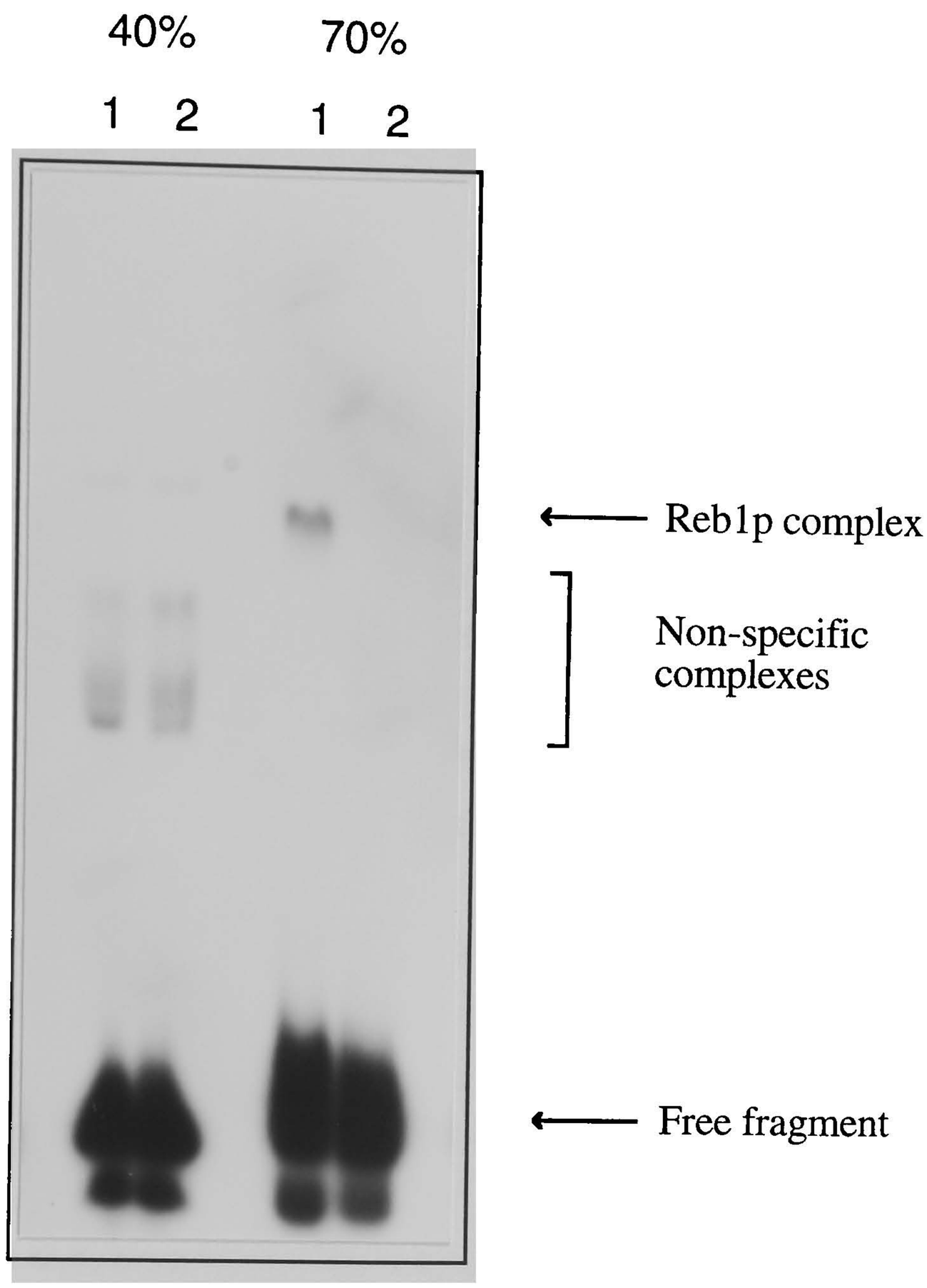


Figure 3.8 A gel retardation assay to determine which fraction of partially purified protein contained overexpressed Reb1p. The 40% and 70% fractions were incubated with a known Reb1p site from the *RAP1* promoter. Lane 1: fragment + protein extract from cells containing the Reb1p overexpressor construct; Lane 2: fragment + protein extract from control cells unable to overexpress Reb1p.

As the 70% fraction contained the partially purified Reb1p it was used as the source of Reb1p for a gel retardation with *PGK* promoter fragment 1A which contains both potential Reb1p binding sites. This retardation assay provided preliminary evidence that there was some Reb1p binding to the *PGK* promoter (Figure 3.9, compare lane 1 with lane 2). However, the high concentration of salts and other proteins in the partially purified preparation of overexpressed Reb1p affected the migration of the Reb1p/DNA complex through the retardation gels. Thus, a clone was constructed which could be used for the production of Reb1p RNA transcripts. This Reb1p RNA could then be translated *in vitro* and the retardation assays repeated with a different source of protein.

PCR primers were designed to the 5' and 3' of the Reb1p coding region with restriction enzyme sites for *Sal1* and *BamH1* incorporated for ease of cloning the PCR product into pSP46. Initial attempts to amplify the Reb1p sequence directly from yeast chromosomal DNA using touchdown PCR were unsuccessful, so a plasmid clone of Reb1p, pBA16, also a gift from Jonathan Warner, was used as a template for PCR. A product of about 2.4kb was amplified after 30 cycles (94°C for 30", 50°C for 40", 72°C for 2'30") followed by 10' at 72°C to make sure that all the product was blunt ended. This PCR product was isolated from an agarose gel and cleaned (Gene Clean, Stratech) before being cloned directly into T-vector (Promega), rather than cutting it with *Sal1* and *BamH1* and cloning into pSP46 as originally intended. A clone was obtained whose orientation allowed for the production of Reb1p RNA using the T7 promoter.

This clone, pAJ141 was linearized with *Sst1* and run-off transcripts were produced using T7 RNA polymerase. Translation was carried out in the presence of ^{35}S methionine and the labelled protein was detected on X-ray film (data not shown). The Reb1p lysate had several bands in it which were not present in the mock lysate (not primed with Reb1p RNA). This suggests that they were Reb1p specific bands, and were probably truncated proteins. These could be due to premature termination in the translation reaction, or to the translation of RNA transcripts which were not full length.

The Reb1p lysate was used as a source of Reb1p in the retardation assays with *PGK* promoter fragments 1A and 1B. Fragment 1A contains both potential Reb1p binding sites, whereas fragment 2 contains only the potential binding site at -623 to -614. The gel retardation assay showed no binding of IVT Reb1p with fragment 1B but clear binding to fragment 1A (Figure 3.10). The multiple bands seen in the retardation assay are probably due to the presence of different sized Reb1 proteins in the Reb1p lysate. Figure 3.10 suggests that the most upstream of the potential Reb1p

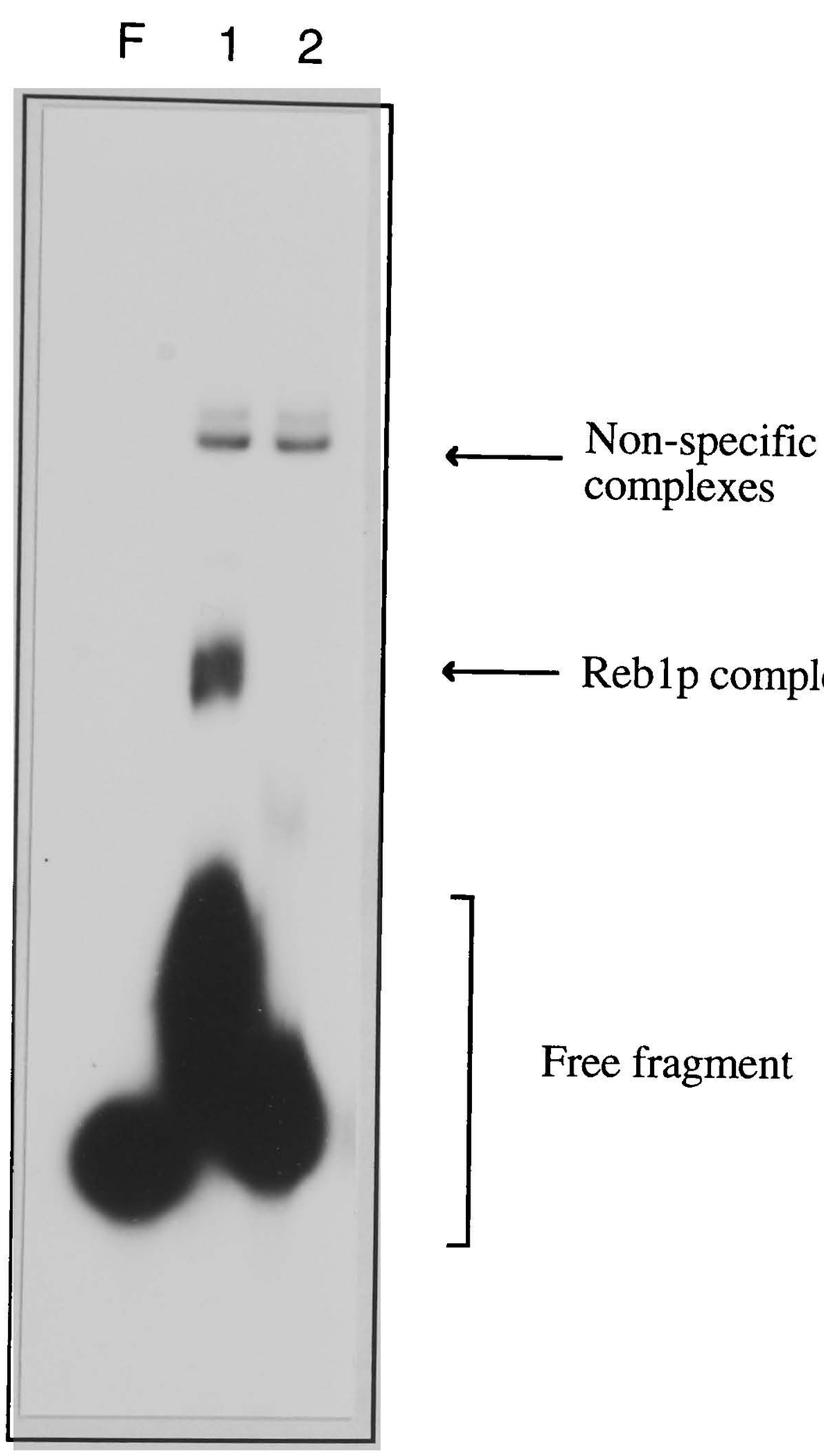


Figure 3.9 A gel retardation assay indicating that Reb1p is able to form a complex with *PGK* promoter fragment 1A. F: fragment alone; 1: fragment + partially purified protein extract containing overexpressed Reb1p; 2: fragment + partially purified protein extract not containing overexpressed Reb1p.

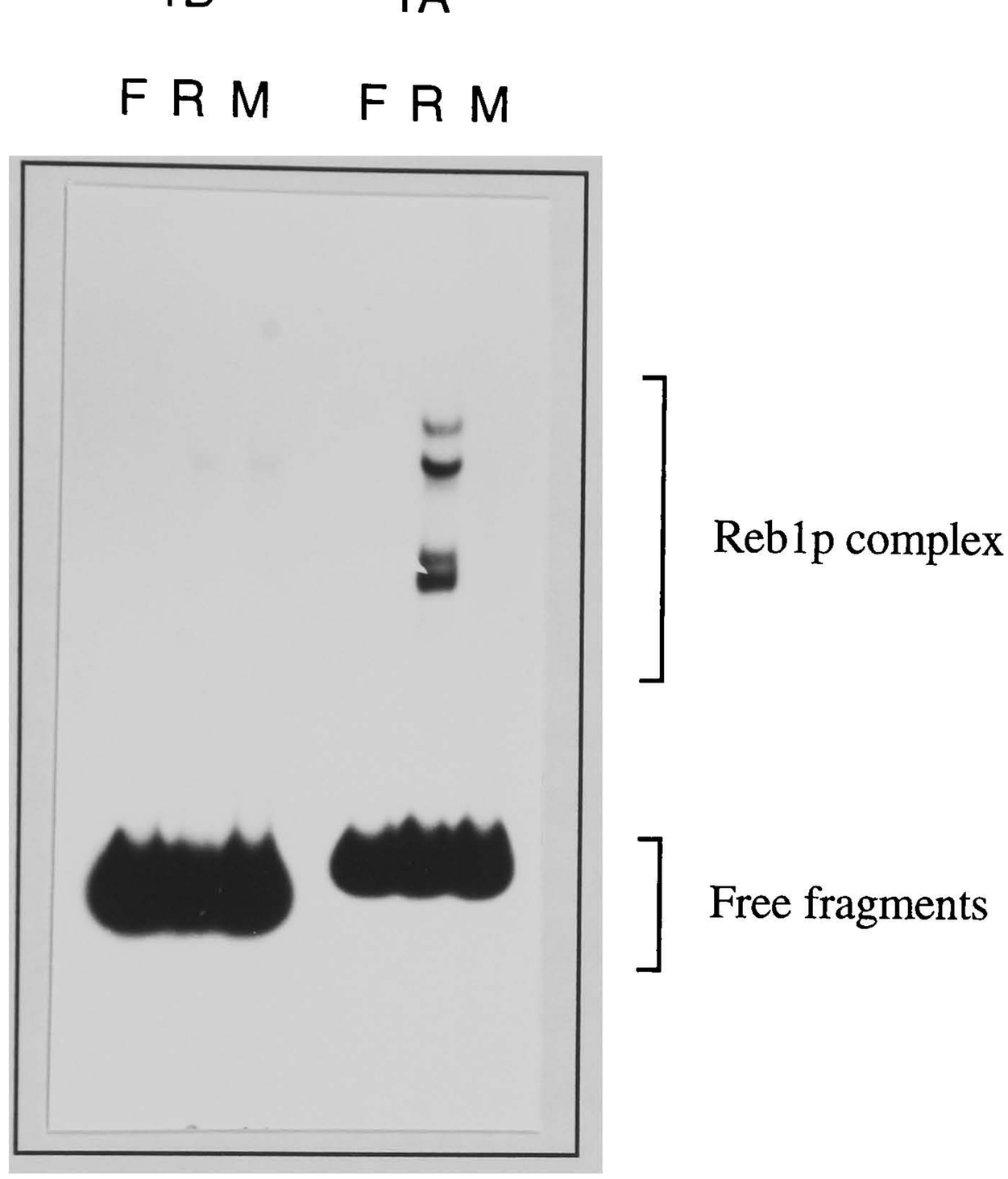


Figure 3.10 A gel retardation to show that IVT Reb1p will form a complex with *PGK* promoter fragment 1A but not with fragment 1B. The multiple bands in the Reb1p + fragment 1A track are due to the presence of truncated forms of Reb1p in the lysate. F: fragment alone; R: fragment + Reb1p lysate; M: fragment + mock lysate.

binding sites is not occupied *in vitro* but that the downstream site (-561 to -552), adjacent to the newly identified Cpf1p binding site, will bind Reb1p.

3.3 Discussion

Whilst the *PGK* promoter contains many potential transcription factor binding sites, only some of these have been shown to interact with proteins *in vitro*. Abf1p, Rap1p and Gcr1p have been shown previously to bind to the region of the *PGK* promoter defined as the UAS, and two new transcription factor binding sites, for Cpf1p and Reb1p, have now been identified upstream of the UAS by gel retardation assays in which fragments from this region of the *PGK* promoter were incubated with *in vitro* translated proteins.

That only two of the five potential transcription factor binding sites identified upstream of the *PGK* UAS actually bind a protein *in vitro* may indicate something about the way in which these proteins interact with their recognition sites. The most upstream of the two potential Abf1p sites (-704 to -692) contains two mismatches to the consensus (Table 3.2). One of these is in the highly conserved 3' element ACG which is separated from the 5' conserved element by a stretch of DNA whose length is more important than its sequence. Mutations in either of the 5' or 3' conserved elements leads to a reduction in, or to the abolition of, binding of Abf1p (Dorsman *et al.* 1989). A mismatch found in both of the potential Abf1p binding sites is a change of the 5' most purine of the consensus being changed to a pyrimidine. An investigation into the interactions of Abf1p with its binding site (McBroom *et al.* 1994a), has shown by missing contact analysis, and by methylation interference, that the protein/DNA interaction at this position is important for Abf1p binding. This same analysis also demonstrated the importance of certain of the bases flanking the consensus sequence for making DNA/protein contacts when Abf1p binds.

The potential Rap1p binding site upstream of the *PGK* UAS (-622 to -608) has three mismatches to the Buchman *et al.* (1988a) consensus and five mismatches to the Graham and Chambers (1994b) consensus (Table 3.2) which is similar to the Buchman consensus, but longer at the 5' end and more constrained at the 3' end. The potential *PGK* sequence matches the proposed core, CACCCA, of both consensus sequences except at the final A residue. A point mutation here (position 9 of the Graham and Chambers consensus) has been shown to have a considerable effect on the strength of Rap1p binding, as has a point mutation at the conserved C, position 12, (Vignais and Sentenac 1989). Since, individually these mutations reduce binding by between 88-96%, in combination they are very unlikely to allow the formation of a complex between Rap1p and the DNA. The other mismatches are found at

Table 3.2 A comparison of the sequences of the potential transcription factor binding sites in the *PGK* promoter, upstream of the UAS, with their consensus binding sites.

Transcription Factor Site	*PGK* Sequence [a]	Consensus [b]	Reference
Abf1p (-704 to -692)	cTCGTGAGTAAgG	RTCRYNNNNNACG	Dorsman *et al.* 1989
Abf1p (-554 to -542)	cTCATAAAGCACG		
Rap1p (-622 to -608)	cTtCACCCtCAtACt	RTRCACCCANNCMCC	Graham and Chambers 1994b
		RMACCCANNCAYY	Buchman *et al.* 1988a
Cpf1p (-547 to -540)	AgCACGTG	RTCACRTG	Hieter *et al.* 1985
Reb1p (-623 to -614)	gCTTCACCCt	YNNYYACCCG	Chasman *et al.* 1990
Reb1p (-561 to -552)	aTGTTACCCt		

[a] Mismatches are indicated by the use of lower case letters.

[b] N: A, T, C or G; R: A or G; Y: T or C; M: C or A

positions 1, 3 and 15, in flanking sequences which may be important for stabilising interactions between Rap1p and the consensus core.

The mismatch between the Cpf1p binding site in the *PGK* promoter and the CDE1 consensus sequence, RTCACRTG, is outside of the core recognition site (CANNTG) for basic Helix-Loop-Helix proteins, the 5' T is replaced with an A in the *PGK* sequence (Table 3.2). The presence of a T in this 5' position inhibits the binding of the transcription factor Pho4p, which recognises the same bHLH consensus, since it places a methyl group in the major groove (Fisher and Goding 1992). The binding of Pho4p to the *PGK* promoter was not investigated in this thesis, and it remains a formal possibility. However, Pho4p is involved in the regulation of *PHO5*, and there are sequences outside of the consensus, apart from the 5' T, which affect Pho4p and Cpf1p binding specificity (Fisher and Goding 1992).

Of the two potential Reb1p binding sites in the *PGK* promoter one does bind Reb1p *in vitro* and the other does not (Figure 3.10). The mismatches to the Reb1p consensus sequence though are all in the same positions (see Table 3.2); a 5' pyrimidine in the consensus is replaced with a G in the most upstream site, which does not bind *in vitro*, and with an A in the downstream site, which does bind Reb1p *in vitro*. In both cases a highly conserved 3' G is replaced with a T. Chasman *et al.* (1990) show that the most 5' pyrimidine is not critical for Reb1p binding, it can be functionally replaced with a G as is the case with the non-binding potential site at -623 to -614. They also suggest that the region of DNA flanking the Reb1p consensus is important since sequences which are identical at every position of the consensus bind Reb1p with different affinities. The flanking sequences are presumably the difference between the two potential Reb1p binding sites in the *PGK* promoter, allowing one to bind Reb1p and the other not.

In the majority of the cases discussed above mismatches to the consensus binding site are not sufficient to explain the lack of binding seen between the *PGK* promoter fragments and *in vitro* translated protein. This emphasises the importance of the DNA flanking the consensus binding site in stabilizing interactions between a transcription factor and its core recognition sequence. This is not surprising bearing in mind that these multifunctional transcription factors have large DNA binding domains (Abf1p: 487 amino acids (aa), Reb1p: 309aa, Rap1p: 235aa, Cpf1p: 131aa) which might be expected to come into contact with a considerable length of DNA, even with bases not immediately adjacent to the consensus site.

A comparison of the proteins which bind upstream of many glycolytic genes shows that the same group of transcription factors are represented at their promoters. Most

glycolytic UAS contain a binding site for Rap1p, frequently close to a "CT block" which is at the core of the Gcr1p binding site. These two transcription factors play a key role in the regulation of transcription from glycolytic genes, usually deletion of the Rap1p site results in a decrease in the level of transcription, and genes with Gcr1p sites show reduced transcription in a *gcr1* background. The roles of Abf1p and Reb1p in glycolytic promoters are less clear; Abf1p appears to have a role in the transcription of *PFK1* but not of *PGK* or *PYK*. At the *ENO2* promoter Abf1p seems to act as an anti-repressor allowing an as yet unidentified factor to activate transcription along with Rap1p and Gcr1p (Willett *et al.* 1993). Reb1p seems to be involved in the activation of *TP1* and also *TDH3* where it is associated with a nucleosome free region. In the *ENO1* promoter Reb1p binds to both the UAS and URS. When Reb1p does bind to a glycolytic promoter, it is in an upstream position in comparison with the other transcription factors, suggesting that one of its functions may be to protect the promoter from upstream interference. In agreement with this, the newly identified Reb1p binding site in the *PGK* promoter is also in the most upstream region of this promoter. The *PGK* promoter is the first glycolytic promoter in which a Cpf1p binding site has been identified and its role here may involve chromatin remodelling, since Cpf1p has not been demonstrated to be a transcriptional activator.

Thus, there are binding sites for four members of the family of multifunctional transcription factors within one promoter. There may be some redundancy of function at the *PGK* promoter since the roles of the multifunctional transcription factors overlap to some extent, and their binding sites are in close proximity to each other. However, the *PGK* promoter makes a good model system in which to study these transcription factors to try to elucidate the molecular mechanisms by which they regulate gene transcription. One way in which this may be done is by studying the effects of deleting individual binding sites from the *PGK* promoter.

Chapter 4

Construction of Yeast Strains to Investigate the Roles of Transcription Factors at the Chromosomal *PGK* Locus

4.1 Introduction

Previous studies of the *PGK* promoter have looked at the effects of deleting transcription factor binding sites from the UAS on the levels of transcription from a multi-copy plasmid-borne construct of the gene. This is a straightforward procedure, requiring little manipulation of the yeast cell, which allows answers to be obtained relatively quickly. Once the required deletion has been made from the promoter, it can be cloned into a multi-copy plasmid either linked to a reporter construct *eg* β-galactosidase, or as a whole gene construct, and the plasmid can be transformed into yeast. After transformants have been selected, the effect of the deletion can be determined by assaying the levels of the reporter, or by Northern blotting, comparing levels of RNA from a plasmid carrying the whole gene with RNA from the deleted gene construct. In such experiments the plasmid copy number must be determined since a construct present in high copy number can appear more active than it really is. This can be overcome to some extent by looking at deletions on single copy number plasmids, but the copy number of these can also vary between transformants. Another disadvantage of looking at the effects of promoter deletions on plasmids is that the higher chromatin structure of plasmids may differ from higher chromosomal chromatin structure. If the transcription factor under investigation is one which exerts some, or all, of its effects by remodelling chromatin then examination of a plasmid-borne deletion may give misleading results.

To minimise the effects of removing a gene from its wild type context, another way of examining the roles of different transcription factors is to look at deletions of their binding sites from a chromosomal copy of the gene. A copy of the gene containing promoter deletions can be integrated at a non-homologous locus giving a yeast strain in which there are two copies of the gene. In this case the promoter can be linked up to a reporter gene, or the coding region can be truncated, allowing the message from the gene carrying the promoter deletion to be distinguished from the message of the wild type gene. Alternatively, the deleted gene construct can be targetted to its chromosomal homologue giving tandem copies of the gene. In some cases the loss of plasmid sequences and one of the gene copies can be selected for so that only the mutated copy of the gene is present in the chromosome. Yeast strains containing deletions from the chromosome take more time to prepare than deletions from

plasmids but the final strain can be identical to wild type except for a specific deletion, and any effects on transcription due to the chromosomal environment of the gene will still be present, thus, any differences between the wild type gene and the mutated version should be a result of the mutation.

Unlike Rap1p, Reb1p and Abf1p, Cpf1p is not encoded by an essential gene and so it is possible to look at the role of Cpf1p binding sites in promoters by looking at the chromatin structure of these genes, and levels of transcription from them, in a *cpf1* null strain. As this is not feasible for the other members of the family of multifunctional transcription factors, yeast strains were made in which binding sites for Rap1p, Reb1p and Abf1p were deleted individually from the chromosomal copy of *PGK* using the Pop-In/Pop-Out method of homologous recombination. A yeast strain was also made in which the potential transcription factor binding site for yATF, identified by Lin and Green (1989) at the very 3' end of the *PGK* UAS (-415 to -402), was deleted.

4.1.1 The "Pop In - Pop Out" Method of Homologous Recombination

Deletions of specific transcription factor binding sites from the *PGK* promoter were introduced into the chromosomal copy of the gene using a method of homologous recombination based on the method of Scherer and Davis (1979). This has advantages over One Step gene disruption (Rothstein 1983), and targeting of *in vitro* mutagenized DNA to the wild type locus (Shortle *et al.* 1984), since it allows the chromosomal copy of the gene to be replaced with a mutated version and all vector sequences are lost. One Step gene disruption replaces one yeast marker, for example, *trp*, with a sequence of DNA containing the mutated gene of interest and a second yeast marker, for example, *ura*. Thus transformed yeast would become trp^- and ura^+, however the chromosomal copy of the target gene is still present. This method is useful for introducing a promoter linked to a reporter gene such as *lacZ* into the chromosome. The method of Shortle *et al.* (1984) uses a truncated version of the gene for targetting and again the wild type copy is not deleted. In both of these cases the plasmid sequences also remain in the genome.

Use of the Pop-In/Pop-Out method of allele replacement (Figure 4.1) means that only one copy of the affected gene is present in the genome of the final strain. This copy is present at the wild type locus, contains no vector sequences and only desired mutations which were introduced into the gene on the integrating plasmid. A *URA3* based integrating vector is used; this contains no ARS sequences so the plasmid cannot exist within the yeast cell unless it has been integrated into the genome. Integration can be targeted to the locus of interest by making a double stranded break

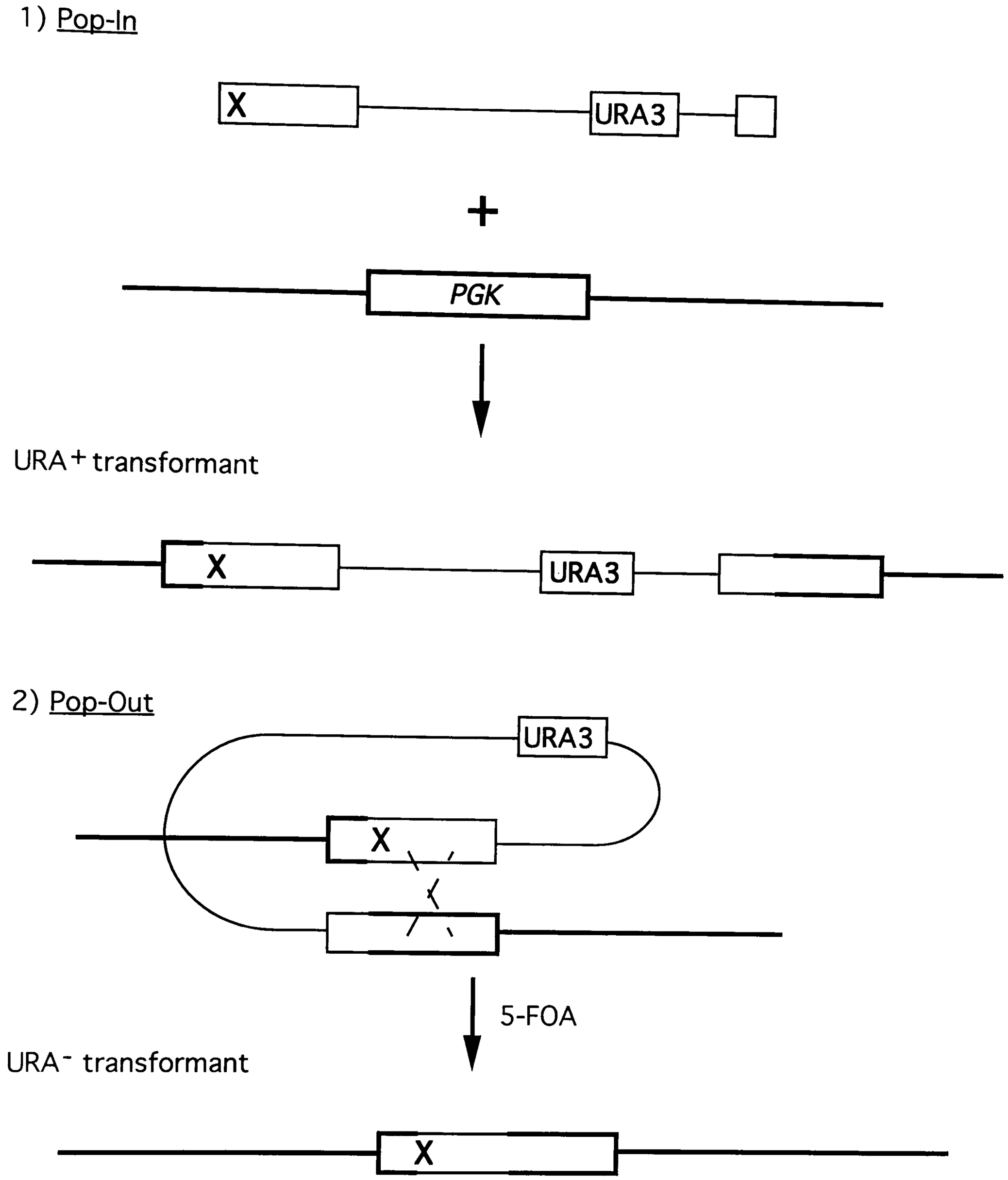


Figure 4.1 Diagram showing the pop-in/pop-out method of allele replacement. During the pop-in a mutation in the promoter of the plasmid-borne copy of *PGK*, indicated by **X**, is introduced into the chromosome. The pop-out results in one copy of *PGK* and the vector sequences being lost, leaving one copy of *PGK* carrying the promoter mutation.

within the target gene on the plasmid. This increases the frequency of recombination (Orr-Weaver *et al.* 1981) and most integration will occur at the region of the genome homologous to the cut site. In some cases there can be recombination between the marker gene and the integrating plasmid but this is reduced if there is no chromosomal copy of the marker, or if it is highly disrupted. The *ura3-52* mutation is a very stable, non-reverting mutation caused by a Ty insertion within the *URA3* coding region (Rose and Winston 1984).

Once the Pop-In has occured there are two copies of the target gene in the genome, one wild type and one mutant, separated by vector sequence including the *URA3* marker (see Figure 4.1). At this stage it is useful to verify that the mutated copy has not undergone gene conversion upon integration; the frequency of gene conversion is highest when the double stranded break used to target integration is closest to the mutation in the incoming gene (Orr-Weaver *et al.* 1988).

The second stage of the proceedure involves selecting for the loss of the *URA3* marker. This happens when a second recombination event takes place between the two copies of the gene present in the chromosome, and it results in the loss of one copy of the gene, and also the intervening plasmid sequences. This step is selected for using 5-fluoroorotic acid (5-FOA; Boeke *et al.* 1984). *URA3* encodes orotidine 5' phosphate decarboxylase which is required for the synthesis of uracil; $URA3^{+}$ cells convert 5-FOA to 5-fluorouracil which is toxic to yeast, but $ura3^{-}$ cells, being unable to catalyse this conversion, are resistant. Thus, $URA3^{+}$ yeast containing the required Pop-In mutation are grown on medium containing 5-FOA and only $ura3^{-}$ cells *ie* those in which a Pop-Out has occured, will survive. Since Pop-Out cells are ura^{-} the medium must contain uracil to allow their growth.

Finally these cells must be screened to determine which copy of the gene has been excised, the mutant or the wild type; if the double stranded break used for targetting the Pop-In recombination is between the mutant site and the shortest stretch of homology then the Pop-Out is more likely to occur in the longer region of homology. Thus it is useful to be able to target an asymmetric insert (with respect to the position of the mutation) to increase the chances of retaining the altered gene after Pop-Out.

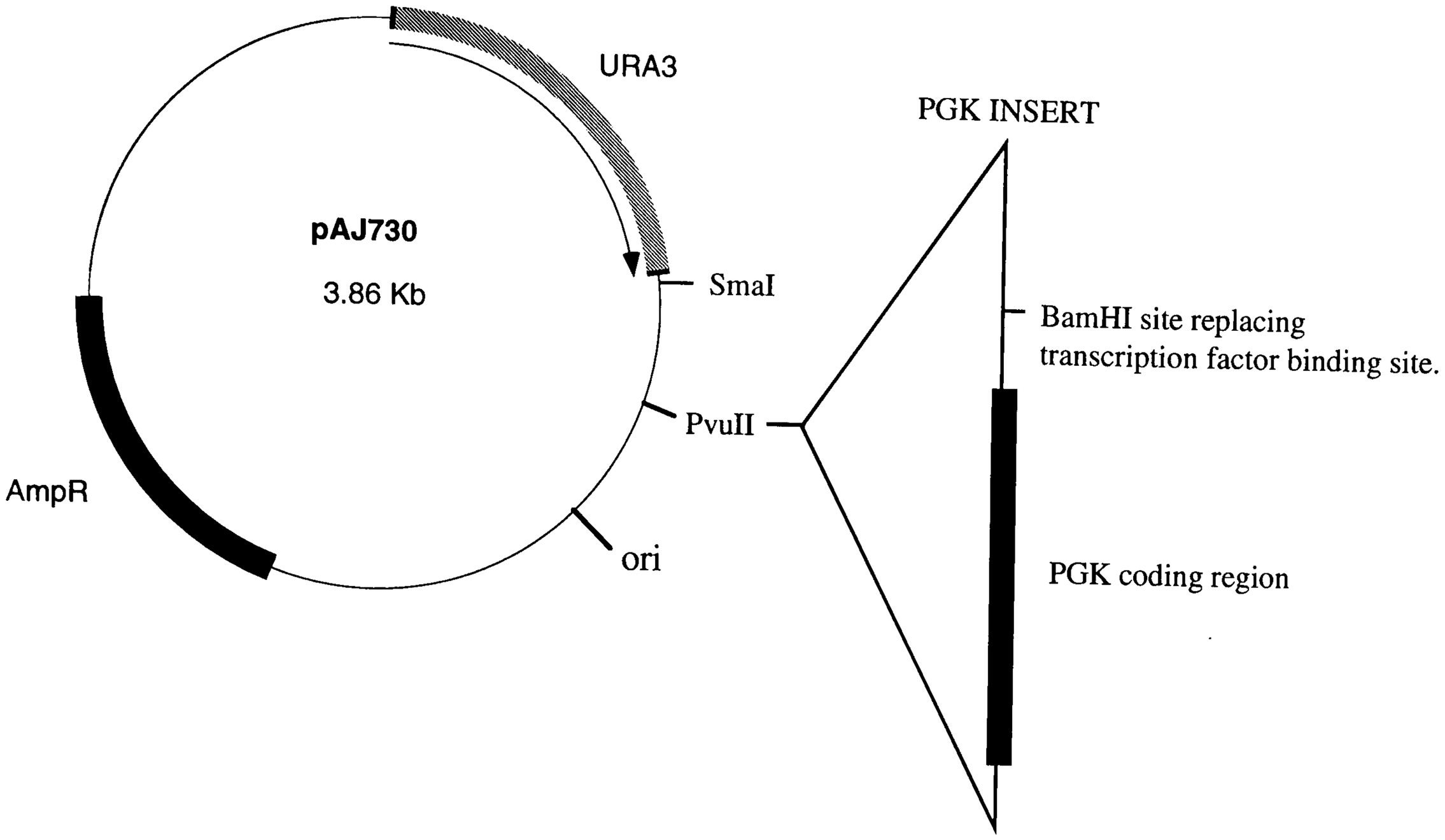
4.2 Results

4.2.1 Constructing Chromosomal Deletions of *PGK* Transcription Factor Binding Sites

Individual window deletions of the Rap1p and Abf1p transcription factor binding sites from the *PGK* gene on a multicopy plasmid had already been constructed (Chambers *et al.* 1988). pKV516 (ΔRap1p, -473 to -464) and pKV502 (ΔAbf1p, -538 to -473) contain the entire *PGK* sequence, with a specific transcription factor binding site deleted from the UAS and replaced with a *BamH1* site, on a *HindIII* fragment. In order to construct the integrating vectors *HindIII/BglII* fragments were isolated from pKV502 and pKV516. This utilised a *BglII* site just upstream of the *PGK* stop codon and facilitated the separation of the *PGK* fragment from vector sequences on an agarose gel. The fragments, which contained the entire *PGK* sequence from the promoter through to +1156 in the coding region, were cloned into the *PvuII* site of pAJ730 (Figure 4.2) generating recombinants pAJ107 (ΔAbf1p) and pAJ108 (ΔRap1p). pAJ730 has the *URA3* gene from YCp50 (Johnston and Davis 1984) subcloned into the polylinker of pSP46 (Ogden *et al.* 1986), but it has no CEN/ARS sequences and so cannot exist in the yeast cell unless it has been integrated into the chromosome.

A deletion of the newly identified Reb1p site was constructed using PCR (Figure 4.3). The entire *PGK* sequence was amplified as two products; an upstream product from immediately 5' to the Reb1p site, as far as the *HindIII* site at -1480 using PCR primers 521B and PGKD, and a downstream product containing the sequence immediately 3' to the Reb1p site through to the *PGK* terminator which was amplified using primers 533B and PGKC2 (see Table 2.1A for primer sequences and co-ordinates). Primers PGKC2 and PGKD had a *BamH1* site incorporated into them, whilst 521B contained a *BglII* site and 533B a *BclI* site. Thus the downstream product could be digested with *BamH1* and *BclI*, and the upstream product with *BamH1* and *BglII*. *BclI* and *BglII* have *BamH1* compatible ends which meant that the PCR products could be cloned directly into pAJ735, this is identical to pAJ730 except that the *SmaI* site has been converted to a *BamH1* site. After the PCR products had been cloned in the correct orientation, generating pAJ105 (ΔReb1p), the integrating vector contained a clone of the *PGK* gene but with the Reb1p binding site replaced with a *BamH1* linker.

The integrating vector carrying *PGK* sequence with the potential yATF binding site deleted was constructed by cloning a PCR product, amplified using primers 521A and 521B, containing the region of *PGK* upstream of the potential yATF site, into pAJ735. The downstream *PGK* sequence, from 3' to the yATF site to +1156 in the coding region, was isolated from pAJ110 as a *BamH1/BglII* fragment; pAJ110 is a multicopy plasmid where the potential yATF site has been deleted from *PGK*, the yATF deletion was generated using PCR in a similar way to the Reb1p deletion (Figure 4.3). Primers 521A, 521B, 533A and 533B (Table 2.1A) were used to



pAJ107: pAJ730 + *PGK*ΔAbf1p

pAJ108: pAJ730 + *PGK*ΔRap1p

pAJ105: pAJ730 + *PGK*ΔReb1p

pAJ106: pAJ730 + *PGK*ΔyATF

Figure 4.2 pAJ730, the *URA* selectable plasmid into which *PGK* constructs with specific transcription factor binding site deletions were cloned. The ΔRap1p and ΔAbf1p constructs were cloned into the *PvuII* site, and the ΔReb1p and ΔyATF constructs at the *SmaI* site after it was converted to a *BamH1* site. pAJ730 and its derivatives are unable to be maintained in yeast unless they are integrated into the chromosome, as they contain no CEN/ARS sequences.

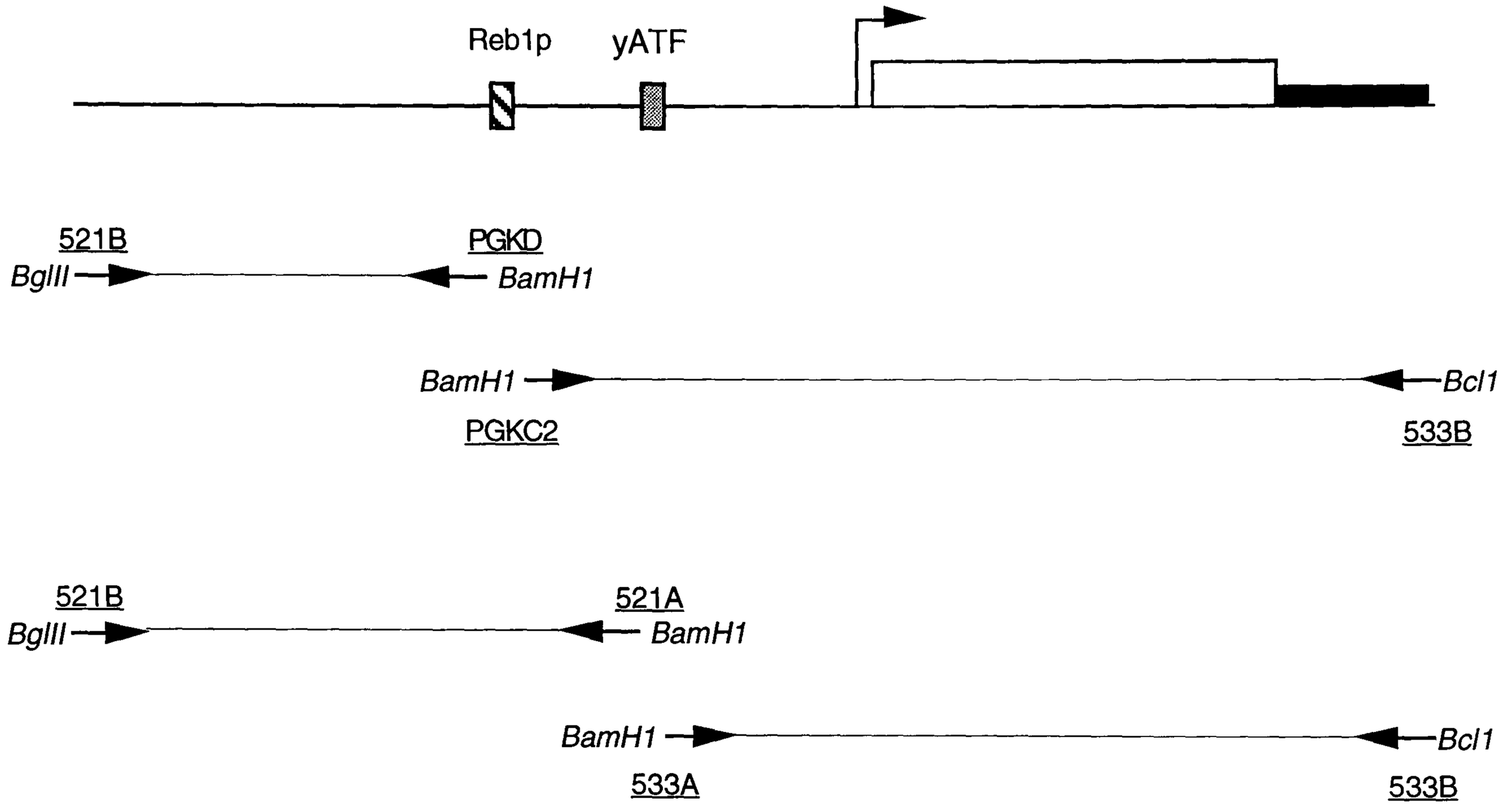


Figure 4.3 PCR strategy for constructing a deletions of the Reb1p binding site, and the potential yATF binding site, from the *PGK* promoter. See text for details of cloning procedure.

amplify the upstream and downstream regions of *PGK* which were cloned into the *BamH1* site of pAJ6 to give pAJ110. When the downstream of the yATF deletion had been cloned next to the upstream deletion fragment pAJ106 was generated. Here the potential yATF binding site had been replaced with a *BamH1* linker in a clone of *PGK* extending from the 5' *HindIII* site to the *BglII* site at +1156 in the coding region.

Since the process of PCR is prone to errors, *Taq* polymerase misincorporates bases at a rate of about 1 per 4 x 10^4 bases (Innis *et al.* 1988), the *PGK* constructs in pAJ105 and pAJ106 were sequenced across their UAS to ensure that the only mutation that had been introduced was that of replacing the transcription factor binding site with a *BamH1* site.

pAJ106, pAJ107 and pAJ108 all contain a unique *Cla1* site at -820 in the *PGK* promoter, the *Cla1* site in the coding region is downstream of the *BglII* site used for subcloning the Rap1p, Abf1p and yATF deletions and therefore no longer present. This unique *Cla1* site was used to linearise pAJ106, pAJ107 and pAJ108 within the *PGK* sequence, before they were transformed into DBY745 using the One Step transformation procedure. This was to help to target the Pop-In integration to the wild type *PGK* locus and to increase the efficiency of recombination. pAJ105 did not contain a suitable site for linearisation within *PGK* so the transformation of DBY745 was carried out with non-linearised plasmid; this did not have a noticeable effect on the number of transformants when compared with the numbers from the pAJ106, pAJ107 and pAJ108 transformations.

After transforming DBY745, URA^+ transformants were selected, since these contain the integrated plasmid, and screened to ensure that the construct had not undergone gene conversion on integration. This screening was achieved with PCR. Oligonucleotide primers, homologous to the 5' end of *PGK* (521B) and to +75 in the coding region (PGK), were used to amplify the promoter region of *PGK* giving a product of about 1.8kb. Chromosomal DNA prepared from the URA^+ transformants using the "Ten minute" prep (Materials and Methods, section 2.6.1) was used as a template. Since, after Pop-In there are two copies of *PGK* in the chromosome, two PCR products are amplified with these primers. One copy should be wild type, the other should contain an extra *BamH1* site. Therefore if gene conversion has not occured one of the PCR products should cut with *BamH1* giving a pattern of three bands on an agarose gel: the full length wild type promoter product which has no *BamH1* site (1.8kb) and two cleavage products from the deleted promoter (~1.4kb and ~0.4kb, see Figures 4.4 and 4.5) If gene conversion has occured then both copies of *PGK* will be wild type and therefore the PCR products will not cut with *BamH1*.

