Molecular interactions of the *GAL* genetic switch

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering

2002

Helen Catherine Ross

School of Biological Sciences

(AGED)

Contents

Figures

Tables

Abstract

Transcription is a necessarily stringently controlled process. In addition to the general transcription machinery and other complexes, specific proteins control the transcriptional initiation of subsets of genes. The control of transcription has been studied extensively at the promoters of the *GAL* genes in *Saccharomyces cerevisiae*, a model eukaryote. The *GAL* genes encode the enzymes of the Leloir pathway, which allows the yeast to metabolise the sugar galactose. The genes are switched on in response to the presence of galactose. This induction is controlled by three proteins, an activator (Gal4p), a repressor (Gal80p) and an inducer (Gal3p), The galactokinase of the Leloir pathway, Gal1p, can also induce the *GAL* genes in the absence of Gal3p. Gal1p and Gal3p share a high degree of similarity throughout their protein sequence. The aim of this project was to investigate the mechanism of the *GAL* genetic switch. Insertion of just two amino acids into Gal3p converts the inducer into a galactokinase while retaining its inducer ability. The kinetics of both Gal1p and the mutant Gal3p (Gal3p+SA) have been investigated. Kinetic constants were determined for each protein with respect to each of its substrates. These data suggest that the site of the insertion may be involved with ATP binding or catalysis. Experiments were subsequently designed to identify the ATP binding site. *In vitro* investigation into the interactions between the proteins of the *GAL* genetic switch yielded data on the state of the repressor Gal80p in solution. An assay was developed to follow transcription of the *GAL* genes *in vivo*, and was used to assess the induction of the *GAL* genes in a variety of genetic backgrounds. The promoters of *GAL1* and *GAL3* were shown to have different effects on transcription of the *GAL* genes, suggesting that the differences observed in transcription by each protein are dependent on promoter sequence rather that protein sequence.

Author's Declaration

No portion of this work has been submitted in support of an application for another degree or qualification at this or any other university or other institute of learning.

Signed

Helen C. Ross

August 2002

Copyright statement

(1) Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made **only** in accordance with instructions given by the Author and lodged in the John Rylands Library of Manchester. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

(2) The ownership of any intellectual property rights which may be described in this thesis is vested in the University of Manchester, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the permission of the University, which will then prescribe the terms and conditions of any such agreement.

Further information on the conditions under which disclosures and exploitation may take place is available from the Head of the School of Biological Sciences.

Acknowledgements

I'd like to thank all members of the Reece lab and the Gene Expression group for help and advice, especially Adam Platt for plasmids and invaluable help at the initial stages of this project and David Timson for advice and discussion, as well as Noel, Cristina, Chris, Gill, Elaine and Steve for varying distractions.

Thanks to my friends, especially Bill, Jude, Luke, Paula and Tim, and to my family for advice, discussion and support.

Finally, thanks to my supervisor, Richard Reece, for advice and encouragement throughout.

This project was funded by the BBSRC.

List of abbreviations

ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
bp	Base pairs
BS^3	Bis(suphosuccinimidyl) suberate
BSA	Bovine serum albumin
CAP	Catabolite activating protein
CDK	Cyclin dependent kinase
Ci	Curie
ChIP	Chromatin immunoprecipitation
cpm	Counts per minute
CTD	Carboxyl terminal domain
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
dNTP	deoxynucleotide triphosphate
EDTA	Ethylenediamine tetra acetic acid
GHMP	Galactokinase, homoserine kinase, mevalonate kinase and phosphomevalonate kinase
GK	Galactokinase
GTF	General transcription factor
HAT	Histone acetyltransferase
H-DAC	Histone deacetylase
HEPES	N-2-Hydroxyethylpiperazine-N'-2-ethensulphonic acid
HSK	Homoserine kinase
kb	Kilobases
kDa	Kilodaltons
LB	Luria bertani
LDH	Lactate dehydrogenase
LTA	Long term adaptation
mRNA	Messenger ribonucleic acid
NAD^+	Nicotinamide adenine dinucleotide, oxidised form
NADH	Nicotinamide adenine dinucleotide, reduced form
oATP	2'3'-dideraldehyde adenosine triphosphate

OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PIC	Preinitiation complex
PEE	Polyethyleneimine
PK	Pyruvate kinase
psi	pounds per square inch
RNA	Ribonucleic acid
RNAP	*Escherichia coli* RNA polymerase
RNA pol II	RNA polymerase II
rRNA	ribosomal ribonucleic acid
RT-PCR	Reverse transcriptase polymerase chain reaction
SAB	Sample added buffer
SDS	Sodium dodecyl sulphate
TAF	TBP associated factor
TBP	TATA binding protein
TCA	Tricarboxylic acid
TdT	Terminal deoxytransferase
TEMED	N,N,N',N'-Tetramethylethyenediamine
TLC	Thin layer chromatography
Tris	Tris-(hydroxymethyl)-methylamine
tRNA	Transfer ribonucleic acid
UAS	Upstream activating sequence
URS	Upstream repressing sequence
UV	Ultraviolet
v/v	Volume/volume
w/v	Weight/volume
YNB	Yeast nitrogen base
YPD	Yeast peptone dextrose

SI units are used throughout, except where indicated above.

Chapter 1

Introduction

1.1 DNA and Transcription

The DNA code is transcribed and translated into proteins that form the basis of every organism. Genes are sequences of DNA encompassing a coding region that will be transcribed and the sequences controlling this process. The initiation control sequences are termed the promoter, which are often, but not always, located upstream of the coding sequence. The promoter includes DNA sequences that are bound by transcriptional activators and repressors. Genes provide a template that is transcribed by an RNA polymerase enzyme, creating a polymer of RNA bases complementary to the DNA sequence. Transcription is necessarily highly regulated, stringently controlled by a wide range of proteins, transcription factors, which in turn are regulated by various triggers including growth conditions and extracellular signals.

Prokaryotic transcription

The process of prokaryotic transcription is less complex than that of eukaryotes, which will be discussed later. The *Escherichia coli* RNA polymerase (RNAP) is made up of four subunits: α, β, β' and σ (Darst *et al.*, 1989). The enzyme function is carried out by a complex of a dimer of α subunits and the β and β' subunits. Different σ factors are employed for the transcription of different sets of genes, but the most common form is σ^{70} (Strickland *et al.*, 1988). Upstream of the start site for each gene there lies two conserved promoter DNA sequences, lying at 10 bp upstream of the transcription start site, '-10', and 35 bp upstream, '-35'. The –35 element is the recognition site for the RNAP, and has the consensus sequence 5'-TTGACA-3'. The –10 element is also known as the Pribnow box and provides a second contact for the RNAP and its consensus sequence is 5'-TATAAT-3' (Rosenberg and Court, 1979). The T and A (thymine and adenine) content of this sequence indicates that the double strands of DNA are not as

tightly hydrogen bonded together as a sequence rich in G (guanine) and C (cytosine), therefore the open complex, a structure which allows one of the DNA strands to be read by RNA polymerase, is formed more easily. In a closed complex, the DNA code is inaccessible to RNA polymerase and cannot be transcribed. The formation of the open complex is an important step for transcription, as it allows the coding strand of DNA to be released temporarily from the double helix and therefore its sequence can be accessed by transcriptional machinery.

Activation of prokaryotic genes is usually mediated by a protein contacting the RNA polymerase directly. For example, the catabolite activator protein (CAP) and cyclic AMP facilitate activation for several genes families in *E. coli*, including the *lac* and *gal* genes, encoding the proteins required for the metabolism of lactose and galactose. In the case of the *gal* genes, CAP contacts the C-terminus of the α subunit of RNAP, bringing the enzyme to the DNA (Niu *et al.*, 1996). Other activators are known to contact this region on the α subunit (Ishihama, 1993). Niu *et al.* showed that two contacts can be made to the RNAP at the *gal* promoter, and when these occur simultaneously, they exhibit synergy (Niu *et al.*, 1996). This is a greater level of transcription than expected from the addition of the two independent activation levels.

1.2 Eukaryotic Transcription

Transcription in eukaryotes is a more complex process than in prokaryotes. Many of the processes and protein complexes involved in transcription are conserved throughout eukaryotes, from yeast to human. To ensure appropriate levels and timing of gene expression a number of problems must be overcome. Each gene contains a promoter, to identify the position of the gene and the start site for transcription. The promoter

sequences contain binding sites for specific transcriptional activators. Activators are proteins that, in response to external or internal signals, direct the transcriptional machinery to individual genes. Eukaryotic DNA is present in the nucleus as chromatin (Kornberg, 1975). Chromatin is a large structure comprising of the DNA and proteins. While allowing the DNA to be tightly packed into the nucleus of the cell, it also forms a barrier to transcription. The relaxing of chromatin structure involves many protein components, allowing the transcriptional machinery to access the gene. Transcriptional activators play a role in the recruitment of the transcriptional machinery to the correct position on the genome. The transcriptional machinery comprises of RNA polymerase, the enzyme that catalyses the transcription of DNA, and a number of protein complexes called the general transcription factors. These complexes perform a number of different roles in transcription, including the regulation of RNA polymerase and ensuring that transcription occurs correctly. Transcription is an involved procedure; the steps outlined above and the complexes involved in each will be discussed briefly, before focusing on the induction of the transcription of a family of genes in yeast.

The Eukaryotic Promoter

Each gene contains control sequences termed the promoter. Promoters contain the DNA sequences required for initiation of transcription. These regions usually include the transcription start site and the TATA box. The TATA box is an AT rich region of DNA found 20 to 30 bp upstream of the transcription start site in higher eukaryotes, and 20 to 40 bp upstream of most yeast genes. This sequence is the binding site for TATA binding protein (TBP), an important general transcription factor that will be discussed later. There are exceptions, such as TATA-less promoters, which usually rely on initiator (Inr) sequences for activation (Martinez *et al.*, 1994). Inr elements can bind

various factors and may facilitate the binding of the transcriptional machinery especially in the absence of the TATA box.

Promoters also contain sequences that bind specific activators. In yeast, these are termed the Upstream Activating Sequences (UAS),and Upstream Repressing Sequences (URS) that bind repressors. The term UAS is usually used to describe nearby binding sites for transcriptional activators. In higher eukaryotes, clusters of binding sites that occur some distance from the gene itself are termed enhancers.

1.3 Transcriptional Activators

Specific activators exact their function on only a subset of genes. Activators invariably consist of a DNA binding domain attached to an activation domain that can contact chromatin modifying proteins and/or the transcriptional machinery to bring about transcription. Activation domains usually fall into one of three classes: acidic, Gln- rich or Pro-rich. There appear to be no conserved motifs in activation domains, suggesting an unstructured protein interaction surface (Triezenberg, 1995). Regulation of gene expression usually involves interplay between activators and repressors. A single activator affecting the transcription of different genes is an efficient way of turning on a set of genes needed at the same time in the cell, for example the different members of a metabolic pathway. It is becoming evident that transcriptional activators can target further complexes involved in transcription. The role of specific activators in recruitment of these factors will be discussed in each section.

1.4 Chromatin

Eukaryotic DNA needs to be packed tightly to fit in the nucleus of the cell. This is accomplished by an arrangement of DNA and proteins that allow the large macromolecules to be packaged. This packaging, the chromatin, must be overcome for transcription to occur, as chromatin naturally represses transcription due to the inaccessibility of the DNA sequence.

Chromatin Structure

The nucleosome is the basic repeating unit of chromatin. It comprises of approximately 146 bp of DNA wrapped 1.65 times around an octomer of histone proteins (Read *et al.*, 1985). The crystal structure of the nucleosome has been determined and shows a cylinder of histone molecules made from two heterodimers of histone proteins H3 and H4, flanked by two heterodimers of histones H2A and H2B. The DNA is wrapped tightly around this structure, forming many contacts with the histone proteins (Luger *et al.*, 1997). There is some irregularity in this wrapping, perhaps suggesting some instability in the rigidity of the structure. This could be important in facilitating the necessary destabilisation of the DNA-nucleosome interactions during transcription. The core nucleosome proteins are among the most conserved proteins found in all eukaryotes. *In vivo* experiments in yeast suggest that if the histone subunits are removed, a general increase in the level of transcription of some genes is observed (Durrin *et al.*, 1992; Han and Grunstein, 1988). Formation of higher order structures facilitated by the histones form a physical barrier to prevent to proteins of the transcriptional machinery accessing the DNA. For a recent review of the nucleosome see Kornberg and Lorch (1999).

Nucleosomal arrays *in vitro* are present as a fibre with a diameter of 10 nm, or are further condensed into a 30 nm fibre. The structure of the 30 nm fibre has long been an area of study. There are two models of the architecture of this fibre. The first suggests a continuation of super-coiling, which is present in the formation of the 10 nm fibre, and the second a zigzag structure where the linker regions cross the fibre in a zigzag matter. A number of recent studies have produced evidence supporting the zigzag theory, and this is currently the preferred model (Cui and Bustamante, 2000; Katritch *et al.*, 2000; Schiessel *et al.*, 2001). Chromatin lacking the amino tail domains of the core histone

proteins cannot form the 30 nm fibre, nor condense into higher order structures that may be present *in vivo* during interphase or meiosis (Carruthers and Hansen, 2000). Acetyl groups on the histone proteins may cause the DNA to "loosen" from the nucleosome structure, perhaps disrupting interactions between the histones and the DNA (Wolffe and Hayes, 1999).

Linker histones (H1) are proteins that are not involved in the core nucleosome, but they are associated with chromatin and bind to the outside of the structure. They have been shown in the past to increase the stability of higher order chromatin structures. It now known that it is the core histone tail domains that are absolutely required for condensation of chromatin into these higher order structures (Carruthers and Hansen, 2000). An important role for the linker histone proteins is in charge neutralisation. DNA is a highly negatively charged molecule, due to charge on the phosphate backbone, and the neutralisation of this charge may allow interaction with other proteins. Linker histones may also play a role in providing binding sites for proteins.

The relationship between the core histone tails and the linker histones has yet to be established. It is clear that the tail regions are essential for condensation and acetylation of these regions inhibits this occurring. Experiments performed *in vitro* suggest that the linker histones suppress acetylation by inhibiting the accessibility of the H3 tail domain to PCAF, a histone acetylase found in humans, and that this inhibition is not due to the folding of the chromatin fibre (Herrera *et al.*, 2000a; Herrera *et al.*, 2000b). *In vitro*, the yeast histone acetylase Gcn5p preferentially acetylates folded nucleosomal arrays (Sendra *et al.*, 2000). This indicates that the histone tails are readily accessible within

the nucleosome structure. The acetylation of these tails is an important mechanism in the control of transcription; this will be discussed in more detail below.

1.4.1 Chromatin remodelling complexes

Chromatin remodelling complexes are a relatively recent discovery; these complexes alter the interactions of DNA with the histone proteins and reposition the nucleosome on the DNA. The best studied of these include Swi/Snf and RSC complexes from yeast, and NURF, CHRAC and ACF from *Drosophila* . All these complexes contain an ATPase subunit essential for remodelling activity. There are three main families of chromatin remodelling complexes: those similar to Swi/Snf and those similar to the ISWI family, and those similar to Mouse CDH-1, each family member contains a related ATPase subunit (Eisen *et al.*, 1995). NURF and ACF are members of the ISWI-like family. Chromatin remodelling complexes and the relationship with transcription has been reviewed recently (Flaus and Owen-Hughes, 2001; Fry and Peterson, 2001).

One of the first chromatin remodelling complexes to be isolated was Swi/Snf in yeast. This complex is non-essential for yeast viability, but is required for the activation of about 3% of non-essential yeast genes, and repression of approximately the same number of genes (Holstege *et al.*, 1998; Sudarsanam *et al.*, 2000). It has been extensively studied at the promoter of the *HO* gene in yeast, which control mating type switching (Cosma *et al.*, 1999; Krebs *et al.*, 1999). The transcription factor Swi5p targets and recruits Swi/Snf to the promoter. This results in hyperacetylation of the DNA surrounding the promoter. It has been suggested that Swi/Snf may expose residues

required for the recruitment of SAGA, and the resulting acetylation may stabilise the decondensed chromatin (Krebs *et al.*, 1999).

All chromatin remodelling complexes that require ATP for their mechanism of action contain proteins that have helicase activity. Helicases catalyse the separation of DNA strands, although none of the chromatin remodelling ATP-dependent subunits that have been identified function in the strand displacement assay for DNA helicase activity. There are other proteins, however, which appear to contain helicase motifs but do not separate the DNA strands, instead they spool the DNA through their structure in an ATP-dependent manner. The DNA repair protein Rad52, which in sequence is closely related to the Swi/Snf complex, has been shown to have the ability to alter local DNA topology. This may suggest that these chromatin remodelling complexes, while lacking classic helicase activity, may react locally to induce changes in DNA twist (Kingston and Narlikar, 1999).

There are differences in activity between different ATP-dependent remodelling complexes. For example, ISWI is not stimulated solely by DNA, but also requires the histone components of nucleosomes, specifically the histone tails (Tsukiyama *et al.*, 1995). Another complex, Mi-2 also requires the histone components of nucleosomes, but is not dependent on the amino-tail domain being present (Brehm *et al.*, 2000; Guschin *et al.*, 2000). A less-related complex, Mot1, is stimulated by the presence of TBP (Auble *et al.*, 1997), a general transcription factor discussed below. This evidence suggests that the catalytic components may interact directly with different chromatin components and interacting proteins. Many remodelling complexes also contain

conserved motifs including bromodomains and chromodomains, which may provide further means of targeting complexes to the correct position on the chromatin. It has also been shown that some remodelling complexes contain components that can interact directly with transcription factors, discussed in more detail below. These different interactions may be used in concert *in vivo* to produce specific targeting of the chromatin modelling complexes to the correct position on the DNA, reviewed by (Peterson and Logie, 2000).

It has been shown by a number of groups that different remodelling complexes can alter the position of the nucleosomes along the DNA (Hamiche *et al.*, 1999; Jaskelioff *et al.*, 2000; Langst *et al.*, 1999). These studies suggest that the repositioning of nucleosomes may be the general method by which chromatin remodelling complexes exert their function. It has been suggested that this mobilisation is due to the application of torsion to the DNA over the surface of the nucleosome. This force could have the effect of pushing the DNA off the nucleosome. Once the DNA has been forced out of the nucleosome structure, reassociation followed by further disassociation may lead to a bulge of unbound DNA being transferred around the structure of the nucleosome, allowing the DNA to be exposed to the transcriptional machinery (Owen-Hughes *et al.*, 1999). RNA polymerase is thought to pass over nucleosomal bound DNA in a similar manner (Studitsky *et al.*, 1995).

It is clear from a number of experiments that different chromatin remodelling complexes undertake different roles in the cell. Another chromatin remodelling complex, the RSC complex is essential for viability in yeast (Cairns *et al.*, 1996). Other

complexes appear to take on different roles outside transcription activation, for example, the ISWI-containing ACF complex can promote the generation of uniformly spaced nucleosome arrays (Ito *et al.*, 1997). Initially chromatin remodelling factors were thought only to play a role in the activation of transcription, but more evidence is emerging to suggest that they may play an equally important role in the repression of transcription, reviewed by (Tyler and Kadonaga, 1999; Wolffe *et al.*, 2000). It may be that this repression is gained *via* higher order chromatin fibre formation. The ability of chromatin to form these fibres may be influenced by the spacing of nucleosomes, which in turn may be influenced by the remodelling complexes. Some chromatin remodelling complexes have been shown to influence nucleosome spacing over large regions of DNA (Sudarsanam *et al.*, 2000).

It is becoming clear that different chromatin modifying complexes utilise different strategies in their contribution to transcriptional activation. These complexes may not only affect transcription, but also affect other cell processes such as DNA repair and replication (Alexiadis *et al.*, 1998; Shen *et al.*, 2000). Chromatin remodelling complexes and the general transcriptional machinery exact their function on many genes, or subsets of genes. A second family of complexes, histone acetyltransferases, which modify the acetylation state of histone tails, also affect the interactions of DNA with histones, and therefore transcription throughout the eukaryotic genome.

1.4.2 Histone Acetyltransferases (HATs)

The acetylation of histone tails leads to a "loosening" of the DNA wrapped around them, and therefore, possibly gene activation. This was first seen with the HAT Gcn5p.

Gcn5p was initially identified as a requirement for full activated transcription of Gcn4p (Georgakopoulos and Thireos, 1992) and later was found have HAT activity (Brownell *et al.*, 1996). This HAT activity was required for Gcn5p to function as a transcriptional co-activator (Candau *et al.*, 1997). Gcn5p was shown to acetylate free histones *in vitro*, but not histones assembled into nucleosomes (Kuo *et al.*, 1996), and it was postulated from this that Gcn5p may act as part of a complex which would allow it to target the nucleosomal histones. Candau *et al.* (1997) also found that the interaction of Gcn5p with Ada2p was also required for activation; Ada2p binds the activation domains of VP16 and Gcn4p. So Gcn5p was thought to form a complex with Ada2p and other Ada proteins, that had previously been shown to interact *in vivo* (Candau *et al.*, 1996). It was found that these proteins also associated with the Spt proteins, and the complete complex was named the SAGA complex, Spt-Ada-Gcn-Acetyltransferase (Grant *et al.*, 1997). The SAGA complex also contains TAFs (TATA- binding protein associated factors) with histone-like folds, perhaps indicating competitive DNA binding between these proteins and other transcription factors. The entire complex is thought to encompass HAT activity, activator interaction and TATA-binding protein interaction and regulation (Grant *et al.*, 1997; Grant *et al.*, 1998; Sterner *et al.*, 1999). Deletion of Gcn5p leads to a relatively minor phenotype in yeast. This may be because of a high level of redundant functions between HATs. It has been observed, however, that if a gene is involved in mitosis it is more likely to have dependence on the SAGA and Swi/Snf complexes (Krebs *et al.*, 2000). The general transcription factor TFIID was found to contain $TAF_{II}250$ that has HAT activity (Mizzen *et al.*, 1996), indicating the link between HAT activity and transcriptional activation.

Since the discovery of Gcn5p, more co-activators have been found that encompass HAT activity. Similar to Gcn5p is the human factor PCAF, other distinct families of HATs have also been discovered, including the $TAF_{II}250$ family and the CBP/p300 family. Each family seems to have a distinct substrate preference. For example, Gcn5p/PCAF HATs appear to work with a subset of transcriptional activators, as well as preferentially acetylating lysine 14 of H3 (Kuo *et al.*, 1996; Kuo *et al.*, 1998; Wang *et al.*, 1998). The preferred substrate differs extensively between families. Another group, the MYST family, have very different biological functions including gene silencing (Reifsnyder *et al.*, 1996) and cell cycle regulation in yeast (Clarke *et al.*, 1999), and prefer H4 as a substrate.

The HAT complex NuA4 contains the HAT Esa1p. It has been found that while the SAGA complex can be deleted without yeast death, Esa1p is essential for yeast viability (Clarke *et al.*, 1999). The targets of NuA4 have been identified as the promoters of ribosomal proteins (Reid *et al.*, 2000), whereas the SAGA complex is associated with a variety of genes of different functions, albeit with some relation to mitosis, NuA4 was the first HAT complex to be linked to a specific function within the yeast cell.