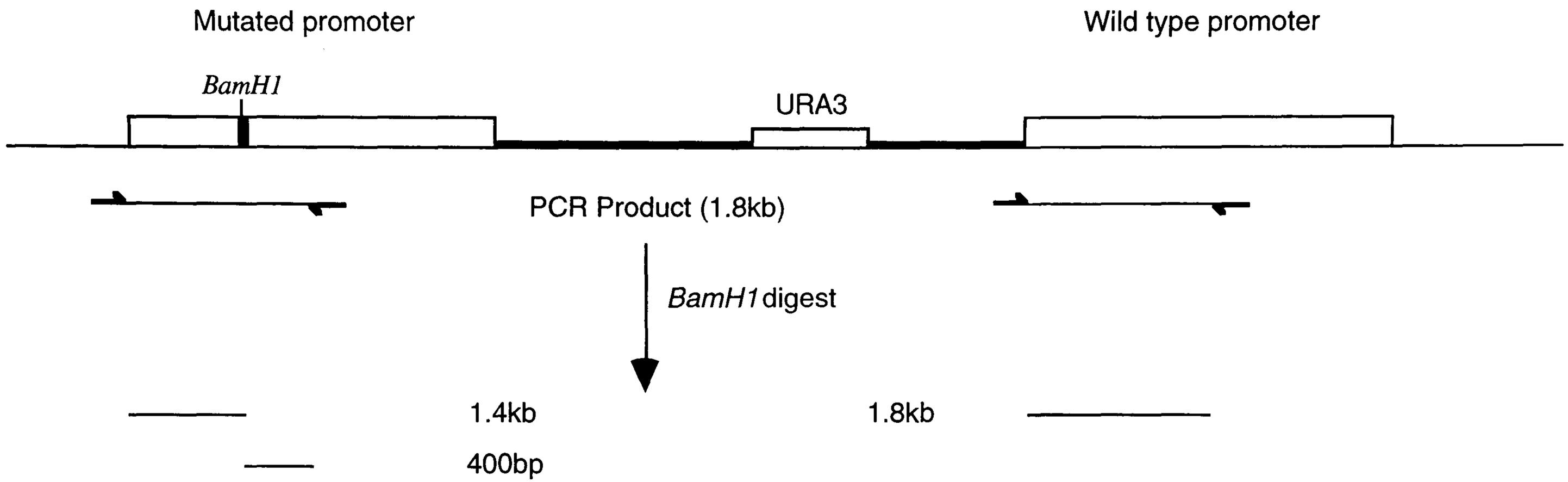
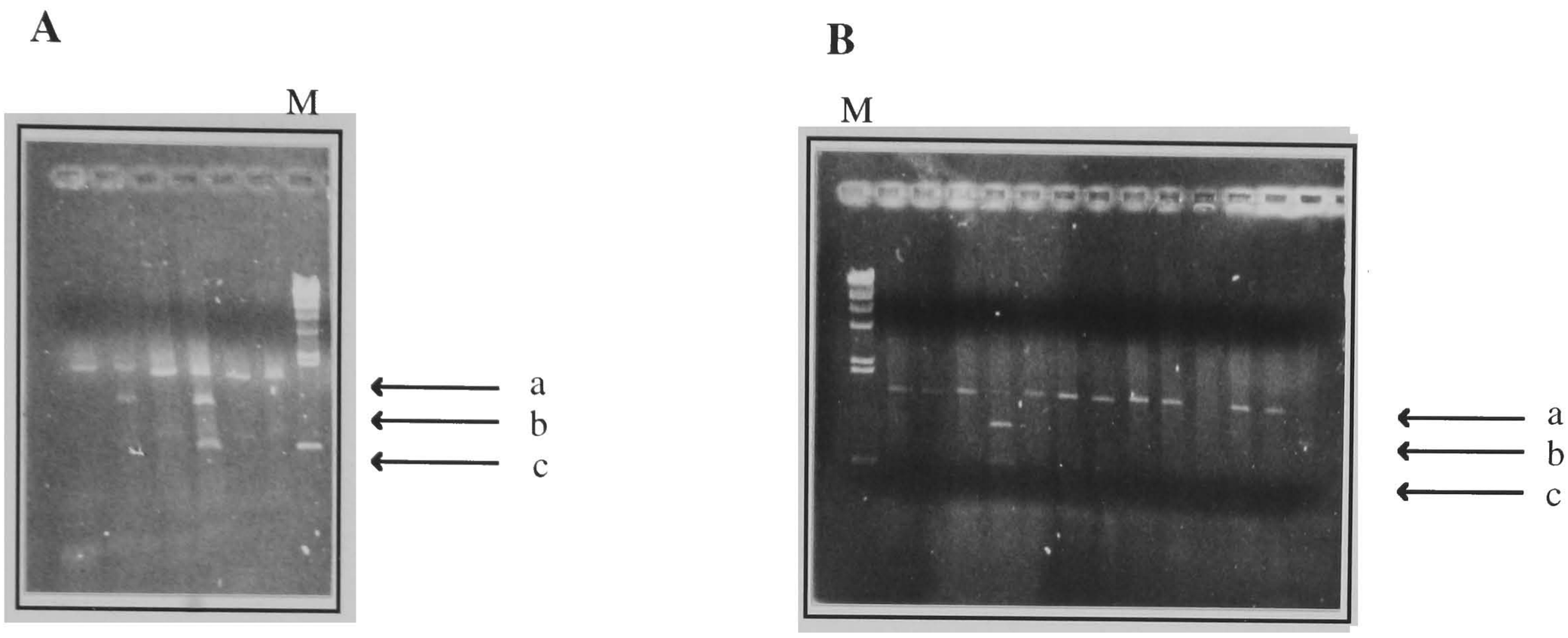


Figure 4.4 The PCR screen designed to ensure that the URA$^+$ Pop-In transformants had retained the transcription factor binding site deletion in the *PGK* promoter. Two *PGK* products were amplified from the two copies of *PGK* in the chromosome, these could be distinguished by a *BamH1* digest if gene conversion had not occurred during the initial integration.

YLP1



YLP2

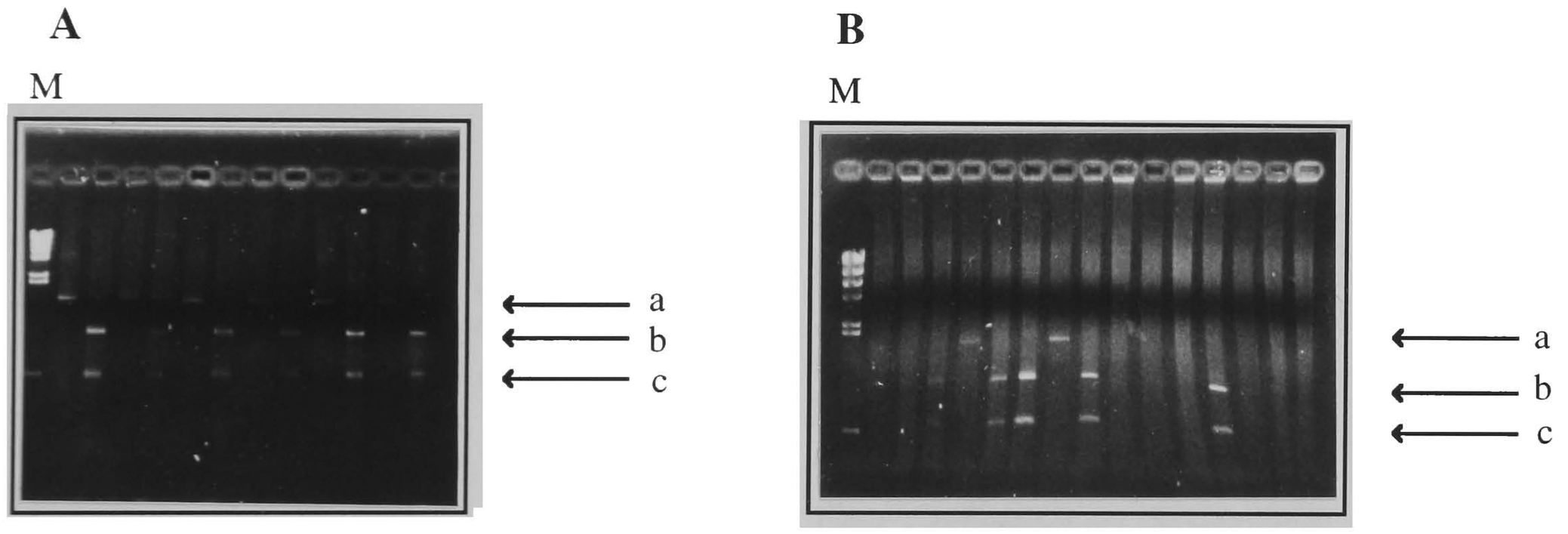
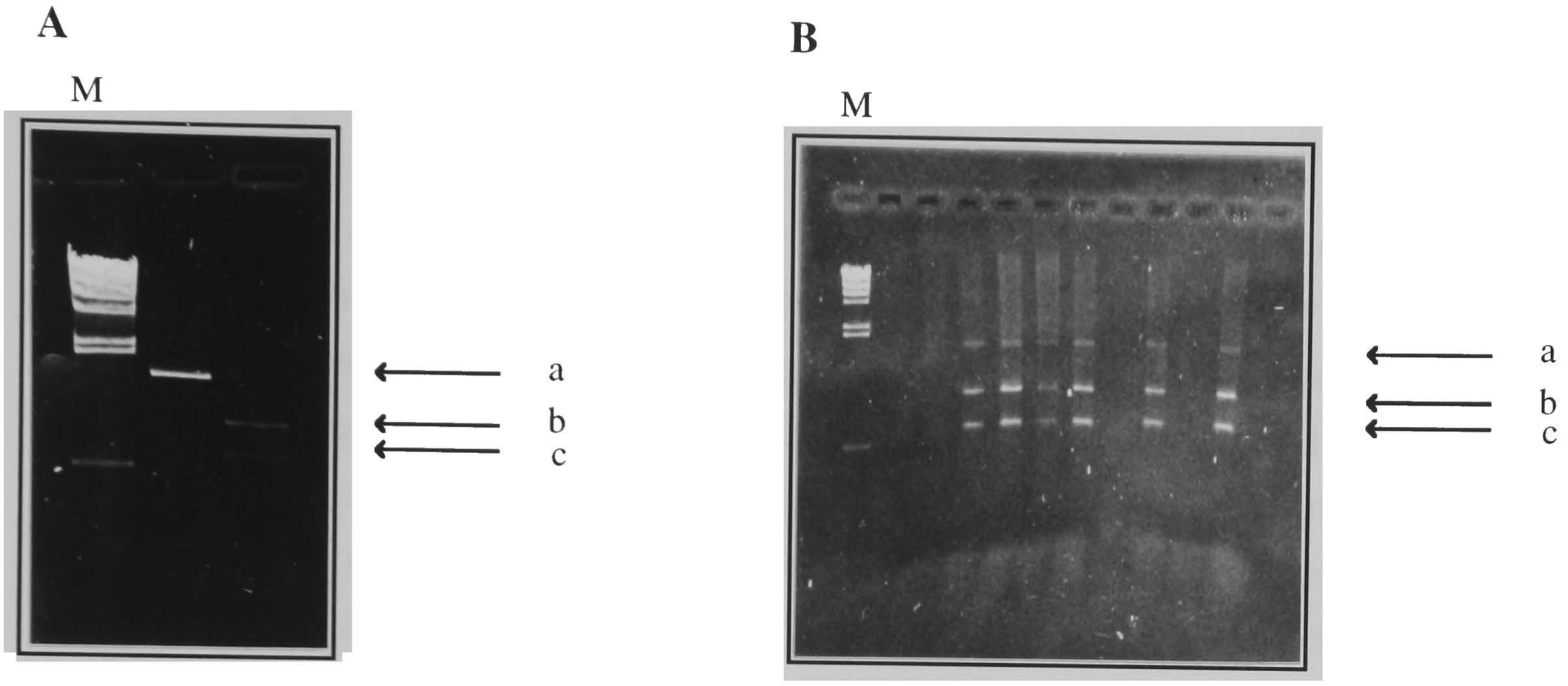


Figure 4.5A The PCR screen which allowed the identification of Pop In and Pop Out events, during the construction of yeast strains YLP1 and YLP2 containing deletions of Rap1p and Abf1p binding sites from the chromosomal *PGK* promoter. For each yeast strain, panel A shows a screen for Pop Ins, and panel B a screen for Pop Outs. Fragment a was cleaved with *BamH1* to give fragments b and c if the transcription factor binding site deletion had been maintained. M: λ/*HindIII* marker (23.1, 9.4, 6.5, 4.3, 2.3, 2.0, 0.5kb)

YLP3



YLP4

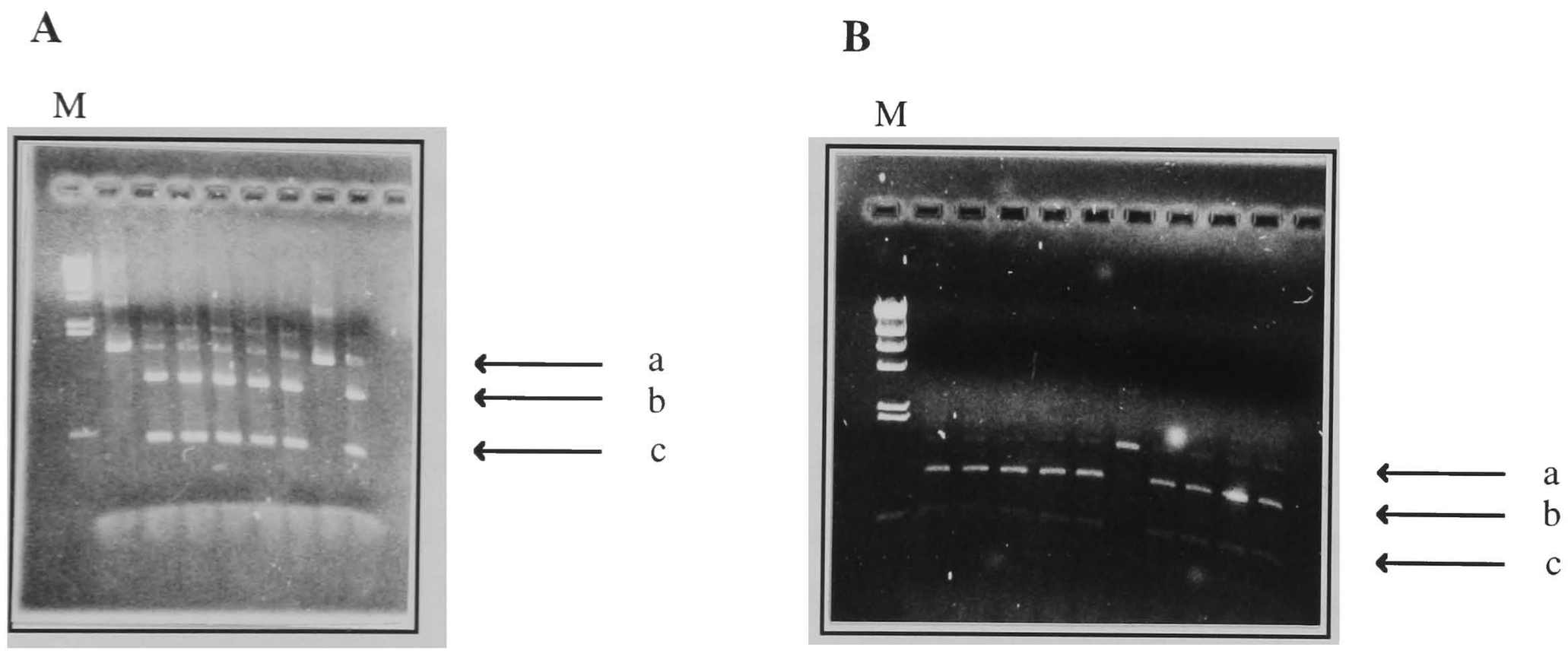


Figure 4.5B The PCR screen which allowed the identification of Pop In and Pop Out events, during the construction of yeast strains YLP3 and YLP4 containing deletions of Reb1p and yATF binding sites from the chromosomal *PGK* promoter. For each yeast strain, panel A shows a screen for Pop Ins, and panel B a screen for Pop Outs. Fragment a was cleaved with *BamH1* to give fragments b and c if the transcription factor binding site deletion had been maintained. M: λ/*HindIII* marker (23.1, 9.4, 6.5, 4.3, 2.3, 2.0, 0.5kb)

Different numbers of URA$^+$ Pop-In colonies had to be screened for each construct in order to find those which had not undergone gene conversion (ΔRap1p: 2 from 3 screened, ΔReb1p: 1 from 8, ΔAbf1p: 6 from 6 and ΔyATF: 9 from 10). This may reflect the increased efficiency with which the procedure of homologous integration was used as problems with the screening procedure were eliminated.

Once a Pop-In event had been identified, using the PCR screen, the cells from this colony were grown at 30°C overnight in 10ml YPD, and 100µl of this was spread on to 5-FOA SC agar. Colonies which grew on the 5-FOA medium had undergone a second recombinational event, losing one copy of *PGK* and also the intervening plasmid sequences (see Figure 4.1). A second round of PCR screening was used to check that the copy of *PGK* which had been lost was the wild type copy (Figure 4.5). The same primers were used as for the first screen but this time there should only be one target in the genome. Thus, if the mutated copy of *PGK* had been retained the amplified product was cleaveable with *BamH1*. As with the Pop-In stage there was a difference in the number of colonies which had to be screened to find one where the Pop-Out had occured correctly, *ie* retaining the mutation; ΔRap1p: 2 from 17, ΔReb1p: 6 from 10, ΔAbf1p: 4 from 10 and ΔyATF: 6 from 8.

As a final check that the strains were correct, chromosomal DNA from the integrant strains was cut with *BamH1* and compared with *BamH1* digested chromosomal DNA from the parent strain, DBY745, in a Southern blot. The blot was probed with the 2.95kb *HindIII* fragment from pB1 (Hitzeman *et al.* 1980) which contains the entire *PGK* sequence. In DBY745 *PGK* is present on a single *BamH1* fragment but in the integrant strains *PGK* spans two fragments since the deletion of the transcription factor binding sites has introduced an extra *BamH1* site. Initially the expected sizes for the fragments generated when chromosomal DNA was cut with *BamH1* were predicted using the sequence of yeast chromosome III which is in the database. The area surrounding the *PGK* locus was examined, and *PGK* was found to lie on a 10.9kb *BamH1* fragment. In the DNA from integrant strains digested with *BamH1*, this fragment should have been cleaved into 8.1 and 2.8kb fragments; all these fragments should have been resolved on a 1% agarose gel which was subequently blotted and probed with the *PGK* clone. However, although the blots showed one band for the DBY745 DNA and two for the integrant strains, the larger of the integrant bands appeared to be no different in size from the DBY745 band. This was because the position of restriction sites in the yeast from which the sequence of chromosome III in the database was obtained is not the same as those in DBY745. In DBY745 *PGK* is present on a fragment larger than 10.9kb; from the λ/*HindIII* marker which was run on the gel with the *BamH1* digested DNAs this fragment was estimated to be about 20kb. The small band released in the integrant

strains was about 5kb which left a fragment of 15kb. Bands of 15 and 20kb are hard to resolve on 1% agarose so a lower percentage gel was prepared. When the DNAs were digested and resolved on a 0.7% agarose gel the differences in size of all the bands was clear (Figures 4.6 and 4.7). Thus the presence of a mutated copy of *PGK*, containing a specific deletion of one of its transcription factor binding sites, at the chromosomal locus was confirmed.

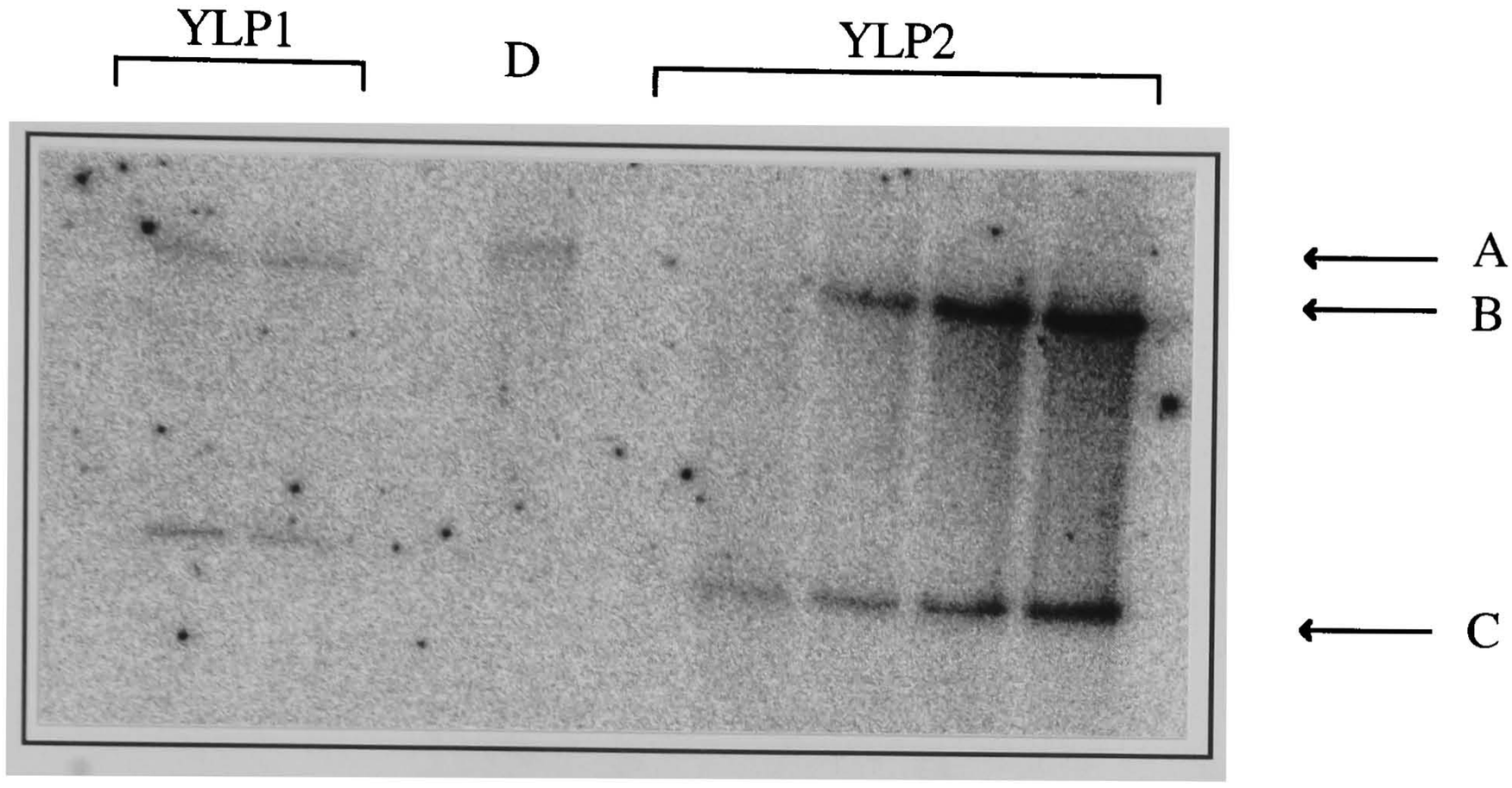
Once constructed these yeast strains, with specific transcription factor binding sites deleted from the promoter of the chromosomal *PGK* gene, were used to investigate the roles of Rap1p, Abf1p, Reb1p and yATF in the regulation of transcription at this locus. The role of Cpf1p was investigated by comparing the level of *PGK* RNA from a *cpf1* null strain, with the level from an isogenic wild type strain.

4.2.2 Rap1p is Central to the Activation of Transcription at *PGK* in the Chromosome

The effect of deleting specific transcription factor binding sites from the *PGK* promoter was determined by looking at the amount of RNA produced by the deleted strain and comparing it with the wild type level produced from the parental yeast DBY745. Strains DBY745, YLP1 (ΔRap1p), YLP2 (ΔAbf1p), YLP3 (ΔReb1p) and YLP4 (ΔyATF) (Table 4.1) were grown to mid-log phase in YPD, usually overnight at 30°C. The cells were harvested and total RNA was extracted (Materials and Methods, section 2.13). This was transfered to nitrocellulose and analysed by Northern blotting. RNA from the parental strain DBY745 was compared with RNA from more than one derivative of each of the deletion strains. Each filter was probed initially with a *PGK* specific probe, the *HindIII* fragment from pB1 which contains the entire *PGK* sequence (Hitzeman *et al.* 1980), and secondly with a probe for ribosomal RNA to act as a loading control (Petes *et al.* 1978; Figures 4.8 and 4.9).

It was clear just from an examination of the autoradiographs by eye (Figures 4.8 and 4.9) that there were differences in the levels of *PGK* RNA produced by the different strains; YLP1 contained very little *PGK* RNA, whilst YLP2, YLP3 and YLP4 contained less than DBY745. These experiments were repeated a number of times, and Figures 4.8 and 4.9 show data from representative autoradiographs. The differences in the levels of *PGK* RNA were quantified using a phosphorimager. First the amount of *PGK* probe hybridised to the filter was determined by scanning the filter. Then this probe was stripped from the filter which was next hybridised to the ribosomal probe. The amount of RNA loaded onto the gel was indicated by the ribosomal probe and thus the *PGK* counts could be normalised (Tables 4.2 and 4.3).

A



B (i)

(ii)

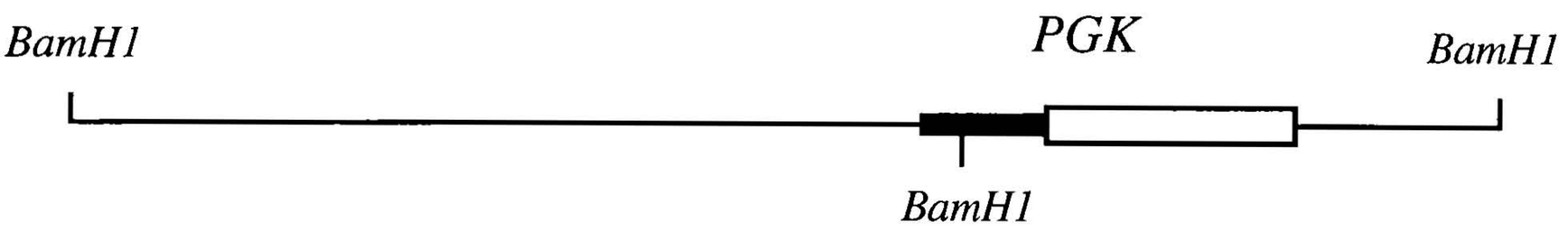


Figure 4.6A A Southern blot of *BamH1* digested chromosomal DNA from DBY745, YLP1 and YLP2, showing the presence of the extra *BamH1* site in the yeast strains constructed with a specific transcription factor binding site deletion from the *PGK* promoter. Band A in the wild type yeast (DBY745), is digested by *BamH1* to give bands B and C in the deletion strains YLP1 and YLP2. Digested DNA was electrophoresed on 0.7% agarose, and the Southern filter was probed with the entire *PGK* sequence. D: DNA from DBY745. **B(i)**: A map of the *BamH1* fragment from chromosome III on which *PGK* is situated, showing **(ii)**, the extra *BamH1* site present after deleting a specific transcription factor binding site.

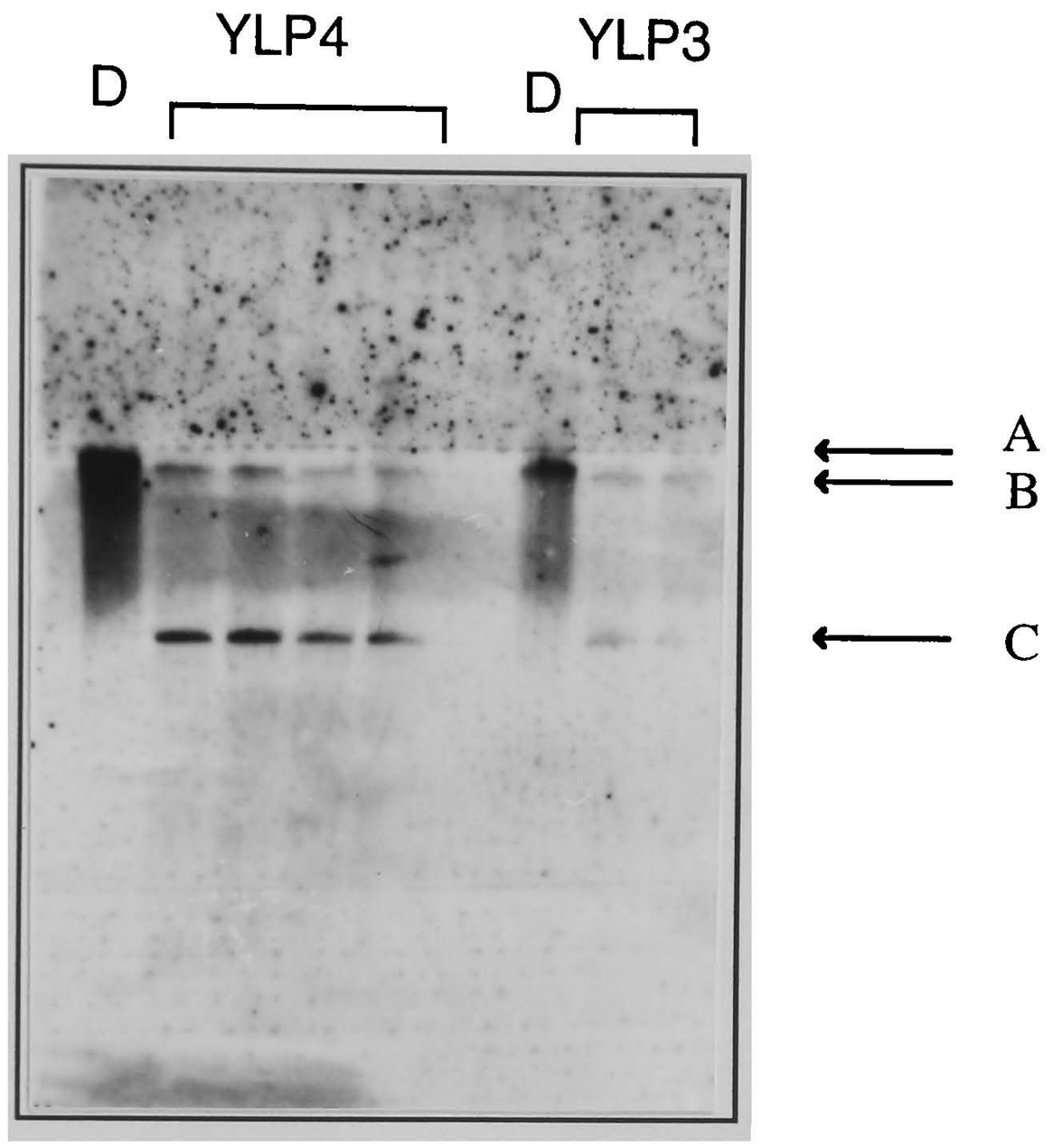


Figure 4.7 A Southern blot of *BamH1* digested chromosomal DNA from DBY745, YLP4 and YLP3, showing the presence of the extra *BamH1* site in the yeast strains constructed with a specific transcription factor binding site deletion from the *PGK* promoter. Band A in the wild type yeast (DBY745), is digested by *BamH1* to give B and C in the deletion strains YLP3 and YLP4. Digested DNA was electrophoresed on 0.7% agarose, and the Southern filter was probed with the entire *PGK* sequence. D: DNA from DBY745.

Table 4.1 Yeast strains constructed from the parental yeast DBY745 by homologous recombination. Specific deletions of transcription factor binding sites from the chromosomal *PGK* promoter are indicated, the co-ordinates of the deletions are relative to the *PGK* ATG.

Strain	*PGK* genotype	Deletion
DBY745	wild type	None
YLP1	ΔRap1p	-473 to -464
YLP2	ΔAbf1p	-538 to -473
YLP3	ΔReb1p	-561 to -552
YLP4	ΔyATF	-427 to -414

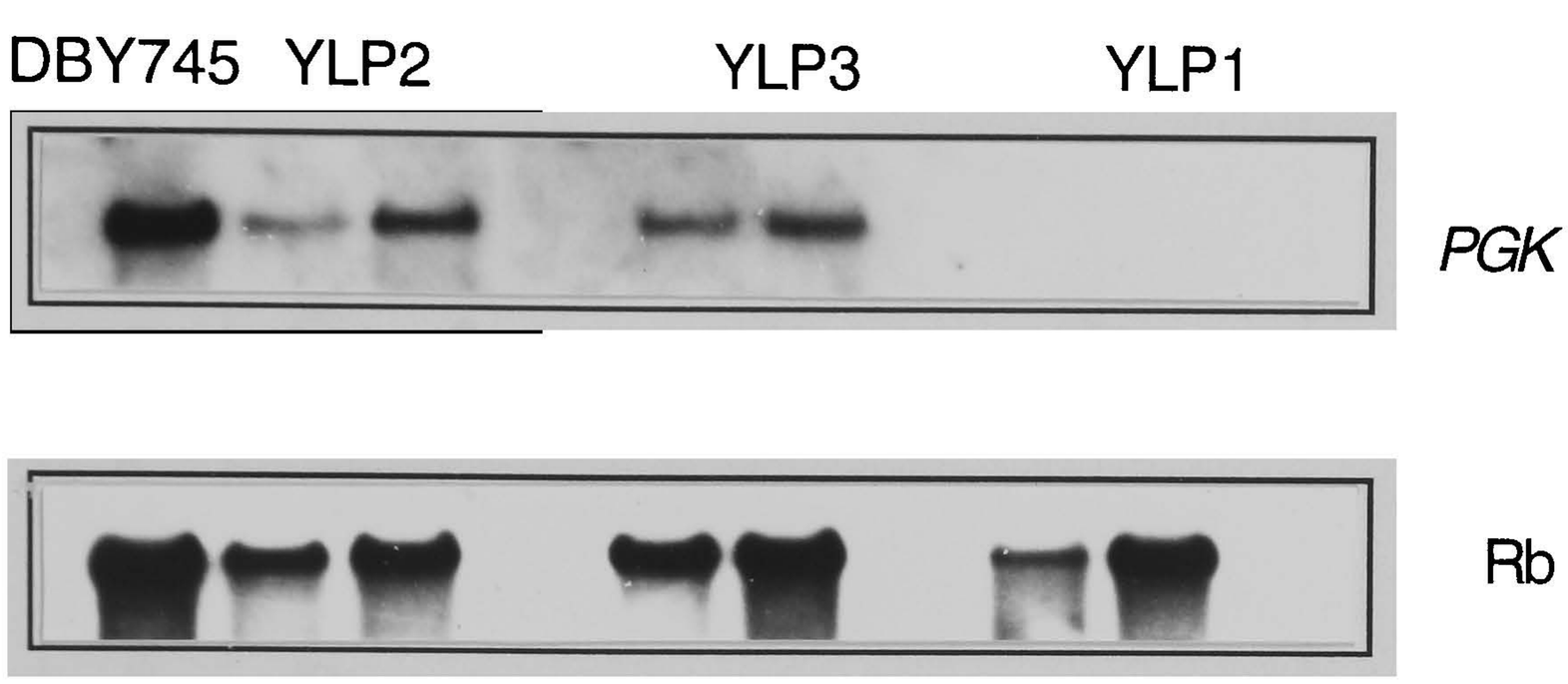


Figure 4.8 A Northern blot comparing the amount of *PGK* RNA from DBY745 with that from yeast strains YLP1, YLP2 and YLP3 which contain specific transcription factor deletions from the chromosomal *PGK* promoter.

Table 4.2 Phosphorimager analysis of Northern blots of chromosomal *PGK* transcription factor binding site deletions. *PGK* levels were normalised for the loading of RNA on the filter as estimated by the ribosomal counts. Results from other experiments using these strains suggest that the differences in *PGK* activity demonstrated are significant.

Strain	*PGK* counts	Rb counts	Corrected *PGK* [a]	% Activity
DBY745	1359455	2156845	1359455	100
YLP1	N/D [b]	698835		-
YLP1	N/D	1661690		-
YLP2	400010	1248563	691002	47
YLP2	430854	1581387	587639	
YLP3	245595	1115899	474693	48
YLP3	551648	1407853	845130	

[a] *PGK* levels were normalised for the loading of RNA on the filter as indicated by the ribosomal counts.

[b] N/D: not detected, the level of *PGK* RNA from YLP1 was so low that it was detected as less than the background of the filter by the phosphorimager. Other experiments using this strain show that in YLP1 the level of *PGK* mRNA is approximately 3-5% of wild type.

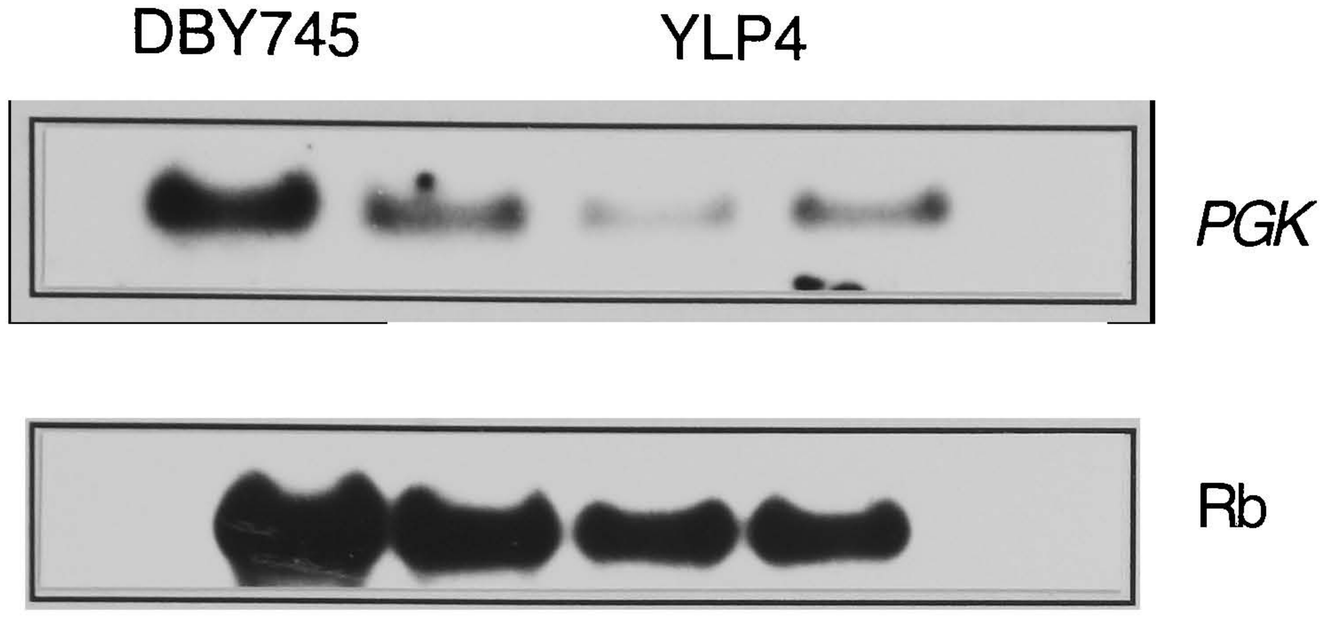


Figure 4.9 A Northern blot comparing the level of *PGK* RNA from DBY745 with that from yeast strain YLP4 which contains a deletion of the potential yATF binding site from the chromosomal *PGK* promoter.

Table 4.3 Phosphorimager analysis of Northern blots of chromosomal *PGK* deleted of its potential yATF site. The results shown are from one experiment using three independently isolated deletion strains. The spread of results from this and other experiments with the same strains suggests that the activity of *PGK* in YLP4 does not significantly differ from wild type *PGK* activity.

Strain	*PGK* counts	Rb counts	Corrected *PGK* [a]	% Activity
DBY745	1370597	3648735	1370597	100
YLP4	763308	2593114	1074040	
YLP4	457526	1770089	943111	79
YLP4	586731	1737933	1231823	

[a] *PGK* levels were normalised for the loading of RNA on the filters as indicated by the ribosomal counts.

The phosphorimager analysis showed that when the binding site for Rap1p was deleted from the *PGK* promoter there was a dramatic reduction in the amount of RNA produced; the level of *PGK* RNA in yeast strain YLP1 was so low that the phosphorimager was unable to detect an increase over the background level in the position where the radioactive *PGK* band should have appeared on the filter. The lack of *PGK* message could have been due to a lack of RNA on the filter but, as the ribosomal probe shows, this was not the case. Indeed, when more than 20µg of YLP1 RNA was electrophoresed, a concentration of RNA at which the gel was almost overloaded, it was possible to detect a small amount of *PGK* mRNA even in the absence of a Rap1p binding site in the promoter (Data not shown). Deletions of the binding site for either Abf1p or Reb1p caused an approximately 50% decrease in the level of transcription from *PGK* (Table 4.2), whilst after correcting for the loading of the gel, deletion of the potential binding site for yATF resulted in a drop of 20% when compared with the wild type level (Table 4.3).

4.2.3 The Role of Cpf1p at the Chromosomal *PGK* Locus

Cpf1p is not encoded by an essential gene so it was not necessary to construct a yeast strain containing a deletion of the Cpf1p binding site from the *PGK* promoter in order to investigate the role that this site plays in regulating transcription at this locus. A *cpf1* null strain, YAG93, has been constructed in which the Cpf1p coding region has been deleted from the chromosomal locus (Kent *et al.* 1993). YAG93 was a gift from Nick Kent (Oxford University) enabling a direct comparison to be made between the levels of *PGK* RNA in DBY745, which has both a wild type *PGK* gene and a wild type *CPF1* gene, and in YAG93, which is isogenic to DBY745 except for the *cpf1* null mutation.

DBY745 and YAG93 were grown to mid-log phase at 30°C overnight in YPD. The cells were harvested and RNA was extracted and analysed by Northern blotting. The filters were probed with the *PGK* specific probe, the *HindIII* fragment from pB1, and with a ribosomal probe to act as a loading control (Figure 4.10). They were then quantified using a phosphorimager (Table 4.4). After correcting for the loading of the filter the *cpf1* null strain was found to have a 29% increase in the level of *PGK* mRNA.

4.2.4 Is the *PGK* Promoter Bidirectional?

When the *Saccharomyces cerevisiae* chromosome III sequence was published (Oliver *et al.* 1992) it was possible to look at the position of *PGK* and its promoter in relation to other genes close to it on the chromosome. As shown in Figure 4.11A the

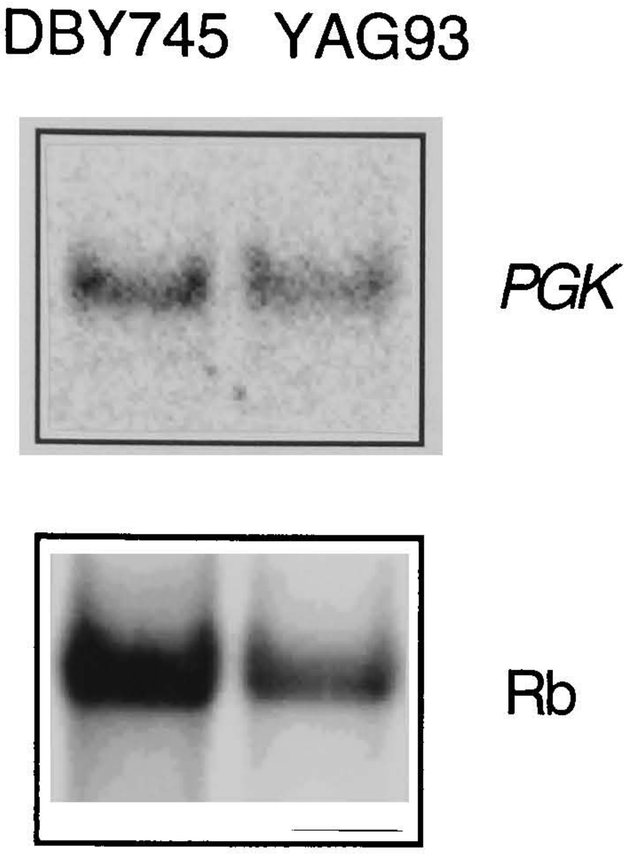


Figure 4.10 A Northern blot comparing the level of *PGK* RNA from DBY745 with that from YAG93 which is *cpf1*$^-$ but otherwise isogenic to DBY745.

Table 4.4 Phosphorimager analysis of Northern blots of chromosomal *PGK* levels in DBY745 and the *cpf1* null strain YAG93. The table shows figures from one representative experiment, each time that levels of chromosomal expression were examined in the *cpf1* null strain a similar increase was seen.

Strain	PGK counts	Rb counts	Corrected PGK	% Activity
DBY745	85845	46435821	85845	100
YAG93	73741	31028160	110358	129

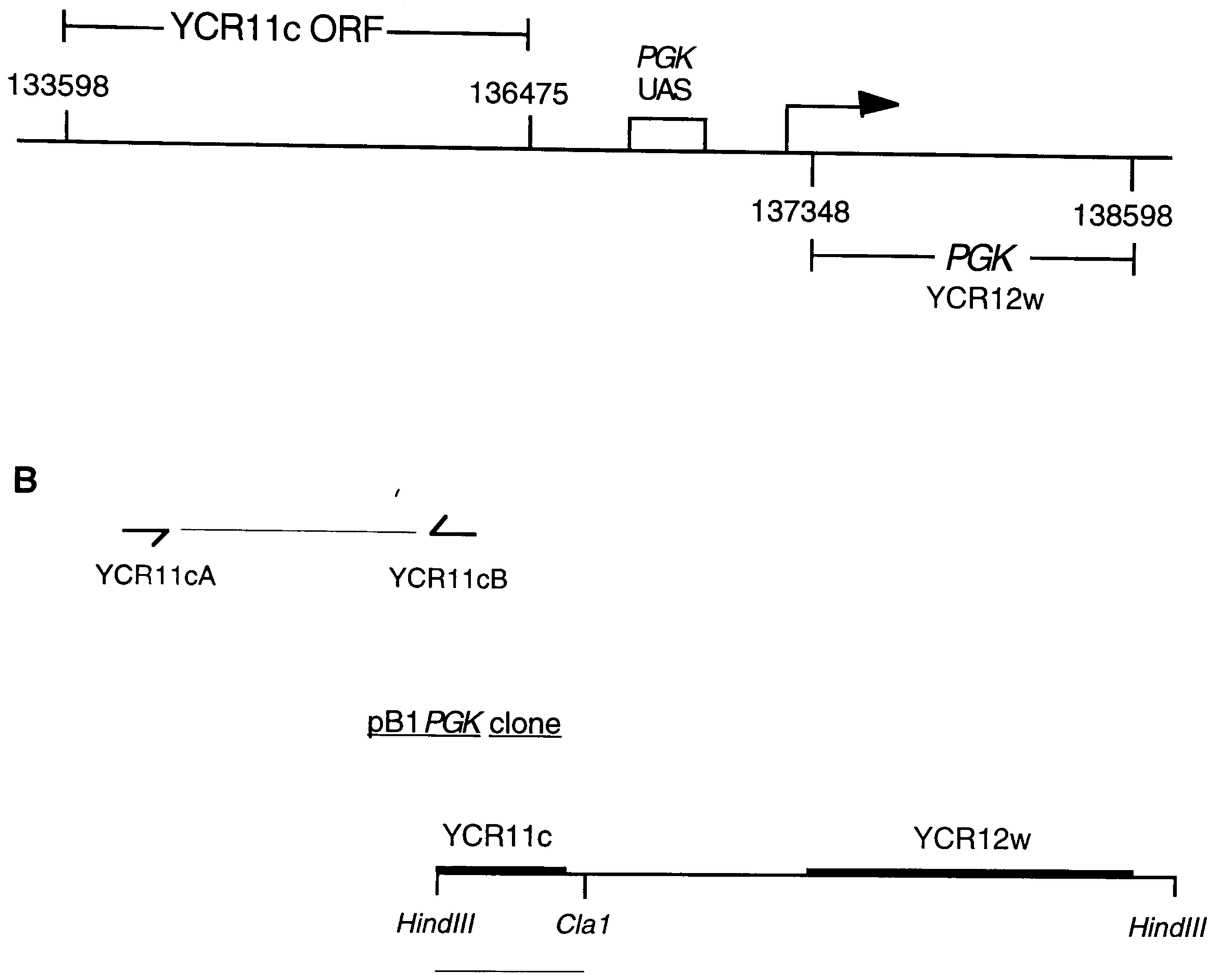


Figure 4.11 **A** shows the region of chromosome III containing YCR12w, the ORF encoding phosphoglycerate kinase, and YCR11c an unassigned ORF upstream of *PGK*. The numbers refer to the positions of these sequences on chromosome III. **B** shows the position of primers used to construct a YCR11c probe by PCR, and the *HindIII/Cla1* digest of pB1 which released a shorter probe containing 600bp of YCR11c. The relative positions of YCR11c and YCR12w on the *HindIII PGK* clone are indicated.