Roles for HATs outside of transcription initiation have also been identified. Elp1p, a protein involved in the elongation process of transcription, contains HAT activity, suggesting a role for histone acetylation after the initiation of transcription has occurred. Acetylated proteins other than histones have been identified, including one HAT, ACTR, which is acetylated by another protein (Chen *et al.*, 1997). This could suggest the existence of acetylation cascades within the cell, which may play a role in transduction of signals.

1.4.3 Histone Deacetylases

Histone deacetylases (H-DACs) were first identified as transcriptional regulators when the purification and sequencing of proteins with known histone deacetylase activity showed similarity to known yeast transcriptional regulators. The yeast protein Rpd3p, already known to affect transcription, was found to share homology with the identified human H-DAC1, thus providing the link between the two functions (Vidal and Gaber, 1991). H-DAC activity plays a role in the repression of transcription, as would be expected of an enzyme with the opposite effect of HATs. Other work in yeast has identified two distinct histone deacetylases activities, termed HDA and HDB. HDB was found to contain Rdp3p as its catalytic component, while this role in HDA was played by another protein, Hda1p (Rundlett *et al.*, 1996). Mutations in either Rpd3p or Hda1p can result in hyperacetylation of the histone tails of the proteins H3 and H4 (Rundlett *et al.*, 1996). In humans, two histone deacetylases HDAC1 and HDAC2 are, along with histone binding proteins, at the core of co-repressor complexes (Zhang *et al.*, 1999). Other components of the complexes confer specificity of substrate to the H-DAC activity.

Many H-DACs identified can be categorised into two groups, Rpd3-like or Hda1-like. A third class of H-DACs has been identified recently. The Sir-2 protein, which is known to be involved in gene silencing, has been identified as a histone deacetylase dependent on NAD^+ for its function. This could be evidence for a role for H-DACs in gene silencing and ageing in yeast and mammals (Imai *et al.*, 2000).

The targeting of H-DAC complexes to the correct regions of chromatin is integral to their function. The co-repressor Sin3-containing complexes, which require H-DAC activity to function, are targeted by a wide variety of methods, including being targeted directly by DNA binding proteins and recruitment by other co-repressor complexes, reviewed by (Knoepfler and Eisenman, 1999). Recruitment of deacetylase activity has also been associated with areas of methylation in the genome (Bird, 2001). Methylation of DNA is a phenomenon linked to gene silencing, which is not found in yeast. In mammals, the sin3 complex can interact with a protein that binds methylated DNA called MeCP2. This protein requires H-DAC activity for repression (Jones *et al.*, 1998; Nan *et al.*, 1997).

Specific activators and recruitment of chromatin modifying complexes

Transcriptional activators can recruit complexes such as Swi/Snf and SAGA to the promoter. Study of the yeast *HO* promoter shows that the activator Swi5p recruits Swi/Snf, followed by SAGA and subsequently a second activator, SBF, recruits the transcription initiation apparatus (Cosma *et al.*, 1999; Krebs *et al.*, 1999). Activation domains containing acidic residues have been shown to interact with a number of different complexes. Gcn4p, a yeast activator involved in the response to amino acid starvation, has been shown to interact independently with members of three classes of complexes involved in transcription activation: Swi/Snf, SAGA and the RNA polymerase II holoenzyme, discussed below. It has been shown at the yeast *PHO8* promoter that the activator Pho4p is able to bind *in vivo* in the absence of either Swi/Snf or SAGA complexes that are required for transcription of the *PHO8* gene (Gregory *et al.*, 1999). This suggests that recruitment takes place, and the activator is indeed causing initiation of transcription. It has also been shown *in vitro* that the activation domain of VP16 can direct the recruitment of SAGA to specific sequences of chromatin template

that contain the relevant binding site. More evidence for this comes again from studies on the activator Gcn4p, in an *in vivo* study at the *HIS3* promoter. Deletion of the Gcn4p binding site leads to a loss of histone acetylation at this locus, but by introducing the binding site at an alternative position, the peak of histone acetylation shifts to this new site (Kuo *et al.*, 2000). The acidic activation domain of Gcn4p is essential for this targeted acetylation. Acetylation alone is not sufficient to activate transcription, the active recruitment of the RNA polymerase II and its associated factors is also required.

1.5 RNA polymerase II

In eukaryotes there are three forms of RNA polymerase, numbered I, II and III (in older literature these are referred to as A, B and C respectively). RNA polymerases I and III are concerned primarily with the transcription of RNA that is not subsequently translated into protein producing, for example, tRNA or rRNA. RNA polymerase II is responsible for the transcription of mRNA that will be translated. This process is controlled by a large number of other protein complexes that ensure correct transcription. The core enzyme of RNA polymerase II contains twelve subunits, all are essential in yeast and all share high homology with their human counterparts, reviewed by (Woychik, 1998). Many of the human polymerase II subunits can substitute functionally for their yeast homologue *in vivo*. RNA pol II in eukaryotes also shares high homology with *E. coli* RNAP. The two largest subunits of yeast RNA pol II, Rpb1 and Rpb2, are homologous to RNAP β' and β, while both Rpb3 and Rpb11 share some homology with the α subunit, which is present as a dimer in RNAP, and Rpb3 and Rpb11 together can form a heterodimer *in vitro*. These four subunits perform the major catalytic functions of RNA polymerase, reviewed by Woychik, (1998); Young, (1991).

The structure of yeast RNA polymerase II has been solved, and shows a marked overall symmetry to the structure of RNAP in *E. coli* in addition to the sequence homology (Cramer *et al.*, 2000; Cramer *et al.*, 2001) (Fig. 1.1). Rpb1 and Rpb2 form a cleft containing the catalytic magnesium ion, and therefore the active site. Other features of the enzyme include grooves; proposed exits for mRNA, pores; channels for nucleotides, jaws; which may help position DNA, and a sliding clamp for stabilisation.

The C terminal domain (CTD)

Rpb1p, the largest RNA pol II subunit, contains a C-terminal domain with the heptapeptide motif sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser, reviewed by Corden, (1990). This highly conserved sequence can become extensively phosphorylated (Dahmus, 1996). This sequence is repeated many times in the CTD, the number of repeats varies between different organisms. The CTD is essential for the viability of yeast (Allison *et al.*, 1988), and the phosphorylation state of this domain is important in transcription. Transcriptional initiation complexes have been found to have an unphosphorylated CTD, and this region is associated with the Mediator complex, discussed below (Kim *et al.*, 1994; Myers *et al.*, 1998). After the switch from initiation to elongation, the CTD is extensively phosphorylated and associated with the Elongator complex, a large complex of proteins associated with RNA polymerase after the initiation of transcription when the enzyme is transcribing the DNA (Otero *et al.*, 1999).

Two kinases which phosphorylate the CTD have been identified as Cdk7 (Kin28p in yeast) (Lu *et al.*, 1992) and Srb10/Cdk8, a kinase/cyclin pair (Liao *et al.*, 1995). Cdk7 is a subunit of the general transcription factor TFIIH, and loss of function mutants of Kin28p in yeast cause a complete loss of transcription of protein-coding genes

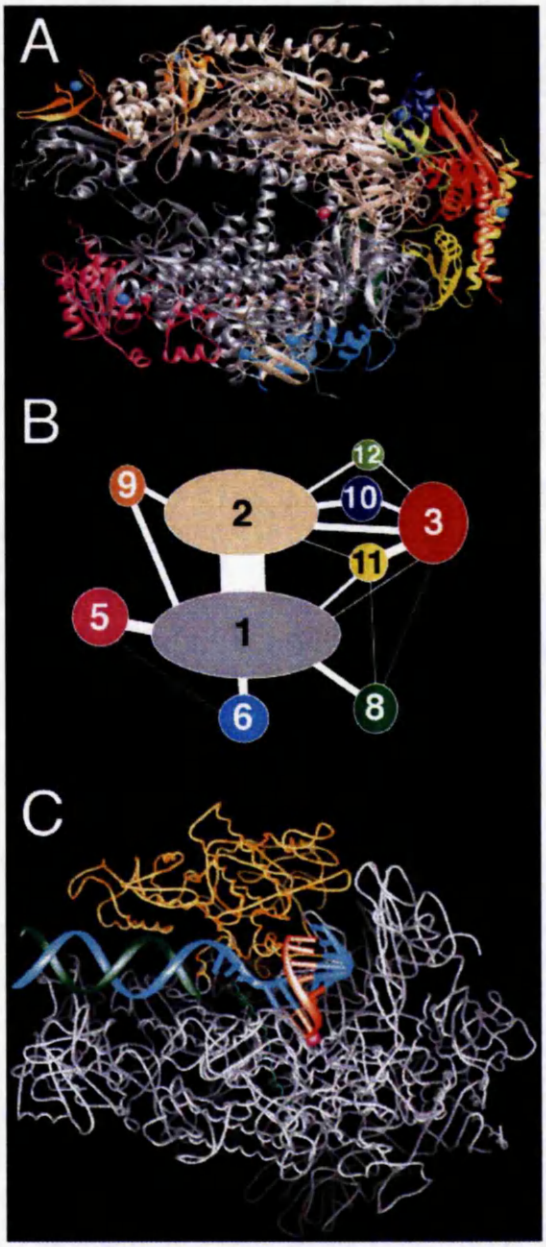


Figure 1.1 The structure of RNA polymerase II from *S. cerevisiae*

The structure of the backbone of RNA polymerase II is shown in A. Each colour represents one subunit of the molecule. B also shows the relationships and interactions between each RNA polymerase II subunit. C shows the position of DNA (blue and green) entering the polymerase, and mRNA (red) being synthesised. The catalytic Mg^{2+} ion is shown in pink.

(Holstege *et al.*, 1998). Cdk7 is thought to be critically important in the switch from initiation to elongation. Analysis of Srb10 suggests that it plays a role in negative regulation of transcription at a genome wide level (Holstege *et al.*, 1998). Another protein, Fcp1p has been identified as a CTD phosphatase in both yeast and humans, although its role is not yet clear (Kobor *et al.*, 1999). The structure of RNA pol II shows that the exit groove for mRNA is positioned near to the CTD (Fig. 1.1). This is consistent with observations of other functions associated with the CTD in RNA processing.

1.5.1 General transcription factors (GTFs)

Basal transcription is the low level of transcription that can occurs *in vivo* in the absence of activated transcription that occurs when the gene is switched on. While RNA pol II is capable of supporting basal levels of transcription *in vitro* (Matsui *et al.*, 1980), it is unable to recognise specific promoter sequences, which direct transcription the correct site, or to support activated transcription. These processes are controlled by the general transcription factors.

GTFs are protein complexes that allow RNA polymerase to recognise promoters and transcription to start at the correct site. These complexes were first discovered *in vitro* after biochemical fractionation of nuclear extracts; GTFs were found in fractions that were able to support activated transcription *in vitro* (Matsui *et al.*, 1980). The assembly of the preinitiation complex (PIC), comprising of template DNA, RNA pol II and five GTFs, begins with the binding of TBP of TFIID to the TATA box, followed by the recruitment of TFIIB, RNA pol II, TFIIF, TFIIE and TFIIH (Ptashne and Gann, 1997). After promoter melting and transcription initiation, the CTD is phosphorylated, and

elongation progresses. Each GTF is a complex of different proteins, many of which have different roles to play in allowing this activated transcription to take place. There has been a large amount of research on each component of the general transcription factors, which will be briefly summarised below.

TFIID

TFIID is composed of the TATA binding protein (TBP) and TBP associated factors (TAFs). TBP is essential for transcription activity from TATA-containing promoters (Cormack and Struhl, 1992), while TAFs are essential for transcription from TATA-less promoters (Burke and Kadonaga, 1996). TBP has been crystallised alone, and as a co-crystal with DNA present. In the presence of DNA, TBP adopts a saddle shaped conformation over the TATA box contacting the minor groove of DNA and bends the DNA accompanied by a partial unwinding of the base pairs (Kim *et al.*, 1993). In the absence of DNA, TBP crystallises as a dimer (Chasman *et al.*, 1993). This has led to speculation that this dimerisation may be a regulatory feature (Jackson-Fisher *et al.*, 1999).

TAFs

TBP-associated factors (TAFs) are found in complex with TBP in the general transcription factor TFIID. *In vitro* it was discovered the TAFs were required for activated transcription from various classes of activators, although they were not required for basal transcription (Pugh and Tjian, 1990). It has also been discovered that TAFs directly contact several transcription factors, perhaps allowing recruitment of TFIID to the promoter. Different TAFs have been found to contact different classes of activator: acidic, Gln-rich and Pro-rich (Verrijzer and Tjian, 1996).

There is a growing body of evidence that TAFs may be involved in other functions in addition to their role in the recruitment of the general transcription machinery by specific activators. TAFs are thought to stabilise TFIID-promoter binding by making specific contacts at initiator elements (Inr) and at downstream promoter elements (DPE) sequences that are present in a number of promoters. These interactions may be especially important at TATA-less promoters (Burke and Kadonaga, 1996; Martinez *et al.*, 1994; Martinez *et al.*, 1995).

TFIIA

Yeast TFIIA comprises of two subunits encoded by the genes *TOA1* and *TOA2*. TFIIA is required for specific transcription *in vitro* (Matsui *et al.*, 1980). *In vivo* it associates with and interacts with TBP and stabilises DNA binding (Imbalzano *et al.*, 1994). It is thought that interaction of TFIIA with specific activators stimulates TFIID-TFIIA-promoter complex assembly (Kobayashi *et al.*, 1995). Extracts from yeast cells with mutant TBP or TFIIA have been found to be defective in recruiting all components of the transcriptional machinery (Holstege *et al.*, 1998). TFIIA also displaces or blocks several negative transcriptional regulators from interacting with TFIID. Holstege *et al* found that yeast extracts from cells with mutations in RNA pol II CTD, Srb proteins or TFIIB were defective in binding all Srb proteins and GTFs apart from TFIIA and TFIID, perhaps indicating that TFIIA and TFIID do not associate with the holoenzyme before transcriptional initiation (Holstege *et al.*, 1998). This information has lead to a proposal that TFIIA and TFIID bind to the promoter independently of the rest of the transcriptional machinery, and are not recruited in the same step as the rest of the transcriptional apparatus.

TFIIB

Yeast TFIIB is a monomer encoded by the *SUA7* gene. Recessive *sua7* mutations shift the transcription start site downstream of normal, indicating a role in start site selection for TFIIB (Pinto *et al.*, 1992). TFIIB interacts with TFIIF (Fang and Burton, 1996), which is part of the RNA pol II holoenzyme. Other evidence shows that specific transcriptional activators contact TFIIB, suggesting a role in recruitment of RNA pol II to the DNA (Lin *et al.*, 1991). Structural studies have suggested that the distance between TFIIB and the RNA pol II catalytic site is approximately 32 bp, the average distance between the TATA box and the transcription start site in humans (Leuther *et al.*, 1996).

TFIIE

TFIIE stimulates the CTD kinase and ATPase activity of TFIIH (Lu *et al.*, 1992; Ohkuma and Roeder, 1994). TFIIE contains a zinc finger motif that is implicated in DNA binding, and can bind to regions of single stranded DNA (Kuldell and Buratowski, 1997). This may suggest that TFIIE may play a role in promoter melting, or maintenance of an open promoter complex. This idea is reinforced by *in vitro* experiments that demonstrate that the need for TFIIE can be bypassed by including a pre-melted promoter complex (Holstege *et al.*, 1995; Pan and Greenblatt, 1994). The requirement for TFIIE *in vivo* varies, perhaps indicating that the requirement is promoter-specific (Holstege *et al.*, 1998; Holstege *et al.*, 1995; Parvin *et al.*, 1992). TFIIE is also thought to play a role in the recruitment of another GTF, TFIIH (Buratowski *et al.*, 1989; Flores *et al.*, 1992).

TFIIF

TFIIF shares many characteristics with the bacterial σ factor (Conaway and Conaway, 1993; Greenblatt, 1991). Crosslinking experiments have identified specific points of contact between TFIIF and template DNA (Robert *et al.*, 1998). TFIIF is critical for tight wrapping of the DNA, perhaps facilitating promoter melting (Coulombe, 1999; Robert *et al.*, 1998).

TFIIH

The nine subunits of TFIIH encompass different enzyme activities including DNA-dependent ATPase, ATP-dependent helicase and CTD kinase, reviewed by (Coin and Egly, 1998). The human helicase activity has been found to be essential for promoter opening *in vivo* (Guzman and Lis, 1999), and that this requirement can, again, be bypassed by using premelted promoter complexes (Pan and Greenblatt, 1994).

As discussed earlier, TFIIH play an important role in the phosphorylation of the CTD of RNA polymerase. The TFIIH kinase/cyclin pairs are Cdk7/cyclinH in humans, and Kin28p/Ccl1p in yeast (Serizawa *et al.*, 1995; Shiekhattar *et al.*, 1995). These CDKs phosphorylate different serine residues in the RNA pol II CTD. In yeast, the cyclin kinase pair Cak/Civ1 regulates cell cycle transitions and also phosphorylates Kin28p, possibly providing a link between cell cycle and transcription (Kimmelman *et al.*, 1999). In mammalian cells Cdk7/cyclin H regulates cell cycle transitions itself, reinforcing the evidence that a link occurs at this point between cell cycle and transcription (Roy *et al.*, 1994).

1.5.2 RNA pol II holoenzyme – The Mediator complex

RNA pol II is present as a large complex within cells, known as the holoenzyme. The proteins associated with RNA pol II are known collectively as the Mediator (Kim *et al.*, 1994; Koleske and Young, 1994). Over twenty proteins have been identified as members of this complex (Myers *et al.*, 1998), but it is not clear if all of these proteins are associated during the transcription of all genes. Much of the work has been done *in vitro*, and the relevant *in vivo* roles remain unclear. Components of the Mediator have different roles in transcription, both positive and negative. This section will outline some of the proteins involved in the yeast mediator, and some of the current thinking on their roles in the cell.

The Srb proteins

Many of the Srb proteins were first identified in a screen designed to isolate Suppressors of RNA polymerase B. *SRB2*, *SRB4*, *SRB5* and *SRB6* genes suppressed mutations associated with loss of function of the CTD (Koleske *et al.*, 1992; Nonet and Young, 1989; Thompson *et al.*, 1993). *SRB4* and *SRB6* are essential for yeast viability (Holstege *et al.*, 1998), while mutations in *SRB2* and *SRB5* show a slow growth phenotype as well as sensitivity to growth temperature (Koleske *et al.*, 1992; Thompson *et al.*, 1993). Similar phenotypes are also seen in yeast which contain mutations in RNA polymerase that result in CTD truncation. *In vitro*, Srb2p and Srb5p are required for basal transcription in a crude extract (Thompson *et al.*, 1993), but this dependence disappears in a more purified system (Svejstrup *et al.*, 1994). This characteristic has also been shown to extend to the Mediator as a whole, suggesting that the CTD and Mediator may be required to overcome general repression of genes. Evidence suggests that these four Srb proteins form a complex *in vivo*; Srb4p interacts directly with Srb2p

and Srb6p, and Srb5p interacts with the complex *via* Srb2p (Koh *et al.*, 1998; Lee *et al.*, 1998). This Srb complex has an affinity with the activation domain of the yeast activator Gal4p, mediated through Srb4p (Koh *et al.*, 1998).

Another essential *SRB* gene, *SRB7* was also identified as being involved in CTD function (Hengartner *et al.*, 1995). The other identified Srb proteins, Srb8p, Srb9p, Srb10p and Srb11p, do not exist in stoichiometric amounts in the Mediator. These proteins are involved in the regulation of a subset of genes and form a sub-complex *in vivo*, reviewed by Myer and Young, (1998). Srb10p and Srb11p form a kinase/cyclin pair and can phosphorylate the CTD of RNA polymerase and cells lacking Srb10p are not fully capable of supporting *GAL* gene expression (Liao *et al.*, 1995). Srb10p can phosphorylate the activator Gal4p and this may affect the ability of Gal4p to activate transcription of the *GAL* genes (Hirst *et al.*, 1999).

The Rgr1 module

The *RGR1* gene is essential for yeast viability. It is required for the negative regulation of some genes involved in meiosis and for the expression of the *HO* gene reviewed by (Carlson, 1997). Studies have shown that Rgr1p is at the centre of a Mediator module, which also contains Sin4p, Gal11p, Pgd11p, and Med2p (Myers *et al.*, 1998). Mutations in these different genes cause similar phenotypes in yeast. Sin4p is not required for yeast viability, but deletion of the gene shows the same defects in repression, derepression and cell morphology as a deletion of the C-terminal of Rgr1. This has led to speculation that Sin4p interacts with the C-terminal of Rgr1p (Carlson, 1997). Mediator purified from Δ*sin4* yeast also lacks Pgd1p and Med2p, suggesting that Sin4p

anchors these two proteins to the module. Mutations in *PGD1* and *MED2* lead to cells unable to support activated transcription from various activators, including VP16 and Gal4p. The last member of this module, Gal11p has been identified in screens for both positive and negative regulators of transcription. Fusion of Gal11p to DNA binding domains shows artificial recruitment that can support activated transcription *in vivo* (Barberis *et al.*, 1995).

Other Mediator Proteins

Other proteins in the Mediator include Rox3p, an essential gene that has deletion phenotypes that overlap with those of *sin4* and *rgr1*, although the Rox3 protein is not part of the Rgr1p module (Myers *et al.*, 1999). *NUT1* and *NUT2* were identified in a screen for mutations affecting the negative regulation of a Swi4-dependent reporter gene. Studies on *NUT2*, which is essential for cell viability, show the protein is necessary for the induction of the histidine biosynthesis genes by the transcriptional activator Gcn4p (Han *et al.*, 1999). Other mediator proteins that have been identified have been named *MED4*, *MED8*, *MED1* and *CSE2*. The *MED* genes are essential, although *CSE2* is not (Han *et al.*, 1999; Myers *et al.*, 1999). The role of these genes is yet to be fully established, although mutations in *MED1* produce similar phenotypes to *SRB10* mutations (Balciunas *et al.*, 1999), and Med8p has been reported to bind directly to the regulatory elements of some genes (Chaves *et al.*, 1999). A general function for the Mediator complex remains ambiguous. The complex appears to have many divergent roles during transcription, but it also shows significant evolutionary conservation within a subset of subunits, with seven of the yeast subunits having orthologs in humans. The Mediator is a modular complex that allows cross talk between the general RNA pol II machinery and gene-specific regulatory proteins.

1.6 Control of nutrient metabolism in yeast

Saccharomyces cerevisiae has been extensively studied with respect to the control of transcription of genes in response to external signals. A specific activator binds to DNA and contacts the general transcription machinery, resulting in transcription of the gene family. Activation of the *PHO* system involves the localisation of the transcription factor inside the nucleus, the *URA* genes are activated by pathway intermediates and the *GAL* genes are transcribed in response to nutrients in the environment. The mechanism of each system will be summarised below.

Regulated cellular localisation

The *PHO* system

The *PHO* genes regulate the metabolism of phosphate in yeast, and the genes are switched on in response to starvation of phosphate. The *PHO5* gene is studied as a paradigm of phosphate regulation. The *PHO5* promoter binds two proteins, Pho4p and Pho2p (Vogel *et al.*, 1989), which act co-operatively to activate transcription (Barbaric *et al.*, 1996; Hirst *et al.*, 1994). Pho4p is a basic helix-loop-helix transcription factor (Ogawa *et al.*, 1993) and Pho2p contains a homeodomain (Berben *et al.*, 1988). Pho4p is phosphorylated at five serine residues by a cyclin/CDK complex, Pho80p/Pho84p, when phosphate is freely available to the cell, but not phosphorylated when phosphate is in short supply (Kaffman *et al.*, 1994). The phosphorylation of Pho4p leads to its exclusion from the nucleus, therefore when phosphate levels drop there is less phosphorylated Pho4p, which is able to remain within the nucleus and activate transcription (Komeili and O'Shea, 1999).

The *PHO* system has also been used to study the effect of chromatin structure on the regulation of transcription. Again, the *PHO5* promoter has been at the centre of this study. There is a change in the nucleosome structure at the promoter between the repressed and the activated state (Svaren and Hörz, 1997), this is thought to be due to the nucleosome remodelling complexes discussed above.

Regulation by pathway intermediates

Ppr1p

The transcription factor Ppr1p controls the transcription of genes in response to uracil (pyrimidine) starvation. Ppr1 binds to a conserved site in the promoter of the *URA* genes (Roy *et al.*, 1990). Ppr1p contains a DNA binding domain and an acidic region thought to be the activation domain (Kammerer *et al.*, 1984; Schjerling and Holmberg, 1996). *In vitro* it has been shown that, although Ppr1p alone cannot activate transcription, addition of one of two pathway intermediates leads to transcriptional activation. The intermediates are orotic acid and dihydroorotic acid , the products of early steps in the pathway. An increase in these intermediates leads to transcription of the later genes in the pathway, and the production of pyrimidines (Flynn and Reece, 1999).

1.7 The *GAL* System

The *GAL* genes produce the enzymes of the Leloir pathway, which metabolise galactose to glucose-6-phosphate, which subsequently enters the glycolytic pathway. The Leloir pathway and the coding genes are shown in Fig. 1.2. Regulation of the *GAL* system in *Saccharomyces cerevisiae* is perhaps one of the best characterised examples of transcriptional control. There are five structural genes, collectively termed the *GAL* genes, which are under the control of three regulatory genes. These regulatory proteins are discussed below.

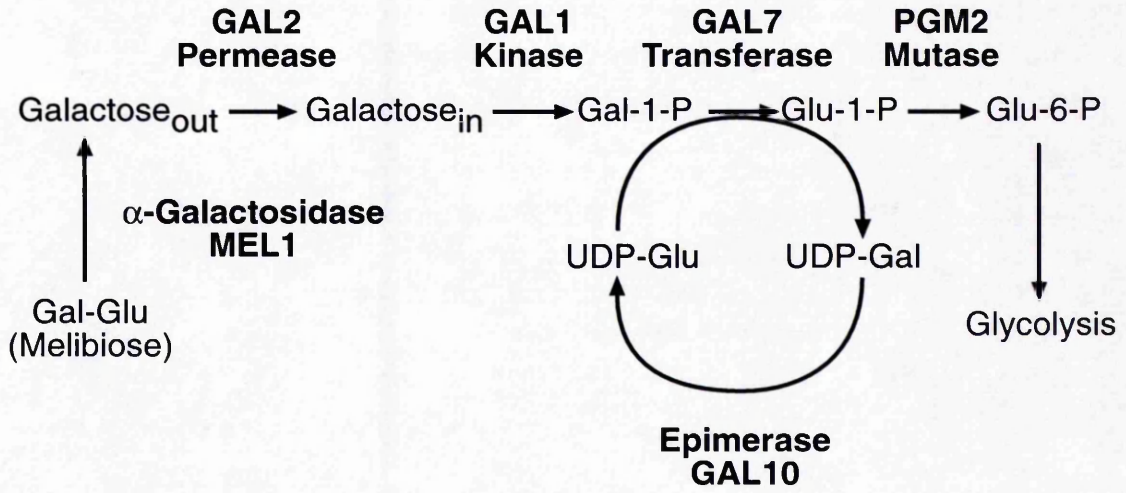


Figure 1.2 The Leloir Pathway

The Leloir pathway metabolises galactose. Gene transcription is under the control of the activator Gal4p, the repressor Gal80p, and Gal3p, the inducer of the *GAL* system

1.7.1 Gal4p

Gal4p is the transcriptional activator of the *GAL* system. It is 881 amino acids in length and encompasses domains responsible for transcription activation, DNA binding, dimerisation and repressor interaction.

DNA Binding

Gal4p binds as a dimer to a 17 bp sequence of DNA that is present in the UAS_{GAL} (Bram and Kornberg, 1985; Carey *et al.*, 1989). The DNA binding domain is localised in amino acids 1 – 100, along with the region responsible for DNA dimerisation (Keegan *et al.*, 1986). These domains are shown in Figure 1.3. The crystal structure of the DNA binding domain has been solved (Marmorstein *et al.*, 1992), and it contains a Zn_2Cys_6 motif common to many yeast transcriptional activators (Carey *et al.*, 1989). The Zn_2Cys_6 motif recognises conserved symmetrical CGG triplets in the 17 bp binding sequence (Marmorstein *et al.*, 1992). Gal4p also contains a nuclear localisation domain at its amino-terminal end (Silver *et al.*, 1986).

Activation

The transcriptional activating potential of Gal4p is independent of the DNA binding domain. Extensive deletion analysis has shown that two areas of Gal4p can activate transcription, residues 148 to 196 and 768 to 881, when attached to the DNA binding domain (Ma and Ptashne, 1987b). Surprisingly, deleting amino acids 768-881 does not allow the other putative transcriptional domain to activate transcription, it may be covered or physically prevented from doing this. If the DNA binding domain of Gal4p is replaced with another, transcription will still be activated from where the molecule

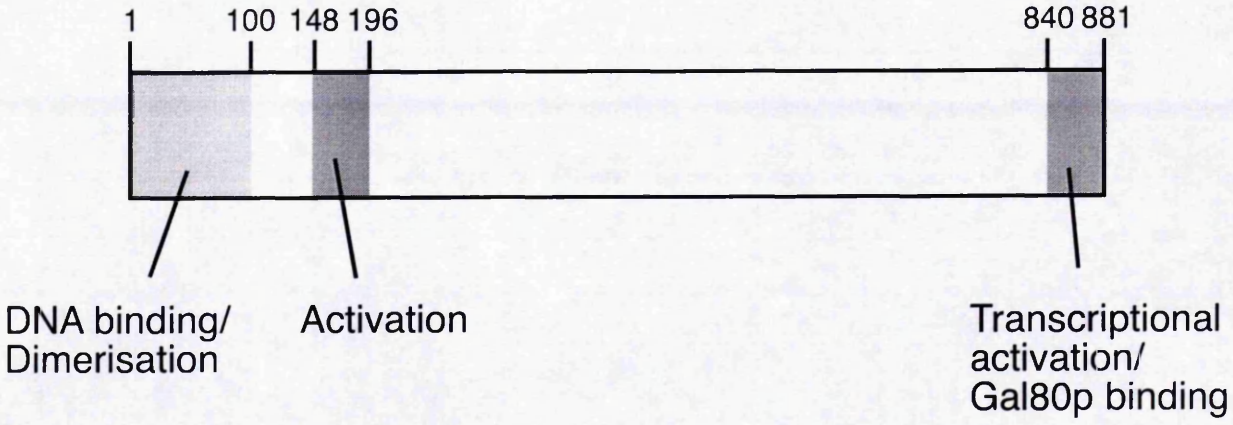


Figure 1.3 Gal4p - The transcriptional activator of the *GAL* genes.

The regions of Gal4p responsible for its function are labelled above. The numbers refer to amino acid sequence.

binds (Brent and Ptashne, 1985), suggesting that there are no specific contacts DNA for each transcriptional activator. The two proposed regions of activation are the most acidic observed in the protein, a common feature of yeast transcriptional activators, and mutants with reduced acidity in the activation regions also display reduced activator function (Gill and Ptashne, 1987).

Phosphorylation

When the yeast cell is growing in galactose Gal4p is present in two forms, both phosphorylated, termed Gal4pII and Gal4pIII (Mylin *et al.*, 1989). It was speculated that the phosphorylation state of Gal4p was important in transcriptional activation. However, Sadowski *et al* showed that the phosphorylation of the Gal4pIII form was dependent on DNA- binding and transcriptional activation, concluding that phosphorylation was a consequence of activation, rather than a factor contributing to it (Sadowski *et al.*, 1991). Gal4p is phosphorylated by the cyclin-CDK pair Srb10p/Cdk8p, it is suggested that phosphorylation is required for maintenance of the activated transcription (Hirst *et al.*, 1999).

1.7.2 Gal80p

Gal80p is the repressor of the *GAL* genes. It is a 435 amino acid protein, and it has been determined that almost all of its sequence is required for repression, amino acids 1 to 321 and 341 to 423. A protein with just these amino acids, 1 to 321 fused to 341 to 423, gives constitutive repression (Nogi and Fukasawa, 1989), indicating that amino acids 322 to 344 may be involved with inducer interaction. Insertion of the entire *E. coli* β-galactosidase gene sequence into this latter region of *GAL80* leads to the production of a protein still capable of repressing Gal4p, although transcription cannot be induced, indicating that there is no spacing requirement for repression (Nogi and Fukasawa,

1989). Gal80p binds directly to the activation domains of Gal4p and blocks activation, probably by preventing access of RNA pol II to the activation domain (Ma and Ptashne, 1987a). Insertion of the acidic activation sequence of VP16 into Gal80p leads to activation of the *GAL* genes (Ma and Ptashne, 1988). Gal80p fused to the activation domain of VP16, in the presence of a activation deficient Gal4p, leads to yeast that grow in galactose and glycerol, indicating non-dissociation of Gal4p and Gal80p (Leuther and Johnston, 1992).

Response to Nutrients by the *GAL* system

Transcription of the *GAL* genes depends on the carbon source the yeast is grown on. There are three different states, and these are outlined below:

1. Glucose

When yeast cells are grown on glucose, the *GAL* genes are not required, and are therefore down-regulated by several mechanisms. Even when galactose is present in addition to glucose, the *GAL* genes are still repressed. Research has concentrated on the promoter of *GAL1*, but is thought to apply to each of the *GAL* structural genes. Glucose repression is mediated in two ways. The first of these involves the specific repressor Gal80p. Repression by Gal80p in a glucose-containing environment accounts for 5 – 10 fold repression of the *GAL1* gene (Johnston *et al.*, 1994; Lamphier and Ptashne, 1992). The second method of glucose repression is facilitated by the Mig1 protein. The localisation of Mig1p is under the control of glucose. When glucose is added to the media, Mig1p is rapidly transported to the nucleus, and when glucose is removed Mig1p is actively exported (De Vit *et al.*, 1997). Mig1p exhibits glucose-dependent DNA binding *in vivo* to the *GAL1* and *GAL4* promoters (Frolova *et al.*, 1999). Mig1p is responsible for a 4-fold repression of the *GAL1* gene (Lamphier and Ptashne, 1992) by

recruitment of the Ssn6p-Tup1p co-repressor to the promoter (Tzamarias and Struhl, 1994). Binding sites for Mig1p are also found in the *GAL4* promoter. Repression of this gene by Mig1p decreases the amount of Gal4p in the cell, therefore there is less Gal4p to bind to the UAS_{GAL} and activate transcription. This mechanism is responsible for 40-fold repression of the *GAL1* promoter (Johnston *et al.*, 1994; Lamphier and Ptashne, 1992).

2. Non-inducing non-repressing carbon source

Raffinose and glycerol are termed non-inducing, non-repressing carbon sources. When yeast is grown in these carbon sources, the *GAL* genes are switched off, but not actively repressed as in the case of repression by glucose. The only mechanism responsible for this repression is Gal80p, which binds to Gal4p and blocks the activation domain (Ma and Ptashne, 1987a).

3. Galactose – activation of the *GAL* genes

When galactose is added to a non-repressing non-inducing media, the *GAL* genes become switched on, achieving maximum transcriptional activation within 30 minutes and producing high levels of mRNA, 1% of the cell total for each *GAL* gene (St John and Davis, 1981). This activation is mediated by the inducer of the *GAL* system, Gal3p.

1.7.3 Gal3p

The *GAL3* gene has upstream binding sites for Gal4p and its transcription can be regulated by both glucose and galactose (Bajwa *et al.*, 1988). Strains lacking Gal3p show a delay in the induction of the *GAL* genes of 3-5 days (Rotman and Spiegelman, 1953). These cells are termed Long Term Adaptation (LTA) mutants. This LTA phenotype led to early confusion of the exact role of Gal3p in the induction process.

Early models of *GAL* gene induction suggested that Gal3p converted galactose, or a product of the Leloir pathway, into an inducing molecule. This would explain why the *GAL* genes are still eventually switched on without Gal3p; levels of the inducer molecule would naturally rise as galactose levels increased in the cell (Broach, 1979). However, Bhat and Hopper showed that activation of the *GAL* genes can occur independently of galactose (Bhat and Hopper, 1992). They overproduced Gal3p *in vivo* and found that this activated *GAL* gene transcription. They also overproduced Gal1p, the galactokinase, which is highly homologous to Gal3p, in a strain lacking in Gal3p, and in the absence of galactose, and found that this overproduction also switches on the *GAL* genes.

Research on the milk yeast, *Kluyveromyces lactis*, gave insight into the mechanism of induction. *K. lactis* contains homologues to Gal4p and Gal80p, but only one molecule Gal1p, which has both galactokinase activity and compensates for *gal3* mutation in *S. cerevisiae* (Meyer *et al.*, 1991). To investigate the interaction between Kl-Gal80p and Kl-Gal1p, his-tagged Kl-Gal80p was bound to a nickel column and cell extract containing Kl-Gal1p was then passed through the column. Kl-Gal1p associated with Kl-Gal80p in the presence of galactose and ATP, or another nucleotide (Zenke *et al,* 1996). In *S. cerevisiae* Gal3p interacts directly with Gal80p and this interaction activates transcription (Suzuki-Fujimoto *et al.*, 1996), and is stabilised by galactose and ATP (Yano and Fukasawa, 1997). It would therefore seem that the inducer Gal3p is the "sensor" of the *GAL* genetic switch.

How does Gal3p activate transcription? Gal3p has no galactokinase activity (Bhat *et al.*, 1990), and *E. coli* galactokinase can complement lack of Gal1p but not Gal3p, indicating that the activity is not required for induction (Bhat *et al.*, 1990). The binding of Gal3p to Gal80p activates transcription. There are two models of how Gal4p-

mediated activation comes about. The first possibility is that the interaction of Gal3p with Gal80p excludes Gal80p from the nucleus allowing Gal4p to recruit the transcriptional machinery (Peng and Hopper, 2000). The second possibility is that Gal3p remains bound, but a conformational change alleviates the repression of Gal4p by Gal80p (Platt and Reece, 1998). The exact mechanism of how the activation domain of Gal4p becomes accessible to RNA pol II machinery is still elusive.

1.7.4 Gal1p and Gal3p

As stated previously, Gal1p compensates for Gal3p as a transcriptional inducer. Gal1p is remarkably similar to Gal3p (Bajwa *et al.*, 1988). The proteins share 90% homology and 70% identity at the amino acid level over the entire length of their sequence. The sequences of both proteins are shown in Fig. 1.4. The similarity is thought to have arisen *via* a duplication of the entire *S. cerevisiae* genome 100 million years ago (Wolfe and Shields, 1997). Other similar "pairs" of genes include *NTH1/NTH2* and *CYC7/CYC1*. The analysis of *K. lactis* supports this hypothesis; as stated previously it has one protein, Kl-Gal1p, that is both the galactokinase and the transcriptional inducer. Gal3p contains five regions of conserved sequence that are found in all known galactokinases. Comparison of these sequences is shown in Fig. 1.5. Region III shows an interesting variation within the Gal3p sequence. There is an apparent "deletion" of two amino acids; a serine and an alanine at amino acid position 164. Analysis of this deletion provides a starting point for investigation into the mechanism of both Gal1p and Gal3p, and the differences between the capabilities of each protein. This in turn could lead to insight into the mechanism of both galactokinase activity and transcriptional control.

```
Gal1p          528 aa vs.
Gal3p          520 aa

        10        20        30        40        50        60        70        80        90       100       110       120
MTKSHSEEVIVPEFNSSAKELPRPLAEKCPSIIKKFISAYDAKPDFVARSPGRVNLIGEHIDYCDFSVLPLAIDFDMLCAVKVLNEKNPSITLINADPKFAQRKFDLPLDGSYVTIDPSV
:. .      :: :.: ...::: . .: ...  :...: .:::.:::::::::::::::::::::::::::::::: ::::::::.:.:::::::: :::::::::::::::::::..:::::
MNTN------VPIFSSPVRDLPRSFEQKHLAVVDAFFQTYHVKPDFIARSPGRVNLIGEHIDYCDFSVLPLAIDVDMLCAVKILDEKNPSITLTNADPKFAQRKFDLPLDGSYMAIDPSV
              10        20        30        40        50        60        70        80        90       100       110

      130       140       150       160       170       180       190       200       210       220       230       240
SDWSNYFKCGLHVAHSFLKKLAPERFASAPLAGLQVFCEGDVPTGSGLSSSAAFICAVALAVVKANMGPGYHMSKQNLMRITVVAEHYVGVNNGGMDQAASVCGEEDHALYVEFKPQLKA
:.:::::::::::::::.:::.::::: ..:::.: :.::..:.:::.::::  :: ::.::::..:::: .. .::..: :::.:::::::::::::::::.:: :::::::::::.:.:::
SEWSNYFKCGLHVAHSYLKKIAPERFNNTPLVGAQIFCQSDIPTGGGLSS--AFTCAAALATIRANMGKNFDISKKDLTRITAVAEHYVGVNNGGMDQATSVYGEEDHALYVEFRPKLKA
   120       130       140       150       160       170       180       190       200       210       220       230

      250       260       270       280       290       300       310       320       330       340       350       360
TPFKFPQLKNHEISFVIANTLVVSNKFETAPTNYNLRVVEVTTAANVLAATYGVVLLSGKEGSSTNKGNLRDFMNVYYARYHNISTPWNGDIESGIERLTKMLVLVEESLANKKQGFSVD
:::::::::::::::::::::: :::::::::::::::::.:::.:::. :.:.:: : ..:.....:::::::::::..:::::.:: . :::::: .::::: ::: :::::.. ::.::.:
TPFKFPQLKNHEISFVIANTLVKSNKFETAPTNYNLRVIEVTVAANALATRYSVALPSHKDNSNSERGNLRDFMDAYYARYENQAQPWNGDIGTGIERLLKMLQLVEESFSRKKSGFTVH
     240       250       260       270       280       290       300       310       320       330       340       350

     370       380       390       400       410       420       430       440       450       460       470       480
DVAQSLNCSREEFTRDYLTTSPVRFQVLKLYQRAKHVYSESLRVLKAVKLMTTASFTADEDFFKQFGALMNESQASCDKLYECSCPEIDKICSIALSNGSYGSRLTGAGWGGCTVHLVPG
... .:::::::::::::: :::::::::::::::::::::::::::.:.::.:.: .::::: .:: ::::::::::::::: : ..:::::.:::.:::::::::::::::.::::.
EASTALNCSREEFTRDYLTTFPVRFQVLKLYQRAKHVYSESLRVLKALKMMTSATFHTDEDFFTDFGRLMNESQASCDKLYECSCIETNQICSIALANGSFGSRLTGAGWGGCTIHLVPS
      360       370       380       390       400       410       420       430       440       450       460       470

      490       500       510       520
GPNGNIEKVKEALANEFYKVKYPKITDAELENAIIVSKPALGSCLYEL
: :::.:.:.::: ..::.:.:: .:: ::..:::::::::.:::::
GANGNVEQVRKALIEKFYNVRYPDLTDEELKDAIIVSKPALGTCLYEQ
     480       490       500       510       520

72.9% identity
```

Figure 1.4 Alignment showing the amino acid sequence similarity and identity between Gal1p and Gal3p of *Saccharomyces cerevisiae*. The upper line represents Gal1p, the lower Gal3p. **:** = identical residues **.** = similar residues

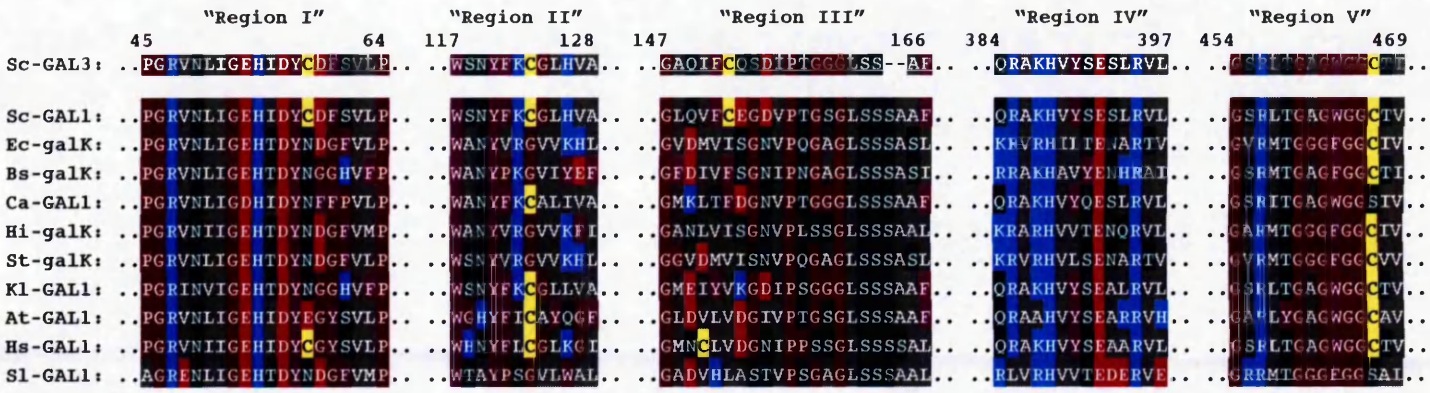


Figure 1.5 Regions of Galactokinase homology

Saccharomyces cerevisiae Gal3p, shown in line one, exhibits five conserved motifs from galactokinases, yet has no enzyme activity. In homology region III there is an apparent "deletion" of two amino acids highly conserved in all other galactokinases.