PGK UAS was found to be situated midway between two open reading frames, YCR12w, encoding phosphoglycerate kinase, and YCR11c which is an unassigned ORF of 2877bp, with homology to *ADP1*, a probable ATP-dependent permease precursor. Interestingly, YCR11c is an open reading frame which runs in the opposite orientation to that of *PGK*. This raised the possibility that the *PGK* UAS is bidirectional, like UAS_{GAL} which controls both *GAL1* and *GAL10*. To determine whether transcription from both *PGK* and YCR11c is controlled by the same UAS, two probes were made to the YCR11c ORF (Figure 4.11B) and these were used to probe Northern filters containing total yeast RNA.

4.2.5 YCR11c Probes

A probe was amplified from chromosomal DBY745 DNA (Copy Number Prep, Materials and Methods, section 2.6.2) using primers YCR11cA and YCR11cB (Table 2.1A) which are homologous to positions +407 and +2943 in the YCR11c ORF. A product of about 2.3kb was produced after a touchdown PCR programme and 30 cycles annealing at 55°C. This YCR11c probe was isolated from an agarose gel prior to random oligolabelling.

A second, shorter, YCR11c probe was isolated from pB1. This was the 660bp between the *Cla1* site upstream of the Reb1p site in the *PGK* UAS and the *HindIII* site at the end of the *PGK* clone. This contains the first 612 bases of the YCR11c ORF. After isolating this fragment from agarose it was labelled by nick translation.

On several occasions Northern filters containing RNA from DBY745, and also the yeast strains with deletions from the *PGK* promoter, which had previously been probed for both *PGK* and ribosomal RNA, were probed with one or other of the YCR11c probes but no message from the YCR11c ORF was detected under conditions in which *PGK* mRNA could be detected. In case the amount of RNA remaining on the filters was considerably reduced after they had been stripped twice, fresh filters were also probed with the YCR11c probes but again no message could be detected (Data not shown).

4.3 Discussion

The yeast strains whose construction has been described above, YLP1-4 (see Table 4.1), were made with the intention of using them to investigate the role of specific transcription factors at the *PGK* promoter. They facilitated the search for answers to questions about what effect the removal of a transcription factor binding site, and hence the transcription factor, would have on transcriptional regulation. Since the

only change to the promoter is the one introduced, and there are no considerations about the effect of plasmid copy number or the effect of plasmid chromatin structure to be made, any difference between the level of transcription from the wild type gene and the mutated gene can be attributed to the lack of a specific factor.

However, these strains are also useful for looking at other aspects of transcription related to the *PGK* promoter. The transcription factors important for carbon source regulation of *PGK* can be investigated at a chromosomal level rather than using plasmids, by growing the deletion strains in different carbon sources. Another use for the deletions which have been made, is to use them to look for evidence of protein-protein interactions between transcription factors which bind to the promoter and mediator proteins thought to affect transcription at this locus. The deletions could be integrated into yeast which are for example *gal11*$^{-}$ or *gcr1*$^{-}$ to see whether a particular transcription factor is required for these proteins to exert their effects.

Alternatively a strain in which the level of transcription from the *PGK* locus was so low as to be barely detectable (YLP1, Figure 4.8) could be used as a parental strain for the integration of other *PGK* related constructs at a non-homologous locus. For example, Dr Ian Graham is trying to identify the regions of Rap1p which are important in its role as a transcriptional activator by performing an experiment in which the Rap1p binding site in the *PGK* promoter has been replaced with a LexA binding site. This construct will be integrated at the *Leu2* locus of YLP1. The resulting strain will then be transformed with plasmid borne LexA::Rap1p fusions and if the region of Rap1p in the fusion is sufficient to activate transcription this will be detected by Northern blotting. As the *PGK* mRNA from the disrupted chromosomal gene is barely detectable it will not be necessary to distinguish between the chromosomal *PGK* message and the message from the reporter *PGK*. It would be possible to perform this experiment in a strain with wild type chromosomal *PGK* and to integrate a truncated reporter, however, a truncated gene might not contain all of the information necessary for making a stable RNA.

The transcription factor binding site deletions from the chromosomal copy of the phosphoglycerate kinase gene confirmed the central role of Rap1p in the control of transcription at this locus, and also raised questions about roles for Abf1p and Reb1p. The amount of *PGK* RNA from strain YLP1 on the filter was so low that it was not possible to quantify it, but since ribosomal RNA was present in those tracks we can be sure that the lack of *PGK* message was not a consequence of there being no RNA present on the filter. Quantification of other filters suggest that the level of *PGK* mRNA in YLP1 is about 3-5% of wild type. Since a deletion of the Rap1p binding site from the *PGK* promoter essentially knocks out the production of *PGK*

mRNA, it might be expected that yeast cells carrying this deletion would not be able to grow as well on glucose medium as wild type cells. In the experiments described above all the strains were grown on a complete yeast medium supplemented with glucose (YPD). Under these conditions no difference in the growth rate of any of the *PGK* promoter mutant strains was apparent when compared with the growth of wild type DBY745. However, when strains were grown on minimal medium (SC) supplemented with glucose, and only adenine, leucine and uracil rather than a rich amino acid mix, the doubling time of YLP1 was found to be slower than that of DBY745 (L. Jenkins, pers comm). This suggests that a lack of *PGK* does affect the yeast cell when glucose is the sole carbon source. YPD medium contains various components which yeast can utilise for growth rather than relying on the metabolism of glucose.

The deletions of the Abf1p binding site from the UAS, and of the Reb1p binding site from the upstream region of the *PGK* promoter also have an effect on the level of RNA, suggesting roles for both of these transcription factors at the PGK locus; deletion of either binding site caused an approximately 50% reduction in the level of *PGK* RNA. The result for the deletion of the Abf1p binding site is in contrast with the findings of earlier experiments which suggest that deletion of the Abf1p binding site from the *PGK* UAS had no detectable effect on the levels of RNA (Chambers *et al.* 1988). However, this earlier study looked at the effects of deletion of the Abf1p binding site when the *PGK* gene was carried on a multicopy plasmid. This situation is more artificial than looking at the chromosomal locus; the gene is present in high copy number which may mask small differences in levels of transcription, and the higher chromatin structure of plasmid DNA may differ from the higher chromatin structure of chromosomal DNA. A difference in chromatin struture may affect the functioning of proteins which cause an alteration in chromatin structure in order to exert their effects. The different conclusions, about the effect of deleting the Abf1p binding site from the *PGK* promoter drawn by the previous study and this work, may be a result of the difference in chromatin structures found between a plasmid and chromosome environment. Alternatively they may simply be a result of the advances that have been made in increasing the sensitivity of methods used for the quantification of DNA and RNA levels. In the earlier study quantification methods were not as sensitive as those available today, this may mean that a small change in the level of RNA was not detected.

In contrast to the deletions of Rap1p, Reb1p and Abf1p binding sites from the *PGK* promoter, examination of transcription in a yeast strain lacking Cpf1p (YAG93) showed an increase in the activity of *PGK*. This difference may be a result of the approaches used to examine the transcription factors; deleting the transcription factor

binding sites from the promoter preventing the protein from binding is not equivalent to looking at transcription from the gene when the binding site is present but the transcription factor is not. The absence of Cpf1p from the yeast cell may affect other genes which in turn affect *PGK*. However, each time transcription of *PGK* in YAG93 was measured there was an increase in activity which could indicate a role for Cpf1p in which it acts as a repressor of transcription.

These results suggest that the Rap1p binding site in the phosphoglycerate promoter is essential for the activation of transcription at this locus, and that the binding sites for Reb1p and Abf1p play an important role in allowing the promoter to activate at the highest level. The small reduction in the amount of transcription seen when the potential binding site for yATF is deleted may indicate that this site does have some role in the regulation of transcription of *PGK*, although the range of transcriptional activities seen in experiments with YLP4 suggest that this is not a significant result. Also the introduction of the *BamH1* linker in place of the potential yATF site has disrupted the 3' flanking sequence of the most downstream of the two Gcr1p binding sites in the UAS. Three of the bases in this region of the Gcr1p binding site no longer match the consensus binding site (Huie *et al.* 1992) This could result in a decrease in the level of transcription, due to the disruption of the Gcr1p interaction.

No YCR11c message was detected by probing Northern filters containing total yeast RNA using conditions that allow the detection of *PGK* RNA. This suggests that either YCR11c is not expressed under the growth conditions used in these experiments *ie* on YPD containing 2% glucose, or that it is expressed at a level too low to be detected by our blotting proceedure. If YCR11c was under control of the *PGK* UAS it might be expected to be expressed at a similarly high level.

However there is an example of two genes under control of the same UAS being expressed at different levels. Expression of *GCY1* and *RIO1* is controlled by a bidirectional promoter containing a Reb1p site and UAS_{GAL}, but whilst *GCY1* expression is regulated by Gal4p, expression of *RIO1* is constitutive and independent of Gal4p. This difference in expression is due to the basal promoter elements rather than the UAS components; the *GCY1* basal promoter contains a consensus TATAA sequence but *RIO1* contains a non consensus TATAGA sequence. YCR11c however appears to contain a consensus TATA sequence 137bp upstream of the ATG.

Chapter 5

The Effect of Transcription Factor Binding Site Deletions on Regulation of a Plasmid-Borne Copy of *PGK*, and the Use of a Minimal Promoter to Determine Roles for Transcription Factors.

5.1 Introduction

As shown in Chapter 4, a deletion of the Abf1p binding site from the chromosomal *PGK* UAS causes a decrease in the level of *PGK* transcription of approximately 50%. This result is in direct contrast with earlier experiments in which the Abf1p binding site deletion was not seen to have any effect on *PGK* transcription (Chambers *et al.* 1988). However, this previous experiment was investigating the effects of the deletion on transcription from a multicopy plasmid-borne *PGK* gene construct. It is possible that the use of the multicopy plasmids masked the effects of the binding site deletion because the high number of plasmids in the cell may have resulted in the production of such high levels of *PGK* RNA that a small reduction in the amount of transcription from each plasmid was not detectable. Alternatively, the difference in requirement of Abf1p for maximal *PGK* transcription could depend on the structural environment of the gene; Abf1p might only be able to exert its effects when *PGK* is in a chromosomal rather than episomal state. A further explanation for these differing results could be a direct consequence of the change in methods used for the quantification of blots. Before densitometric and phosphorimager analyses were available, differences in the intensity of bands on autoradiographs were determined visually, thus small differences between constructs may have been overlooked.

In the light of these possibilities it was important to repeat the experiment in which the effect of deleting the Abf1p binding site from a multicopy plasmid-borne copy of *PGK* was investigated. Since both Abf1p and Reb1p are thought to have functions which involve chromatin remodelling, and both are weak transcriptional activators, they may perform similar functions in the *PGK* promoter. At the *ILV1* promoter Abf1p and Reb1p binding sites have been shown to be functionally interchangable (Remacle and Holmberg 1992). The similarities between these two multifunctional transcription factors prompted an investigation into the effects of a multicopy Reb1p binding site deletion at the same time as repeating the Abf1p experiment. As a deletion of the Rap1p binding site had a dramatic effect on the levels of *PGK* RNA whether it is from a plasmid-borne copy of the gene (Chambers *et al.* 1988) or from the chromosomal locus (this work), this part of the experiment was not repeated.

A second way of looking at the roles of transcription factors to see whether they are capable of activating transcription is to introduce binding sites into a minimal promoter reporter plasmid. A minimal promoter consists of TATA box and RNA initiation site sequences that are sufficient to promote a basal level of correctly initiated transcription. This minimal promoter is usually used to drive expression of a reporter gene. The reporter may be the coding region for an enzyme whose levels within the cell are easy to assay *eg lacZ*, or it could be a heterologous coding region whose expression can be determined by RNA analysis. Such a plasmid can be used to determine the activation potential of DNA fragments containing transcription factor binding sites. This is an artificial system where DNA sequences are cloned immediately upstream of the basal promoter which alters the spacing found in wild type promoters, but it allows a rapid assessment of the strength of transcription factor binding sites. If the reporter gene is *lacZ*, transformants can be grown on indicator plates containing X-gal, which is a substrate for β-galactosidase. In the presence of β-galactosidase, X-gal is metabolised giving a blue colour. The white/pale blue colonies of transformants containing just the minimal promoter plasmid become dark blue if a binding site for a transcriptional activator is cloned upstream of the basal elements. This crude assessment can then be quantified by preparing protein extracts from the transformants, and assaying these for β-galactosidase activity.

A minimal promoter plasmid was used in this work to see whether the two newly identified transcription factor binding sites for Reb1p and Cpf1p in the *PGK* promoter were able to increase the level to which the *PGK* UAS was able to activate transcription. It was also used to investigate a role for the potential transcription factor binding site which had previously been identified for yeast Activating Transcription Factor (yATF, Lin and Green 1989) at the 3' end of the UAS. Finally, a minimal promoter construct was transformed into two strains of yeast, W303-1A and R884-1C, which are *GAL11* and *gal11* respectively, to look for evidence of a protein/protein interaction between Gal11p and Rap1p. Thus one simple system can be used in a variety of ways to provide preliminary evidence which can then be verified in further experiments.

5.2 Results

5.2.1 Multicopy Plasmid-Borne Deletions of the Reb1p and Abf1p Sites from the *PGK* UAS

The Reb1p site was deleted from a multicopy plasmid borne copy of *PGK* by cloning the PCR products amplified using primer pairs 521B/PGKD and 533B/PGKC2 (see Table 2.1A for primer sequences and co-ordinates) into the *BamH1* site of pAJ6, a leucine selectable yeast shuttle vector, giving pAJ112 (Figure 5.1). This was sequenced across its UAS to ensure that mutations had not been introduced by the process of PCR. These PCR products (Figure 4.3) are the same as those which were cloned into pAJ735 to make the ΔReb1p integrating vector. A deletion of the Abf1p site from the *PGK* UAS already existed on the multicopy plasmid pKV502 (Chambers *et al.* 1988). In addition, a control plasmid for these experiments was pMA27, which is a multicopy plasmid bearing a wild type copy of the *PGK* gene (Mellor *et al.* 1983).

The leucine-selectable plasmids pKV502, pMA27 and pAJ112 were transformed into DBY745 using the One Step method (Chen *et al.* 1992). Two transformants of each plasmid were selected at random and grown to mid-log phase in SC-leu. RNA was extracted and examined by Northern blotting. The filters were probed with a *PGK* specific probe which consisted of the entire *PGK* sequence isolated as a *HindIII* fragment from pB1 (Hitzeman *et al.* 1980), then a ribosomal probe (Petes *et al.* 1978) to act as a loading control, and finally with a *leu2* probe. The yeast marker on these plasmids, *leu2d*, is a truncated version of the leucine gene, so the *leu2* probe recognised the *leu2d* RNA message from the plasmids, allowing an estimation of copy number to be made.

The results of the different probings can be seen in Figure 5.2. After the filters were scanned using a phosphorimager, the relative *PGK* and ribosomal counts for each plasmid were determined. Since the plasmid copy numbers were similar, compensating for them had very little effect on the *PGK* counts corrected for ribosomal loading (Table 5.1). After these corrections had been made, the plasmid carrying a deletion of the Reb1p binding site was found to be as active as the wild type construct, and the deletion of the Abf1p binding site resulted in only a slight increase in transcriptional activity.

5.2.2 The Effect of Deleting the Potential yATF site from the *PGK* Promoter

A possible role for the potential yATF site in the *PGK* UAS was examined by making a deletion of this sequence from the *PGK* promoter. As with the deletion of the Reb1p site, this was achieved using PCR (Figure 4.3). PCR primers 521A and 521B (Table 2.1A) were used to amplify the region immediately upstream of the potential yATF site using pKV521 (Chambers *et al.* 1988) as a template. After a touchdown PCR a product of 1.4kb had been amplified. The region of *PGK*

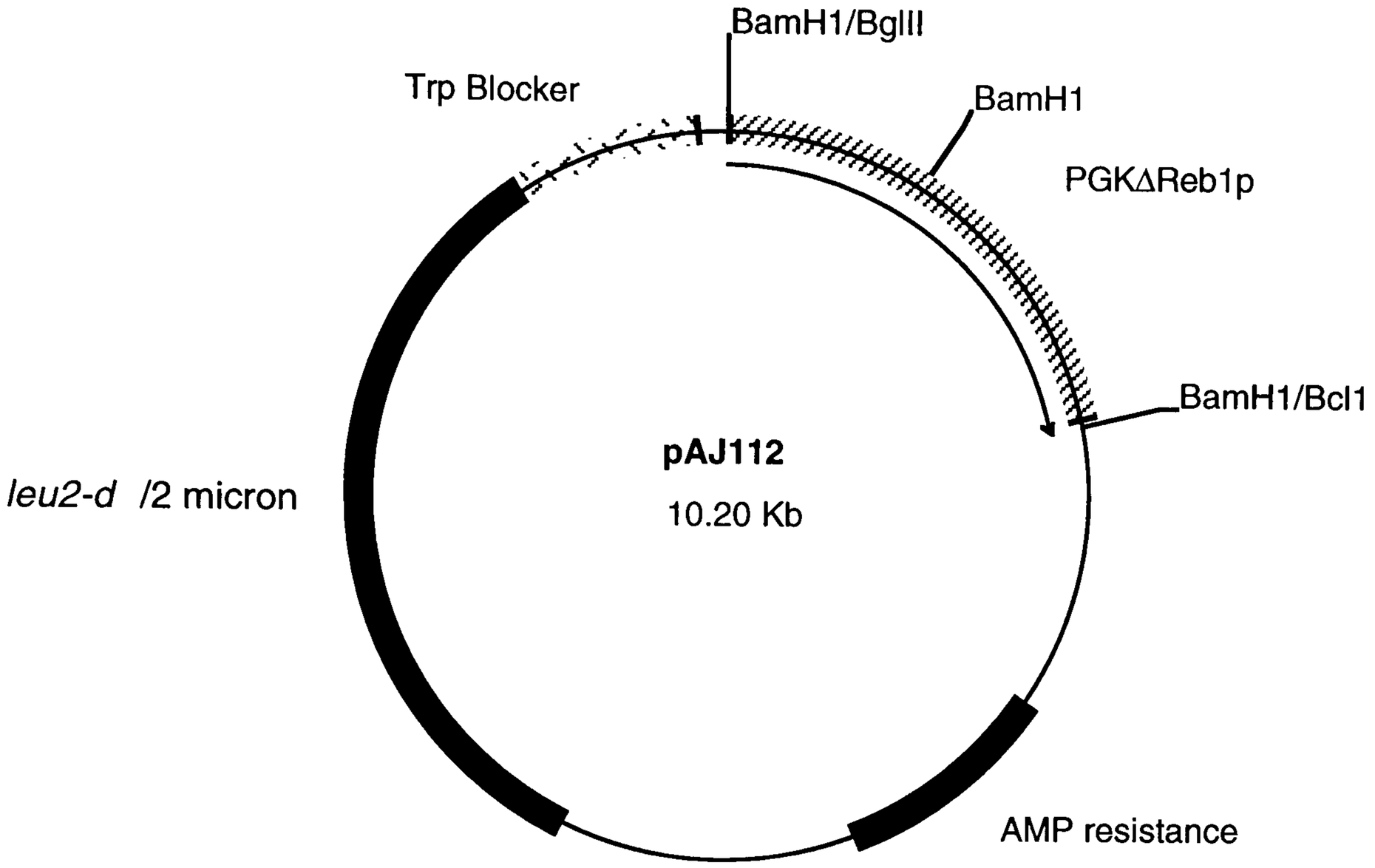


Figure 5.1 Multicopy plasmid pAJ112 which contains the entire *PGK* sequence but with the Reb1p binding site (-562 to -552) deleted and replaced with a *BamH1* linker. This is a shuttle vector with an origin of replication and ampicillin resistance gene for growth in *E.coli*, and *leu2-d*/2 micron sequences for maintenance in yeast.

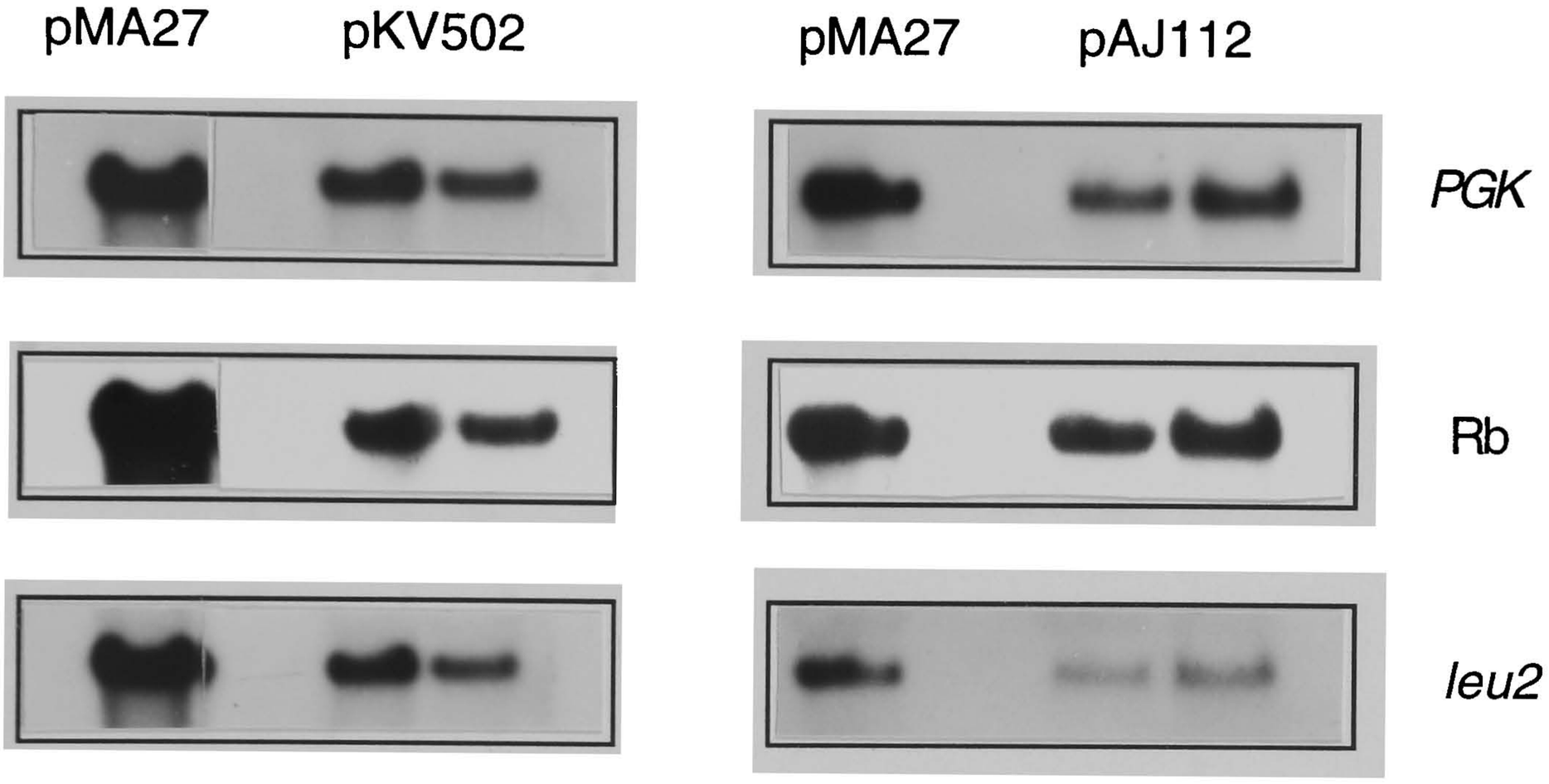


Figure 5.2 Northern blot of RNA from DBY745 transformed with pMA27, pKV502 and pAJ112 probed with *PGK*, a ribosomal probe to act as a loading control, and *leu2* for an estimation of plasmid copy number.

Table 5.1 Phosphorimager analysis of Northern blots of RNA from DBY745 transformed with multicopy plasmids pMA27, pKV502 and pAJ112. The levels of *PGK* RNA were corrected for the loading of the gel and for plasmid copy number.

Transformant	*PGK* counts	Rb counts	Copy Number	Corrected *PGK*	% Activity
DBY745+pMA27	3476124	3642694	1	3476124	100
DBY745+pKV502	1433362	1280813	1.2	3397125	
DBY745+pKV502	810774	817271	0.8	4517169	114
DBY745+pMA27	760731	6803282	1	760731	100
DBY745+pAJ112	312981	3267345	0.8	775283	
DBY745+pAJ112	455725	4978188	0.8	762198	101

downstream of the yATF sequence was amplified on a 1.9kb fragment using PCR primers 533A and 533B, with pKV533 (Chambers *et al.* 1988) as a template. The upstream product was cut with *BamH1* and *Bcl1*. The two products were then cloned into the *BamH1* site of pAJ6. This formed pAJ110, a multicopy plasmid containing the entire *PGK* sequence but with a *BamH1* linker replacing the potential yATF binding site. The plasmid was sequenced in the region of its UAS to ensure that no mutations had been introduced by the process of PCR.

DBY745 was transformed with the multicopy plasmid pAJ110, and with the control plasmid pMA27, containing the wild type *PGK* sequence, using the One Step method of transformation. 100ml SC-leu cultures innoculated with three transformants each of pAJ110 and pMA27 were grown overnight. RNA was prepared from one 50ml aliquot of harvested cells whilst the other was frozen at -80°C for later plasmid copy number analysis. The RNA was electrophoresed, blotted and the filter probed firstly with the *PGK* specific probe from pB1 (Hitzeman *et al.* 1980), and then with a ribosomal probe (Petes *et al.* 1978) to act as a loading control (Figure 5.3). The filters were then scanned with a phosphorimager to determine the relative number of counts in each lane (Table 5.2). The plasmid copy numbers of the constructs were also determined. In this instance, the plasmid specific probe selected was a clone of *PGK*. This will also recognise the chromosomal copy of *PGK*, but as there is only one copy of this compared to 50-100 copies of plasmid-borne *PGK*, the chromosomal DNA was not visible on the probed filter. Also, since the growth conditions for all the cells were the same, the amount of message from the chromosomal *PGK* gene should be constant. The ratio of plasmid to chromosomal counts was then calculated (Table 5.2).

After the *PGK* counts had been corrected for ribosomal loading and for plasmid copy number, an average *PGK* activity was calculated for each plasmid. The construct containing the potential yATF site deletion was found to be about 25% more active than the wild type *PGK* control (Table 5.2). Thus, if the potential yATF binding site within the UAS did play a role in the regulation of *PGK*, it could be mediating a negative effect on transcription. However, the variation in the data in Table 5.2 is too large to draw conclusions about the role of yATF at the *PGK* locus. The overlap in *PGK* counts between pMA27 and pAJ110 means that it is not possible to say that 124% activity is any different from 100% activity.

5.2.3 A Plasmid-Borne Copy of *PGK* in a *cpf1* background

To investigate the effect of a lack of Cpf1p on *PGK* transcription, the equivalent of deleting the Cpf1p binding site from the promoter present on a high copy number

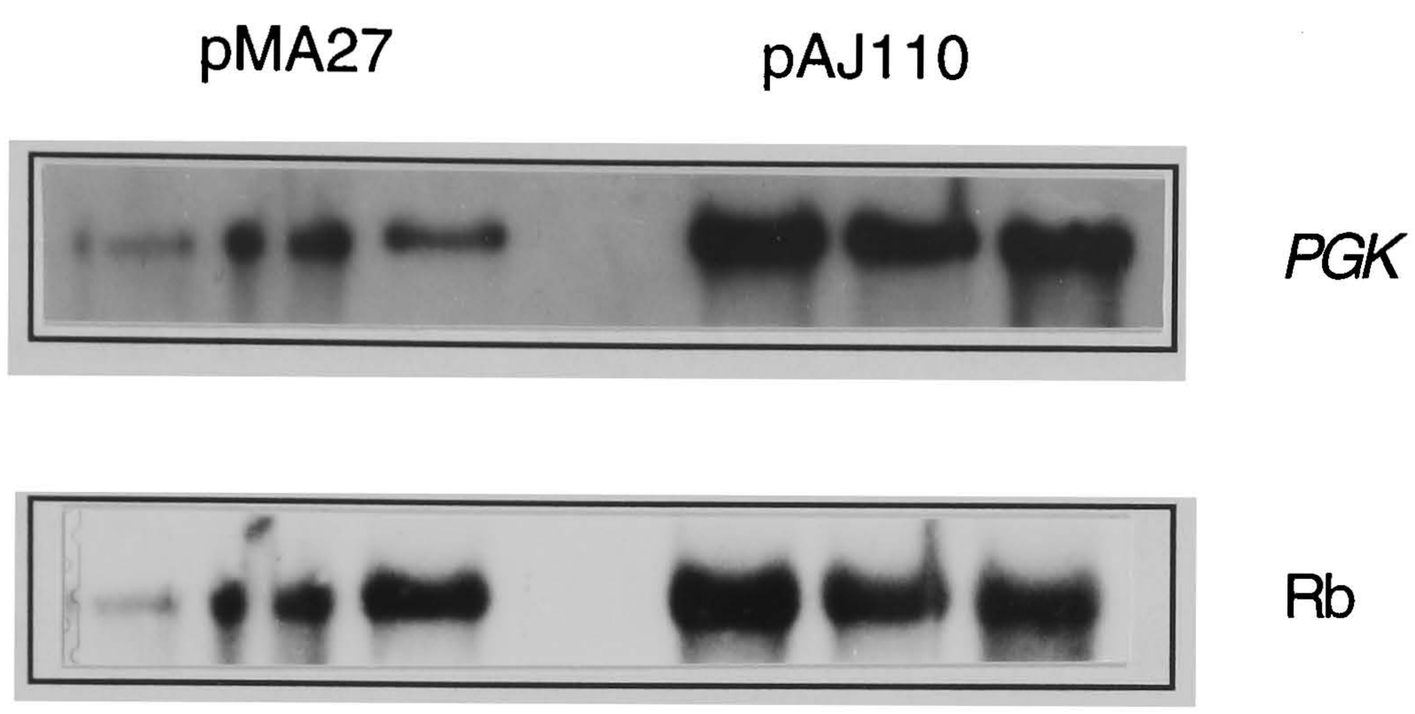


Figure 5.3 Northern blot of RNA from DBY745 transformed with pMA27 and pAJ110. The blot was probed with *PGK*, and a ribosomal probe.

Table 5.2 Phosphorimager analysis of Northern blots of RNA isolated from DBY745 transformed with the multicopy plasmids pMA27, carrying a wild type copy of *PGK*, and pAJ110, which carries a copy of *PGK* deleted of its potential yATF binding site.

Plasmid	*PGK* counts	Rb counts	Copy Number	Corrected *PGK*	% Activity
	18180.77	43530.06	1.9	0.22	
pMA27	69323.31	91009.74	1.7	0.45	100
	41800.28	108252.57	1.2	0.32	
	182427.28	130576.90	2.2	0.635	
pAJ110	107142.12	103906.04	3.4	0.303	124
	122562.21	100036.80	4.3	0.285	

plasmid was to measure the activation from a wild type copy of *PGK* in yeast cells which do not contain Cpf1 protein. A One Step transformation of DBY745 and YAG93 (a yeast strain isogenic to DBY745 but with a *cpf1* null mutation, see Materials and Methods, section 2.1) with the multicopy plasmid pMA27 was carried out. Two transformants of each strain were then grown to mid-log phase in SC-leu medium and RNA was extracted for analysis by Northern blotting. The filters were probed with a *PGK* specific probe, a ribosomal probe and a *leu2* probe (Figure 5.4). This allowed the levels of *PGK* message to be corrected for both RNA loading and for the copy number of the plasmids (Table 5.3).

After correcting the *PGK* counts for the amount of RNA loaded on the gel, as indicated by the ribosomal probe, and for plasmid copy number, as indicated by the *leu2* probe, the activity of the *PGK* promoter in the absence of Cpf1p is about 20% lower than in DBY745. This suggests that if the Cpf1p binding site plays a role in transcription from a plasmid-borne copy of *PGK* it is only a small one.

5.2.4 Transcriptional Activity of the Redefined UAS

The newly identified Reb1p and Cpf1p transcription factor binding sites in the *PGK* promoter are upstream of the previously defined UAS. In order to see whether these two binding sites have an effect on the activating potential of the UAS, this region of the promoter was examined in a minimal promoter plasmid.

The minimal promoter plasmid pAJ28 (L. Jenkins; Figure 5.5) contains the TATA and RNA initiation sequences (T'R) from the *PGK* promoter, which are sufficient to direct a low level of accurately initiated transcription (Stanway *et al.* 1989), upstream of the *lacZ* coding region. This minimal promoter is not able to produce high levels of β-galactosidase. However, the presence of a binding site for a transcriptional activator upstream of the minimal promoter causes an elevated level of transcription from *lacZ*, and cells containing such constructs will appear blue when grown on X-gal indicator plates. Thus, DNA sequences can be cloned into a *BamH1* site upstream of T'R to test whether or not they are able to activate transcription.

A fragment of the *PGK* promoter from -402, the 3' boundary of the UAS, to the *Cla1* site at -820, was isolated as a *BamH1/Cla1* fragment from pKV505 (Chambers *et al.* 1988). This fragment, containing the previously defined UAS and the 5' region of the promoter with the newly identified binding sites for Reb1p and Cpf1p, was cloned into pAJ28. This generated pAJ133, where the orientation of the transcription factor binding sites in relation to the basal promoter was the same as in the wild type —— ?

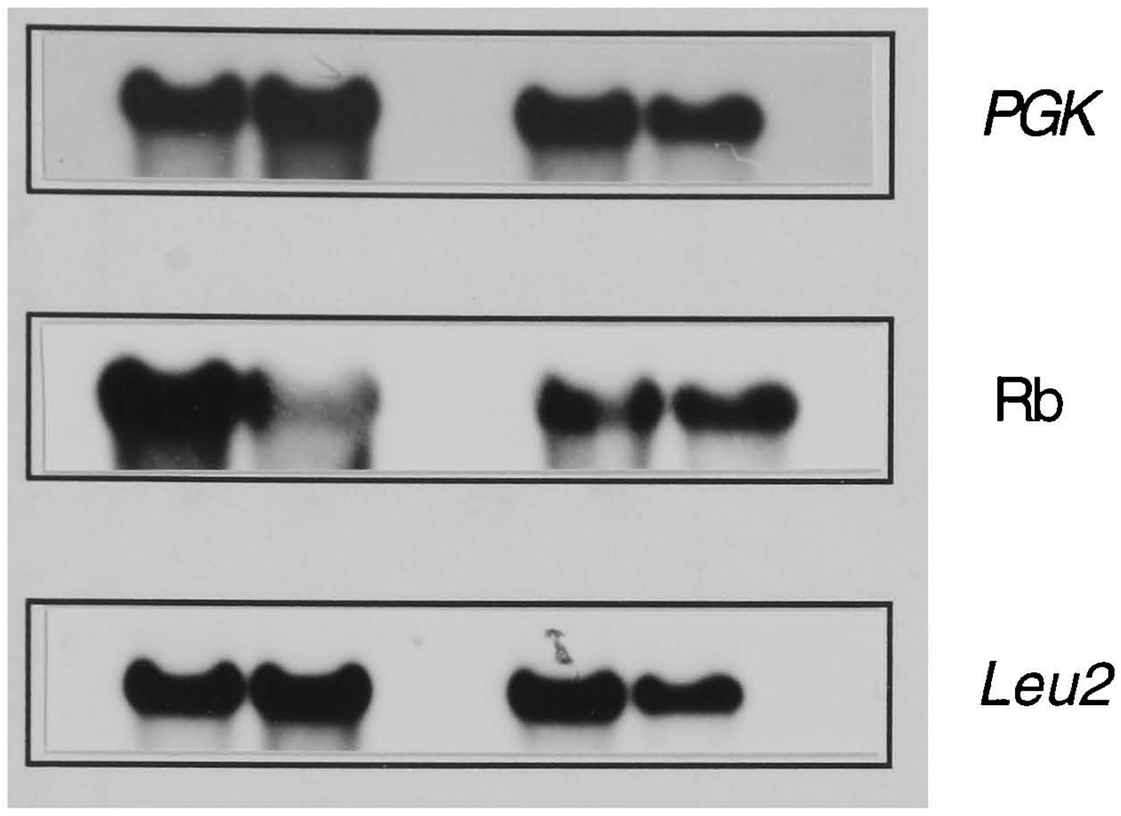


Figure 5.4 Northern blot of RNA from DBY745 and YAG93 transformed with pMA27, probed with *PGK*, a ribosomal probe to act as a loading control, and *leu2* for an estimation of plasmid copy number.

Table 5.3 Phosphorimager analysis of Northern blots of RNA from yeast strains DBY745 and YAG93 both transformed with the multicopy plasmid pMA27. Two transformants of each plasmid were assayed in parallel. The amount of *PGK* RNA was normalised for the loading of the gel using the ribosomal probe, and then corrected for plasmid copy number using the *leu2* probe.

Strain	*PGK* counts	Rb counts	Leu2 counts	Copy number	Corrected *PGK*	% Activity
DBY745	3476124	3642694	308438	1	3476124	100
DBY745	3997344	1446012	342233	2.8	3596369	
YAG93	3169559	2043856	358139	2	2824497	77
YAG93	2047962	1886310	234398	1.5	2636575	

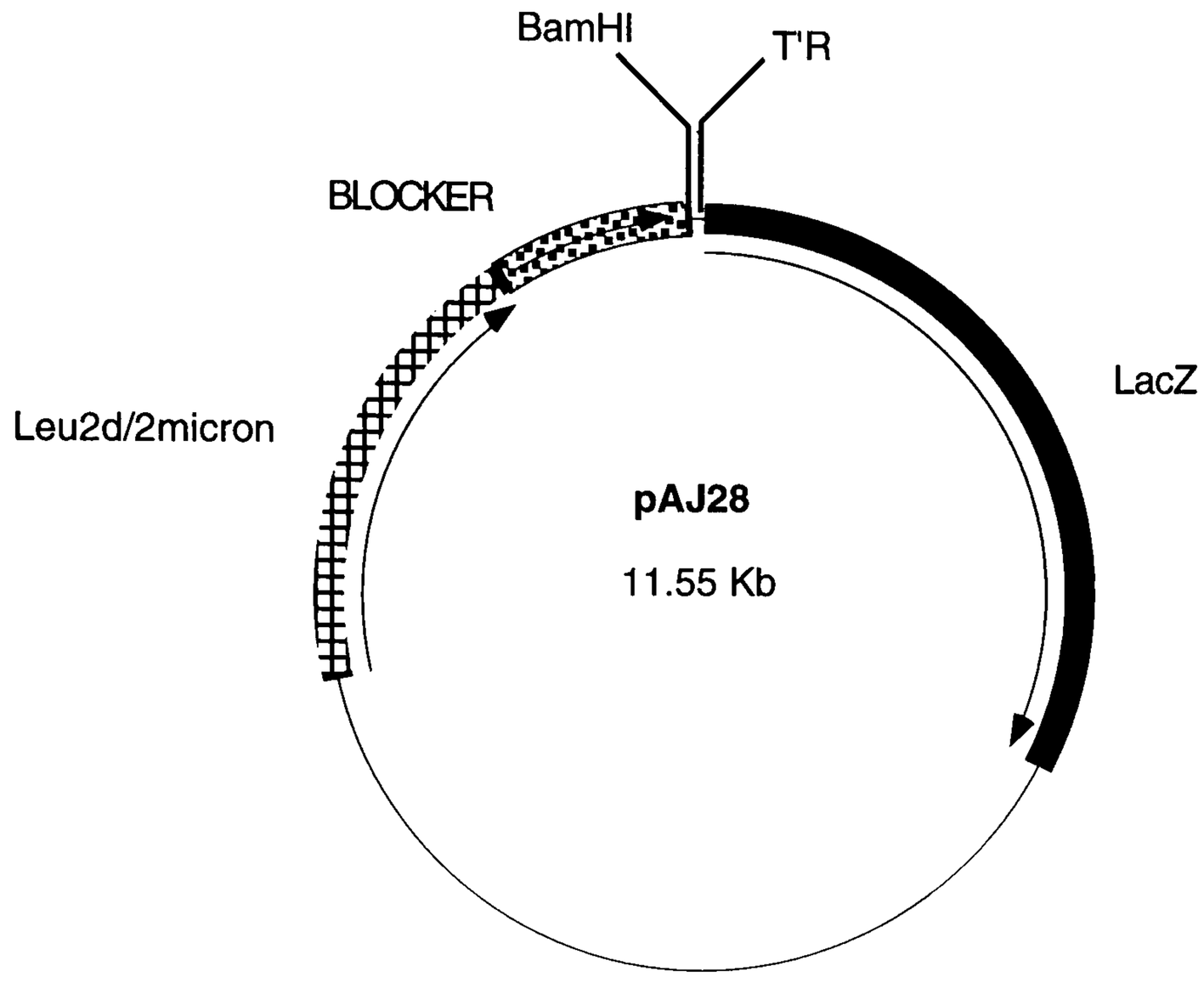


Figure 5.5 Minimal promoter plasmid pAJ28, a multicopy shuttle vector in which the TATA and RNA initiation sequences (T'R) from the *PGK* promoter are able to direct a low level of accurately initiated transcription from the *lacZ* coding region. A unique *BamH1* site upstream of T'R is used for the cloning of DNA fragments containing suspected binding sites for transcriptional regulators.

promoter. pAJ133 was transformed into DBY745 by the One Step method, and transformants were streaked onto X-gal indicator plates and compared with transformants of pAJ28, and pAJ211, which contains the originally defined UAS cloned in the correct orientation upstream of T'R (L. Jenkins). Quantitative liquid β-galactosidase assays were also carried out, using total protein extracts made from each of the transformants.

Each transformant was grown to mid-log phase in SC-leu medium. The cells were harvested as two 50ml pellets one of which was frozen at -80°C, so that an estimation of plasmid copy number could be made, the other of which was used to make the total protein extract as described in materials and methods. The protein concentration of the total protein extracts was determined with a Biorad assay (Bradford 1976) and 5μg of each extract was put into the β-galactosidase assay. Reactions were incubated for 10-20 minutes, and the β-galactosidase activity was expressed as activity per mg protein per minute (Table 5.4). Plasmid copy number was determined as described in materials and methods (2.6.2); the amount of plasmid DNA in each transformant, as indicated by a *lacZ* probe, was compared with the amount of chromosomal DNA, as indicated by a ribosomal probe. The β-galactosidase activities for each transformant were then corrected for differences in plasmid copy number and the average activity of three transformants for each construct was plotted in a Bar Chart (Figure 5.6)

As would be expected, transcription from the plasmids containing the *PGK* UAS fragments is much higher than that from the control plasmid pAJ28. However, although the average β-galactosidase activity for the extended UAS (EUAS), which includes the newly identified Reb1p and Cpf1p binding sites, is higher than the average activity for the original UAS, the range of β-galactosidase activities obtained for these two sets of transformants overlaps. This suggests that there may be no real difference in the activation potentials of these two constructs.

5.2.5 The *PGK* yATF Sequence in a Minimal Promoter Plasmid

The second way in which a possible role for the potential yATF site in the *PGK* UAS was investigated, was to synthesize two oligonucleotides (yATF1 and yATF2, Table 2.1C). When these were annealed they contained the potential yATF sequence from the *PGK* promoter (TTTCGTCACAC), flanked by *BamH1* compatible ends to facilitate cloning. After cloning of these oligonucleotides into a *BamH1* site, a *Pst1* site was also generated which aided screening of miniprep cultures. The oligonucleotides were cloned into the minimal promoter plasmid, pAJ28, to see whether or not they were able to activate transcription.

Table 5.4 A comparison of the β-galactosidase levels of minimal promoter reporter contructs. Three separate transformants of each plasmid were assayed in parallel. The β-galactosidase activities were corrected for plasmid copy number before an average activity was calculated.

Plasmid	Construct	β-gal activity /mg protein/minute	Copy Number	Average Activity (Corrected)
		0.6	1.5	
pAJ28	T'R	0.6	0.9	0.95
		0.88	0.5	
		27.6	2.6	
pAJ211	UAS	34.4	1.6	13.37
		20.8	2.6	
		26	2.3	
pAJ133	EUAS[a]	33.2	2.2	15.55
		32.4	1.6	

[a] the EUAS fragment contains the previously defined UAS and 280bp upstream. See text for details.

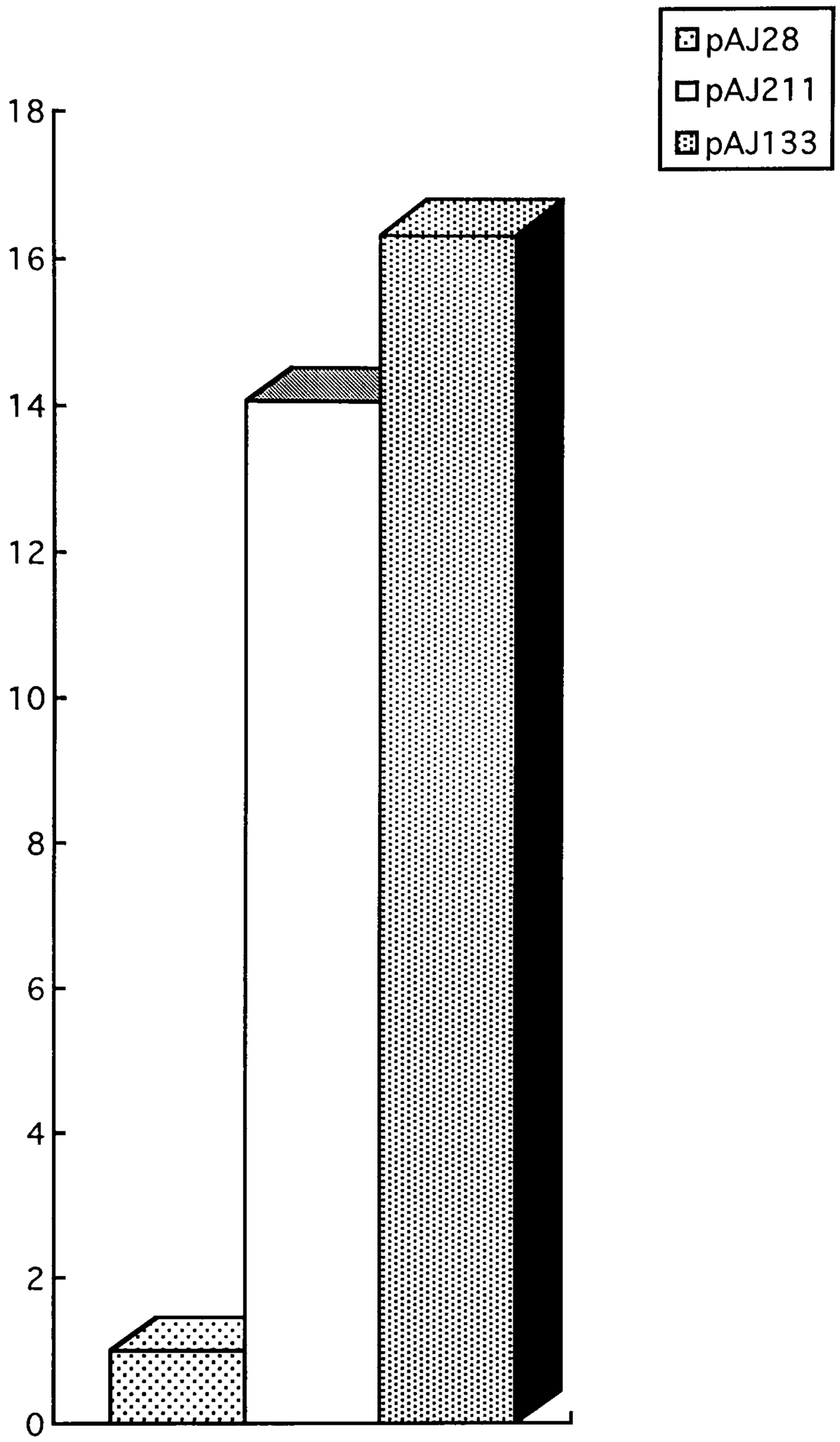


Figure 5.6 A comparison of the β-galactosidase activities of pAJ28, pAJ211 and pAJ133. Three transformants of each construct were assayed in parallel, the average activities of these after correction for plasmid copy number are shown above. pAJ28 was assigned a relative activity of 1.

yATF1 and yATF2 were phosphorylated and annealed (see Materials and Methods, section 2.9) before they were cloned into pAJ28 to give pAJ128 which contained the *PGK* potential yATF sequence upstream of T'R. A tandem insertion of yATF upstream of T'R, pAJ129, was achieved by cloning the yATF oligonucleotides into the reformed *BamH1* site of pAJ128. It had been found previously that the *PGK* UAS containing the CTTCC blocks and potential yATF site alone was not able to activate a minimal promoter, but that the CTTCC blocks and potential yATF sequence in conjunction with the Rap1p site were able to activate (Stanway *et al.* 1989). Thus pAJ128 was also used as a vector to clone the *TEF2* Rap1p site upstream of the yATF sequence to see whether this sequence had any effect on transcription activated by Rap1p. The plasmid containing the Rap1p binding site upstream of yATF was pAJ130. All pAJ28 derived constructs were sequenced to check the number and orientation of the insertions. Sequencing was carried out using a primer which bound to the T'R sequence in pAJ28, just downstream of the *BamH1* site where the insertions had been made.