KEY: Sc = *Saccharomyces cerevisiae*; Ec = *Escherichia coli*; Bs = *Bacillus subtilis*; Ca = *Candida albicans*; Hi = *Haemophilus influemzae*; St = *Salmonella typhimurium*; Kl = *Kluyveromyces lactis*; At = *Arabidopsis thaliana*; Hs = *Homo sapiens*; Sl = *Streptomyces lividans*. Amino acids have been coloured according to their properties. Blue indicates positively charged amino acids (H, K, R), red indicates negatively charged amino acids (D, E), green indicates polar neutral residues (S, T, N, Q), grey indicates nonpolar aliphatics A, V, L, I, M,), and purple indicates nonpolar aromatic residues (F, Y, W). Brown is used to indicate proline and lysine, whereas yellow indicates cysteine.

1.7 Aims of this project

This project aimed to investigate the mechanism of the *GAL* genetic switch, especially with reference to Gal1p and Gal3p and their mechanisms. Experiments were designed to investigate the differences in function between Gal1p and Gal3p. Plasmids have been constructed which encode *GAL3* sequence with insertions from the *GAL1* gene (Platt *et al.*, 2000). These insertions are based around homology region III shown in Fig. 1.5. The initial aim of this project was to characterise the phenotypes of the insertional mutants, and to subsequently assay any galactokinase activity that they may have *in vitro*. The galactokinase activity of Gal1p would also be assessed, and compared to activity of Gal3p mutants. Results from these experiments may lead to proposed functions for the conserved homology region III. Experiments to identify the binding sites of the small molecules associated with both Gal1p, and subsequently Gal3p, may lead to insights into the induction of the *GAL* genes, and relationship of gene induction to galactokinase activity. The induction of the *GAL* genes can be mediated by both Gal3p and Gal1p. *In vivo* the induction appears to occur immediately upon addition of galactose to the growth medium, this induction is mediated by Gal3p. In yeast deleted for *GAL3*, the growth of the yeast is delayed for 2-3 days; this induction of the *GAL* genes is mediated by Gal1p and is known as the LTA phenotype. Investigation of the pattern of expression may give insights into the mechanism of the LTA phenotype, and why the induction of the *GAL* genes differs when induced by Gal1p rather than Gal3p.

The aims of this project were to investigate the molecular interactions involving the proteins employed in the induction of the *GAL* genes. Much work has been done on the *GAL* genetic switch, but the mechanism of action by Gal3p is still unclear. By studying this protein, the highly homologous protein Gal1p, and the interactions these proteins

have with other proteins and small molecules it may be possible to gain further insights into the mechanism of *GAL* gene induction.

Chapter 2

Materials and Methods

All chemicals and solutions were obtained from SIGMA or BDH unless otherwise indicated. All media and solutions were prepared and autoclaved at 15 psi before use.

2.1 *Escherichia coli* strains

Strain	Genotype	Source or reference
XL1-Blue	*recA1, endA1, gyrA96, thi1, hsdR17, supE44, relA1, lac* F' [*proAB, lacI*^{q}*ZDM15, Tn10* (tet^r)]	(Bullock *et al.*, 1987)
DH5α	*F- mcrA, mcrB f*80d*lacZDM15* D(*lacZYA-argF*)U169 *endA1, recA1 hsdR* (rk^-, mk^-) *supE44, thi-1,* l- *gyrA96, relA1*	(Hanahan, 1983)

2.1.1 Growth of *E. coli*

E. coli cells were grown in liquid culture at 37°C with aeration. Growth was monitored by measuring the culture at $A_{600\,nm}$. Cells were grown in Luria Bertani broth (LB) with or without ampicillin (50-100 μg ml^{-1}). *E. coli* were grown on solid media by plating on LB + 1% (w/v) agar with or without selective antibiotics. Plates were inverted and incubated at 37°C. Colonies usually appeared after 12-24 hr (Maniatis *et al.*, 1982).

2.2 Transformation of DNA into *E. coli* cells

Competent cells were prepared by a variation on the method of Hanahan (1983). LB broth (100 ml) was inoculated with the *E. coli* strain to be transformed and incubated at 37°C with aeration until A_{600nm} = 0.6 to 0.9. The cells were harvested by centrifugation at 1800 *g* for 5 min at 4°C and resuspended in 50 ml of ice cold 100 mM $MgCl_2$ and incubated on ice for 60 min. The cells were spun at 1800 *g*, 5 min at 4°C and then resuspended in 40 ml of ice cold 100 mM $CaCl_2$ and incubated on ice for 60 min. The

cells were centrifuged again, 1800 *g*, 5 min at 4°C, and resuspended in 2-3 ml of ice cold 100 mM $CaCl_2$, 10% glycerol. The cells were then snap frozen in liquid N_2 in aliquots of 100 µl and stored at -80°C.

Competent cells, prepared as detailed above, were allowed to thaw slowly on ice. The cells (50 µl) were then mixed with the DNA (100 ng) to be transformed and incubated on ice for 30 min. The cells were then heat shocked at 42°C for 2 min. LB broth (1 ml) was then added to the transformation mixture, followed by incubation at 37°C for 60 min. Aliquots of each culture (200 µl) were then plated onto selective LB agar plates and incubated at 37°C. Colonies usually appeared after 12-24 hours (Maniatis *et al.*, 1982).

2.3 *Saccharomyces cerevisiae* yeast strains

Strain	Genotype	Source or reference
FY250	MAT α *ura3-52 his3Δ200 leu2Δ1 trp1Δ63*	F. Winston, (Harvard Medical School)
JPY5	*FY250::lys2Δ385*	(Barberis *et al.*, 1995)
JPY5::Δ*GAL1*	*FY250::lys2Δ385::Δgal1*	(Platt *et al.*, 2000)
JPY5::Δ*GAL3*	*FY250::lys2Δ385::Δgal*	(Platt *et al.*, 2000)
JPY5::Δ*GAL1*, Δ*GAL3*	*FY250::lys2Δ385::Δgal1Δgal3*	(Platt *et al.*, 2000)
MC2	*MAT a trp1 ura3-52 leu2-3 prc1- 3407 prb1-112 pep4- 3*	C. Stirling. (University of Manchester)

2.3.1 Growth of yeast

Yeast were grown in liquid media, either YPD broth (1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose) or YNB with amino acids for selection plus carbon source (0.7% w/v Yeast Nitrogen Base w/o amino acids (DIFCO), 0.072% w/v amino acid mix, 2% w/v carbon source or 3% (v/v) ethanol). The culture was incubated at 30°C with aeration. Growth was monitored by measuring the absorbance of the culture at 600 nm. Yeast was grown on solid media by plating out from either a single colony or a frozen stock on either YPD +1% agar, or yeast dropout media, carbon source plus 1% agar (w/v). Plates were inverted and incubated at 30°C. Colonies usually appeared after 2 days (Sherman *et al.*, 1986).

2.4 Transformation of yeast

This protocol is adapted from Sherman *et al.*, (1986). The strain to be transformed was grown in 200 ml YPD until an absorbance at 600 nm of 1.0 was reached. The cells were harvested by centrifugation at 1800 *g*, 5 min. The pellet was resuspended in 20 ml TE (10 mM Tris-HCl (pH 7.4), 1mM EDTA), and centrifuged at 1800 *g*, 5 min. The pellet was then resuspended in 2 ml TE including 0.1 M lithium acetate. The solution was agitated at 30°C for 60 min. Then 200 µl of the yeast cells was mixed with 10 µl of 5 $\mu g \mu l^{-1}$ sonicated salmon sperm (New England Biolabs) and approximately 1 µg of the DNA to be transformed. 40% (w/v) PEG4000 in TE including 0.1 M lithium acetate was added and the mixture heat shocked at 42°C for 15 min, then centrifuged for 10 sec. The resulting pellet was resuspended in 0.2 ml TE buffer, and the solution spread onto selection plates. The transformed colonies usually appeared within 2 days.

2.5 Plasmids used in this study

The plasmids listed below are based on pJP139, Figure 2.1. pJP139 is a multicopy plasmid containing origin of replication and markers allow for selection after transformation into *E. coli* and yeast, the gene, which confers ampicillin resistance on to *E. coli*, and the yeast *HIS3* gene. These plasmids expressed chimeric genes expressed from either the *GAL1* or the *GAL3* promoter. Numbers in the protein description refer to amino acids replaced in Gal3p with the corresponding sequence from Gal1p. These plasmids were obtained from Reece lab stocks (Platt *et al.*, 2000).

Plasmid	Encoded Protein	Promoter
pJP139	none	none
pAP102	Gal1p	*GAL1*
pRJR374	Gal1p	*GAL3*
pAP101	Gal3p	*GAL1*
pAP25	Gal3p	*GAL3*
pAP100	Gal3p+130-210	*GAL1*
pAP91	Gal3p+130-210	*GAL3*
pAP109	Gal3p+141-185	*GAL1*
pAP107	Gal3p+160-170	*GAL1*
pAP110	Gal3p+160-170	*GAL3*
pAP104	Gal3p+SA	*GAL1*
pAP114	Gal3p+SA	*GAL3*
pRJR372	Gal3p+A	*GAL3*

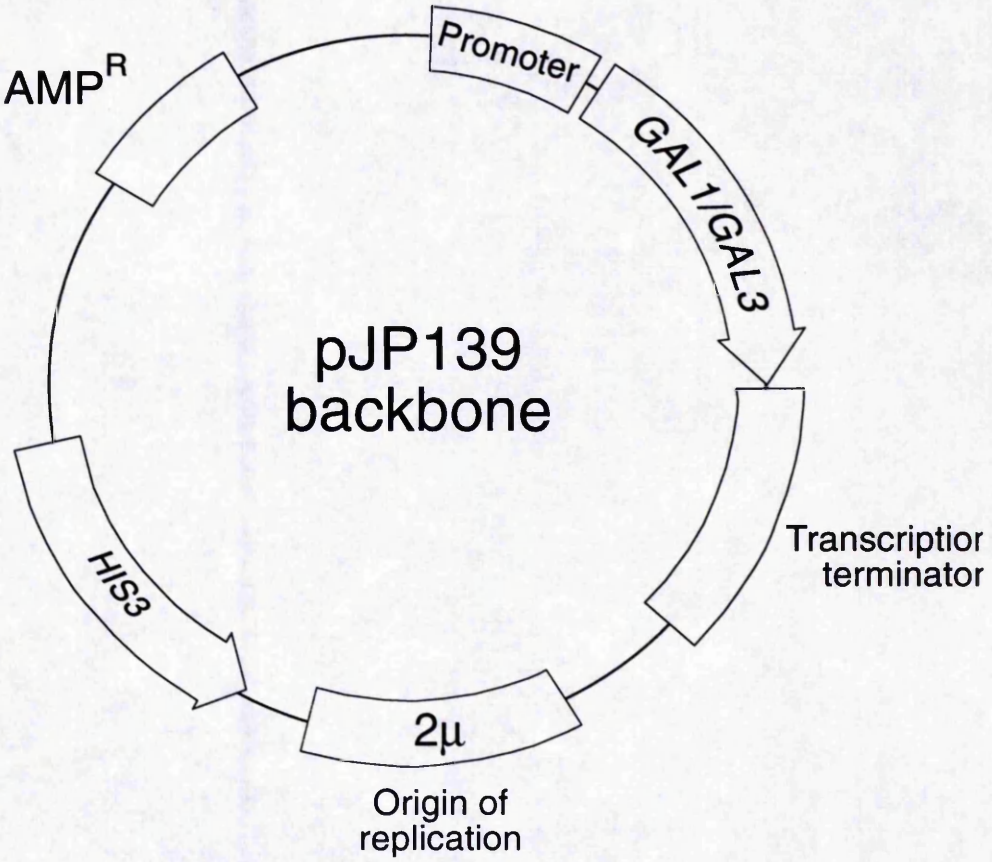


Figure 2.1 Representation of the plasmid pJP139

GAL1, GAL3 or the chimeric proteins were introduced into this plasmid where indicated under the control of either the *GAL1* or *GAL3* promoter. The plasmid also contains a *HIS3* auxotrophic marker, and a gene conferring ampicillin resistance.

The plasmids listed below are based on pYEX-BX (Clontech) shown in Figure 2.2. pYEX-BX is an expression plasmid where the gene expressed is controlled by the *CUP1* promoter. The plasmid contains a copy of the *URA3* gene and a deficient copy of the *LEU2* gene. It also contains the gene to confer ampicillin resistance and has the 2 μ origin of replication, causing the plasmid to be present at high copy number in the cell. These plasmids were obtained from Reece lab stocks(Platt *et al.*, 2000).

pAP45	Gal3p	*CUP1*
pAP60	Gal1p	*CUP1*
pAP112	Gal3p+SA	*CUP1*

2.5.2 Oligonucleotide sequences used in this study

Name	Complementing region	Sequence
HCR3	+1 to +20 *GAL10* coding strand	5' ATGACAGCTCAGTTACAAAG 3'
HCR4	+170 to +190 *GAL10* non-coding strand	5' GAATGTGATGCTTGGTCAAG 3'
HCR5	+374 to +394 *ACT1* coding strand	5' AAACTTTCAACGTTCCAGCC 3'
HCR6	+687 to +703 *ACT1* non-coding strand	5' TTGAAGAAGATTGAGCAGCG 3'
3053A	*GAL1* +1 to +18 coding strand	5' ATGACTAAATCTCATTCA 3'

2.6 Analysis of nucleic acids

2.6.1 Gel electrophoresis of DNA

DNA fragments were separated by electrophoresis on agarose gels. The protocol was adapted from (Maniatis *et al.*, 1982). Agarose gels can be used to separate DNA ranging in size from a hundred base pairs to over 20 kb.

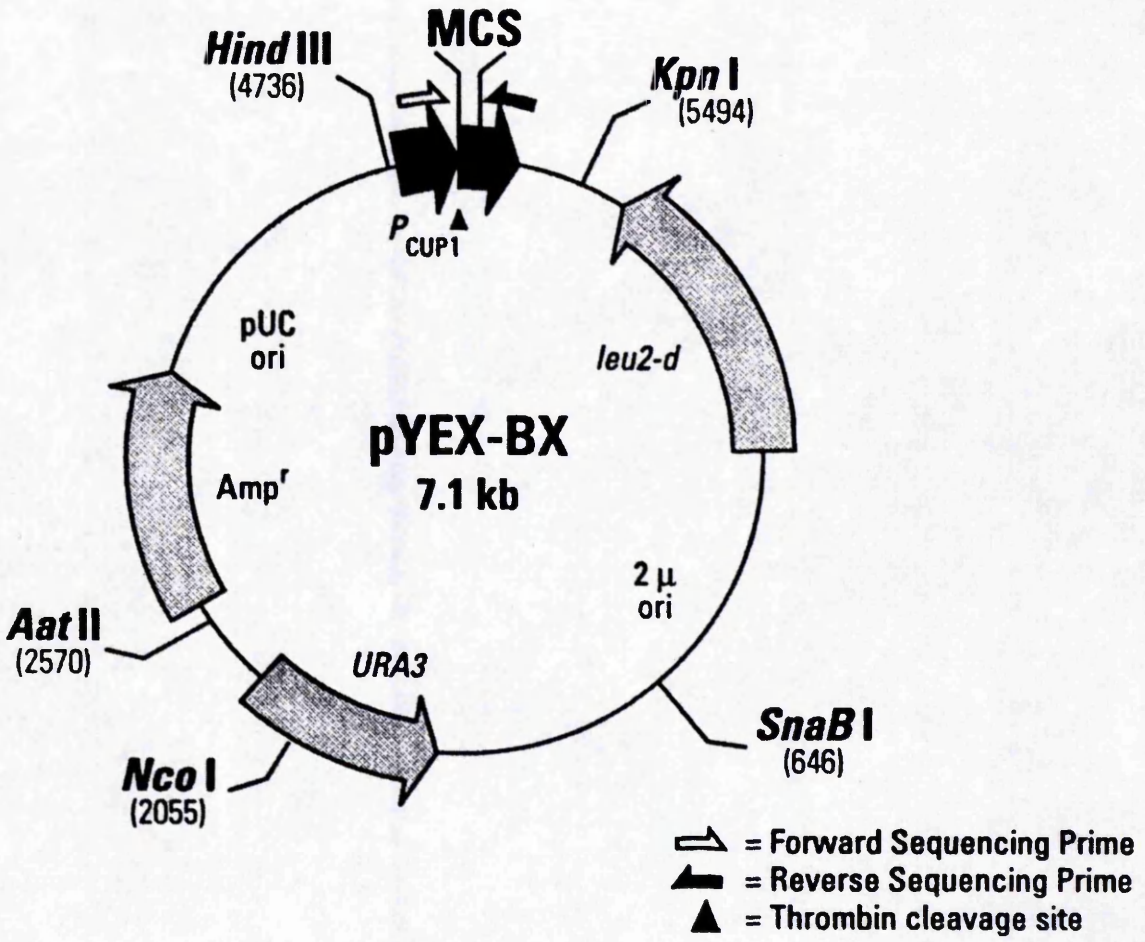


Figure 2.2 Representation of the plasmid pYEX-BX.

pYEX-BX is an expression plasmid for copper inducible protein expression (Clontech). The gene was inserted into the multi-cloning site (MCS), where it was expressed from the *CUP1* promoter. The plasmid also contained a gene conferring ampicillin resistance, and *URA3* and *leu2-d* auxotrophic markers.

The gels are a complex network of polymeric molecules through which negatively charged DNA can migrate under the influence of an electric field. Agarose, at a concentration of 0.5% to 3% (w/v), was melted in TAE buffer (pH 8.0) (40 mM Tris acetate, 2 mM EDTA), and poured into a gel former with a comb to form wells in the gel. DNA was mixed with loading buffer (40% (w/v) sucrose, 100 mM Tris-HCl (pH 8.0), 1 mM EDTA, 0.5 mg/ml bromophenol blue) and loaded into the wells of the set gel. A voltage of 100 V was applied until the Bromophenol Blue had migrated through approximately 75% of the length of the gels. The gels were then stained in TAE, 5 μM ethidium bromide for 30 min. The DNA was visualised with UV light and analysed using Quantity One software (BioRad).

2.6.2 Preparation and purification of DNA

Plasmids were prepared from *E. coli* by two methods. The method chosen was dependent on the amount of DNA required for subsequent manipulations. In general, two sizes of bacterial culture were used, 5 ml and 100 ml. Cells were grown in the desired amount of media with the required selective antibiotics. Plasmids in cultures of 5 ml were isolated using the Mini-Prep kit from Qiagen (according to manufacturer's instructions). Plasmid was extracted from cultures of 100 ml using the Midi-prep kit (Qiagen). Impurities were removed from DNA using PCR purification kit (Qiagen). DNA fragments were purified from agarose gels using the DNA Gel purification kit (Qiagen).

2.6.3 DNA modifying enzymes

Restriction enzymes were purchased from New England Biolabs, Roche and Boehringer Mannheim. Protocols were adapted from (Maniatis *et al.*, 1982). DNA was digested by

incubation with restriction enzyme at 37°C for 1 hr. The DNA fragmentation pattern was observed by electrophoresis of a sample of the digest (approximately 300 ng of DNA) through a 1% agarose gel.

T4 DNA ligase (New England Biolabs) was used for the ligation of both blunt- and sticky-ended DNA fragments. The following buffer used: 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, 5% (w/v) PEG 8000. DNA fragments were incubated at a molar ratio of insert:vector of 3:1, and in a total volume of 20 µl. One unit of ligase was added, and the mixture was allowed to stand at 20°C for 2 hr (cohesive or 'sticky' end ligation), or placed at 15°C overnight (flush or 'blunt' end ligation).

2.6.4 DNA sequencing

After purification of plasmids or PCR products, 100 ng/kb of the DNA to be sequenced was mixed with 4 µl Terminator mix (PE Biosystems) and 3.2 pmol oligonucleotide to a final volume of 10 µl. This reaction mix was then placed in a PCR machine and was heated to 96°C for 4 min, followed by 30 cycles of 98°C (30 s), 50°C (15 s), 60°C (4 min). The samples were then precipitated by addition of 10 µl 3 M NaAc (pH 4.5), 250 µl ethanol and 1µl of 10 $mgml^{-1}$ glycogen, mixed and placed on ice for 10 min. Reactions were then centrifuged at maximum speed in a microcentrifuge for 15 min. After removal of the supernatant, the pellet was washed with 70% ethanol. It was centrifuged again for 5 min and the supernatant thoroughly removed. The pellets were then air dried for 30 min at 20°C.

Sequencing was then carried out using the Big Dye method by the Sequencing Facility, School of Biological Sciences, Stopford Building, Manchester University.

2.7 Purification of protein

Gal80p and Gal4p (1-93+768-881) used in *in vitro* studies were obtained from Reece lab stocks (Gifts from D.J Timson, R.J. Reece). Gal4p (1-93+768-881) was purified from *E. coli* based on the method of (Reece *et al.*, 1993). Gal80p, Gal1p, Gal3p and Gal3p+SA were over-expressed and purified from yeast based on the method of Platt and Reece (1998). MC2 transformed with pAP45, pAP60 or pAP112 were grown at 30°C in yeast nitrogen base plus dropout amino acids until the $A_{600\,nm}$ reached 1.0. Cells were inoculated with 0.5 mM $CuSO_4$ followed by further incubation with aeration 24 hours at 30°C.

2.7.1 Extraction of protein

The cells were harvested by centrifugation at 4400 *g* for 10 min, then washed in protein purification Buffer A (50 mM HEPES (pH 8.0), 500 mM NaCl, 1.4 mM β-mercaptoethanol, 10% glycerol). The solution was centrifuged at 4400 *g* for 5 min and resuspended in protein purification Buffer B (50 mM HEPES (pH 8.0), 500 mM NaCl, 1.4 mM β-mercaptoethanol, 10% glycerol plus 5 ml protease inhibitor solution (1 protease inhibitor tablet (EDTA free) (Boehringer Mannheim) diluted in 15 ml distilled water)). The sample was loaded into a bead beater (Biospec) with an equal quantity of glass beads (Sigma). The cells were broken open at 0°C by nine cycles of 1 min

agitation, followed by 1 min cooling time. The resulting lysate was centrifuged at 23,500 *g* for 20 min, and the supernatant retained and diluted 2:1 with protein purification Buffer A.

2.7.2 Isolation of poly-histidine tagged protein

This procedure was carried out at 4°C. Approximately 2 ml of ProBond Resin (Invitrogen) that contains nitriloacetic acid (NTA), which uses nickel to bind histidine residues, was placed in a 10 ml chromatography column (BioRad) and allowed to settle to a final column volume of 1-2 ml. Using a peristaltic pump, the column was then washed with 10 column volumes of protein purification Buffer A at a flow rate of 1 $ml.min^{-1}$. The diluted cell extract was then loaded onto the column, followed by a wash of at least 20 column volumes of protein purification Buffer C (protein purification Buffer B containing 30 mM imidazole, but without β-mercaptoethanol). The histidine tagged protein was then eluted by protein purification buffer D (protein purification buffer B containing 250 mM imidazole, but without β-mercaptoethanol), and the elutant collected in fractions of approximately 3 ml. If these fractions were not analysed immediately, they were snap-frozen in liquid N_2 and stored at –80°C.

2.8 Analysis of Protein

2.8.1 SDS-polyacrylamide gel electrophoresis

Proteins were analysed on 10% (w/v) discontinuous polyacrylamide gels containing 0.1% (w/v) SDS, with a 4% (w/v) polyacrylamide stacking gel, which were run using a Mini-protean III gel system (BioRad) according to the methods described by (Hames, 1981). The stacking gel (125 mM Tris-HCl (pH6.8), 4% (w/v) bis-Acrylamide solution,

0.1% (w/v) SDS, 0.1% (w/v) Ammonium persulphate, 0.01% (v/v) TEMED) was layered over the separating gel (375 mM Tris-HCl (pH8.8), 10% (w/v) bis-Acrylamide solution, 0.1% (w/v) SDS, 0.1% (w/v) Ammonium persulphate, 0.01% (v/v) TEMED).

Protein samples were prepared by adding an equal volume of sample application buffer (SAB is 125 mM Tris-HCl (pH 6.8), 4% SDS, 20% glycerol, 10% β-mercaptoethanol, 0.002 % bromophenol blue) and incubated at 95°C for 5 min. Gels were run at 200 V for 40 min; stained in 30% (v/v) methanol, 0.01% (w/v) Coomassie blue, 12% (w/v) trichloroacetic acid, 10% (w/v) sulfosalacyclic acid for 30 min; and destained in 7.5% (v/v) acetic acid, 5% (v/v) ethanol. Gels were dried using Gel Drying Film (Promega), according to the manufacturer's instructions.

2.8.2 Concentration of sample

The purified protein samples were concentrated using Centriprep 30, for volumes over 3 ml, and Microcon 30 for smaller volumes (Amicon, according to the manufacturers instructions). The samples are centrifuged over a membrane, removing liquid and small molecules from the protein sample.

2.8.3 Determination of protein concentration

The colour of Bradford's reagent (BioRad) changes on contact with protein (Bradford, 1976). The intensity of the colour is dependent on the amount of protein in the sample. Bovine serum albumin, BSA (New England Biolabs) at was used to construct a standard curve of concentrations each time the assay was used, and the absorbance of the solutions were measured at $A_{595\,nm}$. Reactions were set up in a 1 ml cuvette, containing 0.8 ml Bradford reagent, protein sample of unknown concentration and H_2O to a total

volume of 1 ml. The absorbance of the unknown protein samples was measured, compared to the standard curve and the unknown sample concentration calculated.

2.9 Assaying for galactokinase activity

2.9.1 Stopped enzyme assay for galactokinase

The stopped enzyme assay for galactokinase was based on that used by Schell and Wilson, (1977). Galactokinase converts ATP and galactose to galactose-1-phosphate and ADP. In the stopped enzyme assay for galactokinase, [^{14}C]-galactose is introduced to the reaction and the conversion to [^{14}C]-galactose-1-phosphate measured.

The galactokinase reaction contained 1 mM dithiothreitol, 3.2 mM NaF, 1.6 mM $MgCl_2$, 40 mM Tris-HCl (pH 8.0) and 1 µl [^{14}C]-galactose (ICN) (specific activity 0.2 µCi). Total volume = 150 µl. Galactose, ATP and purified protein were added at concentrations detailed below. The reaction was initiated by addition of the galactokinase enzyme, and the reaction was incubated in a water bath at 30°C. Samples (20 µl) of the reaction mix were removed at time points of 1 min, 5 min, 10 min and 30 min. The samples were spotted onto positively charged Whatman DE81 discs, which were then placed in ice-cold 1% (w/v) galactose solution to quench the reaction. The discs were subsequently washed 3 times for 5 min in 0.1% (w/v) galactose solution to reduce non-specific association by uncharged [^{14}C]-galactose molecules, before oven drying at 75°C for 5 - 10 min. The filters were placed in 10 ml SIGMA-FLUOR scintillation fluid to measure the counts per minute from each sample.

To measure the K_m values for Gal1p with respect to galactose, a set of galactokinase reactions containing varying galactose concentrations was set up. Galactose was included at concentrations of 0.1 - 5 mM, ATP was included at 0.8 mM in each reaction. Four sets of these reactions were set up, and to each set Gal1p was added at a different concentration (2.5, 3.3, 5 and 10 nM). To measure the K_m values for Gal1p with respect to ATP, four sets of galactokinase reactions were set up, each containing a range of ATP concentrations from 0.1 – 5 mM. Galactose was added at a concentration of 2 mM. To each of the four sets a different Gal1p concentration was added: 2.5, 3.3, 5 or 10 nM.

Rates were calculated as follows: The counts per minute from an unwashed DE81 filter spotted with 1 µl [^{14}C]-galactose was measured to assess the maximum counts per minute that would be obtained if every [^{14}C]-galatose molecule was converted to [^{14}C]-galactose-1-phosphate. The counts per minute from each galactokinase reaction was then expressed as a percentage of the value obtained from the unwashed filter. This is the percentage of [^{14}C]-galactose that was transformed to [^{14}C]-galactose-1-phosphate in each reaction. Assuming that the enzyme has no preference for [^{14}C]-galactose over unlabelled galactose, this conversion rate was calculated for each reaction. For example, if 1% of the counts per minute from the control [^{14}C]-galactose filter was present on the filter of the reaction that contained 5 mM unlabelled galactose, it was calculated that the concentration of galactose-1-phosphate present on the filter would be 0.05 mM. The concentration of galactose-1-phosphate in each reaction was calculated in this way. These values were plotted for each enzyme concentration set of reactions, and the rate of production of galactose-1-phosphate calculated. Using these rates, a Lineweaver Burke graph was plotted and the K_m is equal to the negative reciprocal of the intercept of the x-axis, V_{max} is equal to reciprocal of the intercept of the y-axis. The Woolf, or

Hanes plot was also used, the K_m is equal to the intercept with the x-axis, and the intercept with the y-axis is equal to K_m/V_{max} (Cornish-Bowden, 1995). The graphs were drawn and analysed using Microsoft Excel.

2.9.2 Enzyme linked assay of galactokinase activity

The enzyme-linked assay for galactokinase is adapted from Ali *et al* (1993) who used it to measure the ATPase activity of DNA gyrase, and is based on the assay by Norby, JG, (1971). This assay was carried out in microtitre plates in a microtitre plate reader (Ascent). All assays were carried out at 30°C. The level of ADP production is measured utilising the ability of lactase dehydrogenase/pyruvate kinase (LDH/PK) enzyme preparation, to catalyse the conversion of ADP and NADH to ATP and NAD^+. The drop in NADH levels is measured at $A_{340\,nm}$. The rates of galactokinase activity were calculated using the extinction coefficient for NADH (6220 $M^{-1}cm^{-1}$). The concentration of NADH in each reaction was calculated. The assay relies on the premise that for every molecule of ADP released, one molecule of NADH is converted to NAD^+. The decrease of NADH molecules is in direct relation to the amount of ADP released, and therefore the amount of galactose-1-phosphate released. The concentration of galactose-1-phosphate can therefore be calculated, this can plotted on a graph and the rate of production calculated.

The reaction mix of 150 µl contained 5 mM $MgCl_2$, 400 µM phosphoenol pyruvate, 20 mM HEPES (pH 8.0), 150 mM NaCl, 1mM dithiothreitol; 10% (w/v) glycerol and 1.5 µl PK/LDH enzyme mix (equivalent to 1.1 U PK,1.5 U LDH) (SIGMA). Galactose, ATP and galactokinase were added at concentrations detailed below. To calculate the enzyme constants with respect to galactose the rate of enzyme activity was measured

over a range of galactose concentrations, 0.01 mM to 5 mM. ATP was added to each reaction at a concentration of 5 mM. A different concentration of galactokinase was added to each set of reaction, these ranged from 1.44 nM to 7.2 nM for Gal1p and from 220 nM to 680 nM for Gal3p+SA. To measure the K_m with respect to ATP, the rate of galactokinase activity was measured over a range of ATP concentrations, also 0.01 mM to 5 mM. Galactose was included in each reaction at a concentration of 5 mM. Galactokinase was added at a different concentration to each set of reaction. Again, these ranged from 1.44 nM to 7.2 nM for Gal1p and from 220 nM to 680 nM for Gal3p+SA. The Michaelis constant (K_m) for each substrate can be calculated. Using the rates of galactose-1-phosphate produced in each reaction, a Lineweaver Burke graph was plotted and the K_m is equal to the negative reciprocal of the intercept of the x axis, V_{max} is equal to the reciprocal of the intercept of the y-axis. The Woolf, or Hanes plot was also used to calculate the enzymic constants, the K_m is equal to the intercept with the x-axis, and the intercept with the y-axis is equal to K_m/V_{max} (Cornish-Bowden, 1995).

2.10 Affinity labelling of proteins

2.10.1 Synthesis of 2'3'didealdehyde ATP (oATP)

The structure of oATP is shown in Fig. 2.3. ATP (0.1 M) was dissolved in water and the solution pH adjusted to 7.0. Sodium periodate (0.11 mM) was added to the solution, which was allowed to stand at 4°C in the dark, as described by Easterbrook-Smith *et al.*, (1976). The reaction was stopped by addition of 0.05 mM ethylenediol and the reaction mix loaded onto a Sephadex G10 column. The column was then eluted with distilled water, and the leading half of the nucleotide peak was analysed by TLC (method

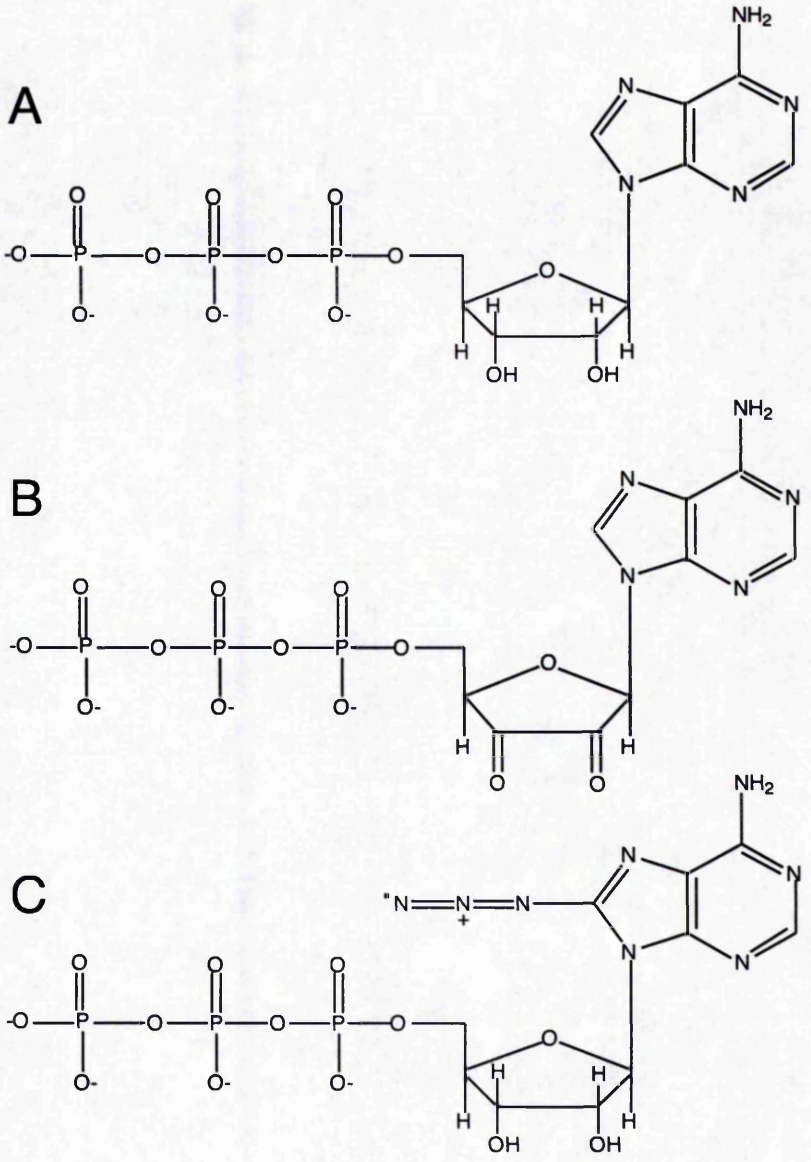


Figure 2.3 The structure of ATP and ATP analogues used in this study

A : ATP

B : oATP

C: 8-azido-ATP

described below), and stored at -80°C. [^{32}P]-oATP was synthesised in the same way from [^{32}P]-ATP (ICN).

2.10.2 Analysis of synthesised oATP

The concentration of oATP was determined by measuring the absorbance at 258 nm, using an extinction coefficient of 14900 $cm^{-1}M^{-1}$ (Easterbrook-Smith *et al*, 1976). A starting concentration of 0.1 M ATP resulted in a solution of 2 mM oATP after dilution by reagents and chromatography. The purity of oATP was checked by thin layer chromatography on polyethyleneimine sheets (Schleicher & Scheull), using 0.8 M NH_4HCO_3 as the developing solution. A sample from each chromatography fraction (5 µl) was spotted onto the PEE sheets approximately 5 cm above the level of the developing solution. The PEE sheet was then placed vertically in 3 cm depth of developing solution, and left until the solution had migrated to approximately 5 cm from the top of the PEE sheet. Ultraviolet light was used to locate the position of the nucleotide. oATP will not migrate up the chromatogram, while ATP will migrate approximately one quarter of the distance of the developing solution, displaying an R_F value of 0.31 (Easterbrook-Smith *et al*, 1976).

2.10.3 Affinity labelling by oATP

When oATP is exposed to strong reducing agents it can form covalent bonds with lysine residues in ATP binding sites (Easterbrook-Smith *et al.*, 1976). The reaction mix (100 µl) contained 1.25 M HEPES (pH 8.0), 1 mM EDTA, 7.5 mM $MgCl_2$, 0.5 µl [^{32}P]-oATP (estimated activity 0.05 µCi), 0.25 µM protein and, where indicated, 1 mM galactose. The reactions were incubated at 30°C for 5 min before addition of 10 µM

$NaBO_4$ and further incubation at 30°C for 35 min. SAB (100 µl) was added, samples were heated to 90°C for 5 min before SDS-PAGE was carried out on 50 µl of each reaction. Gels were exposed to phosphorimager screen for up to 48 hr before visualisation on phosphorimager (BioRad) and analysed using Quantity One software (BioRad).

2.10.4 Affinity labelling of proteins with [^{32}P]-8-azido ATP

On exposure to UV light, azido groups can label lysine residues on proteins (Abraham *et al.*, 1983). The structure of 8-azido-ATP is shown in Fig. 2.3. 8-azido-ATP contains an azido group that can covalently bind to ATP binding sites of proteins *via* active lysine residues. The reaction mix (10-20 µl) contained 20 mM Tris-HCl (pH 7.4), 5 mM $MgCl_2$, 2 µCi (10 – 20 nM) [^{32}P] 8-azido-ATP (ICN) + protein and, if indicated, 3 mM galactose. Gal1p labelling reactions contained 1.5 µM of protein, and TdT was included at approximately the same concentration as measured by SDS-PAGE. Gal4p was included at 1.5 – 2.5 µM. These reactions were incubated on ice for 20 min, followed by 5 min UV irradiation by BLAK-RAY® longwave ultraviolet lamp at a distance of 5 cm. The reactions were then mixed with SAB buffer (Appendix C) and applied to SDS-PAGE. The gels were dried and exposed to phosphorimager screen for up to 48 hr before visualisation by phosphorimager (BioRad) and PC with Quantity One software (BioRad).

2.11 Crosslinking of Proteins

2.11.1 *In vitro* crosslinking of proteins by Bis(sulphosuccinimidyl) suberate

Bis(sulphosuccinimidyl) suberate (BS^3) (Pierce) was used to crosslink Gal4p with either Gal80p or the super-repressor $Gal80^{S-2}p$, a form of Gal80p that contains a single amino acid mutation of Glu^{351}→Lys (Nogi and Fukasawa, 1984). This version of Gal80p can repress the *GAL* genes, and this repression cannot be alleviated by Gal3p. The protocol was based on the manufacturer's instructions. Gal4p (1-93+768-881) (2 μM) was mixed with either Gal80p (1.8 μM) or $Gal80^{S-2}p$ (2 μM) in protein purification Buffer A (Section 2.7.1). The reaction was incubated at 20°C for 15 min. The BS^3 stock solution was dissolved to a concentration of 3 mM, in 5 mM sodium citrate buffer (pH 5.5), and then added to the reaction to a final concentration of 0.3 mM. The reaction was then incubated at 20°C for 30 min. To quench the reaction, Tris-HCl (pH 7.5) was added to a final concentration of 25 mM and the reaction was incubated at 20°C for a further 15 min. The resulting solution was run on a 10% (w/v) SDS polyacrylamide gel. Control reactions that did not contain BS^3 were incubated for the same period, and Tris-HCl (pH 7.5) was added at the appropriate point in the protocol.

2.11.2 Crosslinking by Glutaraldehyde

Glutaraldehyde is a amine-amine crosslinker (Hopwood, 1969). Gal3p (2 μM) was mixed with Gal80p (1.8 μM) in protein purification Buffer A. The reaction also contained 5 mM galactose or 5 mM ATP where indicated. The solution was incubated at 20°C for 30 min. Glutaraldehyde to a final concentration of 0.0025% (v/v) was added, and the reaction was incubated for a further 30 min at 20°C. To quench the

reaction, lysine was added to a final concentration of 0.1 M. The reaction was incubated for 5 min at 20°C, followed by electrophoresis on a 10% (w/v) SDS-PAGE gel.

2.12 Western Blotting

Western blotting protocol was adapted from Harlow and Lane, (1988). Samples to be analysed were separated using SDS-PAGE, described above. They were then transferred onto Trans-Blot® Transfer Medium (BioRad) using the BioRad wet transfer system (according to manufacturer's instructions) with western transfer buffer (48 mM Tris; 39 mM glycine; 0.037 %(w/v) SDS; 20% (v/v) methanol). A voltage of 100 V was applied for 1 hour. The nitrocellulose was then stained for 10 min in Ponceau stain to confirm the presence of marker proteins, and to identify protein bands. The protein size markers were marked in pencil to allow size determination. The membrane was then washed 3 times for 5 min in PBS (pH 7.5). Non-specific binding was blocked by washing the membrane in 10 ml blocking agent (5% (w/v) powdered milk (Marvel)) for 60 min at 20°C. The primary antibody was then added at a 1:1000 dilution to the blocking agent and the membrane incubated overnight at 20°C with agitation. The membrane was then washed 3 times for 5 min in PBS (pH 7.5) before being placed in 10 ml fresh blocking agent plus a 1:5000 dilution of appropriate secondary antibody (SIGMA) in blocking agent. This was incubated for 60 min at 20°C. The membrane was then washed again 3 times for 5 min, in PBS (pH 7.5). The membrane was developed using the ECL kit from Amersham (according to manufacturers instructions) 7.35 ml per blot, and visualised with a Fluor-S™ MultiImager (BioRad) or by exposure to photographic film.

2.12.1 Primary Antibodies

Name	Description	Source Organism	Reference
30C12	Gal3p monoclonal	Mouse	Platt *et al.* (2000)
797	Gal80p polyclonal	Rabbit	A. Platt, Manchester University

2.13 RNA methods

2.13.1 Preparation of total RNA for measuring expression

YNB with selective amino acids plus 2% glucose (w/v) (5 ml) was inoculated with the yeast strain to be analysed, and incubated at 30°C for 24 hr with agitation. It was then diluted with 100 ml YNB with selective amino acids plus 2% raffinose and incubated for a further 24 hours at 30°C with aeration. A 5 ml sample was then removed. The remainder of the culture was used to inoculate a 100 ml culture of yeast dropout with selective amino acids plus 2% galactose (w/v) plus 3% ethanol (v/v) to a final OD of 0.2, which was then incubated at 30°C with aeration. Further 5 ml samples were taken at the time points of 1, 2, 4, 8, 12, 24, 48 and 72 hr. After 24 hr fresh dropout media plus 2% galactose (w/v) plus 3% ethanol (v/v) was re-inoculated to an OD of 0.2. After the 5 ml samples were removed, they were centrifuged 1800 *g*, 5 min, and the media removed. The cells were then washed in 1 ml distilled water and centrifuged again. The cell pellets were then frozen and stored at -20°C until a full set of nine expression samples was obtained. The mRNA was then isolated using RNAEasy kit (Qiagen, according to manufacturer's instructions), and the concentration measured by

absorbance at 260 nm and checked by visualisation on an agarose gel stained with ethidium bromide.

2.13.2 RT-PCR

RT-PCR was carried out using Reverse-iT™ one step RT-PCR kit (AB gene), which contains a PCR mastermix and a reverse transcriptase enzyme solution. mRNA (0.5 μg) was mixed with 25 μl Master Mix, 10 μM of each primer,1 μl reverse transcriptase blend and distilled H_2O to a total volume of 50 μl. The samples were then placed in a PCR machine and heated to 47°C for 30 min, followed by heating to 94°C for 20 s. This was followed by 22 cycles of 94°C (20 s), 62°C (30 s), 72°C (60 s). The sample was then heated to 72°C for 5 min as the final elongation step. The reaction was then added to 10 μl of 5 x loading buffer, and 10 μl of the final solution was run on a 3% (w/v) agarose gel to resolve the PCR products. The agarose gels were analysed using Quantity One software (BioRad).

Chapter 3

Insertion of two amino acids confers galactokinase activity on to Gal3p

3.1 Introduction

Galactokinases from different organisms each contain the conserved sequence GLSSSA within homology region III (Fig.1.5). Gal3p, the inducer of the *GAL* genes in *S. cerevisiae* only shares partial sequence identity in this region when compared to the galactokinase (Gal1p). To investigate affect of the observed sequence differences in homology region III, various plasmids expressing chimeric genes were constructed by A. Platt, University of Manchester. *GAL1* sequences were inserted into *GAL3* in place of the original DNA sequence in the context of a yeast expression vector (Platt *et al.*, 2000). The insertions were centred on position 164 of Gal3p. The sequence insertions ranged from coding for 80 amino acids, to coding for just two amino acids, a serine and an alanine (Fig. 3.1). Gal3p contains the conserved homology region III without these two amino acids. Mutants were also made that inserted either one or two alanine residues or a serine residue directly after amino acid 164 of Gal3p. The chimeric and mutant sequences were inserted into the plasmid pJP139 under the control of either the *GAL1* promoter or the *GAL3* promoter. Many of the genes were inserted into the plasmid under control of both the *GAL1* and the *GAL3* promoters, to investigate whether the promoter has any effect on the expression of the *GAL* genes, which would affect the growth of the yeast strains *in vivo*.