Plasmids pAJ28, pAJ128, pAJ129, pAJ130 and pAJ207 were transformed into DBY745 using the One Step transformation procedure. pAJ207 contains one Rap1p site upstream of T'R and was constructed by L. Jenkins. Transformants were streaked onto X-gal indicator plates to look for gross differences in the amounts of β-galactosidase produced. A more quantitative determination of the levels of β-galactosidase was then obtained by making total protein extracts and carrying out liquid β-galactosidase assays. As before, half of each culture was used for plasmid copy number analysis. Without the benefit of phosphorimager analysis, the *lacZ* and ribosomal probings were compared by eye to look for any gross differences in copy number. As there did not appear to be any (Figure 5.7, lanes 1-10), the β-galactosidase activities for each construct (Table 5.5) were not altered.

As can be seen from Figure 5.8, the minimal promoter plasmid, pAJ28, is only able to activate transcription of the *lacZ* reporter weakly. When this basal level of transcription is compared with that from the plasmids containing either one or two copies of the *PGK* potential yATF binding site, pAJ128 and pAJ129, no increase over the basal level can be seen, suggesting that the yATF sequence is not able to activate transcription of the minimal promoter either on its own, or in tandem. However, when one Rap1p site is upstream of the minimal promoter, pAJ207, the levels of β-galactosidase are about 10X those of pAJ28. A similar increase is seen if the Rap1p site is in conjunction with the yATF sequence, pAJ130.

5.2.6 The CRE Consensus is No More Active than the yATF Sequence

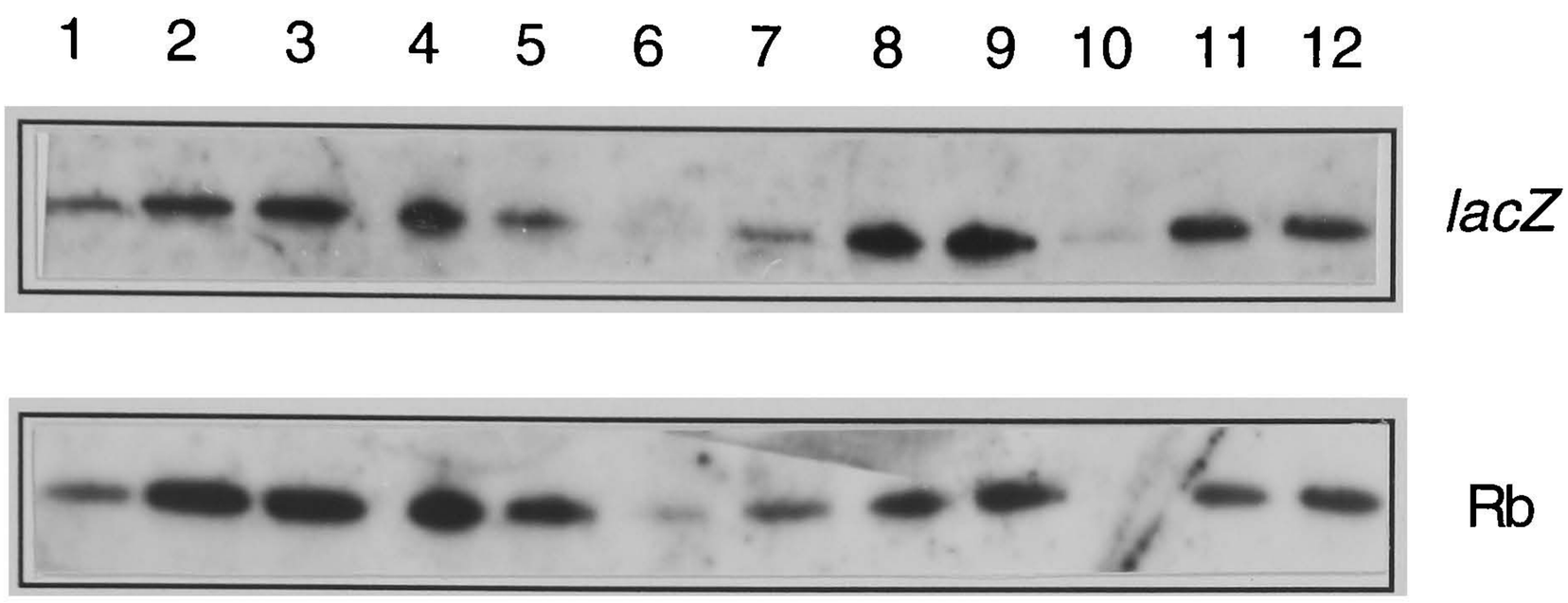


Figure 5.7 Copy Number data for DBY745 transformed with pAJ28 (lanes 1 and 6), pAJ128 (lanes 2 and 7), pAJ129 (lanes 3 and 8), pAJ207 (lanes 4 and 9) and pAJ130 (lanes 5 and 10). Also for W303-1A and R884-1C transformed with pAJ210 (lanes 11 and 12). Blots were probed with *lacZ* to detect plasmid DNA, and with a ribosomal probe to detect chromosomal DNA. Data in lanes 1-10 refer to the plasmids in Table 5.5, and in lanes 11 and 12 to plasmids in Table 5.7.

Table 5.5 A comparison of the levels of transcriptional activity of sequences from the *PGK* promoter cloned into the minimal promoter plasmid pAJ28. The results shown are from two different transformants assayed on separate occasions. For plasmid copy number analysis see Figure 5.7, lanes 1-10.

Plasmid	Construct	β-gal activity /mg protein/minute
pAJ28	T'R	0.32
pAJ128	1 X ATF	0.4
pAJ129	2 X ATF	0.84
pAJ207	RAP1	4.8
pAJ130	RAP1 ATF	5.8
pAJ28	T'R	0.6
pAJ128	1 X ATF	0.52
pAJ129	2 X ATF	0.6
pAJ207	RAP1	3.72
pAJ130	RAP1 ATF	4.8

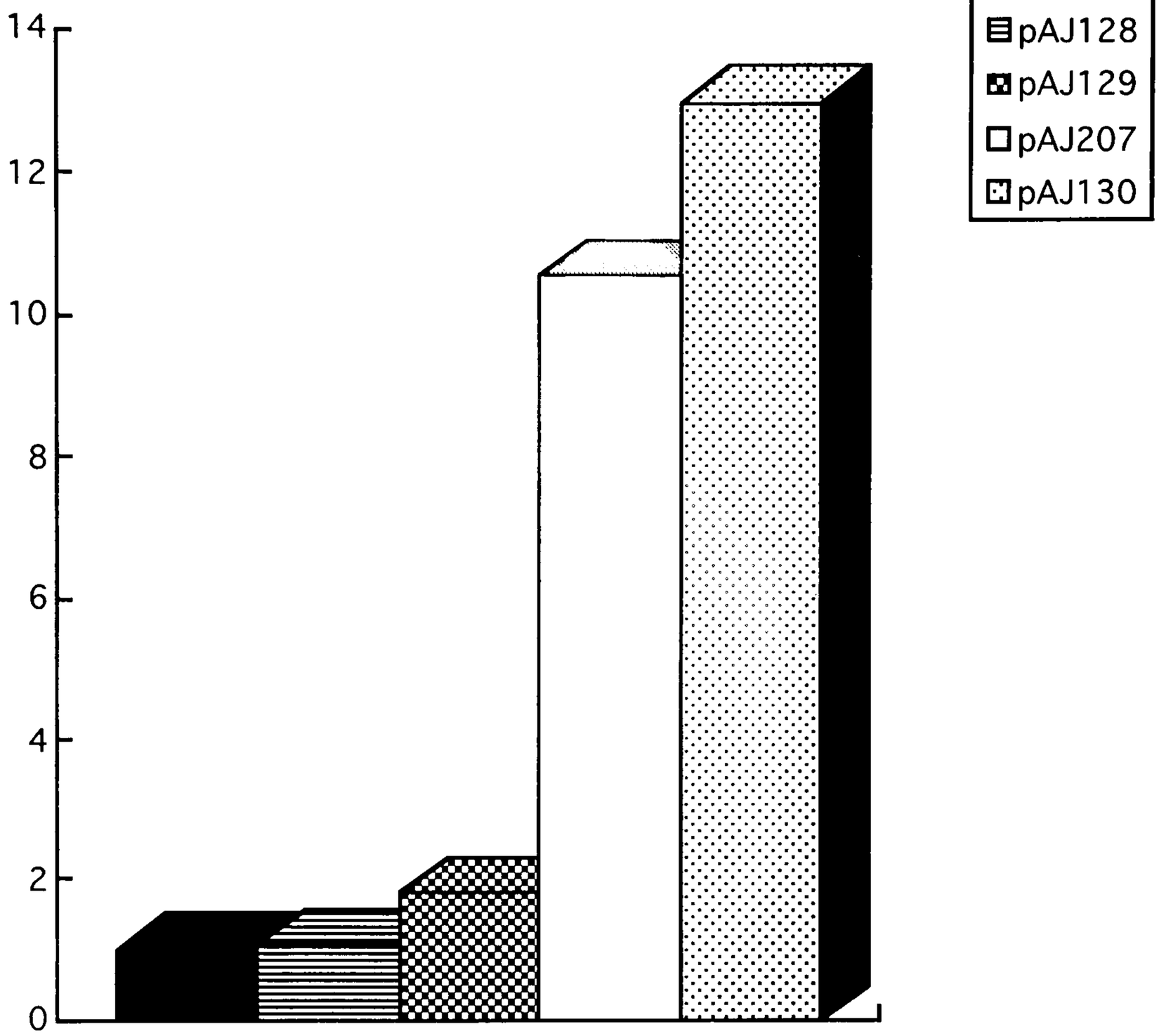


Figure 5.8 A comparison of the β-galactosidase activity of pAJ28 based reporter constructs. The bar chart shows the average for two transformants assayed separately, but with β-galactosidase activities normalised in relation to pAJ28 which was assigned a relative activity of 1.

The potential yATF binding site from the *PGK* promoter appears to be unable to activate the minimal promoter T'R, but as the effect of a true binding site for this transcription factor on the same minimal promoter was not known, this result did not necessarily mean that yATF does not bind to this sequence. In order to clarify this point, a second set of oligonucleotides was synthesised. This set, CRE1 and CRE2 (Table 2.1C) formed the consensus yATF binding site (TGACGTCA; Lin and Green 1988), the CRE (cAMP responsive element), when annealed.

As with the yATF oligonucleotides, CRE1 and CRE2 were phosphorylated and cloned into the *BamH1* site of pAJ28. Sequencing with T'R primer identified clones containing either one or two copies of the CRE consensus (pAJ131 and pAJ132, respectively). The *TEF2* Rap1p binding site was cloned upstream of the CRE consensus, and two clones were obtained in which the orientation of the Rap1p site differed; a minus orientation (pAJ134), and a plus orientation (pAJ135), which is the orientation found in the *PGK* UAS. These new plasmid constructs were transformed into DBY745 using the One Step method and once again transformants were first examined on X-gal indicator plates and secondly in liquid assay (Table 5.6, Figure 5.9).

As before, the plasmid copy number for each of the transformants was determined, this time with the benefit of the phosphorimager. The ratio of ribosomal probe to *lacZ* probe was similar (0.2-1.5) for all the transformants, except for two pAJ28 transformants and one pAJ132 transformant, whose copy numbers were considerably higher than the others. This may have been due to the quality of the initial DNA prep or to the amount of DNA loaded onto the gel. The transcriptional activity of these transformants is in agreement with that from the same constructs in other experiments, in all cases these transformants give a basal level of transcriptional activity. If the high copy numbers are a true reflection of the number of plasmids within the yeast cell, and were affecting the amount of β-galactosidase, then it might be expected that these transformants would be anomalously active. Since the basal level of transcription of these transformants has not been affected by their high plasmid copy numbers it is likely that the anomaly lies in the copy number estimation. Thus, the average activities represented in Figure 5.9 have not been corrected for copy number.

The results of these β-galactosidase assays show that the levels of activation seen with the CRE consensus oligonucleotide constructs are similar to the levels from the yATF constructs. Neither one nor two copies of either the yATF or CRE oligonucleotides is able to activate transcription to more than the basal level of

Table 5.6 A comparison of the levels of β-galactosidase activity of sequences from the *PGK* promoter, and the CRE consensus, when cloned into the minimal promoter plasmid pAJ28. The results shown are from two different transformants assayed on separate occasions.

Plasmid	Construct	β-gal activity /mg protein /min	Copy Number
pAJ28	T'R	0.24	12.83
pAJ207	1 X RAP1	1.81	1.12
pAJ134	RAP1(-) CRE	3.4	1.19
pAJ128	1 X ATF	0.24	0.72
pAJ129	2 X ATF	0.29	1.12
pAJ130	RAP1 ATF	2.13	0.63
pAJ131	1 X CRE	0.27	0.95
pAJ132	2 X CRE	0.4	12.56
pAJ135	RAP1(+) CRE	4.2	1.54
pAJ28	T'R	0.8	6.96
pAJ207	1 X RAP1	2.88	0.59
pAJ134	RAP1(-) CRE	5.8	0.77
pAJ128	1 X ATF	0.96	0.47
pAJ129	2 X ATF	0.72	0.37
pAJ130	RAP1 ATF	2.8	0.2
pAJ131	1 X CRE	0.96	0.26
pAJ132	2 X CRE	0.94	0.52
pAJ135	RAP1(+) CRE	4.4	0.66

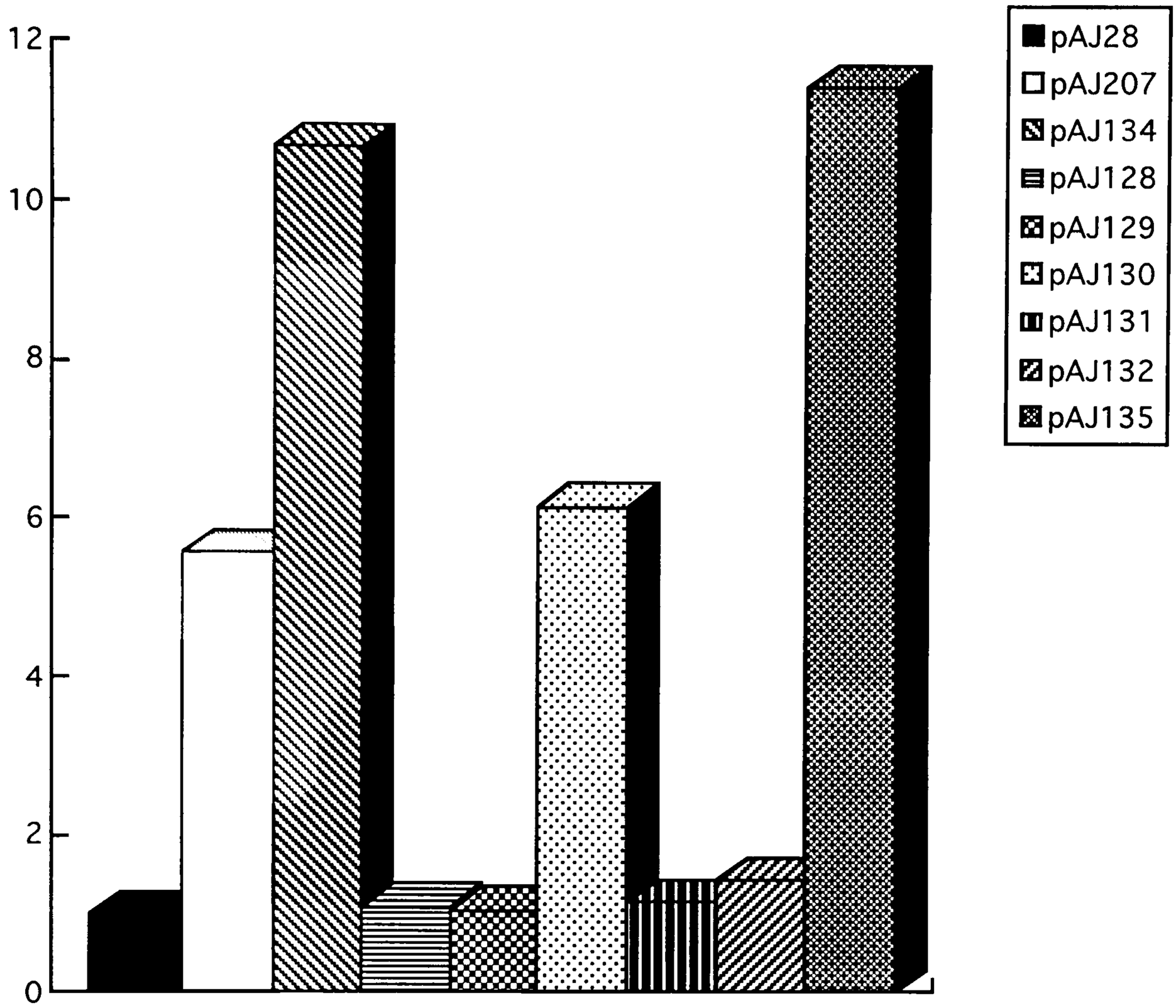


Figure 5.9 A comparison of the β-galactosidase activity of pAJ28 based reporter constructs. The bar chart shows the average for two transformants assayed separately, but with their β-galactosidase activities normalised to pAJ28 which was assigned a relative activity of 1.

pAJ28 (compare pAJ28 with pAJ128, pAJ129, pAJ131 and pAJ132 in Figure 5.9). When the Rap1p binding site is upstream of the yATF sequence (pAJ130), there is a small increase over the level of activation of the Rap1p site alone (pAJ207). However, when the Rap1p binding site is upstream of the CRE consensus (pAJ134 and pAJ135), the level of β-galactosidase produced is twice that of pAJ207, this increase is not affected by the orientation of the Rap1p binding site.

5.2.7 An Interaction between Rap1p and Gal11p

Gal11p has been proposed as a mediator of *PGK* transcription. To determine whether an interaction between Rap1p and Gal11p, two yeast strains W303-1A (*GAL11*) and R884-1C (*gal11*), a gift from Clive Stanway (Oxford University), were transformed with a multicopy plasmid, pAJ210, containing three Rap1p binding sites upstream of the minimal promoter, T'R (L. Jenkins). Transformants were grown in SC medium and total protein extracts were made. After the protein concentration had been determined, β-galactosidase assays were carried out to see whether there was any difference in the ability of the three Rap1p binding sites to activate the minimal promoter in the presence and absence of Gal11p.

The results of this analysis are in Table 5.7 and Figure 5.10 showing that in the absence of Gal11p the plasmid containing three Rap1p binding sites is only about 60% as active as pAJ210 in the presence of Gal11p. As with the other experiments using multicopy plasmids, the copy numbers of the transformants used here were checked by isolating copy number DNA, and looking at the ratio of *lacZ* probe to ribosomal probe (see Figure 5.7, lanes 11 and 12). The drop in activity of pAJ210 in the *gal11* strain could potentially have been due to an effect of growth rate. If one yeast strain does not grow as fast as a second one they will have different doubling times, this could affect the replication of plasmids contained within the strains such that the strain growing more slowly will contain fewer plasmids. As R884-1C appeared to grow more slowly than W303-1A, the growth rates of the two strains were determined to see whether a difference in growth rate could explain the difference in activity of the plasmids. However, when the growth curves were plotted both cultures were seen to grow at the same rate (Figure 5.11).

5.3 Discussion

The results of the binding site deletions from the *PGK* promoter on a multicopy plasmid, and also the level of transcription from pMA27 in a *cpf1*$^-$ background differ from the results obtained when looking at similar constructs in a single copy chromosomal location. Whilst the chromosomal deletion of Reb1p and Abf1p

Table 5.7 A comparison of the levels of β-galactosidase activity between *GAL11* and *gal11* strains of yeast transformed with pAJ210 which contains three Rap1p sites upstream of the minimal promoter T'R. For plasmid copy number analysis see Figure 5.7, lanes 11 and 12.

Yeast Strain	Genotype	Plasmid	β-gal activity /mg protein/minute
W303-1A	*GAL11*$^{+}$	pAJ210	23.2
R884-1C	*gal11*$^{-}$	pAJ210	14

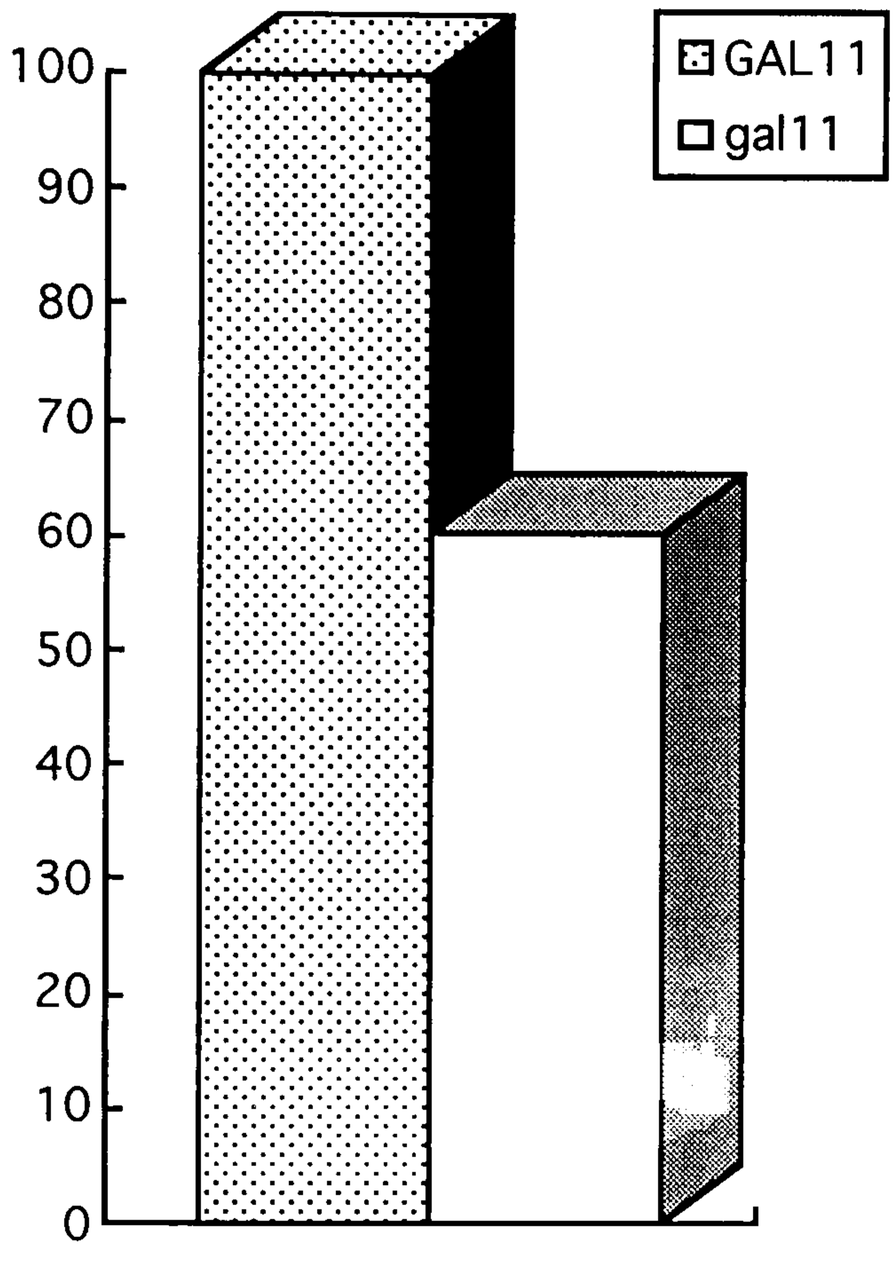


Figure 5.10 The β-galactosidase activity (shown as a percentage) of pAJ210 transformed into GAL11 (W303-1A) and *gal11* (R884-1C) strains of yeast.

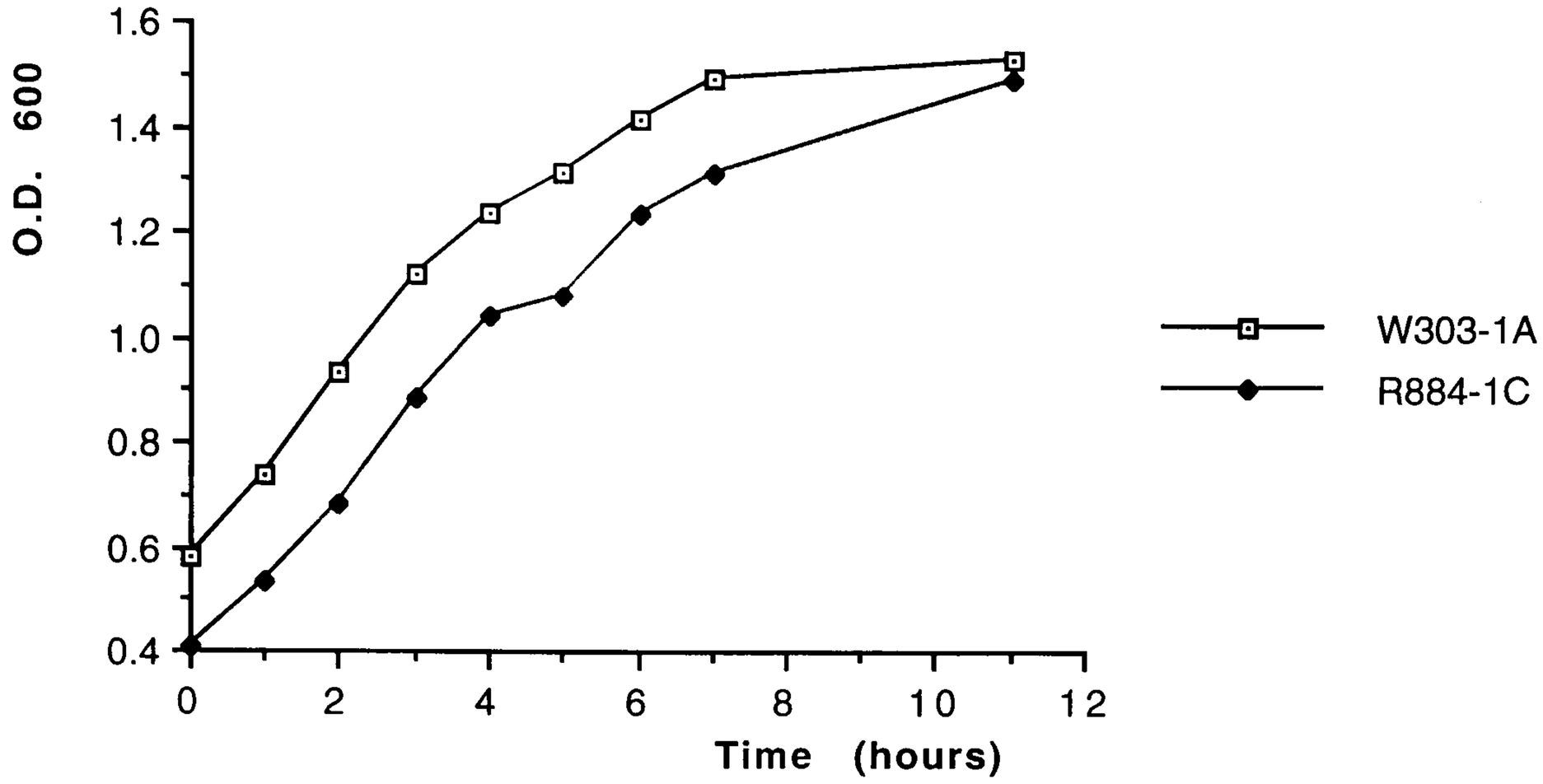


Figure 5.11 Growth curves for W303-1A and R884-1C. SC cultures were innoculated and grown overnight at 30°C after which OD_{600} readings were taken every hour for seven hours, with a final reading eleven hours after the first reading.

binding sites caused an approximately 50% decrease in the level of mRNA produced from *PGK*, no such decrease was seen from the multicopy plasmids. The deletion of the Abf1p binding site, pKV502, showed a 14% increase in levels of transcription over those from the control plasmid pMA27. This figure is within the range of experimental variation and shows that the previous conclusions about the effects of a multicopy plasmid-borne Abf1p binding site deletion were correct. Equally, the deletion of the Reb1p binding site had no effect on the activity of the plasmid-borne *PGK* promoter. Thus, with respect to Abf1p and Reb1p binding sites, the difference between the chromosomal and episomal states of *PGK* is most likely to be one of chromatin structure.

Unlike the absence of Abf1p and Reb1p, a lack of Cpf1p caused the levels of chromosomal *PGK* to increase by about 30%. However, when the effect of the same lack of Cpf1 protein was measured using a multicopy plasmid-borne copy of *PGK* (pMA27), the level of transcription was reduced by about 20% in YAG93, which is the *cpf1*$^-$ strain of yeast. Cpf1p has not been shown to be a transcriptional activator, and it seems likely that these differences in the influence of Cpf1p on plasmid and chromosomal copies of *PGK* are again due to the effects of this protein on chromatin structure.

When the ability of the previously defined UAS to activate transcription from the minimal promoter plasmid pAJ28 was compared with the ability of an extended UAS fragment to activate the same minimal promoter, it was found that although the average activation by the promoter fragment including the two newly identified transcription factor binding sites was slightly higher than that from just the UAS alone, the range of β-galactosidase activities for the two constructs overlapped. This suggests that the role played by Reb1p and Cpf1p during transcriptional activation at the *PGK* promoter is not to provide further strong activation domains. This is not surprising since Reb1p has only a weak activation domain, and Cpf1p has not been demonstrated to activate transcription. Rather they may be necessary to influence changes to chromatin structure. As has already been shown, Reb1p does not play an important part in the activation of transcription from a *PGK* promoter present on a multicopy plasmid and a lack of Cpf1p has only a small effect, so it is perhaps not unexpected that no effect is seen when looking at these new binding sites in a multicopy, minimal promoter plasmid system.

The investigation into a role for the potential yATF binding site at the 3' end of the *PGK* UAS involved determining whether this sequence was able to activate transcription from a minimal promoter plasmid, and also whether a deletion affected transcription from a multicopy plasmid-borne copy of *PGK*. The constructs made by

cloning oligonucleotides matching the *PGK* yATF sequence upstream of the minimal promoter T'R do not show any activation of the promoter over basal level even when two copies are present. This is in direct contrast to earlier work which suggested that even a single copy of a yATF binding site could activate a minimal promoter, and that in multicopy, yATF was able to activate to the level of a UAS (Jones and Jones 1989, Lin and Green 1989). This would suggest that the sequence under investigation here was not a *bona fide* yATF binding site.

However, a consensus binding site for yATF, the cAMP Response Element (CRE) was also cloned upstream of T'R in both single and tandem copies. Neither of these constructs were able to activate the minimal promoter plasmid which poses the question that the design of the minimal promoter may not be optimal for yATF function. With regard to this, the earlier experiments of Lin and Green (1989) have been repeated using a CRE in place of the yATF site (Sellers *et al.* 1990) and only a very low level of activation was obtained. This is similar to the results obtained in this work, where the CRE is unable to activate the minimal promoter. When either the yATF or CRE oligonucleotides were cloned between the Rap1p binding site and the basal promoter, an increase in the level of activation was seen over that caused by a Rap1p binding site alone. If the CRE was cloned adjacent to the Rap1p binding site, this increase was twice that of the Rap1p binding site alone. This may indicate some kind of cooperation between Rap1p and the CRE, and might suggest that the potential yATF binding site in the *PGK* promoter does not act within the cell. The small increase seen when the yATF sequence is in conjunction with the Rap1p binding site may be a result of an increase in distance between the Rap1p binding site and the basal promoter, facilitating activation by Rap1p.

Deleting the potential yATF site from the *PGK* promoter does not provide any evidence for an activating role for this sequence. In fact a deletion of the potential yATF site caused a 24% increase in the level of transcription from *PGK* on a multicopy plasmid, although, as discussed earlier, it is not possible to distinguish this figure from that of 100% for the wild-type gene so this result cannot be regarded as significant. A repressing activity of yATF has been identified in yeast (Nehlin *et al.* 1992, Vincent and Struhl 1992) thus it is possible that yATF is acting to repress transcription at *PGK*. However, a yATF binding site which is responsible for repression of *his3-303* was able to activate transcription when placed upstream of the *gal1* TATA element (Sellers *et al.* 1990). Thus it might be expected that the potential yATF binding site from the *PGK* promoter would activate transcription when taken out of context and inserted upstream of a minimal promoter. A possible explanation for the increase in transcription from a copy of *PGK* deleted of the yATF sequence is that the construction of the deletion has affected the binding site for Gcr1p; although

Gcr1p recognises a short core consensus sequence CTTCC, longer flanking regions are required for its interaction with DNA (Huie *et al.* 1992).

Plasmid pAJ210 containing three Rap1p sites upstream of a minimal promoter is only able to activate transcription to 60% of wild type in the absence of Gal11p. This suggests that Gal11p can interact with Rap1p to activate transcription, and is supported by a recent paper, Stanway *et al.* (1994). This shows that *PGK* mRNA levels are about 50% lower in a *gal11* strain than in *GAL11* yeast, but if the Rap1p binding site is deleted from the *PGK* promoter the *gal11* mutation does not affect *PGK* transcription. While they show that a Rap1p site is required for Gal11p to exert its effects, they cannot rule out the possibility that Gal11p interacts with Gcr1p. Use of the plasmid containing only Rap1p sites (pAJ210), suggests that an interaction between Rap1p and Gal11p is important for the positive influence of Gal11p on *PGK* transcription. An interaction between Rap1p and Gal11p has been demonstrated at the HMR silencer (Sussel *et al.* 1995), who show that a mutation in *GAL11* is able to suppress the silencing defect of a *rap1*^{s} strain. This effect appeared to be silencer specific, and the restoration of silencing was not due to a reduction in the level of transcription.

It has been suggested (Nishizawa *et al.* 1990), that Gal11p interacts with both Rap1p and initiation sequences thus bringing the UAS into contact with the basal promoter. If the UAS is placed immediately adjacent to the minimal promoter, at a distance of 25bp, then there is no reduction in transcriptional activity in the *gal11* strain. The results shown here do not agree with the hypothesis that Gal11p is dispensible when the UAS is in close proximity to the TATA box; the Rap1p sites in pAJ210 are cloned just upstream of the TATA box in the minimal promoter, but the construct shows a decrease in activity in the absence of Gal11p.

However, Sakurai *et al.* (1993) suggest that Gal11p acts to activate transcription as a basal transcription factor. In a construct consisting of tandem Rap1p sites upstream of a *CYC1-lacZ* reporter which has been integrated into the chromosome, they find that Gal11p stimulates the basal promoter activity to the same extent as transcription from the Rap1p sites; the difference in the spacing between the UAS and TATA box does not affect the integrated construct. This would mean that the reduction in β-galactosidase level seen when pAJ210 is in a *gal11* background was due to the effect of Gal11p on T'R, the basal promoter. If Gal11p does act as part of the pre-initiation complex, it is not required at all promoters, in which aspect it resembles TFIIE (see Introduction, 1.3.5). Gal11p dependency could require the basal promoter or the activators present upstream. The presence of Gal11p in a twenty polypeptide complex necessary for allowing basal transcription factors and RNA pol II to

respond to acidic activators (Kim *et al.* 1994) strengthens the argument that Gal11p acts at the basal promoter. However, influence from the activator proteins bound upstream of the basal promoter would not seem unreasonable, particularly in the light of Stanway's results described above.

Chapter 6

Construction of a Genomic Yeast Library

6.1 Introduction

As discussed in the introduction (1.7.1), Rap1p is a member of the family of multifunctional yeast transcription factors, and performs many important functions within the yeast cell. Many studies of Rap1p have been undertaken to define the domains of the protein necessary to carry out these varied functions. The N-terminus is dispensible for cell viability (cited in Hardy *et al.* 1992a), so interest has focused on the C-terminus to find regions for telomere binding, activation and silencing.

A yeast library was constructed, with the intention of using a genetic screen to identify DNA sequences which encode proteins which are able to provide the same C-terminal functions as Rap1p (Figure 6.1). This could identify proteins which have the same functional properties as the C-terminus of Rap1p, or which interact with Rap1p within the cell to mediate some of its functions. To achieve this, strains of yeast with temperature sensitive (*rap1^{ts}*) mutations in the *RAP1* gene were used (Kurtz and Shore 1991). These yeast are able to grow at 25°C, but when moved to the non-permissive temperature of 37°C the DNA binding activity of Rap1p is abolished. Since functional Rap1p is essential for cell viability the *rap1^{ts}* mutant strains cannot grow.

The library of approximately 1kb fragments of the yeast genome was cloned into a single copy plasmid vector containing a C-terminal truncated version of the *RAP1* gene. The resulting library plasmids were then transformed into yeast, and transformants which grew at 37°C were selected. Any cells which grew at 37°C should contain the plasmid in which the insert of yeast DNA complements the Rap1p C-terminus function.

There are many approaches for the construction of libraries in yeast, depending on the nature of the insert required. The first step is to break high molecular weight chromosomal DNA into smaller fragments for cloning. This can be achieved using restriction enzyme digests, or by mechanical shearing of the DNA using needles, or by sonication. Once the DNA has been broken into smaller fragments a fraction of a specific size can be selected if required. DNA which has been mechanically sheared should be treated, for example with Klenow or Mung Bean Nuclease, to ensure that it is blunt ended before any attempt at cloning it is made. At this stage synthetic

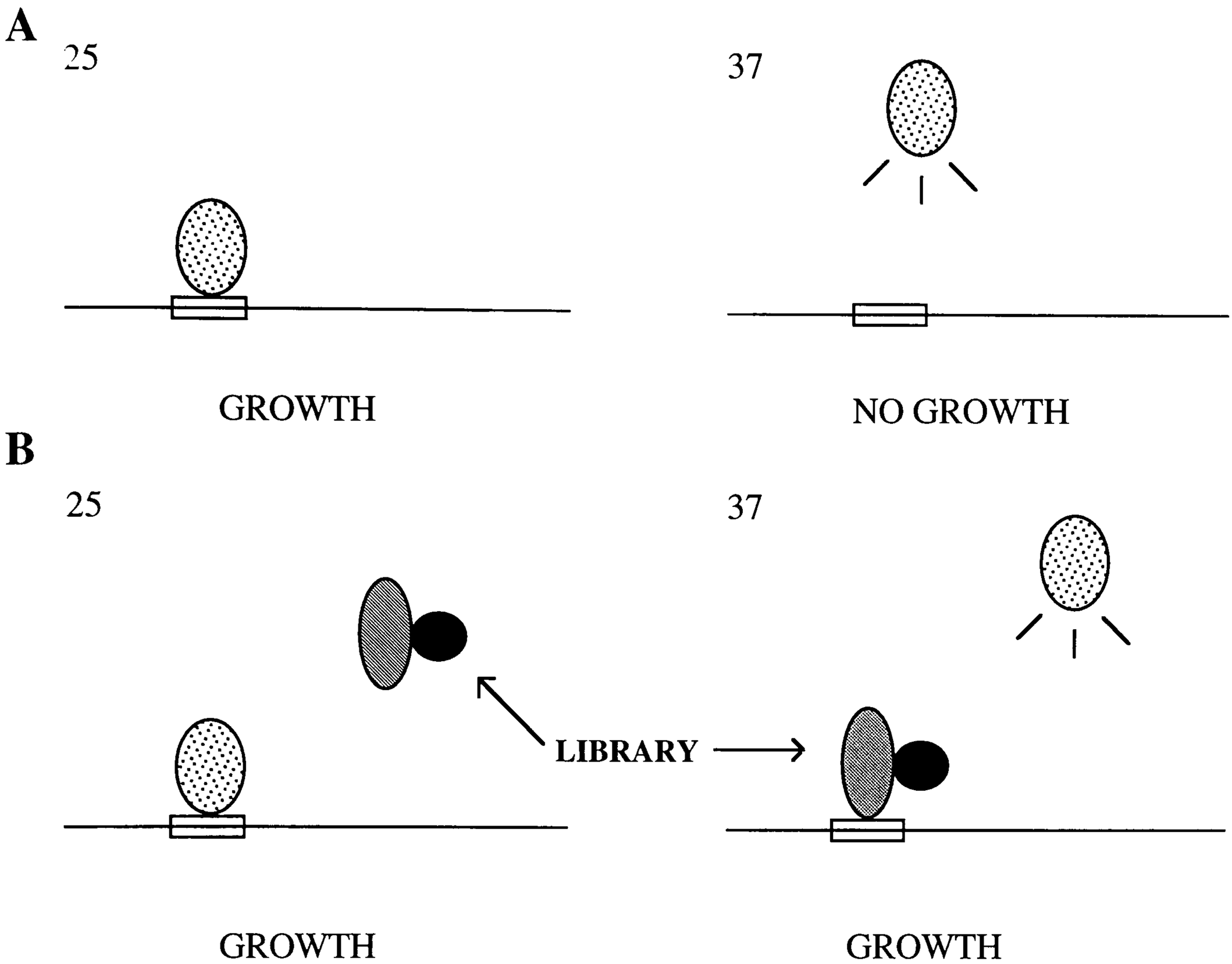


Figure 6.1 The genetic screen used to identify sequences which complement a *rap1*^{ts} phenotype. **A**: The ts yeast will not grow at 37°C because the ts mutation prevents Rap1p from binding to DNA and performing essential functions. **B**: If the library produces a fusion protein which can complement the ts mutation then growth will occur at 37°C.

restriction enzyme linkers can be added. The starting DNA should have a molecular weight which is significantly higher than that of the required insert DNA to ensure that the whole genome is represented in the ligation reactions with phosphatased vector.

Most approaches clone partially restricted yeast DNA; this yields libraries containing relatively large (~6kb) overlapping fragments of DNA. These clones can then be characterised further, for example by blotting, to identify clones of interest. However, partial digestion could not be used in the construction of this library since the resultant fragments would be too large.

Once the yeast fragments had been cloned the library had to be transformed into yeast. The probability of any given DNA sequence being present in the library is given by the following equation:

$$N = \frac{\log_e (1-P)}{\log_e (1-f)}$$

P : desired probability

f : fractional proportion of genome in a single recombinant

N : necessary number of recombinants

If the desired probability of finding a given 1kb fragment of *Saccharomyces* DNA, from a total 14000 kb, is 99%, then 6.45×10^4 recombinants are needed. All of these must then be transformed into yeast in order to have a good chance of selecting any which will allow the *rap1*^{ts} strains to grow at the non-permissive temperature. A large number of yeast transformants should grow at the permissive temperature of 25°C indicating that the transformation proceedure has worked efficiently, and that enough transformed yeast are present at 37°C to detect the very small percentage which contain a plasmid able to complement the *rap1*^{ts} .

The work described in this chapter was carried out in collaboration with Dr Alistair Chambers and Dr. Ian Graham.

6.2 Results

6.2.1 Construction of pAJ800

The vector into which the library was to be cloned was pAJ800 (Figure 6.2), this was constructed by cloning the N-terminus and DNA binding domain of Rap1p into pAJ731, a *URA* selectable, single copy shuttle vector (Graham and Chambers 1994a). PCR primers U/S1 and U/S2 (see Table 2.1A) were used to amplify a region starting in the promoter of the *RAP1* gene, through the N-terminus and as far as the

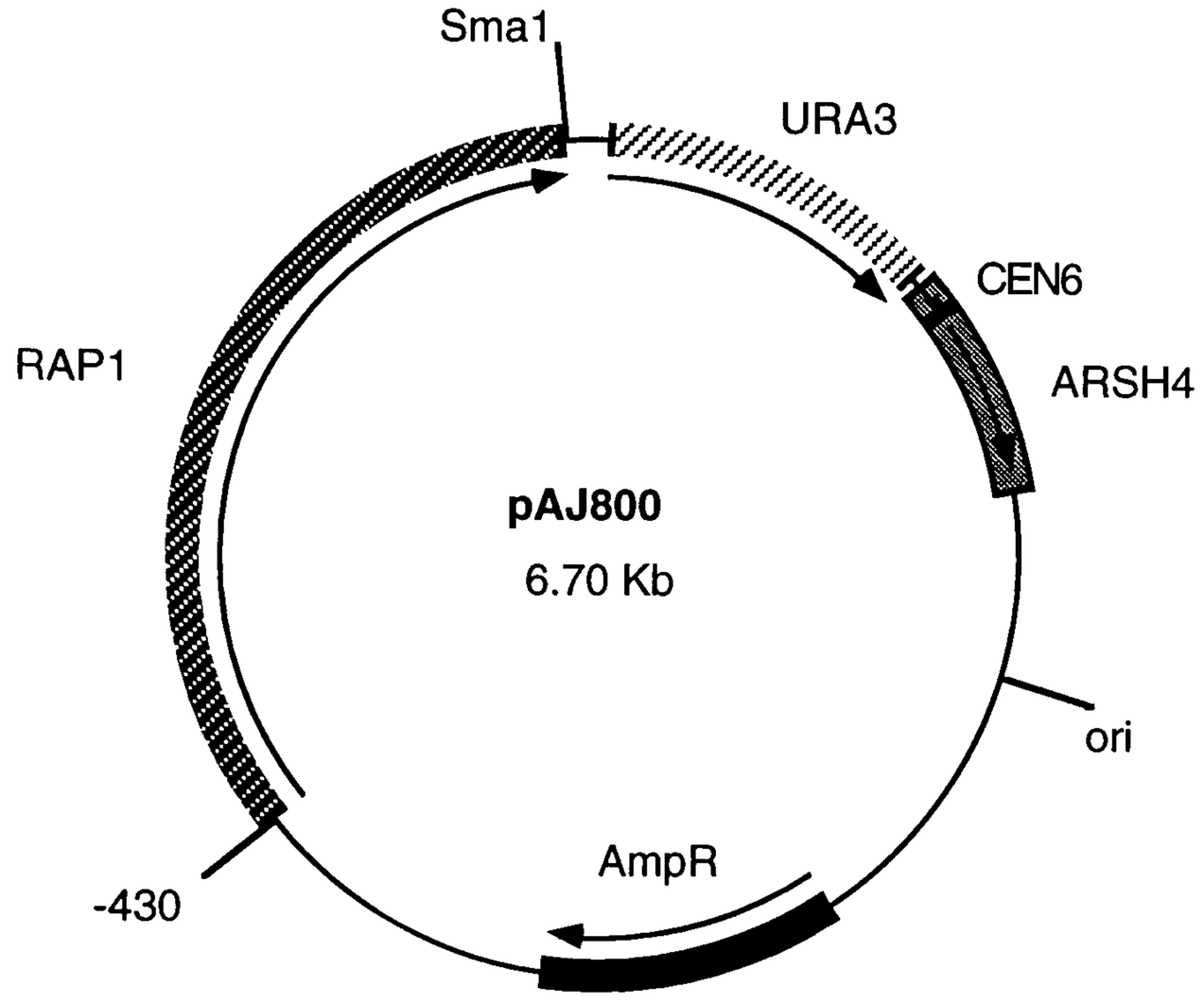


Figure 6.2 The URA selectable, single copy, shuttle vector pAJ800, which contains a region from the *RAP1* gene containing the promoter from -430, and the coding region as far as the 3' end of the DNA binding domain. There is a *Sma1* site at the end of the DNA binding domain into which yeast fragments can be cloned.