The plasmids were transformed into three different yeast strains. These were deleted for *GAL3* alone (Δ*gal3*), *GAL1* alone (Δ*gal1*), or for both *GAL1* and *GAL3* (Δ*gal1*, Δ*gal3*). In a Δ*gal3* strain, cells display the LTA phenotype when grown on galactose (Rotman and Spiegelman, 1953). These colonies do not appear on plates until two to three days after wild type yeast. The LTA phenotype can be overcome by overproduction of Gal1p

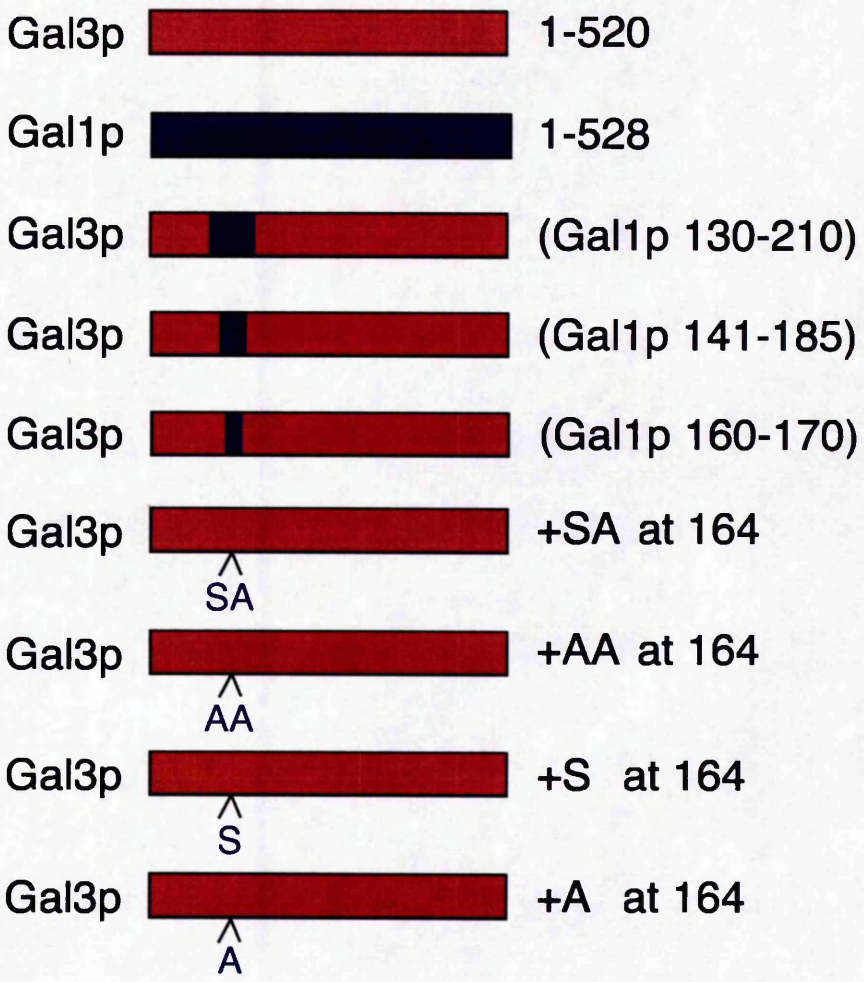


Figure 3.1 Gal1p Gal3p chimeric proteins

This representation shows the site of insertion of Gal1p sequence into Gal3p. Gal3p sequence is represented in red, and the replacement Gal1p sequence in blue, with the residues inserted detailed above. The numbers of the insertion refer to the amino acids replaced in Gal3p.

(Bhat and Hopper, 1992). A Δ*gal1* yeast strain will not grow on media in which galactose is the sole carbon source, as the cell does not have a functional galactokinase and therefore cannot metabolise the sugar.

Previous work has indicated that a Δ*gal3* yeast strain is only partially compensated by addition of a low copy number plasmid encoding *GAL3* (Platt and Reece, unpublished data). This experiment uses high copy number plasmids, therefore encoded proteins are over-expressed in the cell. The use of high copy number plasmids has one further important implication for this experiment. In a Δ*gal3* strain, addition of a high copy number plasmid producing Gal1p will result in the masking of the LTA phenotype. The plasmid-borne copy of *GAL1* will compensate for the deletion of *GAL3* and the yeast growth will not be delayed.

The aim of these experiments was to identify the function of homology region III, and to investigate whether the two amino acids difference between Gal1p and Gal3p in this region has implications for the difference in function between these two proteins.

3.2 Growth of Δ*gal3* yeast containing *GAL3* chimeras

A Δ*gal3* yeast strain was transformed with plasmids containing the chimeric sequences. The yeast cells were then spotted on plates containing either galactose or raffinose, a non-repressing non-inducing carbon source, incubated at 30°C and monitored regularly for growth. The yeast colonies were photographed after twenty hours (Fig. 3.2). The yeast

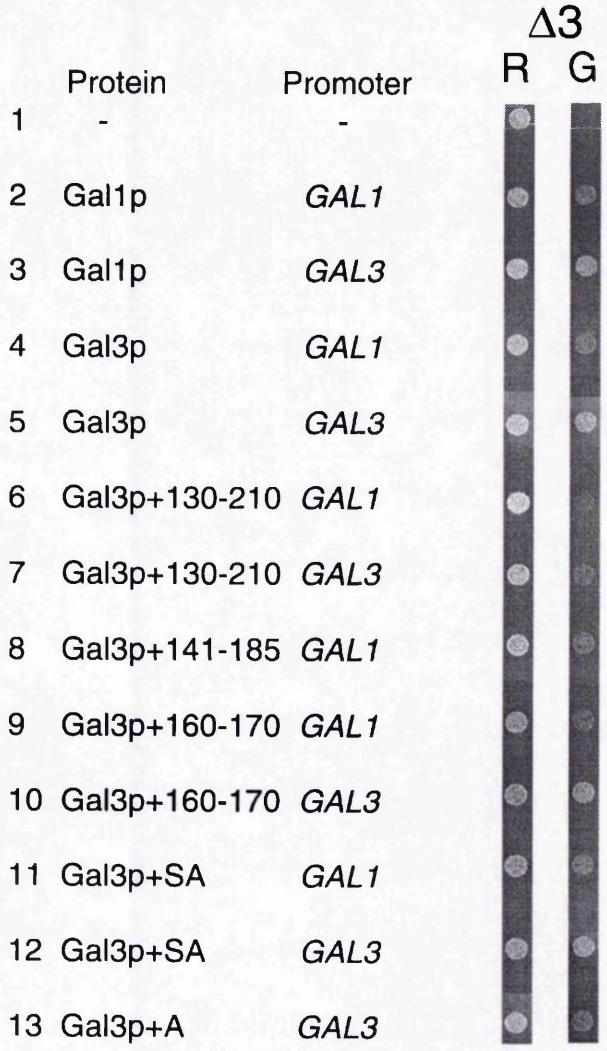


Figure 3.2 The insertion of *GAL1* sequences into *GAL3* does not affect Gal3p activity

Plasmids expressing Gal1p, Gal3p or the indicated chimera were transformed into yeast strains lacking *GAL3* (Δ3). Cells were then plated onto media containing either raffinose (R) or galactose (G) as the sole carbon source. Growth was monitored after 20 hours incubation at 30° C. The chimeras are numbered to indicate the amino acids from Gal3p that have been replaced by the corresponding sequences from Gal1p.

transformed with parental vector, which does not code for a *GAL* gene, grew on raffinose but after twenty hours had not grown on galactose (Fig. 3.2, line 1). The strain did grow after three days (data not shown), displaying the LTA phenotype due to the chromosomal copy of Gal1p. Plasmids coding for Gal3p were able to complement the deletion of *GAL3* (Fig. 3.2, lines 4 and 5). Plasmids containing the *GAL1* sequence were also able to complement the *GAL3* deletion, and the yeast had grown after twenty hours, (Fig. 3.2, lines 2 and 3). This complementation is due to the overproduction of Gal1p because the gene was expressed from a high copy number plasmid. Production of proteins from different promoters, either the *GAL1* or the *GAL3* promoter, appeared to affect the growth of the yeast strain *in vivo*, although the difference could not be measured accurately by this method.

All the chimeric forms of Gal3p shown in Fig. 3.2 were able to complement the deletion of *GAL3*, indicating that these proteins can induce the *GAL* genes *in vivo*, (Fig. 3.2, lines 6 to 13). Plasmids coding for *GAL3* with an insertion of either two alanine residues or one serine residue were also transformed into *Δgal3* yeast, to test if different combinations of amino acids at position 164 would affect the function of the molecule. The mutant proteins coding for Gal3p+AA and Gal3p+S did not allow growth on galactose, indicating that they do not retain the Gal3p function (data not shown). A likely explanation is that these mutations do not allow the protein to fold correctly *in vivo*, or perhaps generally destabilise its structure resulting in defective activity. For this reason, the plasmids coding for these mutant proteins were not transformed into either *Δgal3* or *Δgal1,Δgal3* yeast as an incorrectly folded protein will not give information about the function of homology region III.

3.3 Growth of a Δ*gal1* yeast strain containing *GAL3* insertional mutants

The plasmids containing the inserts shown in Fig. 3.3 were transformed into a Δ*gal1* yeast strain, and the transformants were spotted onto media containing raffinose or galactose and monitored regularly for growth. The plates were photographed after 20 hours (Fig. 3.3). Growth was observed in all transformed strains grown on raffinose. Yeast deleted for *GAL1* and transformed with the parental vector showed no growth on galactose after twenty hours (Fig. 3.3, line 1), and had not grown after five days (data not shown). The presence of endogenous Gal3p may have induced the *GAL* genes, but these cells could not grow because they cannot metabolise galactose due to the lack of a galactokinase. When the cells were transformed with a copy of *GAL1*, growth was observed after twenty hours (Fig. 3.3, lines 2 and 3), the plasmid-based gene complemented for the missing chromosomal copy. In cells transformed with plasmids coding for Gal3p, no growth was observed at twenty hours (Fig. 3.3 lines 4 and 5), or after 5 days (data not shown). Cells transformed with plasmids encoding chimeric Gal3p proteins containing Gal1p sequence showed growth after twenty hours (Fig. 3.3 lines 6 to 12), indicating that these proteins had acquired galactokinase activity from the *GAL1* sequence inserts. Yeast cells transformed with plasmid coding for Gal3p with just two amino acids, a serine and an alanine inserted at position 164 showed growth after twenty hours (Fig 3.3, lines 11 and 12). This indicated that the insertion of just two amino acids into Gal3p confers galactokinase activity onto the protein. Gal3p with an alanine inserted at the same position did not allow growth on galactose (Fig. 3.3, line 13), indicating galactokinase activity was not gained from this insertion.

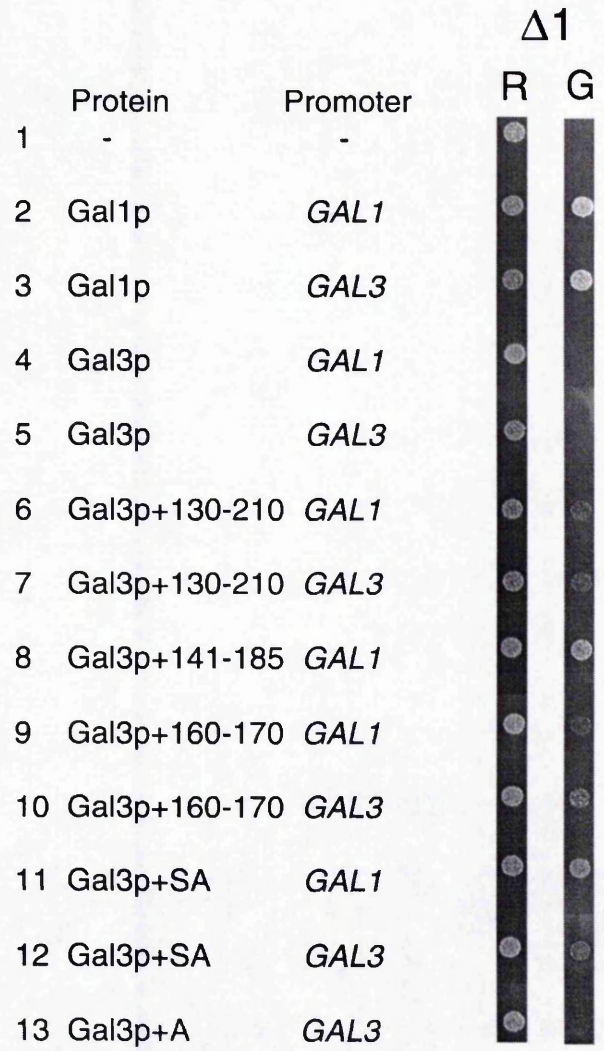


Figure 3.3 The insertion of *GAL1* sequences into *GAL3* confers Gal1p activity onto the protein

Plasmids expressing Gal1p, Gal3p or the indicated chimera were transformed into yeast strains lacking *GAL1* (Δ1). Cells were then plated onto media containing either raffinose (R) or galactose (G) as the sole carbon source. Growth was monitored after 20 hours incubation at 30° C. The chimeras are numbered to indicate the amino acids from Gal3p that have been replaced by the corresponding sequences from Gal1p.

3.4 Growth of a *Δgal1*, *Δgal3* yeast strain containing *GAL3* insertional mutants

The plasmids coding for the chimeric proteins shown in Fig. 3.4 were transformed into a yeast strain deleted for both *GAL1* and *GAL3*, to test the dual function of the chimeric proteins. The transformed strains were spotted on to media containing either galactose or raffinose, and monitored regularly for growth. The resulting cultures were photographed after twenty hours (Fig. 3.4). As expected, growth was observed in cells transformed with the parental vector when they were grown on media containing raffinose, but not on galactose (Fig. 3.4, line 1), and no further growth on galactose was seen after five days (data not shown). Yeast transformed with plasmids encoding Gal1p displayed growth on galactose, because of the bifunctional nature of this protein when over-expressed (Fig. 3.4, lines 2 and 3). A *GAL3* sequence did not permit growth on galactose (Fig. 3.4, lines 4 and 5), this was due to the lack of intrinsic galactokinase activity. When the plasmids containing the chimeras shown in lines 6 to 12 were transformed into the yeast cells, the cells did exhibit growth on galactose after twenty hours, indicating the dual nature of the chimeric proteins. These proteins can induce the *GAL* genes and also contain galactokinase activity. The cells transformed with a plasmid encoding the Gal3p+A protein did not exhibit growth (Fig. 3.4, line 13), confirming that this protein does not have galactokinase activity, and no further growth was observed after five days (data not shown). These experiments clearly demonstrate that the insertion of *GAL1* sequence into *GAL3*, centred on homology region III, convert the inducer Gal3p into a dual function protein, which induces the *GAL* genes and has galactokinase activity. Indeed, just two amino acids inserted into the *GAL3* sequence at position 164 are sufficient to confer galactokinase activity on to the inducer Gal3p.

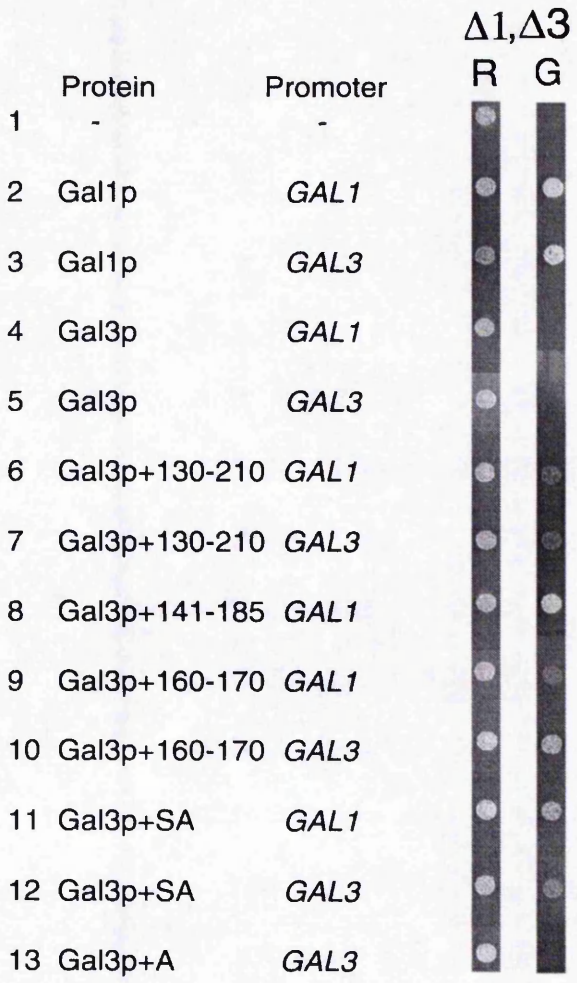


Figure 3.4 The insertion of *GAL1* sequences into *GAL3* converts the protein to a bi-functional molecule

Plasmids expressing Gal1p, Gal3p or the indicated chimera were transformed into yeast strains lacking both*GAL1* and*GAL3* (Δ1, Δ3). Cells were then plated onto media containing either raffinose (R) or galactose (G) as the sole carbon source. Growth was monitored after 20 hours incubation at 30° C. The chimeras are numbered to indicate the amino acids from Gal3p that have been replaced by the corresponding sequences from Gal1p.

3.5 Kinetics of galactokinase activity

Galactokinase enzymes catalyse the reaction between galactose and ATP to produce galactose-1-phosphate and ADP. Gal1p is the galactokinase found in *S. cerevisiae* and the protein is produced in response to the addition of galactose to the growth media. The conversion of galactose to galactose-1-phosphate is the first step of the Leloir pathway, shown in Fig. 1.2, which allows the yeast cell to metabolise galactose when there is no preferred carbon source present. Gal1p is highly similar in amino acid sequence to Gal3p, sharing 90% sequence homology and 70% sequence identity. As shown in Section 3.2, a chimeric mutant of Gal3p, Gal3p+SA, has been produced that can complement a deletion of *GAL*1, indicating that Gal3p gains galactokinase activity from the insertion of these two amino acids. This chimeric protein was also able to complement a deletion of *GAL*3, indicating that the protein has retained its Gal3p function. To test the kinetics of this protein *in vitro*, a suitable assay for galactokinase activity had to be developed. This assay was tested on wild-type Gal1p that contains a polyhistidine tag to aid purification. This experiment would then be repeated on recombinant polyhistidine tagged Gal3p+SA. In this way, the kinetics of the two proteins can be compared and the results will give insights into the mechanism of galactokinase, especially with reference to the differences between the two proteins, Gal1p and Gal3p.

3.5.1 Production of recombinant *GAL* proteins.

Gal1p, Gal3p and Gal3p+SA were expressed in, and purified from *S. cerevisiae*, as described in Materials and Methods. The proteins were isolated using nickel affinity chromatography, and the purification process was assessed using SDS-PAGE. The purification process of Gal1p is shown in Fig. 3.5. Gal1p, Gal3p and Gal3+SA are all

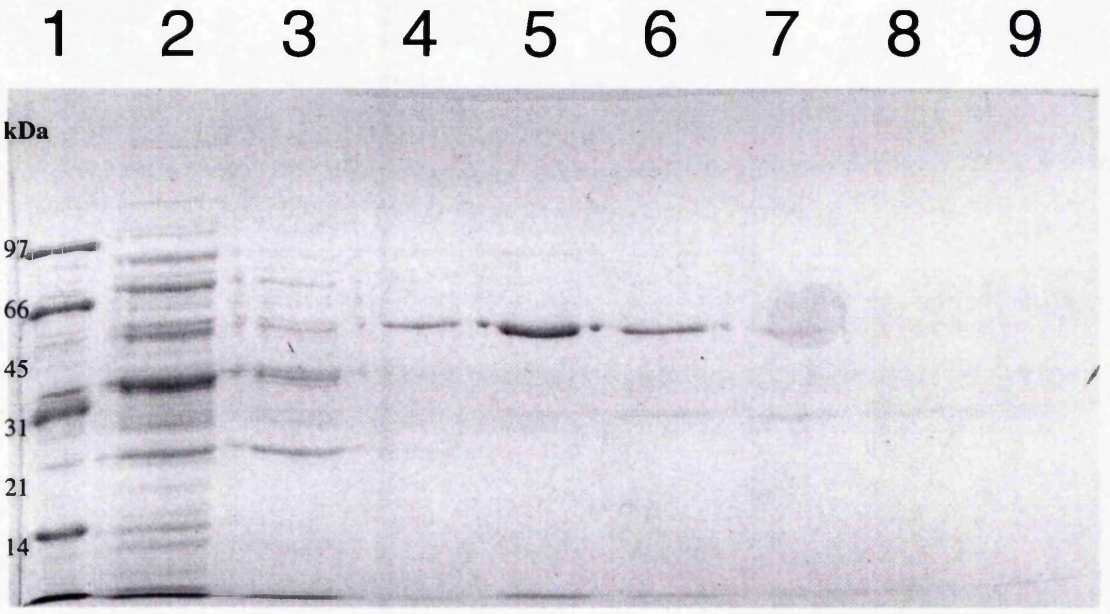


Figure 3.5 SDS-PAGE during the purification of Gal1p

This gel shows the fractions collected from a Gal1p purification from a ProBond resin column, followed by SDS-PAGE on a 10% (v/v) acrylamide gel stained with coomassie blue. Lane 1 contains molecular weight markers, their sizes in kDa are shown, lane 2 the flow through from the column, and lane 3 the low imidazole wash to remove weakly bound proteins. Lanes 4-9 show fractions collected at high imidazole concentration. Gal1p can be seen at its expected size position of 59 kDa.

approximately 59 kDa. The purity of the fraction that contained the majority of desired protein (Fig. 3.5, lane 5) was judged by visually assessing the amount of protein other than that at the desired size position. Fractions that were judged >90% pure were used in galactokinase assays and other *in vitro* manipulations.

3.6 The stopped enzyme assay for galactokinase activity

The stopped enzyme assay is based on the assay used by Schell and Wilson (1977) to measure the galactokinase activity of Gal1p purified from *S. cerevisiae*. A schematic of the assay is shown in Fig. 3.6. The reaction mix was set up, containing galactose, [^{14}C] –galactose and ATP. Galactokinase was added to start the reaction. At time points a sample of the reaction mix was spotted on to positively charged DE81 filter discs (Whatman). Converted [^{14}C]-galactose-1-phosphate should bind to the positively charged filters *via* the negatively charged phosphate group. These filter papers were washed in ice-cold dilute galactose solution to prevent non-specific binding of unconverted [^{14}C]-galactose. The filters were then dried before scintillation counting was performed. Scintillation counts the ^{14}C content of each filter and the amount of converted [^{14}C]-galactose-1-phosphate attached to each filter could be calculated. These values were used to plot the amount of [^{14}C]-galactose-1-phosphate produced over time, from which the rates of galactokinase reaction were calculated.