3' end of the DNA binding domain (-430 to amino acid 595). The primers were designed to incorporate an in frame *SmaI* site 3' of the DNA binding domain, into which blunt ended library fragments could be cloned. The template for the PCR reactions was pPE710 which contains the whole of the *RAP1* gene, and its promoter as far as -436. After touchdown PCR, a product of about 2.2kb was visible, this was reamplified and cleaned with a phenol/chloroform extraction before being digested with *BglII*. The digested product was cloned, by Dr. Ian Graham, into *BglII* cut pAJ731 to give pAJ800, a single copy yeast plasmid containing the N-terminus and DNA binding domain of Rap1p.

Prior to the construction of the yeast library some control experiments were carried out by Dr. I. Graham and Dr. A. Chambers to ensure that it was feasible to look for DNA sequences which were able to complement the C-terminus functions of Rap1p using the approach outlined above. Firstly, the *rap1^{ts}* strains of yeast were transformed with a plasmid which carried the whole *RAP1* gene, to see that the wild type protein could complement the mutant and thus allow growth at 37°C. If this had not been possible then the whole approach of constructing a library would not have worked.

Next *rap1^{ts}* yeast strains YDS409, YDS410 and YDS413 were transformed with the plasmid pAJ800, which contains only the Rap1p N-terminus and DNA binding domain. Growth of transformants was tested at both the permissive, and non-permissive temperatures (25°C and 37°C), but only found at 25°C. This confirmed that pAJ800 alone is unable to produce functional Rap1p, and thus cannot complement the temperature sensitive phenotype. So, if a DNA sequence cloned downstream of the Rap1p DNA binding domain allowed growth of a ts strain at 37°C, that DNA sequence must encode a protein which is able to provide the function of the missing C-terminus, or interact with Rap1p *in vivo*.

Finally, a known yeast activation domain was cloned into the *Sma1* site downstream of the DNA binding domain, to see whether this could replace the Rap1p C-terminus which is thought to contain the activation domain. Preliminary evidence suggests that the Gal4p activation domain is able to complement the ts phenotype of YDS410 and YDS413, allowing growth to occur at 37°C.

6.2.2 Construction of yeast library

Initially, to obtain random fragments of about 1kb, yeast chromosomal DNA was treated with a sonicator until all high molecular weight DNA had been randomly sheared into fragments within the 0.5-2kb range, as seen on an agarose gel.

Fragments thus formed were cloned downstream of the DNA binding domain of Rap1p. Transformants screened from the supposed libraries prepared in this manner were found to contain very few inserts, which may have been due to inefficient formation of blunt ends, or to the instability of the DNA after sonication.

The second approach to cloning small fragments of yeast chromosomal DNA was to select restriction enzymes which cut within a 4bp recognition sequence. Since the recognition site for such enzymes will occur on average every 4^4 bases, fragments of DNA with an average length of 256bp will be generated. However since such restriction enzyme sites are not distributed evenly over the DNA, a range of small fragments should be produced.

20µg chromosomal yeast DNA, prepared using the copy number prep, was digested with *RsaI*, a restriction enzyme which recognises the 4bp sequence GT↓AC. When the DNA had been digested to completion, it was incubated with *SmaI* cut and phosphatased pAJ800 in ligation reactions with molar vector:insert ratios of 1:5, 1:10 and 5:1. To ensure that the whole yeast genome was represented in the library ligations the number of genome equivalents present was calculated (see Table 6.1).

An aliquot of each ligation was transformed into *E. coli* to ensure that the reactions including insert DNA showed an increase in transformation efficiency over a ligation control containing only phosphatased vector. Colonies from these transformations were screened with a *Pst1* digest which showed that about 70% of transformants from the 1:5 and 1:10 ligations contained a library insert. The remainder of these ligations was then transformed into MC1061, and 100µl aliquots were spread on LBAp plates. Transformants from these plates were pooled to inoculate 500ml LBAp cultures for the large scale preparation of DNA using the Promega "Wizard Maxipreps" kit. This provided a stock of concentrated library DNA ready for transformation into yeast.

6.2.3 Transforming the Library into Yeast

The first yeast transformations with 10mg of the library, used the One-Step method of transformation, and the most severe of the $rap1^{ts}$ strains, YDS413. These transformations were not efficient enough to give a large enough number of colonies at 25°C such that there would be a reasonable chance of finding complementation at 37°C. The lithium acetate method of transformation was then used. This was more efficient, but still no complementation was seen at 37°C.

Table 6.1 Calculating the genome equivalents present in each of the pAJ800 ligation reactions with *Rsa1* cut DBY745, in the construction of a yeast genomic library.

Molar Ratio Vector:Insert	pmol Insert	Genome equivalents
1:5	3.33	1.43×10^8
1:10	6.66	2.86×10^8
5:1	0.267	1.15×10^7

<u>Calculating the Number of Genome Equivalents</u>

1pmol ds DNA = 0.75µg/kb

∴ 2.5µg of 1kb insert = 2.5/0.75 pmol = 3.33 pmol

3.33×10^{-12} moles x $6.02 \times 10^{23} = 2 \times 10^{12}$ molecules

each molecule is approximately 1kb

∴ the total number of bp is 2×10^{15}

Yeast genome length is 14 Mbase

∴ $2 \times 10^{15} / 1.4 \times 10^7 = 1.43 \times 10^8$ genome equivalents

All further analysis of the *Saccharomyces cerevisiae* genomic library in pAJ800 was carried out by Dr. I. Graham, who transformed 10µg of the library into the *rap1^{ts}* strain YDS410, using the One-Step method. The *rap1^{ts}* mutation in this strain is not as severe as that in YDS413 (Kurtz and Shore 1991), and after a week of incubation transformants started to appear at 37°C. These took longer to grow on the plate transformed with the library made from the 1:10 vector:insert ligation. Altogether 9 colonies grew on the 1:5 library plate and 2 on the 1:10. The transformation plates which were left at 25°C had approximately 2000 colonies on them, this suggests that 0.28% of the library clones was able to complement the *rap1^{ts}* strain.

All 11 colonies were restreaked on Sc-ura plates and of these 9 grew when returned to 37°C, so they were grown overnight in YPD for plasmid rescue (see materials and methods, 2.29). The rescued plasmids were retransformed into MC1061 and minipreps were made. The miniprep DNA was sequenced with primer 2501⊕ (Table 2.1B) which anneals to the Rap1p DNA binding domain sequence upstream of the *SmaI* site and reads downstream into the cloned DNA sequence.

Sequences obtained were compared with the yeast database (Table 6.2) and having ascertained that each clone contained an insert with homology to yeast sequence, "Wizard" miniprep DNA of each was retransformed into YDS410 and also YDS409. These retransformants grew both at 25°C and 37°C showing that genuine complementation was occuring.

6.3 Discussion

Transformation of *rap1^{ts}* yeast with a genomic yeast library yielded some transformants which were able to grow at the non-permissive temperature of 37°C. The library insert DNA sequences, from plasmids rescued from these transformants, were compared with the yeast database (Table 6.2), and demonstrated to be of yeast origin. Certain of the clones contained multiple insertions of yeast sequence which is probably a result of the library having been prepared from yeast DNA with an average length of 256bp. The DNA sequence was then used to predict the amino acid sequence of the protein which would be produced from the fusion of Rap1p DNA binding domain with the yeast DNA insert (Table 6.3).

At the outset of these experiments, it was predicted that complementation of the *rap1^{ts}* strains would occur if the sequence cloned downstream of the Rap1p DNA binding domain encoded a protein which had the same properties as the Rap1p C-terminus, or was able to interact with Rap1p within the cell. The DNA sequence might even have been part of the Rap1p C-terminus. So, to find that the fusion

Table 6.2 A comparison of library insert sequences, which allowed complementation of the *rap1*^{ts} strain of yeast, YDS410, with the yeast database. Clones 1, 5, 7 and 9 contained multiple insertions.

Clone	Positions	Accession No	% Identity	To Positions	Gene?
1	1-196	U18922	100	60594-60759	ChrV telomere
1	1-196	Z48148	99	6719-6914	ChrX telomere
1	1-196	SCCHRIX	99	6723-6918	ChrIX telomere
1	1-196	Z34098	99	3922-4117	ChrX telomere
1	27-196	U22383	97	21457-21288	ChrXII telomere
1	196-223	MISC13	100	812-782	OXI3
2	1-167	SCE9379	98	24229-24063	ME14-YER046
3	1-167	SCUBA2G	100	20-186	UBA2 (ChrIV)
4	1-157	YSCPRE5A	100	1958-1803	PRE5 (- strand)
5	1-77	SCD9717	100	23002-22926	ChrIV ORF
5	74-169	SCHRIII	98	261489-261584	TUP1 (- strand)
6 and 8	1-230	SCA21	99	1129-900	2μm (?REP1)
7	1-145	SCRDNA02	100	30-174	rDNA NTS
7	146-191	SCRAR1	100	46-1	RAR1 (- strand)
9	51-97	YSCH9986	100	5196-5150	YHR164c
9	100-146	SCGACOP	89	2053-2007	COP1 (- strand)

Table 6.3 The amino acid sequences of the RAP1:: Library fusions. The sequence of wild type Rap1p is shown at the bottom of the table. Potential protein kinase C phosphorylation sites are indicated by bold type.

Library Clone	Amino Acid Sequence
1	AATPTYGA**SDR**QP**SLK**MYYPAE
2	AATHAKALREELSNMPSITL
3	AATTNILDLEIT
4	AA**THR**NK
5	AATHSSLSP
6 and 8	AA**THK**VQKCLVSYTKGTNSISGNRKGLSR
7	AAT**TGR**SRGLV
9	AANMMQIEDVDTNLYDLAIS**TLK**N
RAP1^{+}	AAKRARNYS**SQR**NVQPTANAASA

proteins had only an extra 4 to 26 amino acids added was surprising. Also, none of these fusion proteins shows any homology with any other protein in the database, yeast or otherwise. This leaves the question, how are these short amino acid sequences able to complement ?

Most of the amino acid sequences which have been added onto the Rap1p DNA binding domain have potential protein phosphorylation sites, and there is also a potential phosphorylation site just downstream of the DNA binding domain in Rap1p. It is possible that phosphorylation of this region of the protein has a key role in the interaction of Rap1p with other proteins involved in its essential functions. This arguement could apply to different modifications of Rap1p for which a certain amino acid(s) is necessary.

Alternatively, rather than having an enhanced ability to interact with other proteins, the extra amino acids may increase the stability of the Rap1p/DNA interaction. Rap1p is envisaged as binding DNA via two subdomains, one which binds the core of the consensus, and one binding the 3' flanking region (Rhodes D. unpublished). There is also a tail region of the DNA binding domain which seems to loop back and interact with DNA near the consensus core (Rhodes D. unpublished). The additional amino acids in the fusion proteins could conceivably extend the tail and stabilise Rap1p binding to DNA.

Recently the DNA binding domain of Rap1p, and a region just downstream of it, have been shown to be responsible for the toxicity of *RAP1* overexpression (Freeman *et al.* 1995). Point mutations in this region which give proteins which are not toxic at high concentrations do not bind DNA as well as wild type Rap1p. Since the binding assays were carried out in the presence of wild type Rap1p, this may indicate the presence of a limiting interacting factor, or modification required for efficient DNA binding. This fits with the hypothesis above, that the amino acids added to the Rap1p DNA binding domain may include an essential modified residue, or form a domain for ineracting with other factors within the cell. If toxicity from *RAP1* overexpression is due to the binding of Rap1p to low affinity sites within the cell, the importance of the region C-terminal to the Rap1p DNA binding domain could be in stabilising such weak interactions (Freeman *et al.* 1995). Again, this function could be fulfilled by the short stretches of amino acids found to be added in the transformants growing at 37°C.

The rescue of the C-terminal deletion of Rap1p may be due to a stabilisation of the mRNA or the protein resulting in a rise in the levels of protein within the cell. This predication could be tested by making total protein extracts from ts yeast and library

transformants, and looking at their ability to bind a Rap1p binding site in a gel retardation assay. If the ability of the transformed yeast to grow at 37°C is simply due to the increased levels of Rap1p in the cell, then the complementation observed could be the result of high levels of Rap1p being able to perform the essential function of the C-terminus fortuitously, possibly through the presence of lots of copies of a domain in the N-terminus or DNA binding domain.

The results of this library construction were not what was initially predicted. If it was to be repeated then some changes to the approach could be made. The library was constructed from very short fragments of DNA which were often cloned as multiple insertions. The chances of isolating DNA sequences which encode yeast proteins with the same essential functions as the Rap1p C-terminus, or that interact with Rap1p within the cell might be increased if the insert DNA was larger. Thus, a library could be prepared from yeast DNA which had been cut with *HincII*. This is a restriction enzyme with a 6bp recognition sequence, but this sequence has a two base degeneracy (GTY↓RAC) making it equivalent to a 5bp recognition site. Digestion with *HincII* would yield DNA fragments of approximately 1kb. As *HincII* is a blunt cutter, these could still be cloned into the *Sma1* site engineered at the 3' end of the Rap1p DNA binding domain.

DNA fragments from a restriction digest of yeast DNA would still not be random. In this work, DNA fragments produced by sonication could not be cloned, this may have been due to instability of the DNA after sonication, or because the formation of blunt ends by Klenow and Mung Bean Nuclease was not efficient. If larger DNA fragments were being used to construct the library, shorter sonication times would be used, or the DNA could be sheared by pushing it through hypodermic needles. The DNA fragments need to be blunt ended for cloning, because the *Sma1* site into which they are cloned has been engineered to be in frame with the upstream *RAP1* sequence.

A change could also be made to the vector, pAJ800, into which the library fragments are cloned. Although pAJ800 transformants grow at 25°C, they do not grow well. This could be because truncated Rap1p, from the plasmid, competes for binding sites with the Rap1p produced by the cell. It is also possible that transcriptional readthrough is occuring from the *RAP1* sequence into the *URA* selectable marker. This might prevent proper initiation from occuring at *URA*. This could be prevented by putting a terminator, or a blocker sequence, downstream of the *RAP1* insert.

Chapter 7

General Discussion

The results presented in this thesis demonstrate the existence of two new transcription factor binding sites, for Reb1p and Cpf1p, in the promoter of *PGK*, upstream of the originally defined UAS, and go on to investigate whether these transcription factors are involved in the regulation of transcription at this locus. In addition, deletions of binding sites for two transcription factors already known to bind the *PGK* UAS, Rap1p and Abf1p (Chambers *et al.* 1989, Chambers *et al.* 1990), were made from the chromosomal locus, in order to extend the earlier studies of these proteins in the activation of *PGK*. At the same time, a minimal promoter was used to try to assign a role to the potential yATF binding site identified by Lin and Green (1989).

7.1 Multifunctional Transcription Factors at the *PGK* Promoter

In order to determine the role of Reb1p in the *PGK* promoter, a deletion of its binding site was made from the chromosomal *PGK* gene. This deletion was found to reduce transcription at this locus by about 50%, showing this binding site to be important for the regulation of transcription. This means that the *PGK* UAS is larger than previously thought, and extends upstream to include the Reb1p binding site at -561 to -552. Levels of transcription from a similar deletion of the Reb1p binding site were investigated on a multicopy plasmid construct. In this case, there was no difference between the Reb1p deleted construct and the wild type promoter. This difference between the chromosomal and plasmid context may be the reason that the Reb1p binding site was not identified in the initial deletion analysis of the *PGK* promoter (Ogden *et al.* 1986), which used *PGK* constructs on high copy-number plasmids. This result may indicate that the role of Reb1p depends on the structure of the DNA in which the binding site is present, or could simply be due to the number of copies of the gene present in the cell when using mulitcopy plasmids. Fifty to one hundred gene copies may mask effects observed when only looking at the chromosomal locus.

This difference in the role of a Reb1p binding site has been observed previously. Binding site mutations which prevent Reb1p from binding to the rDNA enhancer had no effect on transcription from a minigene construct on a multicopy plasmid (Kulkens *et al.* 1989). However, when the same construct was integrated at the

chromosomal locus, the mutations in the Reb1p binding site were found to reduce transcription from an adjacent operon (Kulkens *et al.* 1992).

The 50% decrease in transcription when the Reb1p binding site is deleted from the chromosomal locus suggests that it is important for the regulation of *PGK*. Its role may be one of influencing chromatin structure, as has been observed at the *GAL1-10* intergenic region. Here Reb1p seems to cause the repositioning of nucleosomes over a 160bp region allowing increased access to transcription factors (Fedor *et al.* 1988, Fedor and Kornberg 1989). A hybrid promoter consisting of the UAS_{GAL} and a Gcn4p binding site is dependent on Reb1p to position nucleosomes, so that Gcn4p can activate transcription (Brandl and Struhl 1990). Reb1p has also been shown to act as a weak activator of transcription (Buchman and Kornberg 1990, Chasman *et al.* 1990, Wang *et al.* 1990) where it can act synergistically with other weak activators such as those binding to a T-rich region, found in *DED1* or *RAP1* promoters (Chasman *et al.* 1990, Graham and Chambers 1994a).

Another role which has been proposed for Reb1p is that of preventing readthrough transcription from disrupting the formation of initiation complexes at the promoter of a gene downstream. A similar effect is seen in the role of Reb1p as a terminator for RNA pol I transcription (Shultz *et al.* 1993, Lang and Reeder 1993). The presence of Reb1p on its binding site causes all three polymerases to pause, although if the transcript is not released then readthrough can occur (Lang *et al.* 1994).

The Reb1p binding site in the *PGK* promoter is the most upstream of the transcription factor binding sites in this promoter. As such it resembles other glycolytic genes which contain Reb1p binding sites (reviewed in Chambers *et al.* 1995). This suggests that Reb1p may act as a boundary for the promoter. The Reb1p binding site in the *TPI* promoter can be deleted with no effect on levels of transcription if the sequence 5' to the binding site is not from *TPI*. However, if the sequence 5' to the Reb1p binding site is from *TPI*, the deletion causes a five-fold reduction in the level of expression (Scott and Baker 1993). The presence of Reb1p may prevent an inhibitory effect from the upstream sequence from interfering with transcription initiation. In the *TPI* promoter Reb1p is proposed to modulate chromatin structure allowing Rap1p and Gcr1p access to their binding sites (Scott and Baker 1993). In the *ENO1* promoter a Reb1p binding site is found at the 5' of the URS, and a deletion of this site results in a 30% loss of URS activity (Carmen and Holland 1994). The *ENO1* promoter also has a Reb1p binding site in its UAS2 (Carmen and Holland 1994). Thus, in the glycolytic genes Reb1p seems to perform a variety of functions, activation and repression of transcription, and also chromatin modulation.

After finding that the deletion of the Reb1p binding site from the *PGK* promoter had an effect if assayed at the chromosomal locus, but not on a multicopy plasmid, the roles of Rap1p and Abf1p were investigated at the chromosomal locus. The effect of deleting the Rap1p binding site is the same regardless of the location of the deletion. A dramatic reduction in the level of transcription from the *PGK* promoter is seen, thus confirming the central role played by Rap1p at this locus. The yeast strain in which the Rap1p site has been deleted from the chromosomal gene is able to grow on medium containing glucose, this suggests that a low level of PGK is being produced. If no PGK was present, then growth of this strain on 2% glucose would be severely reduced (Ciriacy and Breitenbach 1979).

When the Abf1p site is deleted from the chromosomal copy of *PGK*, the effect on transcription is in contrast to that seen when the same deletion is looked at on a multicopy plasmid. The chromosomal deletion caused an approximately 50% decrease in the level of transcription, whilst no effect is seen on a multicopy plasmid. Thus, like Reb1p, Abf1p exerts its effects at the chromosomal locus. This suggests that the roles of these two proteins at the *PGK* promoter may be similar. Both are weak transcriptional activators, and whilst Reb1p can displace nucleosomes (Fedor *et al.* 1988), Abf1p is able to bend DNA (McBroom and Sadowski 1994b). A related function for Reb1p and Abf1p has been shown at the *ILV1* promoter where binding sites for the two proteins are functionally interchangeable (Remacle and Holmberg 1992). Similarly, the rRNA enhancer is composed of redundant elements, a Reb1p binding site, an Abf1p binding site and a T-rich region, again suggesting similar functions for Reb1p and Abf1p (Morrow *et al.* 1993b).

The role of Cpf1p at the *PGK* promoter was investigated using a *cpf1* null strain of yeast since Cpf1p is not encoded by an essential gene. In this strain the activity of the chromosomal *PGK* promoter increased by 29%, and that of the plasmid-borne promoter decreased by about 20%. This difference in transcriptional activity between the chromosome and plasmid location, in the absence of Cpf1p, is not unusual for genes with CDE1 sites (N. Kent, pers comm). However, these small changes in *PGK* expression make it difficult to assign a role to Cpf1p at this locus.

The role of Cpf1p in the regulation of transcription is unclear. The presence or absence of Cpf1p appears to have no effect on the expression of *MET25*, *TRP1* and *GAL2* (Mellor *et al.* 1991) all of which contain CDE1 motifs. When the Downstream Activation Sequence (DAS) from *LPD1* was cloned upstream of a *CYC*::*lacZ* fusion activated transcription was obtained and the level of this transcription was not affected by mutation of the CDE1 motifs present (Sinclair *et al.* 1994). Similarly,

Mellor *et al.* (1990) saw no activation from a *PGK* minimal promoter after insertion of the CDE1 motif from either *TRP1*, *SAM2* or *CEN3*. However, Bram and Kornberg (1987) suggest that transcription is repressed if a Cpf1p binding site is placed between the UAS and TATA element of a GAL1-HIS fusion, and Thomas *et al.* (1989) have shown that deletion of the CDE1 motifs from *MET25* results in a considerable reduction in the level of transcription. *MET16* and *MET25* have been shown not to be transcribed in the absence of Cpf1p (Thomas *et al.* 1992), and also to be transcribed in the absence of Cpf1p (Kent *et al.* 1994). These differences are due in part to the different growth conditions used by the two groups. More recently transcription from *MET16* has been shown to require Cpf1p, although this is not sufficient for full UAS activity (O'Connell *et al.* 1995)

It is likely that Cpf1p plays a role in modulating chromatin structure; sensitivity to micrococcal nuclease at the *TRP1* promoter is lost in the absence of Cpf1p (Mellor *et al.* 1990), and localized changes to the chromatin structure of *MET16*, which depend on Cpf1p, have also been detected (O'Connell *et al.* 1995). Masison *et al.* (1993) suggest that Cpf1p alters chromatin structure to facilitate the formation of active transcription complexes. This view is supported by evidence from the study of Upstream Transcription Factor (USF), a mammalian transcription factor for which Cpf1p may be the yeast homologue. USF can compete with the assembly of promoter fragments into nucleosomes allowing the formation of stable preinitiation complexes (Workman *et al.* 1990). The promoter can then be bound by other transcription factors. Cpf1p may modulate chromatin structure by interacting with chromatin proteins such as histones. Mutations in *SPT21*, *RPD1*, *RPD2* and *CCR4* could all complement *cpf1*$^-$ methionine auxotrophy, but not the centromere defects (McKenzie *et al.* 1993). This led to the proposal that Cpf1p does not bind directly to CDE1 to function during transcription, but forms part of a complex binding a target sequence.

It is possible that *in vivo* the Cpf1p binding site in the *PGK* promoter is bound by another bHLH transcription factor. Recently mutations in a protein encoded by *SGC1* have been shown to suppress the requirement for Gcr1p at glycolytic promoters, including *PGK* (Nishi *et al.* 1995). Sgc1p contains a region of similarity to the basic-Helix-Loop-Helix motif which makes it a good candidate for interaction at the *PGK* promoter.

7.2 yATF Does Not Have a Role in the Regulation of Transcription at the *PGK* Locus

In mammalian cells, members of the ATF/CREB family of proteins are recognised by their ability to bind the CRE consensus, TGACGTCA (Lin and Green 1988). A protein with a similar binding affinity was purified from yeast (Lin and Green 1989, Jones and Jones 1989), and a potential binding site for yATF was identified in the *PGK* UAS (Lin and Green 1989). Early work suggested that a single yATF binding site could activate transcription from a minimal promoter, and that multiple copies of the yATF binding site could activate transcription to the level of a UAS (Lin and Green 1989, Jones and Jones 1989). The results presented here do not show any such activation, either with the potential yATF binding site or with the optimal ATF binding site, the CRE, neither of which was able to stimulate the level of transcription over a basal level. Sellers *et al* (1990) repeated the earlier experiments of Lin and Green (1989) using the CRE and obtained only a very low level of activation, not strong activation to the level of a UAS. Since, the CRE oligonucleotides were unable to activate the minimal promoter plasmid, pAJ28, the lack of activation seen with the oligonucleotides containing the potential yATF binding site could not be taken to mean that this sequence played no role in *PGK* transcription.

Both activator and repressor activities of ATF have been identified in yeast (Sellers *et al*. 1990, Vincent and Struhl 1992) as in mammals (Hai *et al*. 1988), indicating that there is probably a family of ATF proteins containing both activators and repressors. Although there could be just one protein whose function depends on promoter context, two or three distinct activities that bind to the CRE have been seen in yeast extract fractions from a heparin-agarose column (Sellers *et al*. 1990), these could of course be due to proteolytic degradation, or to differential modification of the same protein.

A repressor protein has been cloned from *S.cerevisiae* and is known as Acr1p (ATF/CREB Repressor, Vincent and Struhl 1992), or Sko1p (Suppressor of Kinase Overexpression, Nehlin *et al*. 1992). Acr1p was identified in a screen for mutations that eliminate repression mediated through the CRE. It has a bZIP DNA binding domain which resembles those of mammalian ATF/CREB proteins (Vincent and Struhl 1992), and binds to the CRE as a homodimer. *SKO1* is identical to *ACR1*, and was isolated as a high copy number suppressor of the lethal overexpression of cAMP dependent protein kinase (cAPK; Nehlin *et al*. 1992), suggesting a possible interaction between Sko1p and cAPK. Sko1p acts as a negative regulator of *SUC2*, interacting positively with Mig1p. The Sko1p binding site in the *SUC2* promoter contains only one half site of the CRE, as such it resembles the potential yATF site in *PGK* which also contains only one half site. Thus, it is possible that if yATF does not activate transcription at the *PGK* locus, then it may have a role in repression.

However Sellers *et al.* (1990) found that the yATF binding site responsible for repression in *his3-303* was able to activate transcription weakly when upstream of the *gal1* TATA element. It might be expected therefore, that if the potential yATF binding site had any function in the control of transcription from *PGK* then it should show weak activation when taken out of context and inserted into a minimal promoter construct.

To ensure that a role for yATF in the transcription of *PGK* had not been overlooked, because binding sites for interacting factors were not present at the minimal promoter, the potential yATF binding site was deleted from the *PGK* UAS in a gene construct on a multicopy plasmid, and also from the chromosomal locus. The multicopy plasmid deletion was no different in activity from the wild type gene, and whilst the deletion of yATF sequence from the chromosomal locus caused a decrease of about 20% in the level of transcription, this is likely to be due to a disruption to the flanking sequence of the upstream Gcr1p binding site, rather than because the yATF binding site is having an effect. The insertion of the *BamH1* linker which removed the potential yATF binding site disrupts three of the bases in the 3' flanking sequence.

Stanway *et al.* (1989) showed that the *PGK* potential yATF binding site is apparently inactive in their minimal promoter plasmid. It was not assayed alone, but in conjunction with the three CTTCC boxes which bind GCR1, this construct was not transcriptionally active, but if the Rap1p site was also included transcriptional activation was observed. In this work, activation by the yATF sequence in conjuction with a Rap1p binding site was looked for, to see whether the yATF binding site had any effect on the stimulation of transcription by Rap1p. The Rap1p/yATF contruct was only slightly more active than a Rap1p binding site alone, and could have been due to an increase in the distance between the Rap1p site and the basal promoter. The Rap1p/CRE constructs however, were about twice as active as a Rap1p binding site on its own. This suggests that a protein which binds the CRE in yeast may be able to interact with Rap1p.

Recently *SKO1/ACR1* was isolated as a high copy number suppressor of the toxicity of Rap1p overexpression (Freeman *et al.* 1995). This is not due to a reduction in the amount of Rap1p within the cell, but may be because it can titrate the excess Rap1p. This model predicts that if this were the case, then very high levels of Sko1p/Acr1p would be toxic due to a lack of Rap1p, and indeed Nehlin *et al.* (1992) found that high expression of *SKO1* inhibits the growth of yeast cells. However, Freeman *et al.* (1995) were not able to detect an interaction between Rap1p and Sko1p/Acr1p using

the two hybrid system. Even if Sko1p/Acr1p does not interact directly with Rap1p, it may provide a link with cAMP for the activation of transcription by Rap1p.

SKO1 was isolated as a suppressor of lethal cAPK overexpression (Nehlin *et al.* 1992), and it has been proposed that cAMP modulates the transcriptional activity of Rap1p (Klein and Struhl 1994). Rap1p was shown to be phosphorylated, and bind DNA more efficiently, in glucose media (Tsang *et al.* 1990), and since transcription from ribosomal protein genes is increased when protein kinase A (PKA) levels are increased by knocking out *BCY1*, which encodes the regulatory subunit of PKA (Toda *et al.* 1987) it is thought that the increase in transcription is due to the phosphorylation of Rap1p. If the Rap1p binding sites are deleted from upstream of the ribosomal protein gene *RPS13*, no increase in the level of transcription is seen in the *bcy1* strain (Klein and Struhl 1994b). However, ribosomal protein genes are induced by both nutrient availability, and cAMP (Neuman-Silberberg *et al.* 1995). These inductions involve different pathways; induction by cAMP requires *de novo* protein synthesis, and also the presence of Rap1p binding sites (Neuman-Silberberg *et al.* 1995). Since in the absence of cAMP, levels of Rap1p within the cell decrease, the protein synthesis required could be that of Rap1p. Thus, cAMP induction is not simply a matter of phosphorylating Rap1p.

7.3 A Model for the Regulation of Transcription at the *PGK* Promoter

The transcription factor with a central role in the activation of *PGK* is Rap1p. Deletion of its binding site causes a large decrease in the level of transcription, presumably because when it is bound to the promoter it provides not only an activation domain, but is also able to interact with Gcr1p, and also Gal11p. Rap1p may help to stabilize the interaction of Gcr1p with its DNA binding site, after which Gcr2p and Gcr3p may be recruited. The role of Reb1p may be to generate a nucleosome free region across the promoter allowing the other transcription factors access to their binding sites. The ability of Abf1p to bend DNA could bring the transcription factors bound to the upstream region of the promoter closer to the basal transcription factors. If Cpf1p does have a role it may be to facilitate the formation of a stable initiation complex. Weak activation domains at the *PGK* promoter are provided by Rap1p, Reb1p, Abf1p, Gcr1p and Gcr2p all of which could make multiple contacts with the basal transcription factors. One link between the upstream and downstream parts of the promoter could be through Gal11p, which interacts with Rap1p, and is also a component of the transcription mediator complex.

7.4 Further Work

Now that two new transcription factor binding sites have been identified in the *PGK* promoter, it is important to find out more about their role at this locus. In order to see whether Reb1p is involved in the positioning of nucleosomes, the micrococcal nuclease DNA footprint of the promoter could be determined. It would be of interest to see whether Sgc1p could bind to the *PGK* promoter. Since it might be predicted that any bHLH proteins which recognise the CACGTG sequence will be able to bind to a fragment from the *PGK* promoter *in vitro*, evidence for an interaction of Sgc1p should be looked for *in vivo*, again using footprinting techniques.

The regulation of *PGK* in response to carbon source can be tested using the yeast strains constructed with transcription factor binding site deletions from the *PGK* promoter. The effects that these deletions have on yeast grown with different carbon sources should be looked at, repression of *PGK* would be expected on non-fermentable carbon souces.

References

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

Allison L. A., Wong J. K., Fitzpatrick V. D., Moyle M. and Ingles C. J. (1988) The C-Terminal Domain of the Largest Subunit of RNA Polymerase II of *Saccharomyces cerevisiae, Drosophila melanogaster* and Mammals: a Conserved Structure with an Essential Function. *Mol Cell Biol* **8**: 321-329

Allison L. A. and Ingles C. J. (1989) Mutations in RNA polymerase II enhance or suppress mutations in GAL4. *Proc. Nat. Acad. Sci. (USA)* ***86***: 2794-2798

Arndt K. T., Styles C. A. and Fink G. R. (1989) A Suppressor of *HIS4* Transcriptional Defect Encodes a Protein with Homology to the Catalytic Subunit of Protein Phosphatases. *Cell* **56**: 527-537

Auble D. T. and Hahn S. (1993) An ATP-dependent inhibitor of TBP binding to DNA. *Genes Dev* **7**: 844-856

Axelrod J. D., Reagan M. S. and Majors J. (1993) GAL4 disrupts a repressing nucleosome during activation of *GAL1* transcription in vivo. *Genes Dev* **7**: 857-869

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

Baker H. V. (1991) *GCR1* of *Saccharomyces cerevisiae* encodes a DNA binding protein whose binding is abolished by mutations in the CTTCC sequence motif. *Proc. Nat. Acad. Sci. (USA)* **88**: 9443-9447

Baker R. E., Fitzgerald-Hayes M. and O'Brien T. C. (1989) Purification of the Yeast Centromere Binding Protein CP1 and a Mutational Analysis of Its Binding Site. *J Biol Chem* **264**: 10843-10850

Baker R. E. and Masison D. C. (1990) Isolation of the Gene Encoding the *Saccharomyces cerevisiae* Centromere Binding Protein CP1. *Mol Cell Biol* **10**: 2458-2467

Bartholomew B., Dahmus M. E. and Meares C. F. (1986) RNA Contacts Subunits II_o and II_c in HeLa RNA Polymerase II Transcription Complexes. *J. Biol Chem* **261**: 14226-14231

Becker D. M. and Guarente L. (1991) High-efficiency Transformation of Yeast by Electroporation. *Methods in Enzymol* **194**: 182-187

Bengal E., Flores O., Krauskopf A., Reinberg D. and Aloni Y. (1991) Role of the Mammalian transcription factors IIF, IIS and IIX during Elongation by RNA Polymerase. *Mol Cell Biol* **11**: 1195-1206

Bennetzen J. L. and Hall B. D. (1982) Codon Selection in Yeast. *J. Biol Chem* **257**: 3026-3031

Berger S. L., Cress W. D., Cress A., Triezenberg S. J. and Guarente L. (1990) Selective Inhibition of Activated but Not Basal Transcription by the Acidic Activation Domain of VP16: Evidence for Transcriptional Adaptors. *Cell* **61**: 1199-1208

Berger S. L., Pina B., Silverman N., Marcus G. A., Agapite J., Regier J. L., Triezenberg S. J. and Guarente L. (1992) Genetic Isolation of ADA2: A Potential Transcriptional Adaptor Required for Function of Certain Acidic Activation Domains. *Cell* **70**: 251-265

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

Bradford M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein-dye binding. *Anal Biochem* **72**: 248-254

Bram R. J. and Kornberg R. D. (1987) Isolation of a *Saccharomyces cerevisiae* Centromere DNA-Binding Protein, Its Human Homolog, and Its Possible Role as a Transcription Factor. *Mol Cell Biol* **7**: 403-409

Brand A. H., Micklem G. and Nasmyth K. (1987) A Yeast Silencer Contains Sequences That Can Promote Autonomous Plasmid Replication and Transcription Activation. *Cell* **51**: 709-719

Brandl C. J. and Struhl K. (1990) A Nucleosome-Positioning Sequence Is Required for GCN4 To Activate Transcription. *Mol Cell Biol* **10**: 4256-4265

Breathnach R. and Chambon P. (1981) Organization and Expression of Eukaryotic Split Genes Coding for Proteins. *Ann Rev Biochem* **50**: 349-383

Brent R. and Ptashne M. (1985) A Eukaryotic Transcriptional Activator Bearing the DNA Specificity of a Prokaryotic Repressor *Cell* **43**: 729-736

Buchman A. R., Kimmerly W. J., Rine J. and Kornberg R. D. (1988a) Two DNA-Binding Factors Recognise Specific Sequences at Silencers, Upstream Activating Sequences, Autonomously Replicating Sequences, and Telomeres in *Saccharomyces cerevisiae*. *Mol Cell Biol* **8**: 210-225

Buchman A. R., Lue N. F. and Kornberg R. D. (1988b) Connections between Transcriptional Activators, Silencers, and Telomeres as Revealed by Functional Analysis of a Yeast DNA-Binding Protein. *Mol Cell Biol* **8**: 5086-5099

Buchman A. R. and Kornberg R. D. (1990) A Yeast ARS-Binding Protein Activates Transcription Synergistically in Combination with Other Weak Activating Factors. *Mol Cell Biol* **10**: 887-897

Buratowski S., Hahn S., Sharp P. A. and Guarente L. (1988) Function of a yeast TATA element-binding protein in a mammalian transcription system. *Nature* **334**: 37-42

Buratowski S., Hahn S., Guarente L. and Sharp P. A. (1989) Five Intermediate Complexes in Transcription Initiation by RNA Polymerase II. *Cell* **56**: 549-561

Buratowski S. (1994) The Basics of Basal Transcription by RNA Polymerase II. *Cell* **77**: 1-3

Cadena D. L. and Dahmus M. E. (1987) Messenger RNA Synthesis in Mammalian Cells Is Catalysed by the Phosphorylated Form of RNA Polymerase II. *J. Biol Chem* **262**: 12468-12474

Cai M. and Davis R. W. (1989) Purification of a Yeast Centromere-Binding Protein That Is Able To Distinguish Single Base-Pair Mutations in Its Recognition Site. *Mol Cell Biol* **9**: 2544-2550

Cai M. and Davis R. W. (1990) Yeast Centromere Binding Protein CBF1, of the Helix-Loop-Helix Protein Family, Is Required for Chromosome Stability and Methionine Prototrophy. *Cell* **61**: 437-446

Carmen A. A. and Holland M. J. (1994) The Upstream Repression Sequence from the Yeast Enolase Gene *ENO1* Is a Complex Regulatory Element That Binds Multiple *Trans*-acting Factors Including REB1. *J Biol Chem* **269**: 9790-9797

Chambers A., Stanway C., Kingsman A. J. and Kingsman S. M. (1988) The UAS of the yeast *PGK* gene is composed of multiple functional elements. *Nuc Acids Res* **16**: 8245-8260

Chambers A., Tsang J. S. H., Stanway C., Kingsman A. J. and Kingsman S. M. (1989) Transcriptional Control of the *Saccharomyces cerevisiae PGK* Gene by RAP1. *Mol Cell Biol* **9**: 5516-5524

Chambers A., Stanway C., Tsang J. S. H., Henry Y., Kingsman A. J. and Kingsman S. M. (1990) ARS binding factor 1 binds adjacent to RAP1 at the UASs of the yeast glycolytic genes *PGK* and *PYK1*. *Nuc Acids Res* **18**: 5393-5399

Chambers A., Packham E. A. and Graham I. G. (1995) Control of glycolytic gene expression in the budding yeast (*Saccharomyces cerevisiae*) *Curr. Genet.* **29**: 1-9

Chasman D. I., Lue N. F., Buchman A. R., LaPointe J. W., Lorch Y. and Kornberg R. D. (1990) A yeast protein that influences the chromatin structure of UAS_G and functions as a powerful auxiliary gene activator. *Genes Dev* **4**: 503-514

Chen C. Y., Oppermann H. and Hitzeman R. A. (1984) Homologous versus heterologous gene expression in the yeast, *Saccharomyces cerevisiae*. *Nuc Acids Res* **12**: 8951-8970

Chen D., Yang B. and Kuo T. (1992) One-step transformation of yeast in stationary phase. *Curr Genet* **21**: 83-84

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

Chesnut J. D., Stephens J. H. and Dahmus M. E. (1992) The Interaction of RNA Polymerase II with the Adenovirus-2 Major Late Promoter is Precluded by the Phosphorylation of the C-terminal Domain of Subunit II_a. *J. Biol Chem* **267**: 10500-10506

Ciriacy M. and Breitenbach I. (1979) Physiological Effects of Seven Different Blocks in Glycolysis in *Saccharomyces cerevisiae*. *J. Bact* **139**: 152-160

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

Clifton D. and Fraenkel D. G. (1981) The *gcr* (Glycolysis Regulation) Mutation of *Saccharomyces cerevisiae*. *J Biol Chem* **256**: 13074-13078

Conaway R. C. and Conaway J. W. (1993) General Initiation Factors for RNA Polymerase II. *Ann Rev Biochem* 62: 161-190

Concino M. F., Lee R. F., Merryweather J. P. and Weinmann R. (1984) The adenovirus major late promoter TATA box and initiation site are both necessary for transcription *in vitro*. *Nuc Acids Res* **12**: 7423-7433

Cormack B. P., Strubin M., Ponticelli A. S. and Struhl K. (1991) Functional Differences Between Yeast and Human TFIID Are Localized to the Highly Conserved Region. *Cell* **65**: 341-348

Cortes P., Flores O. and Reinberg D. (1992) Factors Involved in Specific Transcription by Mammalian RNA Polymerase II: Purification and Analysis of Transcription Factor IIA and Identification of Transcription Factor IIJ. *Mol Cell Biol* **12**: 413-421

Damman R., Lucchini R., Koller T. and Sogo J. M. (1995) Transcription in the Yeast rRNA Gene Locus: Distribution of the Active Gene Copies and Chromatin Structure of Their Flanking Regulatory Sequences. *Mol Cell Biol* **15**: 5294-5303

Della Seta F., Ciafre S., Marck C., Santoro B., Presutti C., Sentenac A and Bozzoni I. (1990) The ABF1 Factor is the Transcriptional Activator of the L2 Ribosomal Protein Genes in *Saccharomyces cerevisiae*. *Mol Cell Biol* **10**:2437-2441

Deminoff S. J., Tornow J. and Santangelo G. M. (1995) Unigenic Evolution: A Novel Genetic Method Localizes a Putative Leucine Zipper That Mediates Dimerization of the *Saccharomyces cerevisiae* Regulator Gcr1p. *Genetics* **141**: 1263-1274

Diffley J. F. X. and Stillman B. (1988) Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. *Proc. Nat. Acad. Sci. (USA)* **85**: 2120-2124

Diffley J. F. X.. Cocker J. H., Dowell S. J. and Rowley A. (1994) Two Steps in the Assembly of Complexes at Yeast Replication Origins In Vivo. *Cell* **78**: 303-316

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*. *Nuc Acids Res* **10**: 2625-2637

Don R. H., Cox P. T., Wainwright B. J., Baker K. and Mattick J. S. (1991) "Touchdown"PCR to circumvent spurious priming during gene amplification. *Nuc Acids Res* **19**: 4008

Dorsman J. C., Doorenbosch M. M., Maurer C. T. C., de Winde J. H., Mager W. H., Planta R. J. and Grivell L. A. (1989) An ARS/silencer binding factor also activates two ribosomal protein genes in yeast. *Nuc Acids Res* **17**: 4917-4923

Dorsman J. C., van Heeswijk W. C. and Grivell L.A. (1990) Yeast general transcription factor GF1: sequence requirements for binding to DNA and evolutionary conservation. *Nuc Acids Res* **18**: 2769-2776

Dowell S. J., Tsang J. S. H. and Mellor J. (1992) The centromere and promoter factor 1 of yeast contains a dimerisation domain located carboxy-terminal to the bHLH domain. *Nuc Acids Res* **20**: 4229-4236

Drapkin R. and Reinberg D. (1994) The multifunctional TFIIH complex and transcriptional control. *Trends Biochem Sci* **19**: 504-508

Drapkin R., Sancar A. and Reinberg D. (1994) Where Transcription Meets Repair. *Cell* **77**: 9-12

Fantino E., Marguet D and Lauquin G. J. M. (1992) Downstream activating sequence within the coding region of a yeast gene: specific binding *in vitro* of RAP1 protein. *Mol Gen Genet* **236**: 66-75

Feaver W. J., Gileadi O. and Kornberg R. D. (1991a) Purification and Characterization of Yeast RNA Polymerase II Transcription Factor b. *J. Biol Chem* **266**: 19000-19005

Feaver W. J., Gileadi O., Li Y. and Kornberg R. D. (1991b) CTD Kinase Associated with Yeast RNA Polymerase II Factor b. *Cell* **67**: 1223-1230

Feaver W. J., Svejstrup J. Q., Bardwell L., Bardwell A. J., Buratowski S., Gulyas K. D., Donahue T. F., Friedberg E. C. and Kornberg R. D. (1993) Dual Roles of a Multiprotein Complex from *S.cerevisiae* in Transcription and DNA Repair. *Cell* **75**: 1379-1387

Fedor M. J., Lue N. F. and Kornberg R. D. (1988) Statistical Positioning of Nucleosomes By Specific Protein-Binding to an Upstream Activating Sequence in Yeast. *J Mol Biol* **204**: 109-127

Fedor M. J. and Kornberg R. D. (1989) Upsteam Activation Sequence-Dependent Alteration of Chromatin Structure and Transcription Activation of the Yeast *GAL1-GAL10* Genes. *Mol Cell Biol* **9**: 1721-1732

Feinberg A. P. and Vogelstein B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132**: 6-13

Fisher F., Jayaraman P. S. and Goding C. R. (1991) c-myc and the yeast transcription factor PHO4 share a common CACGTG-binding motif. *Oncogene* **6**: 1099-1104

Fisher F. and Goding C. R. (1992) Single amino acid substitutions alter helix-loop-helix protein specificity for bases flanking the core CANNTG motif. *EMBO J* **11**: 4103-4109

Flores O., Maldonado E. and Reinberg D. (1989) Factors Involved in Specific Transcription by Mammalian RNA Polymerase II: Factors IIE and IIF Independently Interact with RNA Polymerase II. *J. Biol Chem* **264**: 8913-8921

Flores O., Lu H., Killeen M., Greenblatt J., Burton Z. F., and Reinberg D. (1991) The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. *Proc. Nat. Acad. Sci. (USA)* **88**: 9999-10003

Flores O., Lu H. and Reinberg D. (1992) Factors Involved in Specific Transcription in Mammalian RNA Polymerase II. *J. Biol Chem* **267**: 2786-2793

Foreman P. K. and Davis R. W. (1993) Point Mutations that Separate the Role of *Saccharomyces cerevisiae* Centromere Binding Factor 1 in Chromosome Segregation From Its Role in Transcriptional Activation. *Genetics* **135**: 287-296

Francesconi S. C. and Eisenberg S. (1989) Purification and Characterization of OBF1: a *Saccharomyces cerevisiae* Protein that Binds to Autonomously Replicating Sequences. *Mol Cell Biol* **9**: 2906-2913

FreemanK., Gwadz M. and Shore D. (1995) Molecular and Genetic Analysis of the Toxic Effect of *RAP1* overexpression in Yeast. *Genetics* **141**: 1253-1262

Furter-Graves E. M. and Hall B. D. (1990) DNA sequence elements required for transcription initiation of the *Schizosaccharomyces pombe* ADH gene in *Saccharomyces cerevisiae. Mol Gen Genet* **223**: 407-416

Gilson E., Roberge M., Giraldo R., Rhodes D. and Gasser S. M. (1993) Distortion of the DNA Double Helix by RAP1 at Silencers and Multiple Telomeric-Binding Sites. *J Mol Biol* **231**: 293-310

Goncalves P. M., Maurer K., Mager W. H. and Planta R. J. (1992) *Kluyveromyces* contains a functional ABF1-homologue. *Nuc Acids Res* **20**: 2211-2215

Goncalves P. M., Griffioen G., Minnee R., Bosma M., Kraakman L. S., Mager W. H. and Planta R. J. (1995) Transcription activation of yeast ribosomal protein genes requires additional elements apart from binding sites for Abf1p or Rap1p. *Nuc Acids Res* **23**: 1475-1480

Goodrich J. A. and Tjian R. (1994) Transcription Factors IIE and IIH and ATP Hydrolysis Direct Promoter Clearance by RNA Polymerase II. *Cell* **77**: 145-156

Graham I. R. and Chambers A. (1994a) A Reb1p-binding site is required for efficient activation of the yeast *RAP1* gene, but multiple binding sites for Rap1p are not essential. *Mol Microbiol* **12**: 931-940

Graham I. R. and Chambers A. (1994b) Use of a selection technique to identify the diversity of binding sites for the RAP1 transcription factor. *Nuc Acids Res* **22**: 124-130

Guarente L. and Hoar E. T. (1984) Upstream activation sites of the *CYC1* gene of *Saccharomyces cerevisiae* are active when inverted but not when placed downstream of the "TATA box". *Proc. Nat. Acad. Sci. (USA)* **81**: 7860-7864

Ha I., Roberts S., Maldonado E., Sun X., Kim L., Green M. and Reinberg D. (1993) Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymeraseII. *Genes Dev* **7**: 1021-1032

Hahn S., Buratowski S., Sharp P. A. and Guarente L. (1989) Isolation of the Gene Encoding the Yeast TATA Binding Protein TFIID: A Gene Identical to the *SPT15* Suppressor of Ty Element Insertions. *Cell* **58**: 1173-1181

Hai T., Lin F., Allegretto E. A., Karin M. and Green M. R. (1988) A family of immunologically related transcription factors that includes multiple forms of ATF and AP-1. *Genes Dev* **2**: 1216-1226

Halfter H., Muller U., Winnacker E-L. and Gallwitz D. (1989a) Isolation and DNA-binding characteristics of a protein involved in transcription activation of two divergently transcribed, essential yeast genes. *EMBO J* **8**: 3029-3037

Halfter H., Kavety B., Vandekerckhove J., Kiefer F. and Gallwitz D. (1989b) Sequence, expression and mutational analysis of BAF1, a transcriptional activator and ARS1-binding protein of the yeast *Saccharomyces cerevisiae*. *EMBO J* **8**: 4265-4272

Hamil K. G., Nam H. G. and Fried H. M. (1988) Constitutive Transcription of Yeast Ribosomal Gene *TCM1* is Promoted by Uncommon *cis*-and *trans*-Acting Elements. *Mol Cell Biol* **8**: 4328-4341

Hardy C. F. J., Balderes D. and Shore D. (1992a) Dissection of a Carboxy-Terminal Region of the Yeast Regulatory Protein RAP1 with Effects on Both Transcriptional Activation and Silencing. *Mol Cell Biol* **12**: 1209-1217

Hardy C. F. J., Sussel L. and Shore D. (1992b) A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. *Genes Dev* **6**: 801-814

Harshman K. D., Moye-Rowley W. S. and Parker C. S. (1988) Transcriptional Activation by the SV40 AP-1 Recognition Element in Yeast Is Mediated by a Factor Similar to AP-1 That Is Distinct From GCN4. *Cell* **53**: 321-330

Henry Y. A. L., Chambers A., Tsang J. S. H., Kingsman A. J. and Kingsman S. M.(1990) Characterisation of the DNA binding domain of the yeast RAP1 protein. *Nuc Acids Res* **18**: 2617-2623

Henry Y. A. L., Lopez M. C., Gibbs J. M., Chambers A., Kingsman S. M., Baker H. V. and Stanway C. A. (1994) The yeast protein Gcr1p binds to the *PGK* UAS and contributes to the activation of transcription of the *PGK* gene. *Mol Gen Genet* **245**: 506-511

Hieter P., Pridmore D., Hegemann J. H., Thomas M., Davis R. W. and Philippsen P. (1985) Functional Selection and Analysis of Yeast Centromeric DNA. *Cell* **42**: 913-921

Himmelfarb H. J., Pearlberg J., Last D. H. and Ptashne M. (1990) GAL11P: A Yeast Mutation That Potentiates the Effect of Weak GAL4-Derived Activators. *Cell* **63**: 1299-1309

Hisatake K., Roeder R. G. and Horikoshi M. (1993) Functional dissection of TFIIB domains required for TFIIB-TFIID-promoter complex formation and basal transcription activity. *Nature* **363**: 744-747

Hitzeman R. A., Clarke L. and Carbon J. (1980) Isolation and Characterization of the Yeast 3-Phosphoglycerokinase Gene (*PGK*) by an Immunological Screening Technique. *J Biol Chem* **255**: 12073-12080

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. *Nuc Acids Res* **10**: 7791-7808

Hoffman C. S. and Winston F. (1987) A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene* 57: 267-272

Holland M. J. and Holland J. P. (1978) Isolation and identification of yeast messenger ribonucleic acids coding for enolase, glyceraldehyde-3-phosphate dehydrogenase and phosphoglycerate kinase. *Biochemistry* **17**: 4900-4907

Hoopes B. C., LeBlanc J. F. and Hawley D. K. (1992) Kinetic Analysis of Yeast TFIID-TATA Box Complex Formation Suggests a Multistep Pathway. *J. Biol Chem* **267**: 11539-11547

Hope I. A. and Struhl K. (1986) Functional Dissection of a Eukaryotic Transcriptional Activator Protein, GCN4 of Yeast. *Cell* **46**: 885-894

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

Horikoshi M., Yamamoto T., Ohkuma Y., Weil P. A. and Roeder R. G. (1990) Analysis of Structure-Function Relationships of Yeast TATA Box Binding Factor TFIID. *Cell* **61**: 1171-1178

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

Huie M. A., Scott E. W., Drazinic C. M., Lopez M. C., Hornstra I. K., Yang T. P. and Baker H. V. (1992) Characterization of the DNA-Binding Activity of GCR1: In Vivo Evidence for Two GCR1-Binding Sites in the Upstream Activating Sequence of *TPI* of *Saccharomyces cerevisiae*. *Mol Cell Biol* **12**:2690-2700

Innis M. A., Myambo K. B., Gelfand D. H. and Brow M. D. (1988) DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Nat. Acad. Sci. (USA)* **85**: 9436-9440

Inostroza J., Flores O. and Reinberg D. (1991) Factors Involved in Specific Transcription by Mammalian RNA Polymerase. *J. Biol Chem* **266**: 9304-9308

Jiang W. and Philippsen P. (1989) Purification of a Protein Binding to the CDE1 Subregion of *Saccharomyces cerevisiae* Centromere DNA. *Mol Cell Biol* **9**: 5585-5593

Johnston M. and Davis R. W. (1984) Sequences that regulate the divergent *GAL1-GAL10* promoter in *Saccharomyces cerevisiae*. *Mol Cell Biol* **4**: 1440-1448

Jones R. H. and Jones N. C. (1989) Mammalian cAMP-responsive element can activate transcription in yeast and binds a yeast factor(s) that resembles the mammalian transcription factor ATF. *Proc. Nat. Acad. Sci. (USA)* **86**: 2176-2180

Ju Q., Morrow B. E. and Warner J. R. (1990) REB1, a Yeast DNA-Binding Protein with Many Targets, Is Essential for Cell Growth and Bears Some Resembance to the Oncogene *myb*. *Mol Cell Biol* **10**: 5226-5234

Kelleher R. J., Flanagan P. M. and Kornberg R. D. (1990) A Novel Mediator Between Activator Proteins and the RNA Polymerase II Transcription Factors. *Cell* **61**: 1209-1215

Kent N. A., Tsang J. S. H., Crowther D. J. and Mellor J. (1994) Chromatin Structure Modulation in *Saccharomyces cerevisiae* by Centromere and Promoter Factor 1. *Mol Cell Biol* **14**: 5229-5241

Kim Y-J., Bjorklund S., Li Y., Sayre M. H. and Kornberg R. D. (1994) A Multiprotein Mediator of Transcriptional Activation and Its Interaction with the C-Terminal Repeat Domain of RNA Polymerase II. *Cell* **77**: 599-608

Klein C. and Struhl K. (1994a) Increased Recruitment of TATA-Binding Protein to the Promoter by Transcriptional Activation Domains *in vivo*. *Science* **266**: 280-282

Klein C. and Struhl K. (1994b) Protein Kinase A Mediates Growth-Regulated Expression of Yeast Ribosomal Protein Genes by Modulating RAP1 Transcriptional Activity. *Mol Cell Biol* **14**: 1920-1928

Klein F., Laroche T., Cardenas M.E., Hofmann J.F.X., Schweizer D. and Gasser S.M. (1992) Localization of RAP1 and Topoisomerase II in Nuclei and Meiotic Chromosomes of Yeast. *J Cell Biol* **117**: 935-948

Koleske A. J. and Young R. A. (1994) An RNA polymerase II holoenzyme responsive to activators. *Nature* **368**: 466-469

Kovari L. Z. and Cooper T. G. (1991) Participation of ABF1 Protein in Expression of the *Saccharomyces cerevisiae CAR1* gene. *J Bact* **173**: 6332-6338

Kulkens T., van Heerikhuizen H., Klootwijk J., Oliemans J. and Planta R. J. (1989) A yeast ribosomal DNA-binding protein that binds to the rDNA enhancer and also close to the site of Pol I transcription initiation is not important for enhancer functioning. *Curr Genet* **16**: 351-359

Kulkens T., van der Sande C. A. F. M., Dekker A. F., van Heerikhuizen H. and Planta R. J. (1992) A system to study transcription by yeast RNA polymerase I within the chromosomal context: functional analysis of the ribosomal DNA enhancer and the RBP1/REB1 binding sites. *EMBO J* **11**: 4665-4674

Kurtz S. and Shore D. (1991) RAP1 protein activates and silences transcription of mating type-genes in yeast. *Genes Dev* **5**: 616-628

Lam K. and Marmur J. (1977) Isolation and Characterization of *Saccharomyces cerevisiae* Glycolytic Pathway Mutants. *J. Bact.* **130**: 746-749

Landshulz W. H., Johnson P. F. and McKnight S. L. (1988) The Leucine Zipper: A Hypothetical Structure Common to a New Class of DNA Binding Proteins. *Science* **240**: 1759-1764

Lang W. H. and Reeder R. H. (1993) The REB1 Site Is an Essential Component of a Terminator for RNA Polymerase I in *Saccharomyces cerevisiae*. *Mol Cell Biol* **13**: 649-658

Lang W. H., Morrow B. E., Ju Q., Warner J. R. and Reeder R. H. (1994) A Model for Transcription Termination by RNA Polymerase I. *Cell* **79**: 527-534

Larson G. P., Castanotto D., Rossi J. J. and Malafa M. P. (1994) Isolation and functional analysis of a *Kluyveromyces lactis* RAP1 homologue. *Gene* **150**: 35-41

Lee J. M. and Greenleaf A. L. (1989) A protein kinase that phosphorylates the C-terminal repeat domain of the largest subunit of RNA polymerase II. *Proc. Nat. Acad. Sci. (USA)* **86**: 3624-3628

Lee M. and Struhl K. (1995) Mutations on the DNA-Binding Surface of TATA-Binding Protein Can Specifically Impair the Response to Acidic Activators *In Vivo*. *Mol Cell Biol* **15**: 5461-5469

Leuther K. K., Salmeron J. M. and Johnston S. A. (1993) Genetic Evidence That an Activation Domain of GAL4 Does Not Require Acidity and May Form a β Sheet. *Cell* **72**: 575-585

Lin Y. and Green M. R. (1988) Interaction of a common cellular transcription factor, ATF. with regulatory elements in both E1a- and cyclic AMP-inducible promoters. *Proc. Nat. Acad. Sci.* **85**: 3396-3400

Lin Y. and Green M. R. (1989) Identification and purification of a *Saccharomyces cerevisiae* protein with the DNA binding specificity of mammalian activating transcription factor. *Proc. Nat. Acad. Sci. (USA)* **86**: 109-113

Lin Y. and Green M. R. (1991) Mechanism of Action of an Acidic Transcriptional Activator *In Vitro*. *Cell* **64**: 971-981

Longtine M. S., Wilson N. M., Petracek M. E. and Berman J. (1989) A yeast Telomere Binding Activity binds to two related telomere sequence motifs and is indistinguishable from RAP1. *Curr Genet* **16**: 225-239

Losa R., Omari S. and Thoma F. (1990) Poly(dA).poly(dT) rich sequences are not sufficient to exclude nucleosome formation in a constitutive yeast promoter. *Nuc Acids Res* **18**: 3495-3502

Lui C., Mao X. and Lustig A.J. (1994) Mutational Analysis Defines a C-terminal Tail Domain of RAP1 Essential for Telomeric Silencing in *Saccharomyces cerevisiae*. *Genetics* **138**: 1025-1040

McBroom L. D. B. and Sadowski P. D. (1994a) Contacts of the ABF1 Protein of *Saccharomyces cerevisiae* with a DNA Binding Site at *MAT*a. *J Biol Chem* **269**: 16455 16460

McBroom L. D. B. and Sadowski P. D. (1994b) DNA Bending by *Saccharomyces cerevisiae* ABF1 and its Proteolytic Fragments. *J Biol Chem* **269**: 16461-16468

McKenzie E. A., Kent N. A., Dowell S. J., Moreno F., Bird L. E. and Mellor J. (1993) The centromere and promoter factor 1, CPF1, of *Saccharomyces cerevisiae*

modulates gene activity through a family of factors including SPT21, RPD1(SIN3), RPD3 and CCR4. *Mol Gen Genet* **240**: 374-386

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

Martinez E., Chiang C., Ge H. and Roeder R. G. (1994) TATA-binding protein-associated factor(s) in TFIID function through the initiator to direct basal transcription froma TATA-less class II promoter. *EMBO J.* **13**: 3115-3126

Masison D. C., O'Connell K. F. and Baker R. E. (1993) Mutational analysis of the *Saccharomyces cerevisiae* general regulatory factor CP1. *Nuc Acids Res* **21**: 4133-4141

Matsui T., Segall J., Weil P. A. and Roeder (1980) Multiple Factors Required for Accurate Initiation of Transcription by Purified RNA Polymerase II. *J. Biol Chem* **255**: 11992-11996

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., Dobson M. J., Roberts N. A., Kingsman A. J. and Kingsman S. M. (1985) Factors affecting heterologous gene expression in *Saccharomyces cerevisiae*. *Gene* **33**: 215-226

Mellor J., Dobson M. J., Kingsman A. J. and Kingsman S. M. (1987) A transcriptional activator is located in the coding region of the yeast *PGK* gene. *Nuc Acids Res* **15**: 6243-6259

Mellor J., Jiang W., Funk M., Rathjen J., Barnes C. A., Hinz T., Hegemann J. H. and Philippsen P. (1990) CPF1, a yeast protein which functions in centromeres and promoters. *EMBO J* **9**: 4017-4026

Mellor J., Rathjen J., Jiang W. and Dowell S. J. (1991) DNA binding of CPF1 is required for optimal centromere function but not for maintaining methionine prototrophy in yeast. *Nuc Acids Res* **19**: 2961-2969

Miller J. H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Miller J., McLachlan A. and Klug A. (1985) Repetitive zinc binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* **4**: 1609-1614

Moretti P., Freeman K., Coodly L. and Shore D. (1994) Evidence that a complex of SIR proteins interacts with the silencer and telomere binding protein RAP1. *Genes Dev* **8**: 2257-2269

Morrow B. E., Johnson S. P. and Warner J. R. (1989) Proteins That Bind to the Yeast rDNA Enhancer. *J Biol Chem* **264**:9061-9068

Morrow B. E., Ju Q. and Warner J. R. (1990) Purification and Characterization of the Yeast rDNA Binding Protein REB1. *J Biol Chem* **265**: 20778-20783

Morrow B. E., Ju Q. and Warner J. R. (1993a) A Bipartite DNA-Binding Domain in Yeast Reb1p. *Mol Cell Biol* **13**: 1173-1182

Morrow B. E., Johnson S. P. and Warner J. R. (1993b) The rRNA enhancer regulates rRNA transcription in *Saccharomyces cerevisiae*. *Mol Cell Biol* **13**: 1283-1289

Mulder W., Winkler A. A., Scholten I. H. J. M., Zonneveld B. J. M., de Winde J. H., Steensma H. Y. and Grivell L. A. (1994) Centromere promoter factors (CPF1) of the yeasts *Saccharomyces cerevisiae* and *Kluyveromyces lactis* are functionally exchangeable, despite low overall homology. *Curr Genet* **26**: 198-207

Muller T., Gilson E., Schmidt R., Giraldo R., Sogo J., Gross H. and Gasser S.M. (1994) Imaging the Asymmetrical DNA Bend Induced by Repressor Activator Protein 1 with Scanning Tunneling Microscopy. *J Structural Biol* **113**: 1-12

Murre C., McCaw P. S. and Baltimore D. (1989a) A New DNA Binding and Dimerization Motif in Immunoglobulin Enhancer Binding, *daughterless*, *MyoD*, and *myc* Proteins. *Cell* **56**:777-783

Murre C., McCaw P. S., Vaessin H., Caudy M., Jan L. Y., Jan Y. N., Cabrera C. V., Buskin J. N., Hauschka S. D., Lassar A. B., Weintraub H. and Baltimore D. (1989b) Interactions between Heterologous Helix-Loop-Helix proteins Generate Complexes that Bind Specifically to a Common DNA Sequence. *Cell* **58**: 537-544

Nehlin J. O. and Ronne H. (1990) Yeast MIG1 repressor is related to the mammalian early growth response and Wilms' tumour finger proteins. *EMBO J.* **9**: 2891-2898

Nehlin J. O., Carlberg M. and Ronne H. (1992) Yeast *SKO1* gene encodes a bZIP protein that binds to the CRE motif and acts as a repressor of transcription. *Nuc Acids Res* **20**: 5271-5278

Neuman-Silberberg F. S., Bhattacharaya S. and Broach J. R. (1995) Nutrient Availability and the RAS/Cyclic AMP Pathway Both Induce Expression of Ribosomal Protein Genes in *Saccharomyces cerevisiae* but by Different Mechanisms. *Mol Cell Biol* **15**: 3187-3196

Nishi K., Park C. S., Pepper A. E., Eichinger G., Innis M. A. and Holland M. J. (1995) The *GCR1* Requirement for Yeast Glycolytic Gene Expression Is Suppressed by Dominant Mutations in the *SGC1* Gene, Which Encodes a Novel Basic Helix-Loop-Helix Protein. *Mol Cell Biol* **15**: 2646-2653

Nishizawa M., Suzuki Y., Nogi Y., Matsumoto K. and Fukasawa T. (1990) Yeast Gal11 protein mediates the transcription activation signal of two different transacting factors, Gal4 and general regulatory factor1/ repressor/activator site binding protein1/ translation upstream factor. *Proc. Nat. Acad. Sci. (USA)* **87**: 5373-5377

O'Connell K. F. and Baker R. E. (1992) Possible Cross-Regulation of Phosphate and Sulfate Metabolism in *Saccharomyces cerevisiae*. *Genetics* **132**: 63-73

O'Connell K. F., Surdin-Kerjan Y. and Baker R. E. (1995) Role of the *Saccharomyces cerevisiae* General Regulatory Factor CP1 in Methionine Biosynthetic Gene Trancription. *Mol Cell Biol* **15**: 1879-1888

O'Shea E. K., Rutkowski R. and Kim P. S. (1989) Evidence That the Leucine Zipper is a Coiled Coil. *Science* **243**: 538-542

Ogden J. E., Stanway C., Kim S., Mellor J., Kingsman A. J. and Kingsman S. M. (1986) Efficient Expression of the *Saccharomyces cerevisiae PGK* Gene Depends on an Upstream Activation Sequence but Does Not Require TATA Sequences. *Mol Cell Biol* **6**: 4335-4343

Oliver S. G. *et al.* (1992) The complete DNA sequence of yeast chromosome III. *Nature* **357**: 38-46

Orr-Weaver T. L., Szostak J. W. and Rothstein R. J. (1981) Yeast Transformation- A Model System for the Study of Recombination. *Proc. Nat. Acad. Sci. (USA)* **78**: 6354-6358

Orr-Weaver T. L., Nicolas A. and Szostak J. W. (1988) Gene Conversion Adjacent to Regions of Double-Strand Break Repair. *Mol Cell Biol* **8**: 5292-5298

Pabo C. O. and Sauer R. T. (1992) Transcription Factors: Structural Families and Principles of DNA recognition. *Ann Rev Biochem* **61**: 1053-1095

Palladino F., Laroche T., Gilson E., Axelrod A., Pillus L. and Gasser S. M. (1993) SIR3 and SIR4 Proteins Are Required for the Positioning and Integrity of Yeast Teleomeres. *Cell* **75**: 543-555

Parvin J. D., Timmers H. Th. M. and Sharp P. A. (1992) Promoter Specificity of Basal Transcription Factors. *Cell* **68**: 1135-1144

Peterson M. G., Inostrroza J., Maxon M. E., Flores O., Admon A., Reinberg D. and Tjian R. (1991) Structure and functional properties of human general transcription factor IIE. *Nature* **354**: 369-373

Petes T. D., Hereford L. M. and Konstantin K. G. (1978) Characterization of two types of yeast ribosomal DNA genes. *J. Bact* **134**: 295-305

Pfeifer K., Kim K-S., Kogan S. and Guarente L. (1989) Functional Dissection and Sequence of Yeast HAP1 Activator. *Cell* **56**: 291-301

Pina B., Berger S., Marcus G. A., Silverman N., Agapite J. and Guarente L. (1993) *ADA3*: a Gene, Identified by Resistance to GAL4-VP16, with Properties Similar to and Different from Those of *ADA2*. *Mol Cell Biol* **13**: 5981-5989

Pinto I., Ware D. E. and Hampsey M. (1992) The Yeast *SUA7* Gene Encodes a Homolog of Human Transcription Factor TFIIB and Is Required for Normal Start Site Selection *In Vivo*. *Cell* **68**: 977-988

Piper P. W., Curran B., Davies 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 in yeast. *Nuc Acids Res* **16**: 1333-1348

Poon D. and Weil P. A. (1993) Immunopurification of Yeast TATA-binding Protein and Associated Factors. *J. Biol Chem* **268**: 15325-15328

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

Pugh B. F. and Tjian R. (1990) Mechanism of Transcriptional Activation by Sp1: Evidence for Coactivators. *Cell* **61**: 1187-1197

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

Ranish J. A. and Hahn S. (1991) The Yeast General Transcription factor TFIIA Is Composed of Two Polypeptide Subunits. *J. Biol Chem* **266**: 19320-19327

Rathjen J. and Mellor J. (1990) Characterization of sequences required for RNA initiation from the *PGK* promoter of *Saccharomyces cerevisiae*. *Nuc Acids Res* **18**: 3219-3225

Reardon R. J., Winters R. S., Gordon D. and Winter E. (1993) A peptide motif that recognizes A·T tracts in DNA. *Proc. Nat. Acad. Sci. (USA)* **90**: 11327-11331

Reinberg D. and Roeder R. G. (1987) Factors Involved in Specific Transcription by Mammalian RNA Polymerase II. *J. Biol Chem* **262**: 3310-3321

Remacle J. E. and Holmberg S. (1992) A REB1-Binding Site Is Required for GCN4-Independent *ILV1* Basal Level Transcription and Can Be Functionally Replaced by an ABF1-Binding Site. *Mol Cell Biol* **12**: 5516-5526

Rhode P.R., Sweder K. S., Oegema K. F. and Campbell J. L. (1989) The gene encoding ARS-binding factor 1 is essential for the viability of yeast. *Genes Dev* **3**: 1926-1939

Rhode P.R., Elsasser S. and Campbell J. L. (1992) Role of Multifunctional Autonomously Replicating Sequence Binding Factor 1 in the Initiation of DNA Replication and Transcriptional Control in *Saccharomyces cerevisiae*. *Mol Cell Biol* **12**: 1064-1077

Rigby P. W. J., Dieckmann N. M., Rhodes C. and Berg P. (1977) Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol Biol* **113**: 237-251

Roberts S. G. E., Ha I., Maldonado E., Reinberg D. and Green M. R. (1993) Interaction between an acidic activator and transcription factor TFIIB is required for transcriptional activation. *Nature* **363**: 741-744

Roberts S. G. E. and Green M. R. (1994) Activator-induced conformational change in general transcription factor TFIIB. *Nature* **371**: 717-720

Roeder R. G. (1991) The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. *Trends Biochem Sci* **16**: 402-408

Rose M. and Winston F. (1984) Identification of a Ty insertion Within the Coding Sequence of the *S. cerevisiae URA3* Gene. *Mol Gen Genet* **193**:557-560

Rothstein R. J. (1983) One-Step Gene Disruption in Yeast. *Methods in Enzymology* **101**: 202-211

Sakurai H., Hiraoka Y. and Fukasawa T. (1993) Yeast GAL11 protein is a distinctive type transcription factor that enhances basal transcription *in vitro*. *Proc. Nat. Acad. Sci. (USA)* **90**: 8382-8386

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

Sayre M. H., Tschochner H. and Kornberg R. D. (1992a) Reconstitution of Transcription with Five Purified Initiation Factors and RNA Polymerase II from *Saccharomyces cerevisiae*. *J. Biol Chem* **267**: 23376-23382

Sayre M. H., Tschochner H. and Kornberg R. D. (1992b) Purification and Properties of *Saccharomyces cerevisiae* RNA Polymerase II General Initiation Factor a. *J. Biol Chem* **267**: 23383-23387

Sawadogo M. and Sentenac A. (1990) RNA Polymerase-B (II) and General Transcription Factors. *Ann Rev Biochem* **59**: 711-754

Scherer S. and Davis R. W. (1979) Replacement of chromosome segments with altered DNA sequences constructed *in vitro*. *Proc. Nat. Acad. Sci. (USA)* **76**: 4951-4955

Schleif R. (1992) DNA Looping. *Ann Rev Biochem* **61**: 199-223

Schultz M.C., Choe S. Y. and Reeder R. H. (1993) In Vitro Definition of the Yeast RNA Polymerase I Enhancer. *Mol Cell Biol* **13**: 2644-2654

Scott E. W. and Baker H. V. (1993) Concerted Action of the Transcriptional Activators REB1, RAP1, and GCR1 in the High-Level Expression of the Glycolytic Gene *TPI*. *Mol Cell Biol* **13**: 543-550

Sellers J. W., Vincent A. C. and Struhl K. (1990) Mutations That Define the Optimal Half-Site for Binding Yeast GCN4 Activator Protein and Identify an ATF/CREB-Like Repressor that Recognises Similar DNA Sites. *Mol Cell Biol* **10**: 5077-5086

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

Shore D., Stillman D. J., Brand A. H. and Nasmyth K. A. (1987) Identification of silencer binding proteins from yeast: possible roles in SIR control and DNA replication. *EMBO J* **6**: 461-467

Shortle D., Novick P. and Botstein D. (1984) Construction and genetic characterization of temperature-sensitive mutant alleles of the yeast actin gene. *Proc. Nat. Acad. Sci. (USA)* **81**: 4889-4893

Sigler P. B. (1988) Acid blobs and negative noodles. *Nature* **333**: 210-212

Simpson R. T. (1990) Nucleosome positioning can affect the function of a *cis*- acting DNA element *in vivo*. *Nature* **343**: 387-389

Snyder M., Buchman A. R. and Davis R. W. (1986) Bent DNA at a yeast autonomously replicating sequence. *Nature* **324**: 87-89

Sopta M., Burton Z. F. and Greenblatt J. (1989) Structure and associated DNA-helicase activity of a general transcription initiation factor that binds to RNA polymerase II. *Nature* **341**: 410-414

Stanway C., Mellor J., Ogden J. E., Kingsman A. J. and Kingsman S. M. (1987) The UAS of the yeast *PGK* gene contains functionally distinct domains. *Nuc Acids Res* **15**: 6855-6873

Stanway C., Chambers A., Kingsman A. J. and Kingsman S. M. (1989) Characterization of the transcriptional potency of sub-elements of the UAS of the yeast *PGK* gene in a *PGK* minimal promoter. *Nuc Acids Res* **17**: 9205-9218

Stanway C. A., Gibbs J. M., Kearsey S. E., Lopez M. C. and Baker H. V. (1994) The yeast co-activator GAL11 positively influences transcription of the phosphoglycerate kinase gene, but only when Rap1p is bound to its upstream activation sequence. *Mol Gen Genet* **243**: 207-214

Strathern J. N. and Higgins D. R. (1991) Recovery of Plasmids from Yeast into *Escherichia coli*: Shuttle vectors. *Methods in Enzymol* **194**: 319-329

Struhl K. (1984) Genetic properties and chromatin structure of the yeast *gal* regulatory element: An enhancer-like sequence. *Proc. Nat. Acad. Sci. (USA)* **81**: 7865-7869

Struhl K. (1985) Naturally occuring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc. Nat. Acad. Sci. (USA)* 82: 8419-8423

Struhl K. (1994) Duality of TBP, the Universal Transcription Factor. *Science* **263**: 1103-1104

Studier F. W., Rosenberg A. H., Dunn J. J. and Dubendorff J. W. (1990) Use of T7 RNA Polymerase to Direct Expression of Cloned Genes. *Methods Enzymol* **185**: 60-89

Sussel L. and Shore D. (1991) Separation of transcriptional activation and silencing functions of the *RAP1*-encoded repressor/activator protein 1: Isolation of viable mutants affecting both silencing and telomere length. *Proc. Nat. Acad. Sci. (USA)* **88**: 7749-7753

Sussel L., Vannier D. and Shore D. (1995) Suppressors of Defective Silencing in Yeast: Effects on Transcriptional Repression at the HMR Locus, Cell Growth and Telomere Structure. *Genetics* **141**: 873-888

Swaffield J. C., Bromberg J. F. and Johnston S. A. (1992) Alterations in a yeast protein resembling HIV Tat-binding protein relieve requirement for an acidic activation domain in GAL4. *Nature* **357**: 698-700

Sweder K. S., Rhode P. R. and Campbell J. L. (1988) Purification and Characterization of Proteins that Bind to Yeast ARSs. *J Biol Chem* **263**: 17270-17277

Tan S., Aso T., Conaway R. C. and Conaway J. W. (1994) Roles for Both the RAP30 and RAP74 Subunits of Transcription Factor IIF in Transcription Initiation and Elongation by RNA Polymerase. *J. Biol Chem* **269**: 25684-25691

Thomas D., Cherest H. and Surdin-Kerjan Y. (1989) Elements Involved in S-Adenosylmethionine-Mediated Regulation of the *Saccharomyces cerevisiae MET25* Gene. *Mol Cell Biol* **9**: 3292-3298

Thomas D., Jacquemin I. and Surdin-Kerjan Y. (1992) MET4, a Leucine Zipper Protein, and Centromere-Binding Factor 1 Are Both Required for Transcriptional Activation of Sulfur Metabolism in *Saccharomyces cerevisiae*. *Mol Cell Biol* **12**: 1719-1727

Thompson C. M., Koleske A. J., Chao S. M. and Young R. A. (1993) A Multisubunit Complex Associated with the RNA Polymerase II CTD and TATA-Binding protein in Yeast. *Cell* **73**: 1361-1375

Toda T., Cameron S., Sass P., Zoller M., Scott J. D., McMullen B., Hurwitz M., Krebs E. G. and Wigler M. (1987) Cloning and Characterization of *BCY1*, a Locus Encoding a Regulatory Subunit of the Cyclic AMP-Dependent Protein Kinase in *Saccharomyces cerevisiae*. *Mol Cell Biol* **7**: 1371-1377

Tornow J., Zeng X., Gao W. and Santangelo G. M. (1993) GCR1, a transcriptional activator in *Saccharomyces cerevisiae*, complexes with RAP1 and can function without its DNA binding domain. *EMBO J* **12**: 2431-2437

Trawick J. D., Kraut N., Simon F. R. and Poyton R. O. (1992) Regulation of Yeast *COX6* by the General Transcription Factor ABF1 and Separate HAP2- and Heme-Responsive Elements. *Mol Cell Biol* **12**: 2302-2314

Tsang J. S. H., Henry Y. A. L., Chambers A., Kingsman A. J. and Kingsman S. M. (1990) Phosphorylation influences the binding of the yeast RAP1 protein to the upstream activating sequence of the *PGK* gene. *Nuc Acids Res* **18**: 7331-7337

Tschochner H., Sayre M. H., Flanagan P. M., Feaver W. J. and Kornberg R. D. (1992) Yeast RNA polymerase II initiation factor e: Isolation and identification as the functional counterpart of human transcription factor IIB. *Proc. Nat. Acad. Sci. (USA)* **89**: 11292-11296

Tyree C. M., George C. P., Lira-DeVito L. M., Wampler S. L., Dahmus M. C., Zawel L. and Kadonaga J. T. (1993) Identification of a minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymeraseII. *Genes Dev* **7**: 1254-1265

Uemura H. and Fraenkel D. G. (1990) *gcr2*, a New Mutation Affecting Glycolytic Gene Expression in *Saccharomyces cerevisiae*. *Mol Cell Biol* **10**: 6389-6396

Uemura H. and Jigami Y. (1992a) *GCR3* Encodes an Acidic Protein That Is Required for Expression of Glycolytic Genes in *Saccharomyces cerevisiae*. *J Bact* **174**: 5526-5532

Uemura H. and Jigami Y. (1992b) Role of GCR2 in Transcriptional Activation of Yeast Glycolytic Genes. *Mol Cell Biol* **12**: 3834-3842

Van Hoy M., Leuther K. K., Kodadek T. and Johnston S. A. (1993) The Acidic Activation Domains of the GCN4 and GAL4 Proteins Are Not α Helical but Form β Sheets. *Cell* **72**: 587-594

Verdier J. M., Stalder R., Roberge M., Amati B., Sentenac A. and Gasser S. M. (1990) Preparation and characterization of yeast nuclear extracts for efficient RNA polymerase B (II)-dependent transcription *in vitro*. *Nuc Acids Res* **18**: 7033-7039

Vignais M. and Sentenac A. (1989) Asymmetric DNA Bending Induced by the Yeast Multifunctional Factor TUF. *J. Biol Chem* **264**: 8463-8466

Vincent A. C. and Struhl K. (1992) ACR1, a Yeast ATF/CREB Repressor. *Mol Cell Biol* **12**: 5394-5405

Wang H., Nicholson P. R. and Stillman D. J. (1990) Identification of a *Saccharomyces cerevisiae* DNA-Binding Protein Involved in Transcription Regulation. *Mol Cell Biol* **10**: 1743-1753

Weis L. and Reinberg D. (1992) Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. *FASEB J.* **6**: 3300-3309

Willett C. E., Gelfman C. M. and Holland M. J. (1993) A Complex Regulatory Element from the Yeast Gene *ENO2* modulates GCR1-Dependent Transcriptional Activation. *Mol Cell Biol* **13**: 2623-2633

Workman J. L., Roeder R. G. and Kingston R. E. (1990) An upstream transcription factor. USF (MLTF), facilitates the formation of preinitiation complexes during *in vitro* chromatin assembly. *EMBO J.* **9**: 1299-1308

Wynne J. and Treisman R. (1992) SRF and MCM1 have related but distinct DNA binding specificities. *Nuc Acids Res* **20**: 3297-3303

Xiao H., Pearson A., Coulombe B., Truant R., Zhang S., Regier J. L., Triezenberg S. J., Reinberg D., Flores O., Ingles C. J. and Greenblatt J. (1994) Binding of Basal Transcription factor TFIIH to the Acidic Activation Domains of VP16 and p53. *Mol Cell Biol* **14**: 7013-7024

Young R. A. and Davis R. W. (1983) Yeast RNA Polymerase II Genes: Isolation with Antibody Probes. *Science* **222**: 778-782

Young R. A. (1991) RNA Polymerase II. *Ann Rev Biochem* **60**: 689-715

Zawel L., Kumar K. P. and Reinberg D. (1995) Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev* **9**: 1479-1490

Zhou Q., Boyer T. G. and Berk A. J. (1993) Factors (TAFs) required for activated transcription interact with TATA box-binding protein conserved core domain. *Genes Dev* **7**: 180-187

Curr Genet (1995) 29: 1–9

REVIEW

Alistair Chambers · Elizabeth A. Packham
Ian R. Graham

Control of glycolytic gene expression in the budding yeast (*Saccharomyces cerevisiae*)

Received: 16 March/19 June 1995

Key words Glycolysis · Transcriptional activation · *Saccharomyces cerevisiae* · Chromatin structure · Glucose induction

Introduction

The glycolytic pathway (Fig. 1) is central to the biochemistry of the yeast cell and is essential for many of the processes in which yeast is used for the benefit of mankind. It allows yeast cells growing under aerobic conditions to convert sugars to ethanol and, in the process, produce energy. The genes encoding the enzymes of the pathway have been cloned and intensively studied. Some of them were amongst the first yeast genes to be isolated and because of their high expression levels their promoters have been widely used to construct yeast expression vectors and as model systems to study transcription. As details of the organisation of glycolytic promoters have emerged, it has become clear that these "simple housekeeping genes" have sophisticated molecular mechanisms controlling their expression. The purpose of this review is to bring together results from a wide range of studies in an attempt to produce an overall view of glycolytic gene expression. We will highlight both the similarities and differences between different glycolytic promoters and the potential use of these systems to answer important general questions about yeast gene expression.

A. Chambers (✉) · E.A. Packham · I.R. Graham
Department of Genetics, University of Nottingham, Queen's Medical Centre, Nottingham NG7 2UH, UK

Communicated by L.A. Grivell

Glycolytic promoters: variations on a common theme?

The organisation of glycolytic gene promoters is summarised in Fig. 2. Some glycolytic promoters, such as *PGK*, *ENO2* and *TPI*, contain only positively acting upstream elements [upstream activation sequences (UAS)] (Cohen et al. 1986; Ogden et al. 1986; Scott et al. 1990). Others, such as *ENO1*, *TDH3* and *PGM1*, also contain negatively acting upstream repression sequences (URS) (Cohen et al. 1987; Rodicio et al. 1993; Kuroda et al. 1994). Much attention has focused on determining the sequence organisation of glycolytic UAS and the characterisation of the transcription factors which interact with them. Most of these UAS contain potential binding sites for the multifunctional transcription factor Rap1p (Shore and Nasmyth 1987). Rap1p binds to the promoters of many housekeeping genes, the mating-type silencers *HML* and *HMR*, and the repeat regions of telomeres (Huet et al. 1985; Shore and Nasmyth 1987; Buchman et al. 1988). Rap1p has been shown to bind in vitro to the promoters of the glycolytic genes *TPI*, *TDH3*, *PGK*, *ENO1*, *ENO2*, *PYK1*, *PDC1* and *ADH1* (Chambers et al. 1989; Nishizawa et al. 1989; Brindle et al. 1990; Butler et al. 1990; Santangelo and Tornow 1990; Scott et al. 1990; Bitter et al. 1991). In vivo footprinting experiments have confirmed that the sites in the *TPI*, *PGK* and *PYK1* promoters are occupied in yeast cells (Scott and Baker 1993; Dumitru and McNeil 1994; Stanway et al. 1994). The weight of evidence suggests that Rap1p is a general binding factor at glycolytic promoters. In those promoters where Rap1p binding has been demonstrated it appears to play an important role in transcriptional activation, although the extent of the dependence of different promoters on Rap1p is variable. In the *PGK* promoter, deletion of the single Rap1p-binding site caused a reduction in transcription of more than 80%

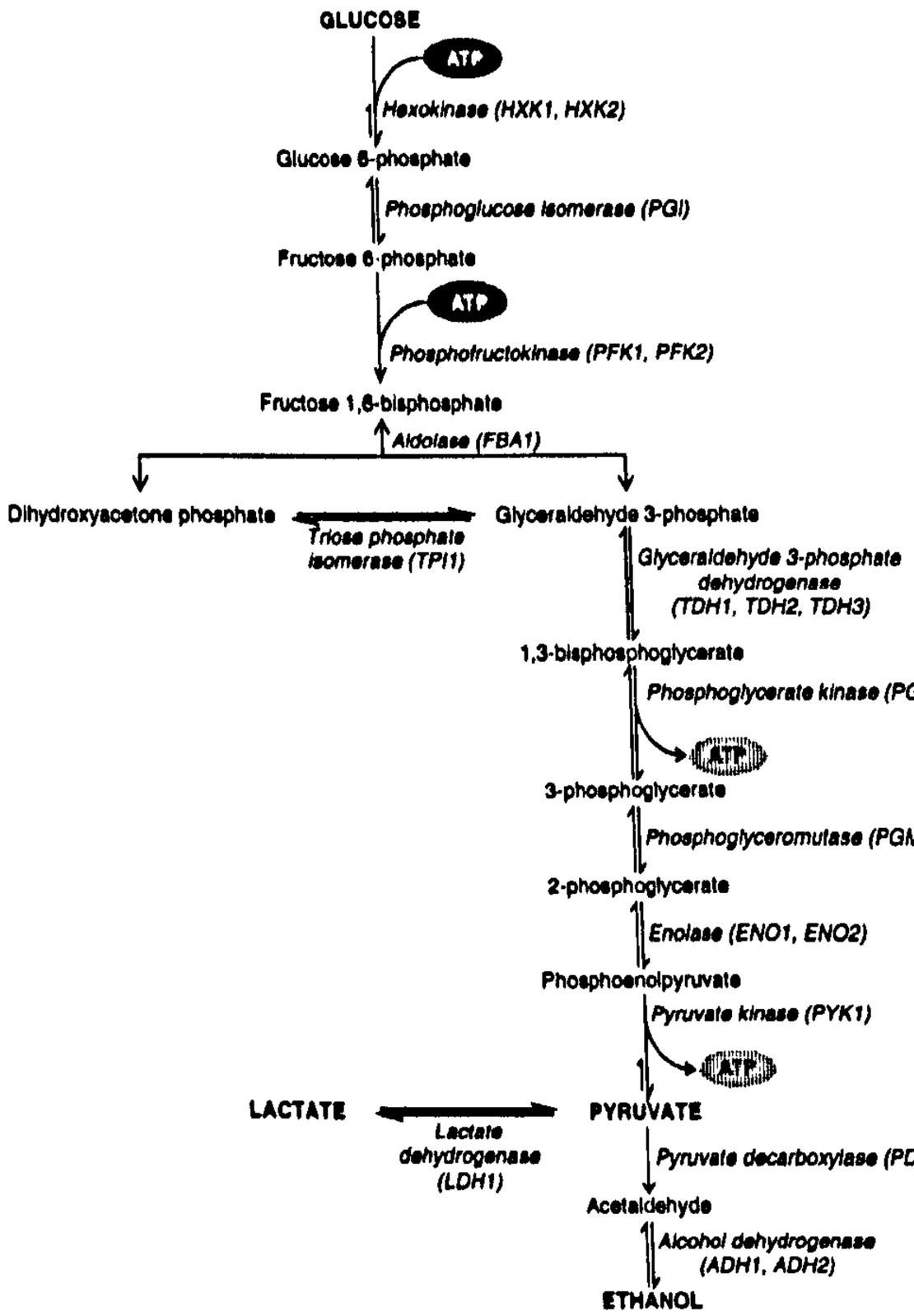


Fig. 1 The glycolytic pathway. Gene designations are shown in *italics* next to the enzymes they encode

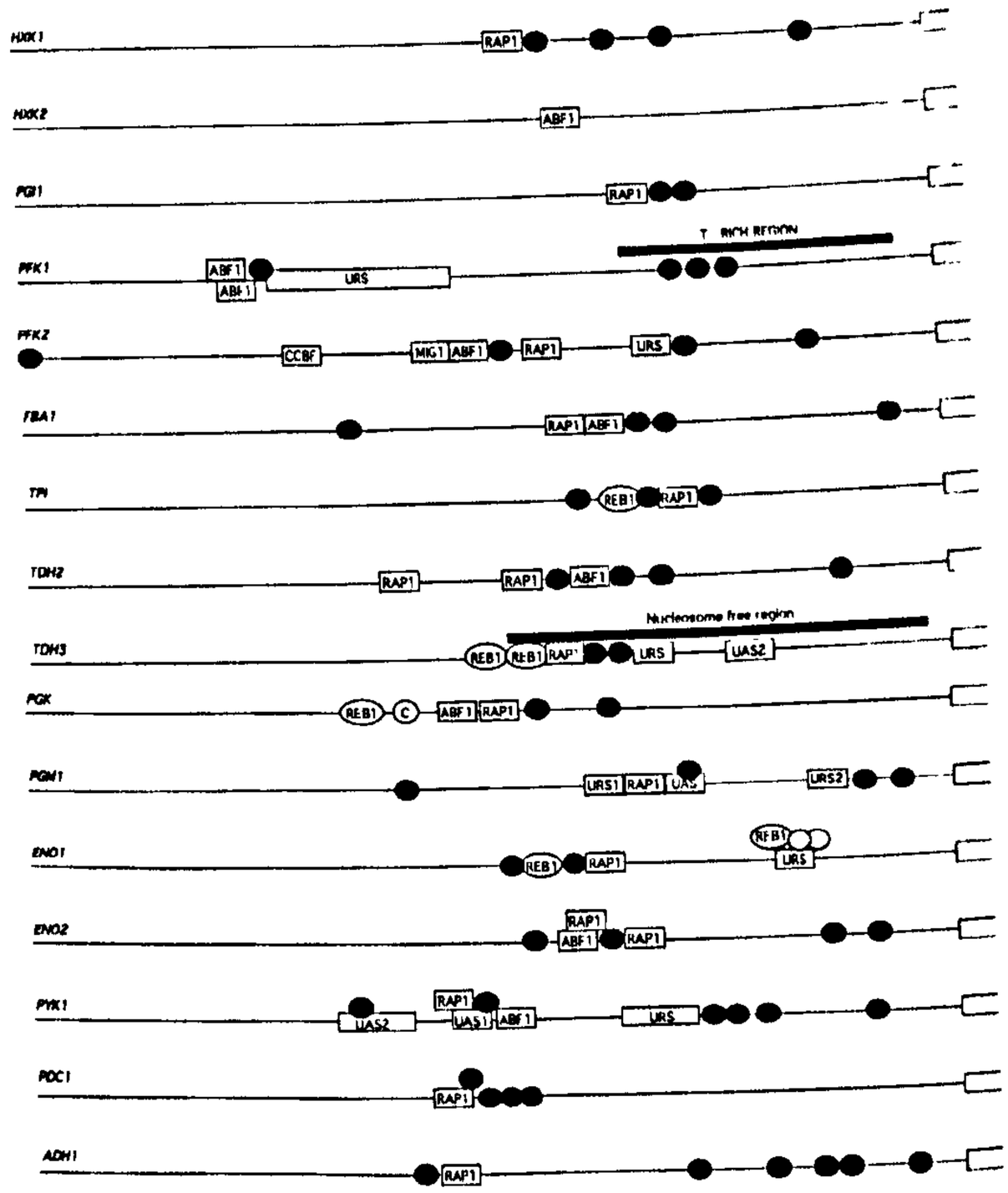


Fig. 2 Organisation of glycolytic gene promoters. The diagrams are approximately to scale and show the positions of sequence elements in the promoters of glycolytic genes. The Rap1p (RAP1)- and Abf1p (ABF1)-binding sites include potential binding sites (matches to consensus recognition sequence), sites shown to bind in vitro and sites shown to bind in vivo, these are discussed in the text. Gcr1p-binding sites are shown as *filled-in circles*, these also include potential binding sites and sites shown to bind in vivo. *UAS* and *URS* show the positions of previously designated regulatory sequences. REB1 indicates the positions of binding sites for Reb1p. *C* shows the position of an in vitro binding site for Cpf1p. *CCBF* and *MIG1* show positions of potential binding sites for the cell-cycle box factor and the Mig1p transcriptional repressor respectively. The *open circles* at the *ENO1* URS represent uncharacterised binding factors

(Chambers et al. 1988). Similarly, mutation of the Rap1p-binding site in the *TPI* UAS caused a ten-fold reduction in transcription (Scott and Baker 1993). In the *PFK2* and *PGM1* promoters, potential Rap1p-binding sites have been identified which play little, if any, role in activation (Heinisch et al. 1991; Rodicio et al. 1993). However, these sites have been less well defined than sites in other promoters and Rap1p binding has not been demonstrated either in vitro or in vivo.