3.6.1 Data from the stopped enzyme assay for galactokinase activity

The stopped enzyme assay was tested first for its ability to measure galactokinase activity. This involved two reactions that contained different galactose concentrations and ATP at an

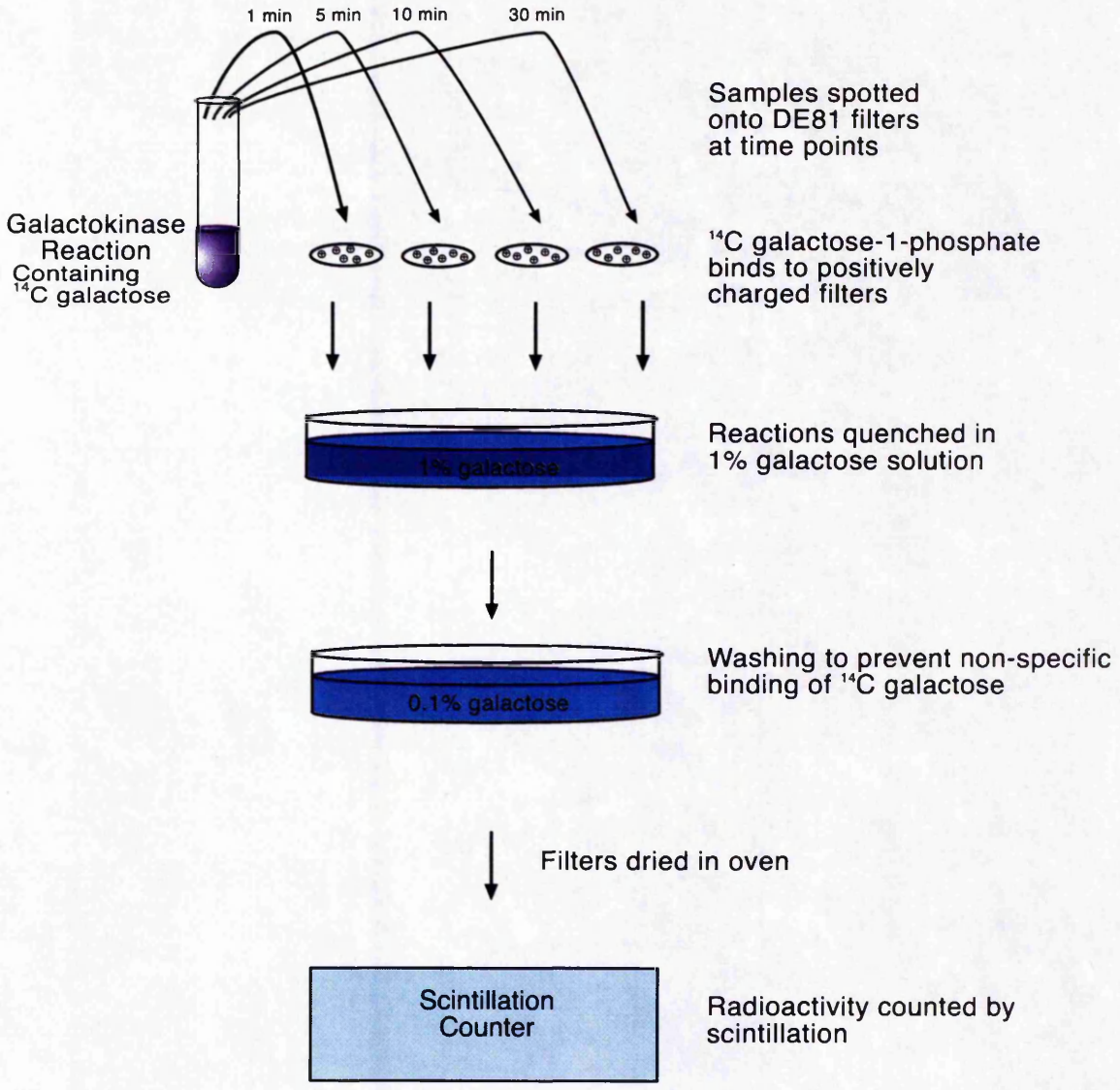


Figure 3.6 Schematic representation of the stopped enzyme assay for galactokinase activity

The stopped enzyme assay for galactokinase was used to assay the activity of Gal1p, Gal3p and Gal3p+SA. The reaction was stopped by spotting on Whatman DE81 filter paper, followed by two washes in galactose solution. Scintillation counting was used to assess the amount of [^{14}C]-galactose-1-phosphate bound to each filter paper disc.

excess concentration. The assay was started by addition of Gal1p to each of the reactions. Control experiments without protein were also carried out, data not shown. After spotting on filter paper and washing in galactose, as described above, the amount of [^{14}C] - galactose included in the product was measured by scintillation counting. The counts per minute measured are shown in Table 3.1 to give an indication of the values obtained. As expected, the counts per minute increased with time, and the reaction appears to be working faster at the higher galactose concentration. This indicated that the assay was measuring galactokinase activity, and it was possible to see variation between enzyme rates.

A set of galactokinase assays, each with a different concentration of galactose, were performed. ATP was included in all reactions at an excess concentration and Gal1p was added to start the reaction. The rates of each reaction were used to plot the graph shown in Fig. 3.6. The shape of the curve of v (the rate) against [S] (substrate concentration) suggests that Gal1p follows Michaelis-Menten kinetics as previously reported by Schell and Wilson (1976). Michaelis and Menten (1913) described the action of enzyme kinetics as an equation, which can be written as

$$v = \frac{V_{max} s}{K_m + s}$$

where v = rate, s = substrate concentration. V_{max} is the limiting rate of the enzyme at saturation values and K_m is the substrate concentration required to give v= 1/2 V_{max}. It is important to note that V_{max} cannot be determined with a finite concentration of substrate, and therefore cannot be calculated using the plot shown in Fig. 3.7. To determine the enzyme constants V_{max} and K_m from experimental data, a number of enzyme rates must be

Time sample taken (min after addition of enzyme)	Scintillation counts per minute		
	No enzyme added 1 mM galactose	+ enzyme 1 mM galactose	+ enzyme 2 mM galactose
1	116	1906	1153
5	120	5915	6171
10	110	8990	9519
30	114	22028	23237

Table 3.1 Counts per minute from stopped galactokinase assay with 20 nM of Gal1p Each column shows the counts per minute obtained from scintillation counting samples taken from galactokinase reactions the times indicated. The second column shows the negative control, a reaction that has all the components without added galactokinase. When the assay includes all the components of the reaction, including galactokinase, the cpm increases rapidly.

.

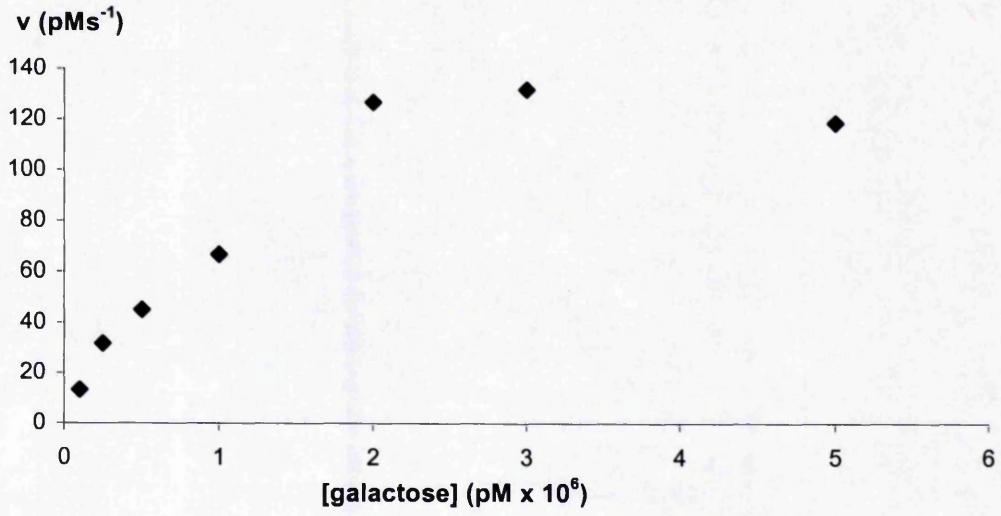


Figure 3.7 Dependence of the rate of enzyme activity on the substrate concentration for Gal1p measured using the stopped enzyme assay. This graph was obtained by plotting the rate (v) against the substrate concentration [galactose] of a galactokinase assays at various substrate concentrations. The initial slope is equal to V_{max} /K_m. This plot confirms that Gal1p obeys Michaelis-Menten kinetics.

measured, varying both substrate and galactokinase concentrations. The resulting data is used to plot the data as straight lines, using two methods detailed below.

Experiments were designed to measure the K_m of Gal1p with respect to either galactose or ATP. The K_m for galactose was investigated first. Four sets of varying galactose concentration reactions were set up. ATP was included in each reaction at an excess concentration. In each of the four sets, a different Gal1p concentration was included. In each set of reactions controls were also included, these contained either no galactose or no enzyme to check that there were no increases over time due to factors other than galactokinase activity. The background of the scintillation counting was also measured, as was the counts per minute obtained from the maximum amount of [^{14}C]-galactose–1-phosphate that it would be possible to generate.

The experimental data was plotted to determine the Michaelis constant of the galactokinase reaction, K_m. This was done in two ways. Lineweaver and Burk took reciprocal values of each side of the Michaelis-Menten equation, and demonstrated that a plot of 1/v against 1/[S] gives a straight line with the slope K_m/V_{max}, the intercept on the y axis $1/V_{max}$, and the intercept on the x axis $-1/K_m$ (Lineweaver and Burk, 1934). This plot is also known as the double reciprocal plot and is used extensively to calculate enzyme constants (Cornish-Bowden, 1995). The rates from each set of the galactokinase reactions were used to plot a Lineweaver-Burk plot (Fig. 3.8). Each line on this graph represented one set of reactions (one Gal1p concentration). The K_m for Gal1p with respect to galactose was calculated using the condition $-1/K_m = x$ when $y = 0$. The values calculated for each enzyme concentration is shown in Table 3.2, the average value was calculated to be 0.86 ± 0.12 mM. Previously published work on untagged Gal1p using the stopped enzyme assay for galactokinase has calculated the K_m of Gal1p

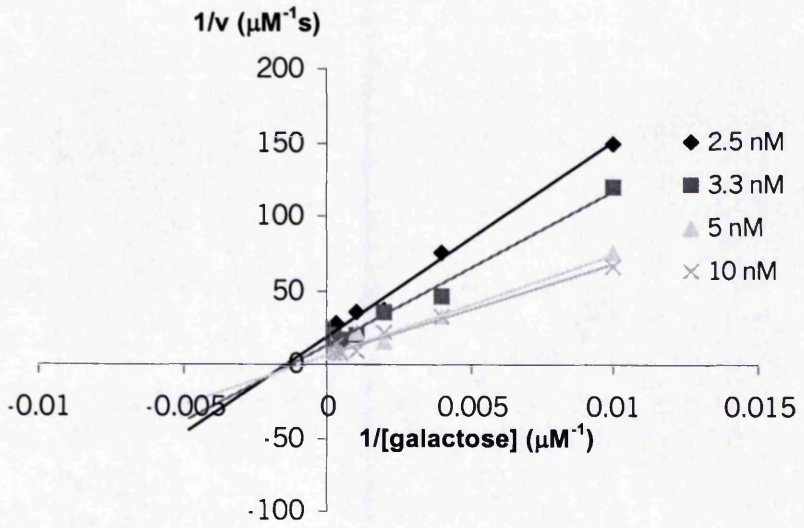


Figure 3.8 Lineweaver-Burk plot of Gal1p activity, varying substrate galactose, measured using the stopped enzyme assay. This Lineweaver-Burk plot was obtained by measuring Gal1p galactokinase rate at varying galactose concentrations. These rates were used to plot 1/rate (1/v) against 1/substrate concentration (1/[galactose]). Each series on the graph is obtained from one Gal1p concentration, as indicated. The intercept with the x-axis is equal to $-1/K_m$, and the intercept on the y-axis is equal to $1/V_{max}$.

with respect to galactose to be 0.6 mM (Schell and Wilson, 1977). In using reciprocal values, the Lineweaver-Burk method can compound small errors, meaning that errors at low substrate concentrations may be emphasised. This has to be taken in to account when assessing the kinetic results (Cornish-Bowden, 1995).

The second method used to calculate the Michaelis constants was the Woolf, or Hanes plot. This gives an even measure of scatter throughout the range of [S] values. Any errors in substrate concentration that have occurred in the assay will occur throughout the graph, as [S] is used on both sides of the graph. The plot is obtained by plotting [S] against [S] / v, concentration of substrate against concentration of substrate divided by the rate of galactokinase activity. The slope of the line is $1/V_{max}$, the y-intercept = K_m/V_{max} and the x-intercept = $-K_m$ (Hanes, 1932). Data obtained in this experiment were used to produce a Woolf plot (Fig. 3.9). The K_m values calculated from each enzyme concentration are shown in Table 3.2. The average K_m for Gal1p with respect to galactose from the Woolf plot is 0.64 ± 0.37 mM. This value shows a greater degree of variation than the value obtained from the Lineweaver-Burk plot, but both values are the same order of magnitude. Both values are also the same order of magnitude to the previously published K_m value for Gal1p with respect to galactose.

V_{max} is the limiting rate of the enzyme at saturation values. V_{max} values from the stopped enzyme assay Gal1p data were also calculated and are included in Table 3.2. The V_{max} values increased as the enzyme concentration increased, with the exception of the 10 nM figures, which decreased slightly. The V_{max} values should have a linear relationship with enzyme concentration, and are used to calculate the catalytic constants of enzyme activity (Cornish-Bowden, 1995).

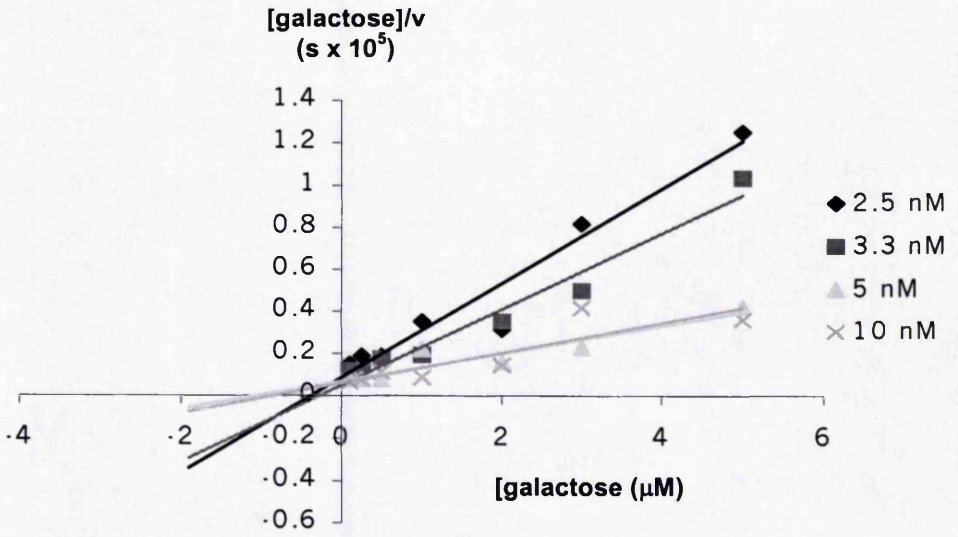


Figure 3.9 Woolfe plot of Gal1p activity, varying substrate galactose, measured using the stopped enzyme assay. This Woolfe, or Hanes plot was obtained by measuring Gal1p galactokinase rate at varying galactose concentrations. These rates are used to plot substrate concentration/rate ([galactose]/v) against substrate concentration ([galactose]). Each series on the graph is obtained from one Gal1p concentration, as indicated. The intercept with the x-axis is equal to $-K_m$, and the intercept with the y-axis is equal to K_m/V_{max}.

Enzyme Concentration (nM)	Lineweaver-Burk		Woolf	
	K_m (mM)	V_{max} (μMs^{-1})	K_m (mM)	V_{max} (μMs^{-1})
2.5	0.85	0.06	0.84	0.005
3.3	1.02	0.08	1.05	0.054
5	0.84	0.15	0.28	0.15
10	0.72	0.14	0.38	0.14
Average value (K_m only)	0.86 ± 0.12		0.64 ± 0.37	

Table 3.2 Catalytic constants of Gal1p with respect to galactose measured using the stopped enzyme assay. This table shows the K_m and V_{max} values for each Gal1p concentration measured. The values labelled Lineweaver-Burk were calculated from the graph shown in Figure 3.7, Woolf or Hanes plot figures were calculated from the graph shown in Figure 3.8. The bottom line shows the average K_m values obtained from each method. V_{max} is dependent on enzyme concentration.

Galactokinase has two substrates, galactose and ATP. After calculating the K_m with respect to galactose, experiments were undertaken to measure the K_m with respect to ATP. Galactokinase reactions were set up, each containing a different ATP concentration. The concentration of galactose-1-phosphate produced in the reaction was measured *via* scintillation counting to follow the rate of the reaction. This set of reactions was repeated four times, each set with a different enzyme concentration. The enzyme rate for each reaction was calculated and these rates were used to plot a Lineweaver-Burk graph (Fig. 3.10). The intercept of each line of best fit crosses on the positive side of the x-axis. This would result in a negative value for K_m. K_m is the substrate concentration required by the enzyme to be working at half V_{max}, and this cannot be a negative value.

This data was also used to produce a Woolf plot (Fig. 3.11). Lines of best fit have not been plotted. It can be seen from this graph that lines of best fit would not form straight lines with positive gradients. This pattern of data is indicative of an enzyme saturated with substrate. The line of best fit would not intercept with the x-axis (data not shown). The experiment had been designed to include the previously observed K_m of 0.15 mM within the range of ATP concentrations tested (Schell and Wilson, 1977), although this value was at the low end of the range of ATP concentrations tested. To test for the presence of an alternative source of ATP in the reaction mix two parallel reactions, one with ATP, one without were performed. When ATP was excluded from the reaction, there was no increase in scintillation counts observed (data not shown), indicating that there was no ATP

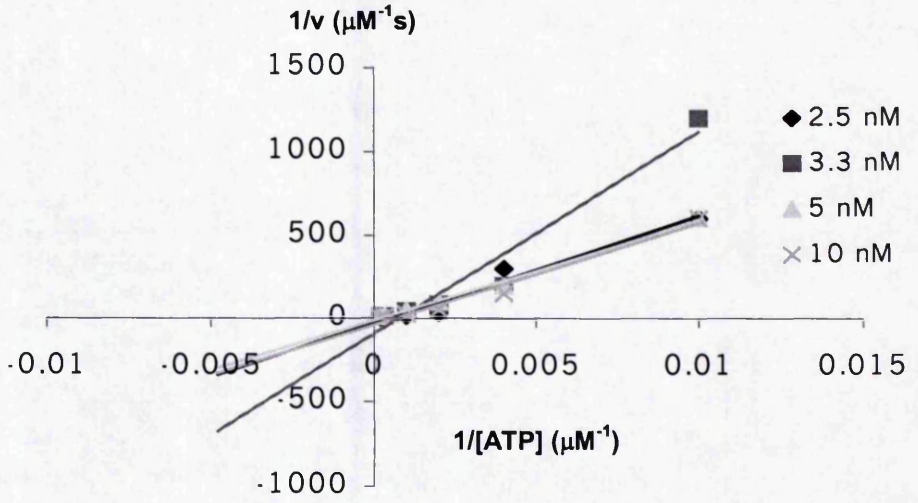


Figure 3.10 Lineweaver-Burk plot of Gal1p activity, varying substrate ATP, measured by the stopped enzyme assay. This Lineweaver-Burk plot was obtained by measuring Gal1p rate at varying ATP concentrations. These rates are used to plot 1/rate (1/v) against 1/substrate concentration (1/[ATP]). Each series on the graph was obtained from one Gal1p concentration, as indicated. The intercept with the x-axis is equal to $-1/K_m$, and the intercept with the y-axis is equal to $1/V_{max}$.

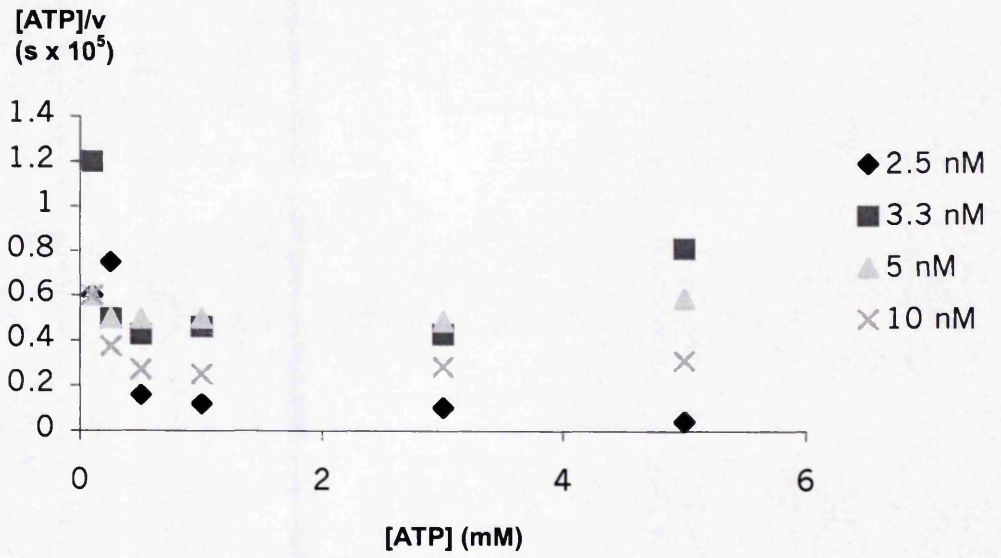


Figure 3.11 Woolfe plot of Gal1p activity, varying substrate ATP, measured using the stopped enzyme assay. This Woolfe, or Hanes plot was obtained by measuring Gal1p rate at varying ATP concentrations. These rates are used to plot substrate concentration/rate ([ATP]/v) against substrate concentration ([ATP]). Lines of best fit have not been drawn in this case. The values do not tend to increase as [ATP] increases, indicating that the galactokinase reaction is saturated with substrate.

contamination of the protein. The experiment to assay the K_m for Gal1p with respect to ATP was repeated, and a similar result was obtained (data not shown). Work on Gal3p+SA, discussed below in Section 3.5.2, had shown that the stopped enzyme assay was not suitable for determination of the Michaelis constants, and another assay was developed to investigate the kinetics of galactokinase activity.

3.6.2 Testing the galactokinase activity of Gal3p+SA with the stopped enzyme assay

The galactokinase activity of Gal3p+SA was measured using the stopped enzyme assay. The reaction included both substrates, galactose and ATP. The galactokinase activity of Gal3p+SA, Gal1p and Gal3p was tested. The result of scintillation counting is shown in Table 3.3. It can be seen from the results that while Gal1p caused a large increase in the measured counts per minute, Gal3p does not. This was as expected, as Gal3p cannot complement a *GAL1* deletion, (Fig. 3.2 and 3.3). Gal3p+SA, the mutant protein, showed a small increase in the scintillation counts per minute, indicated by a rise in [^{14}C] galactose-1-phosphate. This meant that Gal3p+SA was converting [^{14}C]-galactose into [^{14}C]-galactose-1-phosphate, but at a much slower rate than Gal1p.These data confirmed the *in vivo* result showing that Gal3p+SA can complement a *GAL1* deletion, and therefore has galactokinase activity.