The second major element found in glycolytic promoters is one or more binding sites for the glycolytic regulatory protein Gcr1p. This protein interacts with a recognition sequence containing the pentamer 5′ CT/ATCC3′ at its core (Baker 1991). This so-called "CT block" was first recognised as an important element in the *PGK* promoter and subsequently in the *TDH3* and *TPI* gene promoters (Ogden et al. 1986; Chambers et al. 1988; Scott et al. 1990; Bitter et al. 1991). Although it was long regarded as a potential binding site for Gcr1p, it was extremely difficult to demonstrate Gcr1p binding to this sequence in vitro, because the interaction between this protein and DNA is very weak. Sequence-specific DNA binding by Gcr1p was demonstrated in 1991 and a consensus DNA recognition site was derived in 1992 (Baker 1991; Huie et al. 1992). The Gcr1p-binding sites shown in Fig. 2 are mainly potential binding sites based on matches to the in vitro consensus. Gcr1p has been demonstrated to bind in vivo at the promoters of the *TPI*, *PYK1* and *PGK* genes (Scott and Baker 1993; Dumitru and McNeil 1994; Stanway et al. 1994). It is possible to find potential Gcr1p-binding sites at a number of locations in every glycolytic promoter; however, it is likely that only those sites in reasonable proximity to Rap1p-binding sites are functionally important in vivo (Huie et al. 1992).

Binding sites for the other multifunctional transcription factors Abf1p and Reb1p are also found within some glycolytic promoters. Abf1p binds to ARS sequences (yeast origins of replication), mating-type silencers, and a number of gene promoters (Buchman

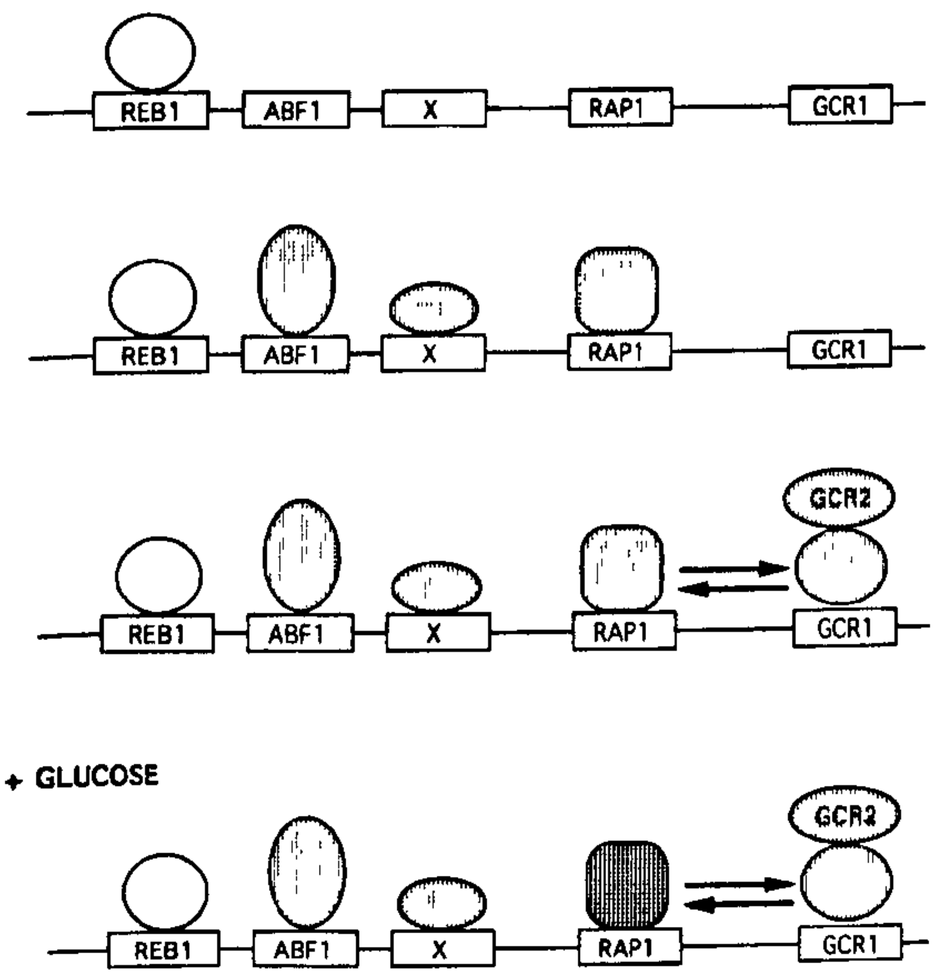


Fig. 3 A model for transcriptional activation of glycolytic genes. The diagram shows a hypothetical glycolytic promoter containing binding sites for Reb1p (*REB1*), Abf1p (*ABF1*), another transcriptional activator (*X*), Rap1p (*RAP1*) and Gcr1p (*GCR1*). Reb1p may bind to the promoter and generate a nucleosome-free region to aid the binding of the other factors. Rap1p binding stabilises binding of Gcr1p, which in turn interacts with Gcr2p. Each factor alone contains a weak transcriptional-activation domain (*light grey shading*) and these domains work together to stimulate transcription. In the presence of glucose, phosphorylation of the activation domain of Rap1p leads to an increase in the activation potential of the protein (*dark grey shading*) and an overall increase in transcriptional activation

et al. 1988; Diffley Stillman 1988, 1989; Dorsman et al. 1989; Rhode et al. 1989; Dorsman et al. 1990; Rhode et al. 1992). Glycolytic promoters containing Abf1p-binding sites include those of the *PFK1*, *PFK2*, *FBA1*, *PGK*, *ENO2* and *PYK1* genes (Brindle et al. 1990; Chambers et al. 1990; Heinisch et al. 1991). The binding sites in the *PFK1* and *PGK* promoters appear to play some role in transcriptional activation (Heinisch et al. 1991; E.A.P., I.R.G. and A.C., manuscript submitted). In contrast, the site within the *PYK1* promoter plays little, if any, role (Nishizawa et al. 1989). In the *ENO2* promoter, Abf1p acts as anti-repressor, allowing an unknown activator to work in combination with Rap1p and Gcr1p (Willett et al. 1993).

Reb1p was first identified as a factor binding to a RNA polymerase I enhancer (Morrow et al. 1989; Ju et al. 1990). It was subsequently shown to be important in RNA polymerase I transcription termination and to bind to a number of polymerase II promoters (Chasman et al. 1990; Lang et al. 1994). Binding sites for this factor have been identified in the *TPI*, *TDH3*, *PGK* and *ENO1* promoters (Scott et al. 1990; Carmen and Holland 1994; Yagi et al. 1994; E.A.P., I.R.G. and A.C., manuscript submitted). In the *TPI* promoter, the Reb1p-binding site is important for transcriptional activation. A mutation in this site led to a five-fold reduction in gene expression (Scott and Baker 1993). The sites in *TDH3* also appear to be important and it is interesting to note that these sites correspond with the 5′ end of a nucleosome-free region which extends over most of the rest of the promoter (Pavlovic and Horz 1988; Yagi et al. 1994). The *ENO1* promoter contains binding sites for Reb1p as components of both UAS and URS elements, which emphasises the multifunctional nature of this transcription factor (Carmen and Holland 1994).

Overall, glycolytic promoters show many similarities in organisation: all appear to contain binding sites for Rap1p and Gcr1p. They also show differences in details: some are bound by Reb1p and Abf1p and some contain URS as well as UAS. The basic mechanisms which control the activation of these genes are likely to be similar (Fig. 3) but imposed on these are gene-specific differences which lead to the variations in expression levels and induction characteristics which these genes display.

Downstream activation sequences in glycolytic genes

Several glycolytic genes contain sequences within their coding regions which appear to be important for maximal gene expression. Such sequences have been termed downstream activation sequences or DAS. Genes in which DAS have been identified include *PGK* and *PYK1* (Mellor et al. 1987; Purvis et al. 1987). Recently the presence of a DAS within the *TDH3* gene has also been suggested, although in this case the evidence is not conclusive (Jung et al. 1995). The *PGK* DAS has been mapped to positions +37 to +236 relative to the ATG of the gene (Mellor et al. 1987). The *PYK1* DAS has been less-well defined, but is thought to lie between positions +516 and +870. An intriguing feature of DAS is the presence of potential binding sites for multifunctional transcription factors. The *PGK* DAS contains an Abf1p-binding site between positions +79 and +91 (I.R.G. and A.C., unpublished) and there is also a good potential binding site for Rap1p towards the end of the *PGK* coding region (Fantino et al. 1992). The *TDH3* coding region contains an Abf1p-binding site within a DNA fragment extending from position +77 to +200 (Jung et al. 1995). This DNA fragment has been shown to activate expression of a reporter gene when inserted upstream of basal promoter elements, but a role in *TDH3* expression has yet to be established (Jung et al. 1995). The *PYK1* coding region also contains a number of potential binding sites for both Rap1p and Abf1p. In the few other genes where DAS have been identified, binding sites for multifunctional transcription factors are also present. The *SRP1*

Table 1 Effects of *gcr1*, *2* and *3* mutations on the activities of glycolytic enzymes in yeast cells grown on different carbon sources. Activities shown are percentages compared to enzyme activity in a wild-type strain (Clifton and Fraenkel 1981; Uemura and Fraenkel 1990; Uemura and Jigami 1992 a)

Enzyme	*gcr1* (glycerol/lactate)	*gcr1* (glycolytic media)	*gcr2* (glycerol/lactate)	*gcr3-2* (glycerol/lactate)	*gcr3-2* (glucose)
HXK	174	226	134	88	76
PGI	9	20	14	58	63
PFK	57	72	58	63	59
FBA	8	31	6	85	139
TPI	4	22	7	106	64
TDH	26	45	7	92	64
PGK	21	55	11	82	62
PGM	1	2	4	118	49
ENO	2	5	3	44	29
PYK	13	32	3	80	82

gene encodes a highly expressed serine-rich protein (Marguet and Lauquin 1986). This gene contains a Rap1p-binding site between positions +710 and +743, which has been shown to be important for maximal gene expression (Fantino et al. 1992). The *LPD1* gene encoding lipoamide dehydrogenase contains both DAS and downstream repression sequences (DRS) (Sinclair et al. 1994). In this case a binding site for the centromere/promoter-binding factor Cpf1p is associated with one DAS, and Abf1p- and Rap1p-binding sites are found associated with the DRS.

Mutations which affect glycolytic gene expression

The notion that glycolytic genes may share a common mechanism of transcriptional activation arose from the isolation of mutations which affect the expression of almost the whole set of these genes. Perhaps the most important of these is the *gcr1* mutation (Clifton and Fraenkel 1981). Yeast strains which possess *gcr1* mutations show reduced expression of most glycolytic enzymes on both fermentable and non-fermentable media (Table 1) (Clifton and Fraenkel 1981). Generally, the effects are more pronounced on non-fermentable media and the different glycolytic enzymes are affected to differing degrees. HXK enzyme activity increases in *gcr1* mutants but all the rest of the glycolytic enzymes decrease in activity. The most severely affected are the enzymes PGI, FBA, TPI, PGM and ENO. These variations in enzyme levels are probably due to effects at the level of transcription. Differences in mRNA levels have been demonstrated for the *ENO1*, *ENO2*, *TDH1,2,3*, *TPI*, *PGK* and *ADH1* genes (Holland et al. 1987; Santangelo and Tornow 1990; Scott et al. 1990; Henry et al. 1994). More recently two other mutations, *gcr2* and *gcr3*, which affect the expression of most glycolytic genes have been isolated (Uemura and Fraenkel 1990; Uemura and Jigami 1992 a). Of these, *gcr2* is the most severe. This mutation has a significantly greater effect in cells grown on non-fermentable carbon sources (Uemura and Fraenkel 1990). In cells grown on medium containing glucose, the growth defect is only partial, a less severe phenotype than that seen in *gcr1* mutants. In general, the genes most affected by *gcr2* are the same as those most affected by *gcr1*. *gcr3* is a less severe mutation which has more effect in glucose-grown cells than in cells grown on non-fermentable carbon sources (Uemura and Jigami 1992 a). However, even in glucose-grown cells, the effects are much less dramatic than those of *gcr1* and *gcr2* mutations. Mutations in three other genes, designated *SGC1*, *SGC2* and *SGC5*, can suppress the requirement of glycolytic genes for the *GCR1* gene product (Nishi et al. 1995). Dominant mutations in *SGC1* and *SGC2*, and recessive mutations in *SGC5*, restore normal growth to a *gcr1* null strain. The mutation in *SGC1* has been shown to mediate effects at the levels of transcription. In a *SGC1/gcr1* double-mutant strain, the level of *PGK*, *PGM1*, *PYK1* and *TPI* mRNAs were all found to be higher than in a *gcr1* single-mutant strain. The *SGC1* gene has been cloned and sequenced and has been shown to encode a transcription factor of the basic-helix-loop-helix class (Nishi et al. 1995). Mutations have also been identified which affect the expression of individual glycolytic genes. One such mutation is *pop2* which affects the expression of the *PGK* gene (Sakai et al. 1992). This mutation causes an increase in activity of the *PGK* promoter on non-fermentable carbon sources. In the *pop2* mutant strain *PGK* expression is no longer inducible by glucose. This suggests that the *POP2* gene influences the mechanism by which the *PGK* promoter is maintained in a repressed state on non-fermentable carbon sources. The gene encodes a protein with glutamine-rich, proline-rich and serine/threonine-rich regions. It may be a transcription factor, but it contains no obvious DNA-binding motifs. No information is currently available regarding the effect of the *pop2* mutation on other glycolytic genes. Mutations in a gene designated *PDC2* specifically affect the expression of *PDC1* and the second structural gene for pyruvate decarboxylase, *PDC5* (Hohmann 1993;

Raghuram et al. 1994). Pdc2p appears to be a transcription factor with an activation domain rich in asparagine residues (Raghuram et al. 1994). Other glycolytic genes do not appear to be affected by *pdc2* mutations (Hohmann 1993).

The role of Rap1p at glycolytic gene promoters

Central to the functioning of glycolytic promoters are the transcription factors Rap1p and Gcr1p. Neither of these alone is sufficient to allow high-level gene expression, but they work synergistically when combined (Stanway et al. 1989; Bitter et al. 1991). Rap1p was first identified as the protein factor TUF which was shown to interact with RPG box sequences in the promoters of the *TEF1*, *TEF2*, *EF-1α* and *RP51A* genes (Huet et al. 1985). TUF was suggested to be a general factor which coordinates expression of many genes encoding different protein components of the translational machinery. The *RAP1* (*TUF*) gene was cloned in 1987 and was shown to encode a protein of 827 amino acids, essential for the survival of yeast cells (Shore and Nasmyth 1987). Extensive studies on the promoters of ribosomal protein genes have provided important insights into the role of Rap1p in transcriptional activation. Many ribosomal protein gene promoters contain two tandemly arranged RPG boxes, often in combination with a downstream T-rich element (Rotenberg and Woolford 1986; Vignais et al. 1987; Woudt et al. 1987). These motifs work together to mediate high-level, regulated, gene expression. Other ribosomal protein gene promoters contain the T-rich element, but instead of RPG boxes possess a single Abf1p-binding site (Hamil et al. 1988; Herruer et al. 1989; Goncalves et al. 1995). Again the Abf1p site works not alone, but in combination with the T-rich region (Goncalves et al. 1995). These studies suggest that the roles of Rap1p and Abf1p at gene promoters may be similar and that both only work efficiently to stimulate gene expression when combined with other factors. The mechanisms by which Rap1p and Abf1p increase transcription have been hard to define. An activation domain has been characterised within Rap1p using Gal4p/Rap1p fusions and a reporter gene containing multiple Gal4p-binding sites upstream of a basal promoter. This approach defined the Rap1p activation domain as the region between amino acids 630 and 695, although even greater activation was obtained when the region between amino acids 630 and 727 was tested (Hardy et al. 1992). More recent experiments using LexA/Rap1p fusions demonstrated that the region between amino acids 635 and 827 of Rap1p is not an activator, but that the region between amino acids 630 and 827 is a strong activator (Moretti et al. 1994). This indicates that the five amino acids which are different between these two constructs are critical for transcriptional activation in the context of the reporter system used.

A set of *rap1* mutant alleles, known as the *rap1*^{t} alleles, cause telomere elongation and instability (Kyrion et al. 1992). One of these alleles encodes a protein which lacks the C terminus from position 664. Cells containing this allele are viable, although they grow more slowly than normal. As activation is almost certainly the essential role of Rap1p, this allele must contain a functional activation domain. If this is correct the activation domain can be further localised to positions 630 to 664. Domain swapping experiments are open to the criticism that the regions identified may not function in the same way within the context of the complete protein and that reporter-gene assays may examine the ability to activate transcription under unrealistic conditions. For these reasons it would be useful to have *rap1* mutations which affect only the activation function of the protein. However, within the large collection of mutant *rap1* alleles described, there are none specifically defective in transcriptional activation. Strains containing temperature-sensitive alleles of *rap1* (*rap1*^{ts} alleles) show decreased activation of the *MATα* locus at non-permissive temperatures, but this effect is thought to be due to lack of binding of Rap1p to the *MATα* alpha UAS, rather than a defect in transcriptional activation (Kurtz and Shore 1991). The *rap 1-5* allele has a mutation at amino acid position 694, at one end of the activation domain of Rap1p. Although it is possible that this mutation could effect activation by Rap1p this is unlikely because the *rap 1-5* allele is one of the least severe of the *rap*^{ts} alleles isolated (Kurtz and Shore 1991). In addition, the mutant Rap1p produced has a heat-labile DNA-binding activity, despite the fact that the mutation is not within the DNA-binding domain of the protein (Henry et al. 1990). The presence of an activation domain within Rap1p might suggest a function similar to that of more conventional transcription factors such as Gal4p and Gcn4p. However, the context-dependent functioning of the protein and its multifunctional nature hint at a more unusual role. This role might involve a re-modelling of chromatin structure around UAS. Evidence for such a role comes from the *HIS4* promoter where Rap1p binding is required for the action of Bas1p, Bas2p and Gcn4p (Devlin et al. 1991). The presence of a Rap1p-binding site causes two adjacent regions of the UAS to show increased sensitivity to micrococcal nuclease. These regions contain the binding sites for Bas1p and Bas2p, and Gcn4p, respectively. This work suggested a role for Rap1p in maintaining accessibility of transcription factor-binding sites within chromatin. A recombination initiation site upstream of the *HIS4* gene also requires binding of Rap1p, Bas1p and Bas2p (White et al. 1993). Interestingly, either two Rap1p sites, or a short region of telomeric DNA, can functionally replace the normal site, perhaps by altering chromatin structure to give increased accessibility to the recombination machinery. The product of the *SIN4* gene is required for full expression of *HIS4* and for *Ty1*, *MATα* and *CTS1*, all of which

contain Rap1p sites within their promoters (Jiang and Stillman 1995). However, Sin4p appears not to work through Rap1p because expression of a *CYC1-lacZ* reporter gene driven by Rap1p-binding sites was unaltered in a *sin4* mutant strain (Jiang and Stillman 1995). The ability of Rap1p to bend DNA may also play some role in Rap1p function at promoters. Rap1p induces a DNA bend upstream of its binding site (Vignais and Sentenac 1989; Gilson et al. 1993). The domain of Rap1p required for this bending appears to be located in the N terminus, a region of the protein not thought to be essential in vivo (Shore and Nasmyth 1987; Gilson et al. 1994). DNA bending cannot therefore be critical for Rap1p function.

Other more radical suggestions have also been made to explain the role of Rap1p in transcriptional activation. One possibility is that interactions between Rap1p and other nuclear components localise gene promoters to a particular domain within the nucleus (Klein et al. 1992). Interactions between the C terminus of Rap1p, Sir2p, Sir3p, Sir4p and Rif1p have been shown to be important in transcriptional silencing and in localising telomeres to the nuclear periphery (Klein et al. 1992; Moretti et al. 1994; Cockell et al. 1995). Factors involved in activation could compete for interactions with Rap1p and sub-localise promoters to a different region of the nucleus.

The roles of Gcr1p, Gcr2p and Gcr3p at glycolytic promoters

Although the mechanism by which Rap1p exerts its effects at glycolytic promoters is unclear, one its roles in this context appears to be to enhance the targeting of Gcr1p to UAS. Gcr1p is a transcription factor which interacts directly with glycolytic gene promoters. It contains an activation domain in the N-terminal one-third of the protein and a sequence-specific DNA-binding domain at the C-terminus (Huie et al. 1992; Tornow et al. 1993). The interaction between Gcr1p and DNA in vivo appears to be stabilised by an interaction between Gcr1p and Rap1p bound to an adjacent DNA sequence (Scott and Baker 1993). In some situations Gcr1p may be targeted to promoters via a protein: protein interaction with Rap1p, without the necessity for direct DNA binding by Gcr1p (Tornow et al. 1993). This is unlikely to be the case in glycolytic promoters because in these promoters most Rap1p sites have Gcr1p sites located close by. Furthermore, Gcr1p has been shown to interact with DNA in vivo at the *TPI*, *PGK* and *PYK1* promoters (Scott and Baker 1993; Dumitru and McNeil 1994; Stanway et al. 1994). Gcr1p is thought to work in combination with Gcr2p (Uemura and Jigami 1992 b). A direct interaction between Gcr1p and Gcr2p has been demonstrated using the two-hybrid system and mutations in *GCR1* can suppress *gcr2* mutations (Uemura and Jigami 1992 b, 1995). Gcr2p could provide an additional transcriptional activation domain at UAS via its interaction with Gcr1p. Little is known about the role of Gcr3p at glycolytic UAS. It may be a DNA-binding protein and it has certain characteristics of a transcriptional activator (Uemura and Jigami 1992 a).

Abf1p and Reb1p

Some glycolytic promoters also contain binding sites for Abf1p and Reb1p. Abf1p is a protein of 731 amino acids encoded by an essential yeast gene (Diffley and Stillman 1989; Halfter et al. 1989 b; Rhode et al. 1989). It contains an unusual type of zinc finger at the N terminus and within the C terminus there are regions rich in acidic amino acids. The amino-acid sequence of Abf1p contains several blocks of similarity with Rap1p. The two proteins are 30% identical and 40% conserved over 60% of their sequence (Diffley and Stillman 1989). Both Rap1p and Abf1p share a region of homology with a protein called San1p. Mutations in the gene encoding San1p can suppress weak *sir4* mutations, implying an interaction between San1p and Sir4p. By extrapolation it was suggested that both Abf1p and Rap1p could also interact with Sir4p via the conserved San1p homology. This was confirmed for Rap1p when it was shown that the San1p homology is within the region of Rap1p required for the interaction between the C terminus of Rap1p and Sir4p (Diffley and Stillman 1989; Moretti et al. 1994). This strongly implies that Abf1p also interacts with Sir4p. Such an interaction is likely to be important in silencing, rather than activation. An isolated binding site for Abf1p inserted upstream of a reporter gene is an extremely weak activator (Buchman and Kornberg 1990; Goncalves et al. 1995). A strong Abf1p-binding site has been shown to give only as much activation as a weak Rap1p-binding site (Buchman and Kornberg 1990). In promoters of ribosomal protein genes Abf1p sites are often found in combination with T-rich elements. Similar situations are found in the *DED1* promoter, where two Abf1p sites work synergistically with a T-rich element, and in the intergenic region between the *YPT1* and *TUB2* genes which also contains a T-rich element (Halfter et al. 1989 a; Buchman and Kornberg 1990). T-rich regions are generally not found in the promoters of glycolytic genes, except for the *PFK1* gene. The arrangement here is probably not analogous to that seen in ribosomal protein gene promoters because the T-rich region is almost certainly too far away from the Abf1p-binding sites to work synergistically with Abf1p. In genes encoding mitochondrial components, Abf1p-binding sites are often found in combination with a binding site for the Hap2/3/4 complex (Dorsman and Grivell 1990; De Winde and Grivell 1992). In many glycolytic gene promoters Abf1p-binding sites are found close to

Rap1p-binding sites to form a combination which may be analogous to those described above. Abf1p may perform its function at glycolytic promoters by organising the architecture of the chromatin around the UAS. Abf1p bends DNA by about 120°, with the centre of the bend 7 bp 5′ of its consensus binding site (McBroom and Sadowski 1994). This DNA bending may be important for the assembly or functioning of the complex of transcription factors bound at the UAS.

Reb1p has also been implicated in determining chromatin structure. At the *GAL* UAS Reb1p binds to DNA and excludes nucleosomes from a region of about 230 bp (Fedor et al. 1988). The protein itself is a weak activator but when combined with the T-rich element from the *DED1* promoter a strong synergistic effect was observed, leading to the suggestion that Reb1p potentiates the action of other activators by providing a more favourable chromatin environment (Chasman et al. 1990). At glycolytic UAS the main role of Reb1p may be to facilitate the binding of other transcription factors to the DNA. In most cases where Reb1p has been shown to interact with glycolytic UAS it binds upstream of the other transcription factors so it could also have a role in defining the boundary of the UAS. Reb1p is important in transcription termination by RNA polymerase I where it causes polymerase pausing (Lang et al. 1994). As Reb1p also causes pausing by RNA polymerase II it may help to prevent readthrough transcription from adjacent ORFs.

Co-ordinate induction of glycolytic gene expression by glucose

The enzymes which comprise the yeast glycolytic pathway are expressed to a high level. In yeast cells growing on a fermentable carbon source, such as glucose, they comprise about 30% of the total soluble protein (Hess et al. 1969; Fraenkel 1982). In yeast cells growing on non-fermentable carbon sources, such as ethanol or glycerol/lactate, many of the glycolytic enzymes may not be required. It is tempting to look at glycolysis in the same way as other microbial metabolic pathways, which switch on only when required to metabolise a particular substrate. In this case, the pathway could be turned on in the presence of glucose or other sugars and turned off in the absence of glucose and the presence of non-fermentable carbon sources. However, the situation is not that simple because some of the glycolytic enzymes are additionally involved in gluconeogenesis during growth on non-fermentable carbon sources. Some glycolytic enzymes are induced by glucose, but others appear to be constitutively expressed, irrespective of the carbon source (Maitra and Lobo 1971; Cohen et al. 1987; Kuroda et al. 1994). Constitutively expressed enzymes are generally those involved in gluconeogenesis as well as glycolysis. For those enzymes that are induced, the apparent degree of induction depends on whether the enzyme activity itself, or the level of mRNA encoding the enzyme, is assayed. Early experiments using a hybrid yeast (*Saccharomyces fragilis* × *Saccharomyces dobzhanskii*) suggested that the specific activities of many glycolytic enzymes increase in response to glucose or galactose. Most increased between three- and seven-fold but some increased 50–70 fold (Maitra and Lobo 1971). In contrast, more recent experiments, using specific probes to detect glycolytic mRNAs, suggested that most glycolytic enzymes are induced less than two-fold and the greatest degree of induction is about four-fold (Moore et al. 1991). These attempts to simultaneously assess the changes in activity of most of the enzymes within the pathway have been supplemented by studies on individual enzymes. Such studies suggest that expression of some glycolytic genes, for example *ENO1* and *TDH3*, is clearly constitutive, whereas other genes, such as *PGK*, *PYK1* and *PDC1*, are induced by glucose (Tuite et al. 1982; Burke et al. 1983; Schmitt et al. 1983; Cohen et al. 1987; Chambers et al. 1989; Kuroda et al. 1994). These results are summarised in Table 2.

Mechanism of glucose induction

Glucose induction of some glycolytic promoters may be analogous to the increase in expression of many

Table 2 Inducibility of glycolytic gene expression by glucose. Induction of glycolytic gene expression by glucose is variable and may depend on the strain assayed and whether enzyme activity or mRNA level is measured. A further difficulty is that the level of induction may be fairly low and some workers define a two-fold increase as induction, whereas others may regard this level of increase as constitutive expression. We have categorised a gene as inducible only where there is clear evidence of a significant change in mRNA level. Data was obtained from the following publications; *PGI* (Aguilera and Zimmermann 1986; Green et al. 1988), *PEK1* and *PFK2* (Heinisch et al. 1991), *FBA1* (Compagno et al. 1991), *TPI* (Scott et al. 1990.), *TDH2* (Moore et al. 1991), *TDH3* (McAlister and Holland 1985 a; Kuroda et al. 1994), *PGK* (Tuite et al. 1982; Chambers et al. 1989), *PGM* (Rodicio et al. 1993), *ENO1/ENO2* (McAlister and Holland 1982; Cohen et al. 1986, 1987), *PYK1* (Nishizawa et al. 1989), *PDC1* (Schmitt et al. 1983), *ADH1* (Santangelo and Tornow 1990)

Gene	Glucose induction
PGI1	No
PFK1	No
PFK2	No
FBA1	No
TPI	No
TDH2	No
TDH3	No
PGK	Yes
PGM	No
ENO1	No
ENO2	Yes
PYK	Yes
PDC1	Yes
ADH1	Yes

ribosomal protein genes which occurs when yeast cells are shifted from non-fermentable media to medium containing glucose. Transcription of these genes increases about four-fold on shifting cells to a glucose medium (Herruer et al. 1987; Kraakman et al. 1993). This effect occurs irrespective of whether the promoter contains an RPG box or an Abf1p-binding site, but in both cases the T-rich element present in these promoters is also required (Kraakman et al. 1993; Goncalves et al. 1995). The sequences involved in glucose induction of many glycolytic promoters have not been well defined, but the response could be imposed on these promoters via changes in the activity of Rap1p and/or Abf1p. At the *PGK* promoter the Rap1p-binding site is necessary for increased transcription in medium containing glucose as a carbon source. However, the Abf1p site at the promoter is not essential for the response to occur (Chambers et al. 1988). The mechanism involved is independent of the interaction between Rap1p and Gcr1p, because glucose induction occurs to about the same degree in a *gcr1* mutant strain (Henry et al. 1994). Changes to the activity of both Rap1p and Abf1p may be mediated via changes in their phosphorylation states. There is evidence that the activation function of Rap1p can be increased as a result of phosphorylation by cAMP-dependent protein kinase (PKA), triggered by the presence of glucose (Klein and Struhl 1994). The pathway which leads to activation of PKA in this situation may not involve the well-characterised RAS pathway, because mutations which effect this pathway do not abolish glucose induction of ribosomal protein gene expression (Kraakman et al. 1993). We previously suggested that DNA binding by Rap1p is reduced in nuclear protein extracts made from yeast cells grown on non-fermentable carbon sources (Chambers et al. 1989). Other workers have subsequently shown that in total protein extracts, the DNA-binding activity of Rap1p does not vary (Santangelo and Tornow 1990; Kraakman et al. 1993). This difference may be the result of post-translational modifications to Rap1p which render the protein more susceptible to proteolysis during the preparation of nuclear protein extracts, but not during the simpler total protein extract procedure.

At other glycolytic promoters Abf1p may have a role in the response. This protein has been implicated in carbon source control of the *COX6* gene, encoding subunit VI of cytochrome c oxidase (Silve et al. 1992; Trawick et al. 1992). However, this gene is regulated by the carbon source in the opposite way to glycolytic and ribosomal protein genes. The regulatory mechanism may involve phosphorylation of Abf1p because at least four different forms of the protein have been detected using electrophoresis. The relative amounts of the different forms varies depending on the yeast growth conditions. In cells grown on medium containing a non-fermentable carbon source Abf1p is mainly in a phosphorylated form, leading to the suggestion that more phosphorylation of Abf1p leads to greater transcription of *COX6* (Silve et al. 1992). If phosphorylation changes to Abf1p are involved in mediating transcriptional effects in response to changes in carbon source, it remains to be explained how these changes can lead to different transcriptional effects at different promoters.

Acknowledgements Work in our laboratory on glycolytic promoters and their mechanisms of action is supported by the UK BBSRC and the University of Nottingham Research Fund. E.A.P is an MRC post-graduate student.

References

Aguilera A, Zimmermann FK (1986) Mol Gen Genet 202:83–89
Baker HV (1991) Proc Natl Acad Sci USA 88:9443–9447
Bitter GA, Chang KKH, Egan KM (1991) Mol Gen Genet 231:22–32
Brindle PK, Holland JP, Willett CE, Innis MA, Holland MJ (1990) Mol Cell Biol 10:4872–4885
Buchman AR, Kornberg RD (1990) Mol Cell Biol 10:887–897
Buchman AR, Kimmerly WJ, Rine J, Kornberg RD (1988) Mol Cell Biol 8:210–225
Burke RL, Tekamp-Olson P, Najarian R (1983) J Biol Chem 2193–2201
Butler G, Dawes IW, McConnell DJ (1990) Mol Gen Genet 223:449–456
Carmen AA, Holland MJ (1994) J Biol Chem 269:9790–9797
Chambers A, Stanway C, Kingsman AJ, Kingsman SM (1988) Nucleic Acids Res 16:8245–8260
Chambers A, Tsang JH, Stanway C, Kingsman AJ, Kingsman SM (1989) Mol Cell Biol 9:5516–5524
Chambers A, Stanway C, Tsang JSH, Henry Y, Kingsman AJ, Kingsman SM (1990) Nucleic Acids Res 18:5393–5399
Chasman DI, Lue NF, Buchman AR, LaPointe JW, Lorch Y, Kornberg RD (1990) Genes Dev 4:503–514
Clifton D, Fraenkel DG (1981) J Biol Chem 256:13074–13078
Cockell M, Palladino F, Laroche T, Kyrion G, Liu C, Lustig AJ, Gasser SM (1995) J Cell Biol 129:909–924
Cohen R, Holland JP, Yokoi T, Holland MJ (1986) Mol Cell Biol 6:2287–2297
Cohen R, Yokoi T, Holland JP, Pepper AE, Holland, MJ (1987) Mol Cell Biol 7:2753–2761
Compagno C, Ranzi BM, Martegani, E (1991) FEBS Let 293:97–100
Devlin C, Tice Baldwin K, Shore D, Arndt KT (1991) Mol Cell Biol 11:3642–3651
De Winde JH, Grivell LA (1992) Mol Cell Biol 12:2872–2883
Diffley JFX, Stillman B (1988) Proc Natl Acad Sci USA 85:2120–2124
Diffley JFX, Stillman B (1989) Science 246:1034–1038
Dorsman JC, Grivell LA (1990) Curr Genet 17:459–464
Dorsman JC, Doorenbosch MM, Maurer CTC, de Winde JH, Mager WH, Planta RJ, Grivell LA (1989) Nucleic Acids Res 17:4917–4923
Dorsman JC, van Heeswijk WC, Grivell LA (1990) Nucleic Acids Res 18:2769–2776
Dumitru I, McNeil JB (1994) Nucleic Acids Res 22:1450–1455
Fantino E, Marguet D, Lauquin GJM (1992) Mol Gen Genet 236:65–75
Fedor MJ, Lue NF, Kornberg RD (1988) J Mol Biol 204:109–127
Fraenkel DG (1982) In:Strathern JN, Jones EW, Broach JR (eds) The molecular biology of the yeast *Saccharomyces*:metabolism and gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 1–37

Gilson E, Roberge M, Giraldo R, Rhodes D, Gasser SM (1993) J Mol Biol 231:293–310
Gilson E, Muller T, Sogo J, Laroche T, Gasser SM (1994) Nucleic Acids Res 22:5310–5320
Goncalves PM, Griffioen G, Minnee R, Bosma M, Kraakman LS, Mager WH, Planta RJ (1995) Nucleic Acids Res 23:1475–1480
Green JBA, Wright APH, Cheung WY, Lancashire WE, Hartley BS (1988) Mol Gen Genet 215:100–106
Halfter H, Muller U, Winnacker EL, Gallwitz D (1989a) EMBO J 8:3029–3037
Halfter H, Kavety B, Vandekerckhove J, Kiefer F, Gallwitz D (1989b) EMBO J 8:4265–4272
Hamil KG, Nam HG, Fried HM (1988) Mol Cell Biol 8:4328–4341
Hardy CFJ, Balderes D, Shore D (1992) Mol Cell Biol 12:1209–1217
Heinisch J, Vogelsang K, Hollenberg CP (1991) FEBS Lett 289:77–82
Henry YAL, Chambers A, Tsang JSH, Kingsman AJ, Kingsman SM (1990) Nucleic Acids Res 18:2617–2623
Henry YAL, Lopez MC, Gibbs JM, Chambers A, Kingsman SM, Baker HV, Stanway C (1994) Mol Gen Genet 245:506–511
Herruer MH, Mager WH, Woudt LP, Nieuwint RTM, Wassenaar GM, Groeneveld P, Planta RJ (1987) Nucleic Acids Res 15:10133–10144
Herruer MH, Mager WH, Doorenbosch TM, Wessels PLM, Wassenaar TM, Planta RJ (1989) Nucleic Acids Res 17:7427–7439
Hess B, Boiteux A, Kruger J (1969) Adv Enzyme Regul 7:149–169
Hohmann S (1993) Mol Gen Genet 241:657–666
Holland MJ, Yokoi T, Holland JP, Myambo K, Innis MA (1987) Mol Cell Biol 7:813–820
Huet J, Cottrelle P, Cool M, Vignais ML, Thiele D, Marck C, Buhler JM, Sentenac A, Fromageot P (1985) EMBO J 4:3539–3547
Huie MA, Scott EW, Drazinic CM, Lopez MC, Hornstra IK, Yang TP, Baker HV (1992) Mol Cell Biol 12:2690–2700
Jiang YW, Stillman DJ (1995) Genetics 140:103–114
Ju Q, Morrow BE, Warner JR (1990) Mol Cell Biol 10:5226–5234
Jung SY, Yoo HY, Kim YH, Kim J, Rho HM (1995) Curr Genet 27:312–317
Klein F, Laroche T, Cardenas ME, Hofmann JFX, Schweizer D, Gasser SM (1992) J Cell Biol 117:935–948
Klein C, Struhl K (1994) Mol Cell Biol 14:1920–1928
Kraakman LS, Griffioen G, Zerp S, Groeneveld P, Thevelein JM, Mager WM, Planta RJ (1993) Mol Gen Genet 238:196–204
Kuroda S, Otaka S, Fujisawa Y (1994) J Biol Chem 269:6153–6162
Kurtz S, Shore D (1991) Genes Dev 5:616–628
Kyrion G, Boakye KA, Lustig AJ (1992) Mol Cell Biol 12:5159–5173
Lang WH, Morrow BE, Ju Q, Warner JR, Reeder RH (1994) Cell 79:527–534
Maitra PK, Lobo Z (1971) J Biol Chem 246:475–488
Marguet D, Lauquin GJM (1986) Biochem Biophys Res Commun 138:297–303
McAlister L, Holland MJ (1982) J Biol Chem 257:7181–7188
McAlister L, Holland MJ (1985a) J Biol Chem 260:15013–15018
McAlister L, Holland MJ (1985b) J Biol Chem 260:15019–15027
McBroom LDB, Sadowski PD (1994) J Biol Chem 269:16461–16468
Mellor J, Dobson M, Kingsman AJ, Kingsman SM (1987) Nucleic Acids Res 15:6243–6259
Moore PA, Sagliocco FA, Wood RMC, Brown AJP (1991) Mol Cell Biol 11:5330–5337
Moretti P, Freeman K, Coodly L, Shore D (1994) Genes Dev 8:2257–2269
Morrow BE, Johnson SP, Warner JR (1989) J Biol Chem 265:20778–20783
Nishi K, Park CS, Pepper AE, Eichinger G, Innis MA, Holland MJ (1995) Mol Cell Biol 15:2646–2653
Nishizawa M, Araki R, Teranishi Y (1989) Mol Cell Biol 9: 442–451
Ogden JE, Stanway C, Kim S, Mellor J, Kingsman AJ, Kingsman SM (1986) Mol Cell Biol 6:4335–4343.
Pavlovic B, Horz W (1988) Mol Cell Biol 8:5513–5520
Purvis IJ, Loughlin L, Bettany AJE, Brown AJP (1987) Nucleic Acids Res 15:7963–7974
Raghuram V, Lobo Z, Maitra PK (1994) J Genet 73:17–32
Rhode PR, Sweder KS, Oegema KF, Campbell JL (1989) Genes Dev 3:1926–1939
Rhode PR, Elsasser S, Campbell JL (1992) Mol Cell Biol 12:1064–1077
Rodicio R, Heinisch JJ, Hollenberg CP (1993) Gene 125:125–133
Rotenberg MO, Woolford JL (1986) Mol Cell Biol 6:674–687
Sakai A, Chibazakura T, Shimizu Y, Hishinuma F (1992) Nucleic Acids Res 20:6227–6233
Santangelo GM, Tornow J (1990) Mol Cell Biol 10:859–862
Schmitt HD, Ciriacy M, Zimmermann FK (1983) Curr Genet 192:247–252
Scott EW, Baker HV (1993) Mol Cell Biol 13:543–550
Scott EW, Allison HE, Baker HV (1990) Nucleic Acids Res 18:7099–7107
Shore D, Nasmyth K (1987) Cell 51:721–732
Silve S, Rhode PR, Coll B, Campbell J, Poyton RO (1992) Mol Cell Biol 12:4197–4208
Sinclair DA, Kornfeld GD, Dawes IW (1994) Mol Cell Biol 14:214–225
Stanway CA, Chambers A, Kingsman AJ, Kingsman SM (1989) Nucleic Acids Res 17:9205–9218
Stanway CA, Gibbs JM, Kearsey SE, Lopez MC, Baker HV (1994) Mol Gen Genet 243:207–214
Tornow J, Zeng X, Gao W, Santangelo GM (1993) EMBO J 12:2431–2437
Trawick JD, Kraut N, Simon FR, Poyton RO (1992) Mol Cel Biol 12:2302–2314
Tuite MF, Dobson MJ, Roberts NA, King RM, Burke DC, Kingsman SM, Kingsman AJ (1982) EMBO J 1:603–608
Uemura H, Fraenkel DG (1990) Mol Cell Biol 10:6389–6396
Uemura H, Jigami Y (1992a) J Bacteriol 174:5526–5532
Uemura H, Jigami Y (1992b) Mol Cell Biol 12:3834–3842
Uemura H, Jigami Y (1995) Genetics 139:511–521
Vignais ML, Sentenac A (1989) J Biol Chem 264:8463–8466
Vignais ML, Woudt LP, Wassenaar GM, Mager WH, Sentenac A, Planta RJ (1987) EMBO J 6:1451–1457
White MA, Dominska M, Petes TD (1993) Proc Natl Acad Sci USA 90:6621–6625
Willett CE, Gelfman CM, Holland MJ (1993) Mol Cell Biol 13:2623–2633
Woudt LP, Mager WH, Nieuwint RTM, Wassenaar GM, Van Der Kuyl AC, Murre JJ, Hockman MFM, Brockhoff PGM, Planta RJ (1987) Nucleic Acids Res 15:6037–6048
Yagi S, Yagi K, Fukuoka J, Suzuki M (1994) J Vet Med Sci 56:235–244

Mol Gen Genet (1996) 250:348–356

ORIGINAL PAPER

Elizabeth A. Packham · Ian R. Graham
Alistair Chambers

The multifunctional transcription factors Abf1p, Rap1p and Reb1p are required for full transcriptional activation of the chromosomal *PGK* gene in *Saccharomyces cerevisiae*

Received: 30 May 1995 / Accepted: 2 October 1995

Abstract We have identified two new transcription factor binding sites upstream of the previously defined UAS within the phosphoglycerate kinase (*PGK*) gene promoter in *Saccharomyces cerevisiae*. These sites are bound in vitro by the multifunctional factors Cpf1p and Reb1p. We have generated targeted deletions of Rap1p, Abf1p and Reb1p binding sites in the promoter of the chromosomal copy of the *PGK* gene. Northern blot analysis confirmed that most *PGK* promoter activity is mediated through the Rap1p binding site. However, significant effects are also mediated through both the Reb1p and Abf1p sites. In contrast, when the promoter is present on a high-copy-number plasmid, both the Abf1p and Reb1p sites play no role in transcriptional activation. The role of Cpf1p was examined using a *cpf1* null strain. Cpf1p was found to have little if any, effect on activation of either the chromosomal or plasmid-borne *PGK* gene.

Key words Phosphoglycerate kinase · Cpf1p · Rap1p · Reb1p · Abf1p

Introduction

The phosphoglycerate kinase (*PGK*) gene is a highly expressed glycolytic gene in the budding yeast *Saccharomyces cerevisiae*. The promoter of the gene contains a UAS located between positions −538 and −402, relative to the translational start codon (Ogden et al. 1986). The UAS contains binding sites for the multifunctional transcription factors Abf1p and Rap1p, as well as the glycolysis-specific transcription factor Gcr1p (Fig. 1) (Chambers et al. 1989, 1990). Abf1p is a multifunctional protein which interacts with yeast DNA replication origins (ARS) and transcriptional silencers (*HMLI*, *HMRE*, *HMRI*) (Diffley and Stillman 1988, 1989; Buchman et al. 1988; Rhode et al. 1989). In addition, it is a weak transcriptional activator that binds to the promoters of a number of genes, including the bidirectional promoter of the *YPT1* and *TUB2* genes, several nuclear genes encoding mitochondrial proteins (*COX6*, *QCR8*), ribosomal protein genes *S33*, *L2* and *L3* and the glycolytic genes *PYK1* and *ENO2* (Halfter et al. 1989; Dorsman et al. 1988, 1989; Della Seta et al. 1990; Chambers et al. 1990; Brindle et al. 1990; Trawick et al. 1992; de Winde and Grivell 1992). Rap1p is also a multifunctional transcription factor. Like Abf1p, it binds both to transcriptional silencers (*HMLE*, *HMRE*) and a number of gene promoters (Shore et al. 1987; Shore and Nasmyth 1987), mainly those of housekeeping genes. They include the glycolytic genes *TPI*, *TDH3*, *ENO1*, *ENO2*, *PYK1*, *PDC1* and *ADH1*, genes encoding ribosomal proteins such as *S24*, *L46* and *L25*, and genes such as *TEF1* and *TEF2*, which encode translation elongation factors (Huet et al. 1985; Woudt et al. 1987; Nishizawa et al. 1989; Scott et al. 1990; Brindle et al. 1990; Butler et al. 1990; Santangelo and Tornow 1990; Bitter et al. 1991). Rap1p also interacts with the $C_{(1-3)}A$ repeat region of yeast telomeres, where it plays a role in controlling telomere length (Longtine et al. 1989; Conrad et al. 1990; Lustig et al. 1990; Sussel and Shore 1991).

The gene encoding Gcr1p was first identified in a search for mutations which affect expression of glycolytic genes (Clifton and Fraenkel 1981). *gcr1* mutant strains show reduced expression of most glycolytic genes on both fermentable and non-fermentable media (Clifton and Fraenkel 1981). Gcr1p was subsequently shown to be a DNA binding protein that interacts with a consensus binding site containing 5′-CT/ATCC-3′ at its core (Huie et al. 1992). The *PGK* UAS contains three good matches to the Gcr1p consensus recognition site,

Communicated by C. P. Hollenberg

E. A. Packham · I. R. Graham · A. Chambers (✉)
Department of Genetics, University of Nottingham,
Queen's Medical Centre, Nottingham NG7 2UH, UK

downstream of the Rap1p site (Ogden et al. 1986). In vivo footprinting experiments suggested that only two of these sites are actually occupied in yeast cells (Henry et al. 1994). Gcr1p appears to work in combination with Rap1p, which may stabilise the interaction between Gcr1p and DNA. There is also some evidence that the interaction between Gcr1p and Rap1p may be sufficient to target Gcr1p to the UAS, without an interaction between Gcr1p and DNA (Tornow et al. 1993).

The role of these transcription factor binding sites in *PGK* promoter function has been investigated previously using deleted promoters carried on a multicopy, 2 μm-based plasmid (Chambers et al. 1988). Deletion of only the Rap1p binding site from the *PGK* promoter resulted in an 80% reduction in *PGK* mRNA level, suggesting a central role for this factor in UAS function (Chambers et al. 1988). Deletion of individual Gcr1p binding sites also caused a significant reduction in *PGK* transcription. In contrast, a similar deletion of the Abf1p site had no effect on *PGK* expression. These results suggested that Abf1p plays little role in *PGK* UAS function. When individual binding sites for Rap1p, Abf1p and Gcr1p were placed upstream of a minimal promoter, none was found to be a strong activator of transcription (Stanway et al. 1989). However, combination of either the Abf1p site, or Gcr1p sites, with the Rap1p site was sufficient for transcriptional activation.

Two other multifunctional transcription factors, Reb1p and Cpf1p, have also been shown to interact with some glycolytic gene promoters. Reb1p is encoded by an essential gene (Ju et al. 1990). It has binding sites in many gene promoters including *RAP1*, *ACT*, *TRP1*, *TRP3*, *TRP5*, *PYK1* and *GAL10* (Chasman et al. 1990). It also binds at the centromere *CEN4* and in subtelomeric regions (Chasman et al. 1990). Binding sites for Cpf1p are found both in promoters and at centromeres. Promoters which contain Cpf1p binding sites include *MET25*, *TRP1* and *GAL2* (Mellor et al. 1990, 1991). The role of Cpf1p in transcriptional activation is controversial; it is not a conventional transcriptional activator, but it may affect chromatin structure (Kent et al. 1994). Unlike the other multifunctional transcription factors, Cpf1p is not encoded by an essential gene (Mellor et al. 1990).

We have now identified new binding sites for both Reb1p and Cpf1p in the *PGK* promoter, upstream of the previously defined UAS. We have made deletions of the Rap1p, Abf1p and Reb1p binding sites, within the promoter of the chromosomal copy of the *PGK* gene and determined their effects on *PGK* expression. The effects of these deletions have been compared with their effects on *PGK* expression from a multicopy plasmid. The role of Cpf1p, at both the chromosomal locus and the plasmid-borne promoter, has been examined in a *cpf1* null strain.

Materials and methods

Strains and media

All plasmid manipulations were carried out in *Escherichia coli* MC1061 [F⁻ *araD139* Δ(*ara-leu*)*7696* Δ(*lac*)*X74 galU galK hsdR2* (r_K^- m_K^+) *mcrA mcrB1 rpsL* (Str^r)]. Experiments were performed using the yeast strains *Saccharomyces cerevisiae* DBY745, or YAG93 (Kent et al. 1994), isogenic to DBY745 but *cpf1* null and therefore phenotypically Met⁻ (see Table 2). Transformed yeast strains were grown in synthetic complete medium with 2% glucose, lacking either uracil (SC-ura) or leucine (SC-leu) (Hawthorne and Mortimer 1960).

Plasmid construction

Uracil-selectable plasmids were constructed containing the *PGK* promoter and coding sequence, with deletions in the promoter region. The promoters of these constructs were identical to the wild-type promoter except that a transcription factor binding site was replaced with a *Bam*HI linker. Integrating plasmids were based on pAJ730, in which the *URA3* gene from YCp50 (Johnston and Davis 1984) was subcloned into the polylinker of pSP46 (Ogden et al. 1986). pAJ735 is a variant of pAJ730, in which the *Sma*I site was replaced by a *Bam*HI site. Multicopy plasmids were generated by cloning the *PGK* gene and its promoter into pAJ6 (Graham and Chambers 1994).

A deletion of the Reb1p transcription factor binding site was made using the polymerase chain reaction (PCR). In order to delete the Reb1p binding site, four primers were used to amplify the upstream and downstream regions of *PGK*; (521B/PGKD with 533B/PGKC2, Table 1) using pKV521 and pB1 as template DNA (Chambers et al. 1988; Hitzeman et al. 1980). After amplification, the upstream PCR product was cut with *Bgl*II and *Bam*HI, and the downstream PCR product was cut with *Bcl*I and *Bam*HI. The resultant fragments were cloned into the unique *Bam*HI site of pAJ735, generating an intact promoter containing a *Bam*HI linker in place of the Reb1p transcription factor binding site (pAJ105). The construct deleted for the potential Reb1p binding site was also cloned into the *Bam*HI site of pAJ6 to give pAJ112. *PGK* constructs deleted for the Rap1p and Abf1p transcription factor binding sites already existed on multicopy plasmids (pKV516, pKV502) (Chambers et al. 1988); *Hin*dIII-*Bgl*II fragments from these plasmids were cloned into pAJ730, generating uracil-selectable plasmids pAJ107 and pAJ108.

In vitro protein production

Abf1p, Rap1p, Cpf1p and Reb1p were produced by in vitro transcription and translation of their respective cloned genes. Abf1p and Rap1p were made as previously described (Chambers et al. 1989); Cpf1p was produced using plasmid pSP73-22, a gift from J. Mellor.

Table 1 Sequences of primers used for deleting the Reb1p transcription factor binding site from the *PGK* promoter and for amplifying the *REB1* coding region

Primer	Sequence
521B	5′ GCTTTCTAACAGATCTATCC 3′
533B	5′ CCGCATTAAAGCTGATCAGAAACGCAG 3′
PGKD	5′ CTTTATGAGGGGATCCTCAATTCAAG 3′
PGKC2	5′ TTGATGTTGGATCCATAAAGCACG 3′
REB1A	5′ TATAGGTCGACCAATATGCC 3′
REB1B	5′ TTTTCCGGATCCAATTTTCTG 3′

A clone of the *REB1* gene was a gift from J. Warner. Primers (REB1A and REB1B, Table 1) were used to amplify the *REB1* coding region using pRS316 as template DNA (Morrow et al. 1993). The amplified *REB1* gene was subsequently cloned into the pGEM-T vector (Promega). The resulting plasmid, pT7-REB1, was linearized downstream of the REB1 coding region using *Sst*I and run-off transcripts were produced using T7 RNA polymerase. To produce the proteins, 1 μg RNA was translated in a rabbit reticulocyte lysate (Promega) in the presence of [^{35}S]methionine (Amersham). As controls, mock reactions, containing no added mRNA, were performed in parallel.

Preparation of yeast total protein extracts and gel retardation analysis

Total protein extracts were prepared from yeast cells as previously described (Graham and Chambers 1994). DNA fragments, labelled using [γ-^{32}P]ATP (Amersham) and T4 polynucleotide kinase, were incubated with either 1–2 μg protein from a yeast total protein extract or 3–5 μl of rabbit reticulocyte lysate primed with *RAP1*, *ABF1*, *CPF1* or *REB1* mRNA. 1–3 μg poly[dI-dC] (Pharmacia) was added to each binding reaction as a non-specific competitor. After a 45 min incubation at room temperature, DNA/protein complexes were resolved by electrophoresis on a 5% polyacrylamide gel containing 0.5X TBE.

RNA analysis

RNA was prepared from yeast by the method of Dobson et al. (1982) for Northern analysis. Filters were probed with a *PGK*-specific probe, to detect *PGK* mRNA, and a ribosomal probe to measure rRNA as a loading control (Chambers et al. 1988). Where necessary, plasmid copy number was determined by probing filters with a *leu2d* probe.

Homologous recombination

Homologous recombination was achieved using the Pop In-Pop Out method (Scherer and Davis 1979). Plasmids pAJ105 (ΔREB1), pAJ107 (ΔABF1) and pAJ108 (ΔRAP1) were transformed into DBY745 and Ura$^+$ transformants were selected. After screening, to find transformants in which the inserted gene had not undergone gene conversion, loss of plasmid sequences was selected for by growth on SC plates containing 1 mg/ml 5-FOA (Boeke et al. 1984). Constructs were verified by restriction analysis of a PCR product amplified from the altered chromosomal DNA, and Southern blotting.

Results

Identification of transcription binding sites upstream of the *PGK* UAS

An examination of the promoter sequence upstream of the *PGK* UAS revealed close matches to the consensus binding sites for the multifunctional transcription factors Abf1p, Cpf1p, Reb1p and Rap1p (Fig. 1). Rap1p and Abf1p have already been shown to bind within the *PGK* UAS, whilst Reb1p binding sites have been identified in the promoters of other glycolytic genes (Chambers et al. 1989, 1990; Scott and Baker 1993; Carmen and Holland 1994). In order to determine if any of these proteins bind upstream of the *PGK* UAS, Abf1p, Rap1p, Cpf1p and Reb1p were all synthesised by in vitro translation (IVT) and tested in gel retardation assays with radioactively labelled promoter fragments (I, IA, IB), isolated from the *PGK* promoter. Fragment I extends from position −624 to position −538 and contains all five potential binding sites (Fig. 1B). When IVT Abf1p or Rap1p was assayed with fragment I, no DNA:protein complexes were detected (Fig. 2A, B). This suggested that fragment I contains neither an Abf1p, nor a Rap1p binding site. Control experiments, using known binding sites, demonstrated that both IVT Abf1p and IVT Rap1p were capable of binding to DNA in our assay (Fig. 2A, B). IVT Cpf1p generated a single complex with fragment I, indicating that Cpf1p can interact with this DNA fragment in vitro (Fig. 2C). The site of the interaction was localized to the 3′ end of fragment I because when 10 bp were removed from the 3′ end of this fragment, complex formation with IVT

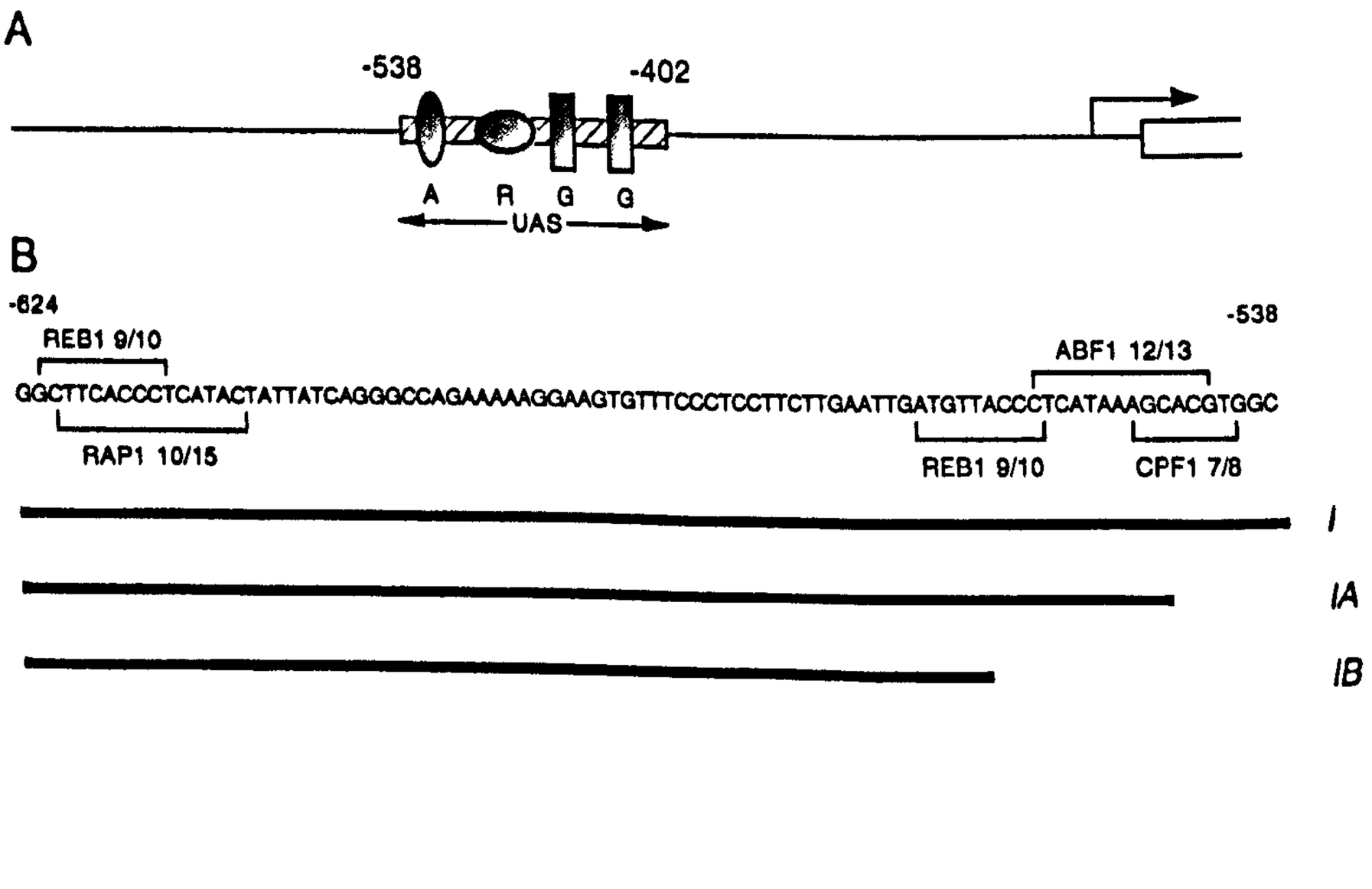


Fig. 1A, B The *PGK* promoter showing (A) the UAS with binding sites for Abf1p (A), Rap1p (R) and Gcr1p (G), and in B the sequence of the promoter upstream of the UAS with potential transcription factor binding sites. Indicated in B are the promoter fragments (I, IA, IB) used in gel retardation assays

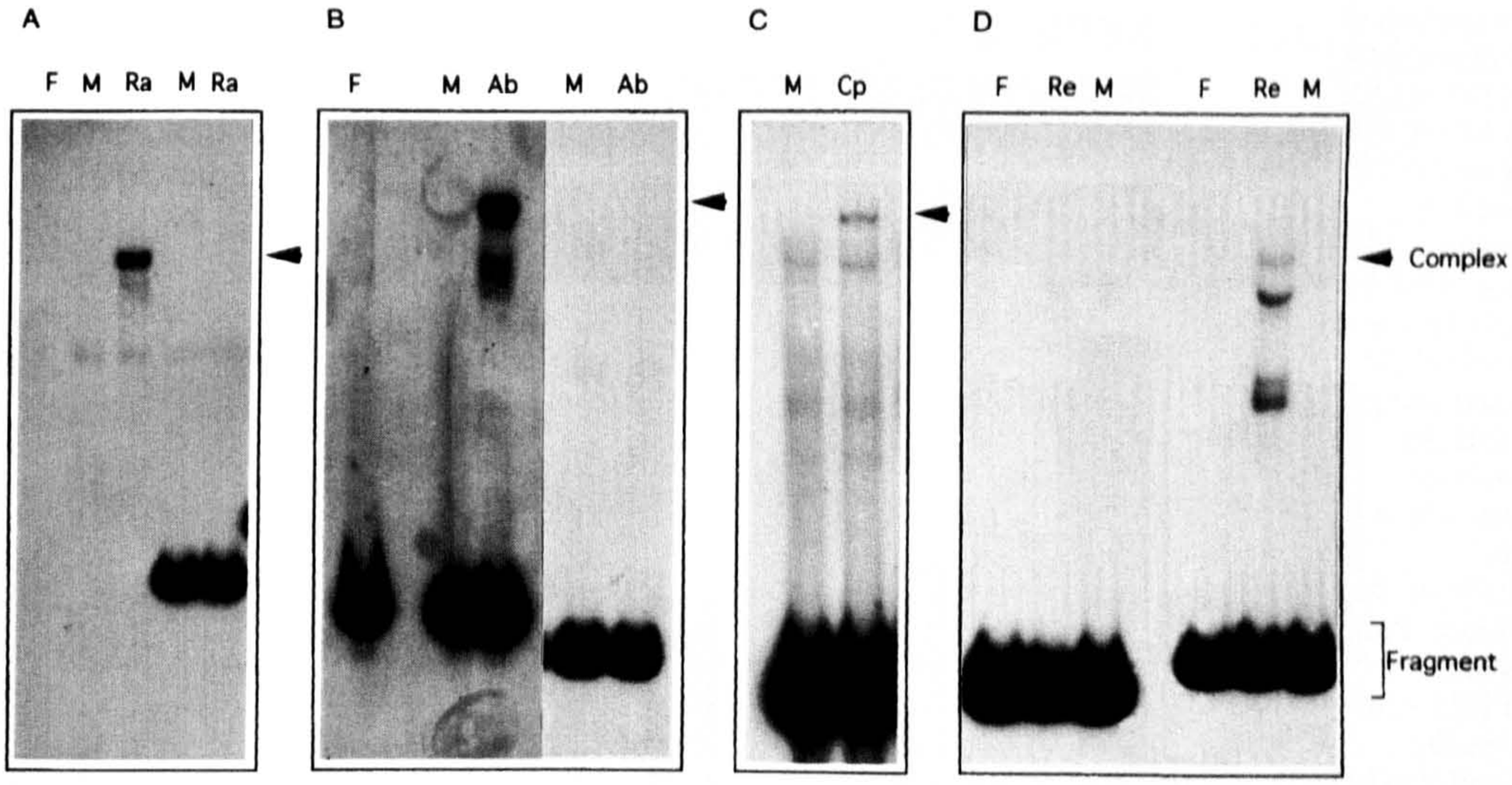


Fig. 2A–D Gel retardation assays using reticulocyte lysates primed with the appropriate mRNAs as sources of the corresponding proteins. **A** Rap1p incubated with a control fragment containing a known Rap1p binding site (first three lanes) and with *PGK* promoter fragment I (final two lanes). **B** Abf1p incubated with a control fragment containing a known Abf1p binding site (first three lanes) and with *PGK* promoter fragment I (final two lanes). **C** *PGK* promoter fragment I incubated with Cpf1p. **D** Reb1p incubated with *PGK* promoter fragments IB (first three lanes) and IA (final 3 lanes). Abbreviations: F, fragment alone; M, rabbit reticulocyte lysate not primed with template mRNA; Ra, Ab, Cp, Re, rabbit reticulocyte lysate primed with mRNA for Rap1p, Abf1p, Cpf1p or Reb1p

Cpf1p was abolished (data not shown). This 10 bp region contains a good match to the Cpf1p consensus binding site. Fragment I also contains two potential binding sites for Reb1p. To test these sites, two truncated versions of fragment I, designated IA and IB, were used. Fragment IA contains both of the potential Reb1p binding sites, fragment IB contains only the upstream site. When each of these fragments was tested with Reb1p in gel retardation assays, complex formation was only seen with fragment IA (Fig. 2D). No complexes were detected with fragment IB. Fragment IB has lost the good match to the consensus Reb1p binding site, spanning positions −552 to −562. As removal of this site abolished Reb1p binding it must be the only Reb1p binding site within fragment I.

The effect of deletions of transcription factor binding sites on transcription of the chromosomal *PGK* gene

We have extended earlier studies of the *PGK* promoter by generating a series of yeast strains containing transcription factor binding site deletions in the promoter of the chromosomal copy of the *PGK* gene (Table 2). The newly identified Reb1p binding site, plus the Abf1p and Rap1p binding sites within the UAS, were removed individually, and the effects on *PGK* expression determined. In each of the strains, the only *PGK* gene present was driven by a promoter lacking one of the specific transcription factor binding sites. The parental yeast strain, DBY745, containing a complete *PGK* promoter and gene, was used as a control. Each yeast strain was grown to mid-log phase and the cells harvested. Total RNA was then isolated from the yeast cells and analysed using Northern blotting (Fig. 3A). Each blot was probed with a *PGK*-specific probe, and with a probe to detect ribosomal RNA, as a loading control.

Clear differences in *PGK* expression were observed between the four strains. The strain in which the Rap1p binding site had been deleted from the promoter contained very little *PGK* mRNA. The strains in which the *PGK* promoter lacked either the Abf1p or Reb1p binding site, contained considerably less *PGK* mRNA than DBY745. In order to quantify these differences, Northern filters were scanned using a phosphorimager and the *PGK* mRNA signals were normalised for differences in RNA loading. These corrected values are represented graphically in the lower part of Fig. 3. Deletion of the Rap1p site in the centre of the UAS (strain YLP1) reduced the amount of *PGK* mRNA below detectable levels. Deletion of either the Abf1p alone (strain YLP2), or the Reb1p site alone (strain YLP3), resulted in a reduction of approximately 52% in the level of *PGK* mRNA. These results confirm the importance of Rap1p as the key activator of the *PGK* promoter, but also indicate that both Abf1p and Reb1p play a role in transcriptional activation. The result for Abf1p is in contrast to the results of previous

Table 2 Yeast strains used in this work

DBY745	*α ade1-100 leu2-3 leu2-112 ura3-52*
YAG93	*α ade1-100 leu2-3 leu2-112 ura3-52 cpf1Δ10-351*
YLP1	*α ade1-100 leu2-3 leu2-112 ura3-52 PGK Δ-463/-475*
YLP2	*α ade1-100 leu2-3 leu2-112 ura3-52 PGK Δ-503/-516*
YLP3	*α ade1-100 leu2-3 leu2-112 ura3-52 PGK Δ-552/-562*

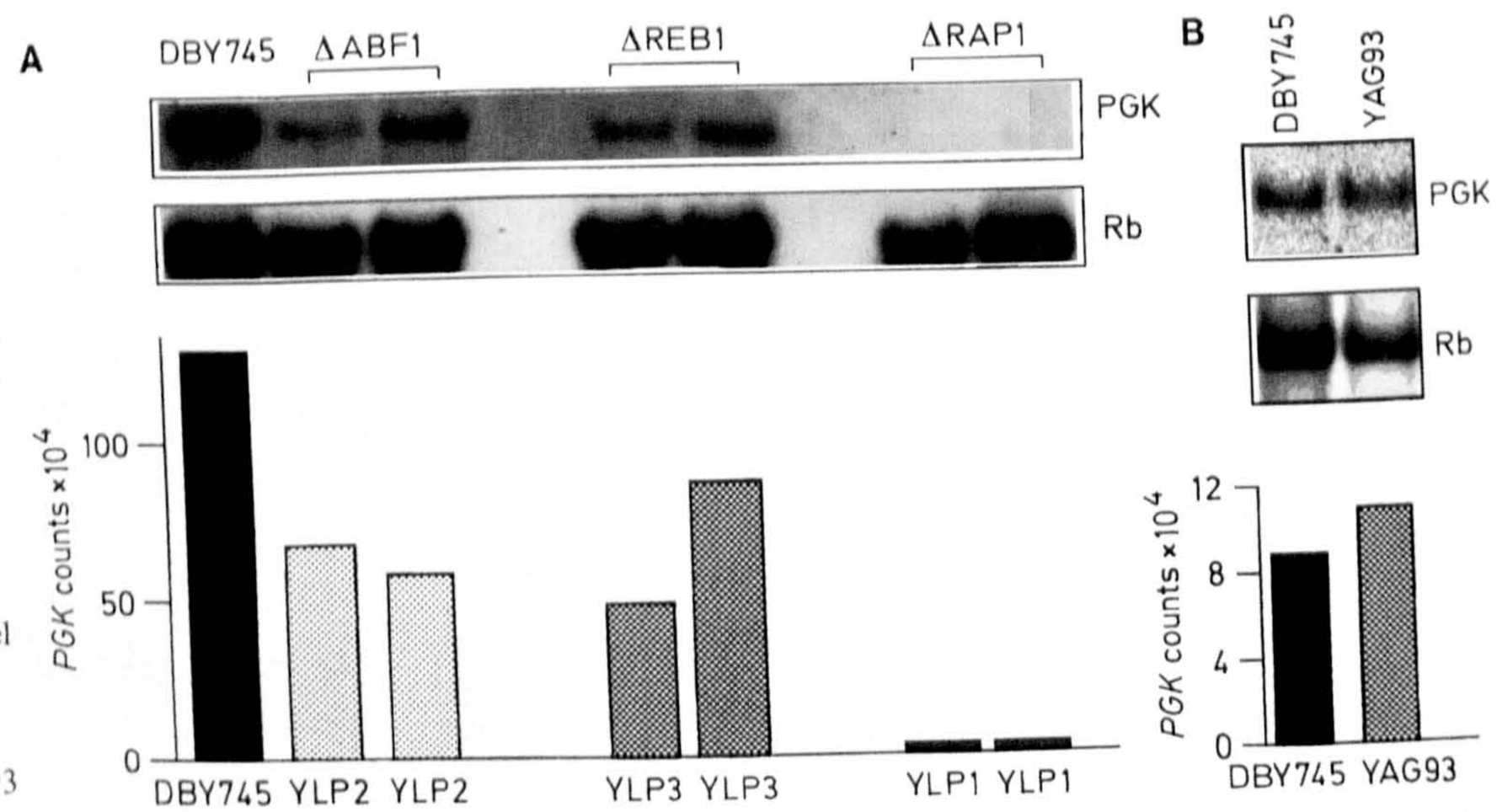


Fig. 3A, B Northern analysis of chromosomal *PGK* expression in yeast strains DBY745, YLP1, YLP2, YLP3 and YAG93. Total RNA was isolated from each strain, transferred to nitrocellulose and probed with a *PGK*-specific probe (PGK) and a probe to detect ribosomal RNA as a loading control (Rb). Each blot was scanned using a phosphorimager and the *PGK* signals corrected for differences in loading. These corrected values are represented graphically in the lower part of the Figure. Panel **A** shows the results for DBY7245, YLP1, YLP2 and YLP3. Panel **B** shows the results for DBY745 and YAG93

experiments using the *PGK* promoter on a high-copy-number plasmid (Chambers et al. 1988).

The activities of chromosomal and plasmid-borne *PGK* promoters in a *cpf1* strain

To investigate the contribution of Cpf1p to the activity of the *PGK* promoter we have made use of a *cpf1* null strain (YAG93). This strain is isogenic with DBY745 except for the *cpf1* mutation (Kent et al. 1994). The activity of the chromosomal *PGK* locus in YAG93 and DBY745 was determined by measuring *PGK* mRNA levels in mid-log phase cultures of each strain. Northern blots of total RNA from these cells were probed using a *PGK*-specific probe and a probe to detect rRNA as a loading control (Fig. 3B). The blot was analysed by phosphorimager and the difference in *PGK* expression between the two strains was determined, after correcting for RNA loading. These corrected values are represented graphically in the lower half of Fig. 3. In the absence of functional Cpf1p there was an increase of 29% in the level of *PGK* mRNA. This increase may be within the bounds of experimental error and may not be significant. Alternatively, it may suggest that Cpf1p plays a minor negative role at the *PGK* promoter. The activity of a plasmid-borne *PGK* promoter was also compared between YAG93 and DBY745. Plasmid pMA27, a multicopy plasmid carrying the wild-type *PGK* promoter and gene, was transformed into each strain and two transformants of each were grown to mid-log phase in selective medium. Cells were then harvested and used to prepare total RNA for analysis on Northern blots. The blots were probed with *PGK*-, ribosomal RNA- and *leu2d*-specific probes (Fig. 4A, and data not shown). The ribosomal RNA probe was used to measure the amount of each RNA loaded, the *leu2d* probe was used to detect transcription of the plasmid-borne *leu2d* gene, as a measure of plasmid copy number. Initially, it appeared that the activity of the *PGK* promoter was increased in YAG93, because these samples contained relatively more *PGK* mRNA. However, the copy number of plasmid pMA27 was also found to have increased in this strain. This makes it difficult to assess the importance of Cpf1p in determining promoter activity. The Northern filters were analysed quantitatively using a phosphorimager and the *PGK* signals normalised for loading and copy number. The results are represented graphically in the lower half of Fig. 4. If a linear relationship between copy number and RNA level is assumed, the normalised results suggest that the activity of the *PGK* promoter was reduced by approximately 20% in YAG93, compared to DBY745. This suggests that Cpf1p plays, at most, a very minor role in the activity of the *PGK* promoter on a high-copy-number plasmid.

The effect of Reb1p and Abf1p binding site deletions on the activity of a plasmid-borne *PGK* promoter

Previous experiments suggested that an Abf1p binding site deletion had no effect on *PGK* promoter function on a high-copy-number plasmid (Chambers et al. 1988). These experiments were performed without the benefits of phosphorimager technology, which allows very accurate quantification of hybridization data. In the light of our result suggesting a role for Abf1p at the UAS of the chromosomal *PGK* gene, we have repeated previous experiments using a *PGK* promoter with an Abf1p binding site deletion on a high-copy-number plasmid (pKV502). We have also constructed and tested a similar plasmid in which the newly identified Reb1p binding site is deleted (pAJ112). As a control we

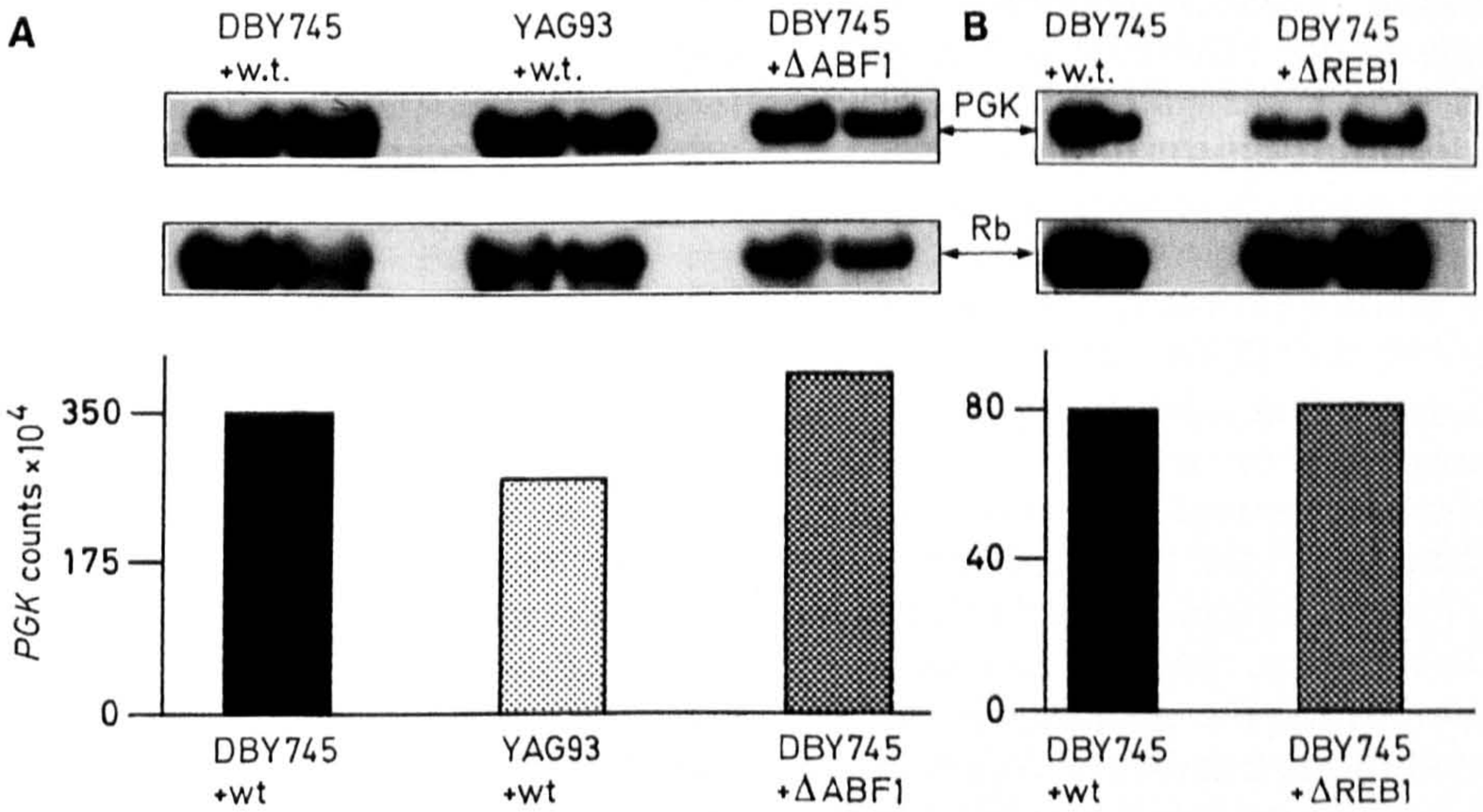


Fig. 4A, B Northern analysis of plasmid-borne *PGK* promoter activity in DBY745 and YAG93. Total RNA was isolated from each strain, transferred to nitrocellulose and probed with a *PGK*-specific probe (PGK) and a probe to detect ribosomal RNA as a loading control (Rb). The blots were also probed with a *leu2d* probe to measure relative plasmid copy number, as described in the text. Each blot was scanned using a phosphorimager and the *PGK* signals corrected for differences in loading and plasmid copy number. These corrected signals are represented graphically in the lower half of the Figure. Panel **A** shows the activity of the wild-type promoter (pMA27) in DBY745 and YAG93, plus the activity of the promoter with the Abf1p site deleted (pKV502/ΔABF1) in DBY745. Panel **B** compares the activity of the wild-type promoter with that of the promoter with the Reb1p site deleted (pAJ112/ΔREB1)

have used plasmid pMA27, which contains the complete *PGK* promoter (Mellor et al. 1983). Two transformants containing each plasmid were grown to mid-log phase and the cells harvested. The harvested cells were used to prepare total RNA which was analysed by Northern blotting; blots were probed with a *PGK* probe and a ribosomal RNA specific probe as a loading control (Fig. 4A, B). The blots were also probed with a probe to detect expression of the plasmid-borne *leu2d* gene, to allow us to compensate for any differences in plasmid copy number (data not shown). On examination of the blots, it appeared that the deletions had very little effect on the activity of the promoter in this situation; differences in the *PGK* signal intensity corresponded to differences in loading, as revealed by the ribosomal loading control. The signals from the different probes were then quantified using a phosphorimager and the *PGK* signals normalised for loading and plasmid copy number. The results are represented graphically in the lower part of Fig. 4. In this case the plasmid copy numbers were all very similar and normalising for copy number had very little influence on the final results. These clearly show that the previous conclusions regarding the effects of the Abf1p site deletion on the plasmid-borne promoter were correct. The activity of the *PGK* promoter increased by 14% when the Abf1p site was deleted, a value within the range of experimental variation. The results also show that deletion of the Reb1p binding site had no effect on the activity of the plasmid-borne *PGK* promoter; indeed, the activities of the promoters plus and minus the Reb1p site were within 1% of each other.

Discussion

Prior to this work, the *PGK* UAS had been defined as the region spanning positions −538 to −402 in the *PGK* promoter (Ogden et al. 1986). This region contains binding sites for the multifunctional transcription factors Rap1p and Abf1p, as well as the glycolytic-specific transcription factor Gcr1p (Chambers et al. 1989, 1990; Henry et al. 1994). Rap1p and Gcr1p had been shown to be involved in transcriptional activation of the promoter but no clear role for Abf1p had been identified (Chambers et al. 1989, 1990). We have now identified binding sites for two other members of the multifunctional transcription factor family, Reb1p and Cpf1p, immediately upstream of the previously defined UAS.

In order to determine the role of Reb1p, we generated a targeted deletion of the Reb1p binding site within the promoter of the chromosomal copy of the *PGK* gene. The Reb1p site was found to be important for full activation of the promoter. When this site was deleted, the level of *PGK* mRNA was reduced by about 50%. The UAS therefore extends over a larger region than previously thought and includes the Reb1p site at −562.

The original studies which identified regions of the *PGK* promoter important for activating transcription concluded that sequences 5′ of −479 were not required (Ogden et al. 1986; Chambers et al. 1988). The involvement of Reb1p was probably missed because

these studies utilised *PGK* promoter deletions on high-copy-number plasmids. In our experiments, deletion of the Reb1p binding site had no effect on the activity of the promoter in this context. This may be because the role of Reb1p depends on the chromosomal context of the promoter. A similar observation regarding the role of Reb1p has been made previously. Disruption of a Reb1p binding site in an rDNA enhancer had no effect on the transcription of the adjacent operon when assayed using a minigene construct on a multicopy plasmid; however, when disruptions of the same Reb1p binding site were examined in the chromosome, a decrease in the levels of transcription was seen (Kulkens et al. 1989, 1992). Alternatively, the difference between the effect of the deletion within the chromsome and the plasmid may be related to the number of copies of the gene present in each case. There is just one copy of the chromosomal *PGK* gene and between 100 and 200 copies of the plasmid-borne gene. The large number of copies of the gene in the latter situation may obscure the effect of Reb1p observed at the chromosomal locus. The 50% reduction resulting from the deletion of the Reb1p binding site at the chromosomal locus suggests that Reb1p has an important role. Reb1p is known to displace nucleosomes from the *GAL1-10* intergenic region and has been shown to be a very weak activator of transcription (Chasman et al. 1990). It can act synergistically with other weak transcriptional activators or T-rich regions such as those found in the *DED1* and *RAP1* promoters (Chasman et al. 1990; Graham and Chambers 1994). Another possible function for Reb1p may be in preventing polymerases from disrupting preinitiation complexes if transcription is carried through from an upstream gene. This effect is also seen when a Reb1p binding site acts as a transcriptional termination site for RNA polymerase I (Lang et al. 1994). The presence of Reb1p at its binding site causes all three RNA polymerases to pause, although if the transcript is not released, read-through can occur (Lang et al. 1994). Reb1p binds upstream of the other transcription factors which interact with the *PGK* promoter, a position similar to that in which it is found in the promoters of *TPI* and *ENO1*; it may therefore have a role in defining the boundary of the promoter, or perhaps impose directionality on the UAS (Scott and Baker 1993; Carmen and Holland 1994).

The observation that Reb1p is important in the chromosomal, but not the plasmid, context prompted us to re-investigate the roles of the Rap1p and Abf1p binding sites within the originally defined UAS. These sites have been studied previously, using high-copy plasmids. Deletion of the Rap1p binding site from the UAS of the chromosomal *PGK* gene confirmed the central importance of Rap1p. Removal of this binding site caused a dramatic reduction in *PGK* promoter strength. The yeast strain in which the only copy of the *PGK* gene has a Rap1p binding site deletion within the promoter is still able to grow on medium containing glucose as the carbon source. The low level of *PGK* transcription presumably allows production of sufficient PGK enzyme for the glycolytic pathway to operate. The result for the Abf1p site deletion in the chromosomal context was in contrast to that previously reported (Chambers et al. 1988). This site is required for full activation of the *PGK* gene. Removal of the site reduced *PGK* expression by approximately 50%, the same magnitude of effect as that caused by the Reb1p site deletion. Therefore, both Reb1p and Abf1p exert their effects at the chromosomal locus, but not when the promoter is on a high-copy-plasmid. Rap1p is important in both situations. These results suggest that the function of Abf1p at the *PGK* promoter may be related to the function of Reb1p. Perhaps both proteins work by mediating particular changes in chromatin structure. Similar roles for Abf1p and Reb1p have previously been suggested by the observation that at the *ILV1* promoter binding sites for the two proteins are functionally interchangeable (Remacle and Holmberg 1992).

The role of Cpf1p at the *PGK* promoter was investigated using a *cpf1* null strain. In this strain the activity of the chromosomal *PGK* promoter increased by 29%, and that of the plasmid-borne promoter was reduced by about 20%. These small changes, which may be within the range of experimental variation, make it hard to draw any conclusions regarding the role of Cpf1p. The role of Cpf1p as a transcription factor remains elusive. The presence or absence of Cpf1p appears to have no effect on the expression of *MET25*, *TRP1* and *GAL2*, all of which contain CDE1 motifs (Mellor et al. 1991). Similarly, it has been reported that no activation of transcription is detected from a *PGK* minimal promoter after insertion of the CDE1 motif from either *TRP1*, *SAM2* or *CEN3* (Mellor et al. 1990). More recently, transcription from *MET16* has been shown to require Cpf1p, although this protein in itself is not sufficient for full UAS activity (O'Connell et al. 1995). It is likely that Cpf1p plays a role in modulating chromatin structure; sensitivity to micrococcal nuclease at the *TRP1* promoter is lost in the absence of Cpf1p and localized changes to the chromatin structure of *MET16*, which depend on Cpf1p, have also been detected (Mellor et al. 1990; O'Connell et al. 1995). Cpf1p may alter chromatin structure to facilitate the formation of active transcription complexes. This view is supported by evidence from a study of the mammalian transcription factor USF, of which Cpf1p may be the yeast homologue. USF can compete with the assembly of promoter fragments into nucleosomes allowing the formation of stable preinitiation complexes (Workman et al. 1990). The promoter can then be bound by other transcription factors. It is also possible that in vivo the Cpf1p site in the *PGK* promoter is bound by another transcription factor of the basic helix-loop-helix class. The protein encoded by the recently identified *SGC1* gene is a good candidate

(Nishi et al. 1995). Mutations in *SGC1* can suppress the requirement for Gcr1p at glycolytic promoters, including *PGK*.

The newly characterised transcription factor binding sites in the *PGK* promoter extend the similarity between *PGK* and other glycolytic promoters. Most glycolytic promoters contain binding sites for Rap1p and Gcr1p, and some have Reb1p and Abf1p binding sites. These promoters are complex in organisation, and it remains a major challenge to determine the precise roles played by the many different transcription factors in controlling the expression of these "simple" housekeeping genes.

Acknowledgements We thank Jane Mellor for her gift of a *CPF1* clone, Jane Mellor and Nicholas Kent for the yeast strain YAG93, and Jonathan Warner for providing a *REB1* clone. This work was supported by the UK BBSRC and a Medical Research Council postgraduate studentship to E.A.P.

References

Bitter GA, Chang KKH, Egan KM (1991) A multi-component upstream activation sequence of the *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase gene promoter. Mol Gen Genet 231:22–32

Boeke JD, LaCroute F, Fink GR (1984) A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol Gen Genet 197: 345–346

Brindle PK, Holland JP, Willett CE, Innis MA, Holland MJ (1990) Multiple factors bind the upstream activation sites of the yeast enolase genes *ENO1* and *ENO2*: ABF1 protein, like repressor activator protein *RAP1*, binds cis-acting sequences which modulate repression or activation of transcription. Mol Cell Biol 10:4872–4885

Buchman AR, Kimmerly WJ, Rine J, Kornberg RD (1988) Two DNA-binding factors recognize specific sequences at silencers, upstream activating sequences, autonomously replicating sequences, and telomeres in the yeast *Saccharomyces cerevisiae*. Mol Cell Biol 8:210–225

Butler G, Dawes IW, McConnell DJ (1990) TUF factor binds to the upstream region of the pyruvate decarboxylase structural gene (*PDC1*) of *Saccharomyces cerevisiae*. Mol Gen Genet 223: 449–456

Carmen AA, Holland MJ (1994) The upstream repression sequence from the yeast enolase gene *ENO1* is a complex regulatory element that binds multiple trans-acting factors including REB1. J Biol Chem 269:9790–9797

Chambers A, Stanway C, Kingsman AJ, Kingsman SM (1988) The UAS of the yeast *PGK* gene is composed of multiple functional elements. Nucleic Acids Res 16:8245–8260

Chambers A, Tsang JSH, Stanway C, Kingsman AJ, Kingsman SM (1989) Transcriptional control of the *Saccharomyces cerevisiae PGK* gene by RAP1. Mol Cell Biol 9:5516–5524

Chambers A, Stanway C, Tsang JSH, Henry Y, Kingsman AJ, Kingsman SM (1990) ARS binding factor 1 binds adjacent to RAP1 at the UASs of the yeast glycolytic genes *PGK* and *PYK1*. Nucleic Acids Res 18:5393–5399

Chasman DI, Lue NF, Buchman AR, LaPointe JW, Lorch Y, Kornberg RD (1990) A yeast protein that influences the chromatin structure of UAS_G and functions as a powerful auxiliary gene activator. Genes Dev 4:503–514

Clifton D, Fraenkel DG (1981) The *gcr* (glycolysis regulation) mutation of *Saccharomyces cerevisiae*. J Biol Chem 256:13074–13078

Conrad MN, Wright JH, Wolf AJ, Zakian VA (1990) RAP1 protein interacts with yeast telomeres in vivo: overproduction alters telomere structure and decreases chromosome stability. Cell 63:739–750

De Winde JH, Grivell LA (1992) Global regulation of mitochondrial biogenesis in *Saccharomyces cerevisiae*: ABF1 and CPF1 play opposite roles in regulating expression of the *QCR8* gene, which encodes subunit VIII of the mitochondrial ubiquinol-cytochrome *c* oxidoreductase. Mol Cell Biol 12: 2872–2883

Della-Seta F, Ciafre SA, Marck C, Santoro B, Presutti C, Sentenac A, Bozzoni I (1990) The ABF1 factor is the transcriptional activator of the *L2* ribosomal protein gene in *Saccharomyces cerevisiae*. Mol Cell Biol 10:2437–2441

Diffley JFX, Stillman B (1988) Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. Proc Natl Acad Sci USA 85:2120–2124

Diffley JFX, Stillman B (1989) Similarity between the transcriptional silencer binding proteins ABF1 and RAP1. Science 246: 1034–1038

Dobson MJ, Tuite MF, Roberts NA, Kingsman AJ, Kingsman SM (1982) Conservation of high-efficiency promoter sequences in *Saccharomyces cerevisiae*. Nucleic Acids Res 10:2625–2637

Dorsman JC, van Heeswijk WC, Grivell LA (1988) Identification of two factors which bind to the upstream sequences of a number of nuclear genes coding for mitochondrial proteins and to genetic elements important for cell division in yeast. Nucleic Acids Res 16:7287–7301

Dorsman JC, Doorenbosch MM, Maurer CTC, de Winde JH, Mager WH, Planta RJ, Grivell LA (1989) An ARS/silencer binding factor also activates two ribosomal protein genes in yeast. Nucleic Acids Res 17:4917–4923

Graham IR, Chambers A (1994) A Reb1p-binding site is required for efficient activation of the yeast *RAP1* gene, but multiple binding sites for Rap1p are not essential. Mol Microbiol 12: 931–940

Halfter H, Muller U, Winnacker EL, Gallwitz D (1989) Isolation and DNA-binding characteristics of a protein involved in transcriptional activation of two divergently transcribed, essential yeast genes. EMBO J 8:3029–3037

Hawthorne DC, Mortimer RK (1960) Chromosome mapping in *Saccharomyces cerevisiae*: centromere-linked genes. Genetics 45:1085–1110

Henry YAL, Lopez MC, Gibbs JM, Chambers A, Kingsman SM, Baker HV, Stanway C (1994) The yeast protein Gcr1p binds to the *PGK* UAS and contributes to the activation of transcription of the *PGK* gene. Mol Gen Genet 245:506–511

Hitzeman RA, Clarke L, Carbon J (1980) Isolation and characterization of the yeast 3-phosphoglycerokinase gene (*PGK*) by an immunological screening technique. J Biol Chem 255: 12073–12080

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

Huie MA, Scott EW, Drazinic CM, Lopez MC, Hornstra IK, Yang TP, Baker HV (1992) Characterisation of the DNA-binding activity of GCR1: in vivo evidence for two GCR1-binding sites in the upstream activating sequence of *TPI* of *Saccharomyces cerevisiae*. Mol Cell Biol 12:2690–2700

Johnston M, Davis RW (1984) Sequences that regulate the divergent *GAL1-GAL10* promoter in *Saccharomyces cerevisiae*. Mol Cell Biol 4:1440–1448

Ju Q, Morrow BE, Warner JR (1990) REB1, a yeast DNA-binding protein with many targets, is essential for cell growth and bears some resemblance to the oncogene *myb*. Mol Cell Biol 10: 5226–5234

Kent NA, Tsang JSH, Crowther DJ, Mellor J (1994) Chromatin structure modulation in *Saccharomyces cerevisiae* by centromere and promoter factor 1. Mol Cell Biol 14:5229–5241