The slow rate of reaction presented a problem for further studies on Gal3p+SA using the stopped enzyme assay. As has been shown, the assay tended to show large variation, which could interfere with rate measurement especially at low levels of galactokinase activity. This relatively low sensitivity meant that it would be difficult to measure accurate rates of the reaction catalysed by Gal3p+SA. This, along with the problems encountered measuring

Sample Time (min)	Scintillation counts per minute		
	Gal1p	Gal3p	Gal3p+SA
1	4313	218	335
5	17500	205	385
10	29688	215	811
30	39536	201	834

Table 3.3 Cpm of galactokinase assay on three *GAL* proteins. The table shows the result of the galactokinase assay carried out with three of the *GAL* proteins. The assay contained ^{14}C labelled galactose, galactose (5mM) and ATP (5mM), proteins were included at 10 – 20 nM The cpm give an indication of the rate of conversion of substrates by measuring the amount of [^{14}C]-galactose-1-phosphate produced at the times indicated.

the K_m of Gal1p with respect to ATP, indicated that the stopped enzyme assay was not suitable for this experiment. Attempts to increase the sensitivity of the assay by increasing the stringency of filter washing were unsuccessful (data not shown), so an alternative assay was developed that could take into account the need for sensitivity, especially at lower rates of galactokinase activity.

3.7 The enzyme-linked assay for galactokinase

Enzyme-linked assays have been developed to measure ATPase activity (Norby, 1971). It was used here to assay galactokinase activity. A schematic of the reaction is shown in Fig. 3.12. When galactokinase converts galactose to galactose-1-phosphate, ATP is converted to ADP. It is this ATP conversion that can be measured. The ADP released indirectly causes a decrease in the absorbance of the reaction mix at 340 nm *via* the enzymes pyruvate kinase and lactate dehydrogenase (LDH/PK). These enzymes converted ADP and NADH to ATP and NAD^+. NADH absorbs lights at 340 nm, whereas NAD^+ does not. As the NADH included in the reaction was depleted by conversion to NAD^+, it resulted in a decrease in the absorbance of the reaction mix at 340 nm. The rate of this decrease was dependent on the rate of ADP released by the galactokinase reaction. Therefore the rate of the galactokinase reaction could be measured indirectly by observing the rate of decrease of absorbance of the reaction mix. This assay uses ATP recycling; ATP is released as a product in addition to being an original substrate. The released ATP does not affect the initial rates of galactokinase, and the conversion of ADP to ATP cannot affect the total concentration of ATP in the reaction. The outcome of having ATP recycling is that the assay will continue in a linear manner, until another of the assay components is depleted.

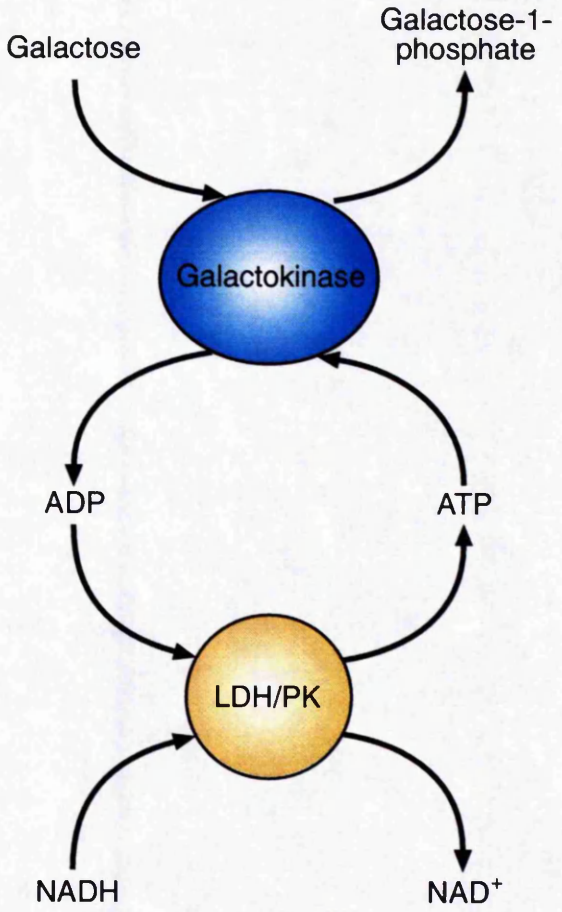


Figure 3.12 Schematic of the enzyme linked assay for galactokinase activity

NADH absorbs light at 340 nm. The decrease in NADH levels is dependent on ADP production by the galactokinase. Therefore the rate of NADH decrease is directly related to the rate of galactokinase activity.

The observed rates were then used to calculate catalytic constants for galactokinase activity using the extinction co-efficient for NADH. The enzyme-linked assay was performed in microtitre dishes, allowing many reactions to be carried out simultaneously.

3.7.1 Kinetics of Gal1p using the enzyme-linked assay

The enzyme-linked assay was first tested for its ability to measure the rate of galactokinase activity of Gal1p. Recombinant Gal1p, containing a poly-histidine tag, from the same purification fraction as that tested by the stopped enzyme assay was used. The stopped enzyme assay demonstrated that the Gal1p is an active galactokinase. The enzyme-linked assay was tested by setting up reactions with a range of galactose concentrations, and a steady concentration of ATP and Gal1p. This experiment tested the ability of the assay to measure galactokinase activity, while checking that the kinetics followed the pattern of the reaction described by the Michaelis-Menten equation. A decrease in absorbance was observed; the rate of decrease was greater at high concentrations of galactose. This decrease did not occur when there was no galactokinase included in the reaction (data not shown). The rates of the reactions were plotted against substrate concentrations (Fig. 3.13). The curve shown is typical of Michaelis-Menten kinetics, with the enzyme rate showing a rapid increase as the substrate concentration increases, followed by a levelling of the enzyme rate as the rate tends towards V_{max}.

Reactions were performed to measure the catalytic constants of Gal1p using the enzyme-linked assay. Each of the four sets of galactose concentration assays included a different Gal1p concentration. The absorbance of each reaction was measured over

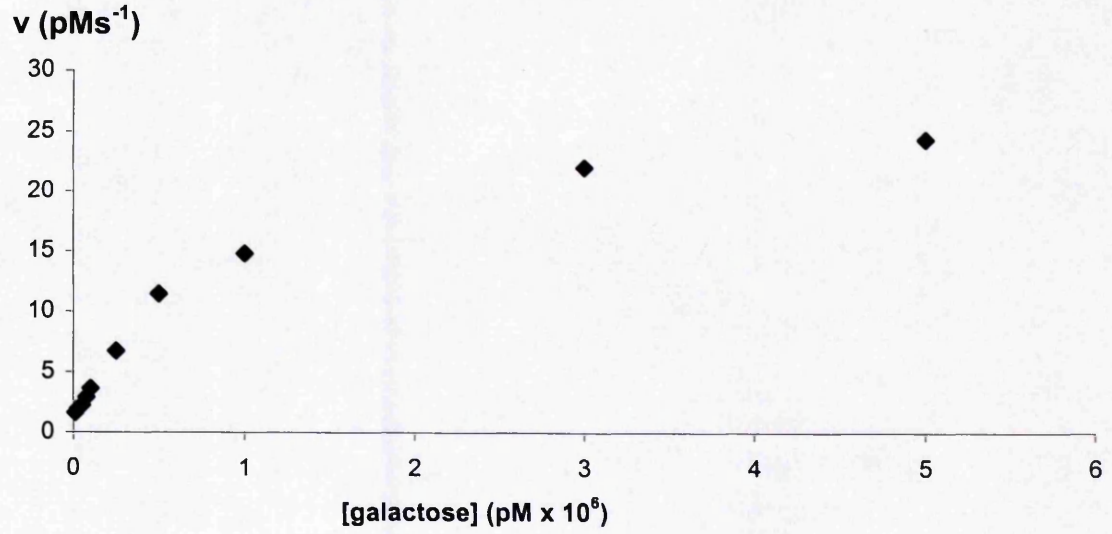


Figure 3.13 Dependence of the rate of galactokinase activity on the substrate concentration for Gal1p measured using the enzyme linked assay. This graph was obtained by plotting the rate (v) against the substrate concentration [galactose] of galactokinase assays with varying substrate concentrations. The initial slope equals V_{max} /K_m.

time, and the rate of the galactokinase reaction calculated. These rates were used to construct Lineweaver-Burk plots, one for each galactokinase concentration (Fig. 3.14). This graph was used to calculate K_m values for each enzyme concentration using the condition $x = -1/K_m$ when $y = 0$, and V_{max} values using the condition $y = 1/V_{max}$ when $x = 0$. The values obtained from each set of assays are shown in Table 3.4. The average K_m value was 0.20 mM ± 0.16 mM; at a galactose concentration of 0.20 mM, Gal1p is working at half its maximum rate (Cornish-Bowden, 1995). The V_{max} values are dependent on the concentration of enzyme in the reaction mix, and it can be seen in Table 3.4 that as the enzyme concentration increases, so do the values of V_{max}.

The data was replotted as a Woolf plot (Fig. 3.15). The K_m values obtained from each set of reactions are also shown in Table 3.4. The average K_m value from this method was calculated to be 0.34 ± 0.19 mM. In both the Lineweaver-Burk method, and the Woolf method, the variation between the K_m values of each enzyme set is large compared to the average value, but all the K_m values obtained are of the same order of magnitude. These compare to a previously published figure of 0.6 mM (Schell and Wilson, 1977). The V_{max} values calculated from each set of reactions are also shown in Table 3.4. With one exception, these values increase as the concentration of enzyme increases.

The K_m for Gal1p with respect to ATP was then investigated. Four sets of varying substrate concentration reactions were set up; each set contained a range of ATP concentrations. Each of the four sets contained a different concentration of Gal1p. The

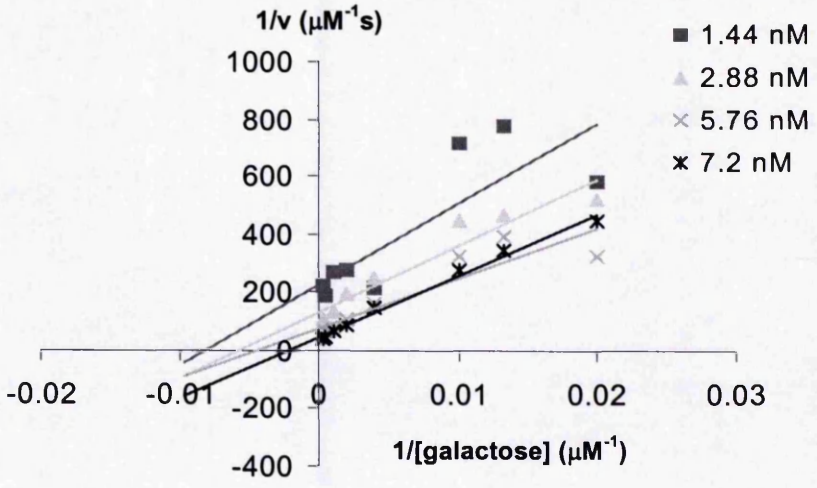


Figure 3.14 Lineweaver-Burk plot of Gal1p activity, varying substrate galactose, measured using the enzyme linked assay. This Lineweaver-Burk plot was obtained by measuring Gal1p rate at varying galactose concentrations. These rates are used to plot 1/rate (1/v) against 1/substrate concentration (1/[galactose]). Each series on the graph is obtained from one Gal1p concentration, as indicated. The intercept with the x-axis is equal to $-1/K_m$, and the intercept with the y-axis is equal to $1/V_{max}$.

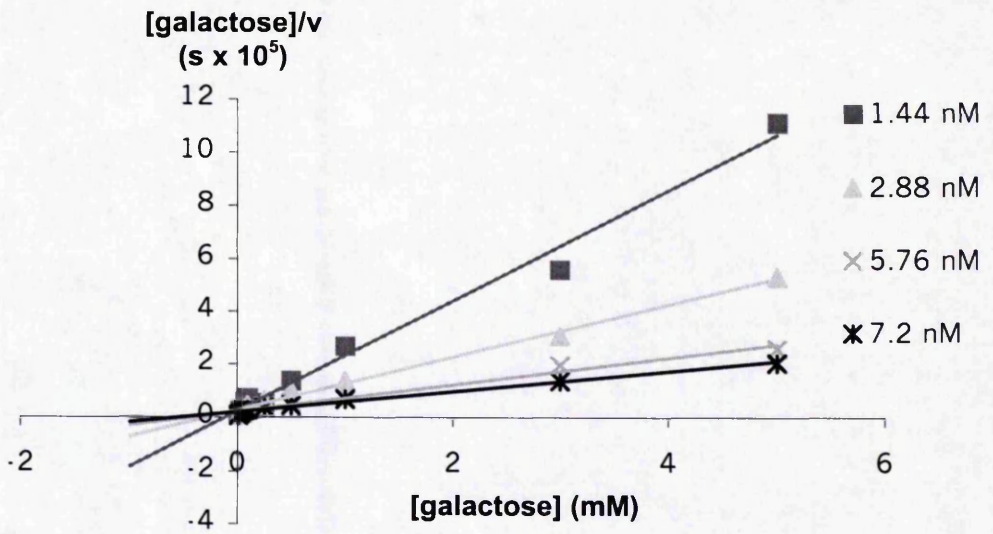


Figure 3.15 Woolfe plot of Gal1p activity, varying substrate galactose, measured using the enzyme linked assay. This Woolfe, or Hanes plot was obtained by measuring Gal1p rate at varying galactose concentrations. These rates are used to plot substrate concentration/rate ([galactose]/v) against substrate concentration ([galactose]). Each series on the graph is obtained from one Gal1p concentration, as indicated. The intercept with the x-axis is equal to $-K_m$, and the intercept with the y-axis is equal to K_m/V_{max}.

Enzyme Concentration (nM)	Lineweaver-Burk		Woolf	
	K_m (mM)	V_{max} (μMs^{-1})	K_m (mM)	V_{max} (μMs^{-1})
1.44	0.06	0.003	0.11	0.003
2.88	0.43	0.008	0.27	0.010
5.76	0.18	0.001	0.42	0.018
7.20	0.12	0.020	0.55	0.025
Average value (K_m only)	0.20 ± 0.16		0.34 ± 0.19	

Table 3.4 Michaelis constant (K_m) and V_{max} values of Gal1p with respect to galactose measured using the enzyme linked assay. This table shows the V_{max} values and K_m values for each Gal1p concentration measured, shown in the first column. The values are calculated from the Lineweaver-Burk plot shown in Figure 3.13, and from the Woolf or Hanes plot shown in Figure 3.14. The bottom line shows the average K_m values obtained from each method with their standard deviation. V_{max} is dependent on enzyme concentration and an average value was not calculated.

rates from each reaction were calculated after following the absorbance of the reaction over time. These rates were used to plot a Lineweaver-Burk graph (Fig. 3.16). Using the condition $-1/K_m = x$ when $y = 0$, K_m values were calculated for each enzyme concentration, and these are shown in Table 3.5. The average value was 0.021 ± 0.003 mM. Again, this data was replotted as a Woolf plot (Fig. 3.17). The K_m values for each enzyme concentration were also calculated by this method are shown in Table 3.5. The average value in this case was 0.028 ± 0.023 mM. V_{max} values were calculated for each enzyme concentration and these are shown in Table 3.5. These values generally increase as the enzyme concentration increases.

In the data obtained from these experiments to investigate the K_m of Gal1p with respect to ATP, the Lineweaver-Burk data shows much less variation between the observed K_m values for each enzyme concentration than the Woolf method. It is clear that the K_m for Gal1p with respect to ATP is lower than that of galactose, indicating that a lower concentration of ATP is required for the enzyme to be functioning at half maximal rate. Both methods of calculation were used to gain confidence in the figures produced. Each method for calculation of Michaelis constants produced different values, but these are of the same magnitude. Compare this with the previous work using the stopped enzyme assay. At first glance, the Lineweaver-Burk plot appeared to be reliable, with the enzyme rates plotted as straight lines. It could be seen immediately from the Woolf plot that the data was flawed (Fig. 3.10 and 3.11). In this case, the data was plotted using both methods and the calculated constants all lay within the same range, therefore confidence was gained that the values obtained are of the correct magnitude.

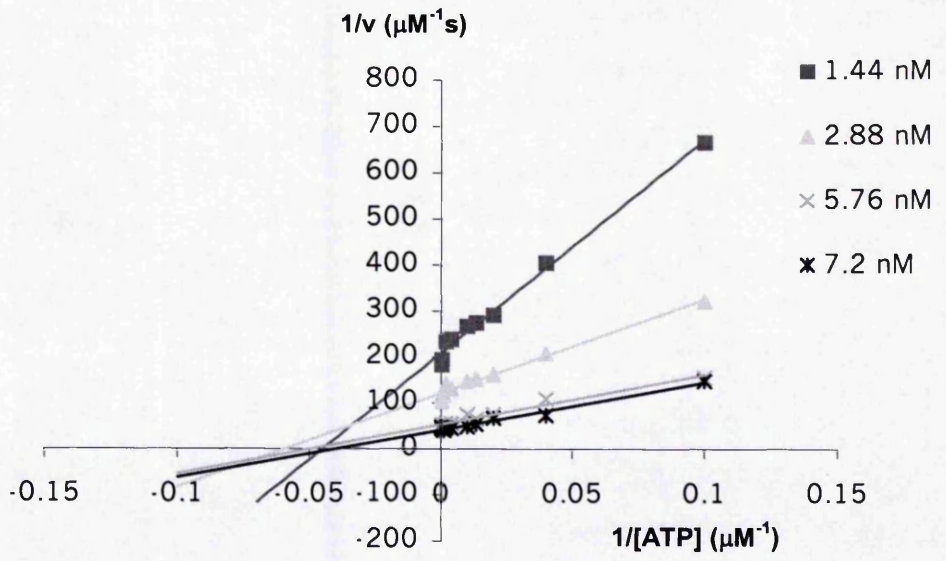


Figure 3.16 Lineweaver-Burk plot of Gal1p activity, varying substrate ATP, measured using the enzyme linked assay. This Lineweaver-Burk plot was obtained by measuring Gal1p rate at varying ATP concentrations. These rates are used to plot 1/rate (1/v) against 1/substrate concentration (1/[ATP]). Each series in the graph is obtained from one Gal1p concentration, as indicated. The intercept with the x-axis is equal to $-1/K_m$, and the intercept with the y-axis is equal to $1/V_{max}$.

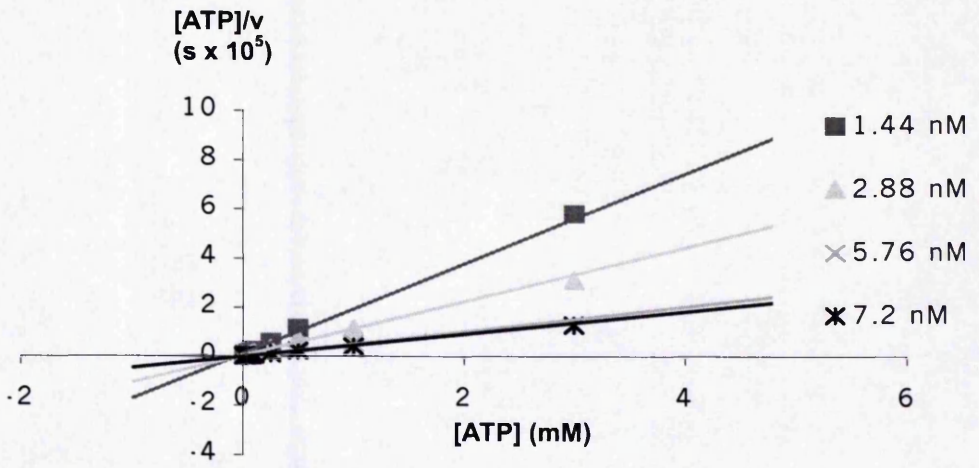


Figure 3.17 Woolfe plot of Gal1p activity, varying substrate ATP, measured using the enzyme linked assay. This Woolfe, or Hanes plot was obtained by measuring Gal1p rate at varying ATP concentrations. These rates are used to plot substrate concentration/rate ([ATP]/v) against substrate concentration ([ATP]). Each series is obtained from one Gal1p concentration, as indicated. The intercept with the x-axis is equal to $-K_m$, and the intercept with the y-axis is equal to K_m/V_{max}.

Enzyme Concentration (nM)	Lineweaver-Burk		Woolf	
	K_m (mM)	V_{max} (μMs^{-1})	K_m (mM)	V_{max} (μMs^{-1})
1.44	0.022	0.005	0.014	0.001
2.88	0.017	0.008	0.06	0.015
5.76	0.02	0.018	0.03	0.066
7.20	0.025	0.023	0.08	0.022
Average value (K_m only)	0.021 ± 0.003		0.028 ± 0.023	

Table 3.5 Michaelis constant (K_m) and V_{max} values of Gal1p with respect to ATP measured using the enzyme linked assay. This table shows the V_{max} and K_m values for each Gal1p concentration measured, shown in the first column. The values were calculated from the Lineweaver-Burk plot shown in Figure 3.15, and from the Woolf or Hanes plot shown in Figure 3.16. The bottom line shows the average K_m values obtained from each method with their standard deviation. V_{max} is dependent on the enzyme concentration.

The values calculated here for K_m for Gal1p with respect to ATP are lower than that observed by Schell and Wilson (1977), who reported a value of 1.5 mM. The1977 study on Gal1p used the stopped enzyme assay. This may have had a bearing on the higher value observed. It is a possibility that the lack of sensitivity seen in this study using the stopped enzyme assay may have led the previous study to a higher figure due to an inability to measure low galactokinase rates effectively.

The catalytic constant (k_{cat}) was calculated using the data obtained from each method. The value gives a measure of how quickly the enzyme is converting the substrates, and is also known as the turnover number. It is calculated by dividing V_{max} by the concentration of enzyme in the reaction. The k_{cat} was obtained for each set of reactions, varying both galactose and ATP and using the data from both the Lineweaver-Burk method and the Woolf method. The average value was calculated to be $3.1 \pm 0.77\ s^{-1}$, meaning that one mole of enzyme produced, on average, 3.1 moles of product every second.

3.7.2 Kinetics of Gal3p+SA using the enzyme-linked assay

The enzyme-linked assay was used to quantify the galactokinase activity of the mutant protein Gal3p+SA. The galactokinase activity of this protein was sufficient to complement a deletion of galactokinase *in vivo* when over-expressed, (Fig. 3.4), but to compare the Gal3p+SA galactokinase activity to Gal1p galactokinase activity *in vitro*, it was necessary to determine the catalytic constants of the mutant protein. It was established in Section 3.6.2 that Gal3p does not contain detectable galactokinase activity when measured by the stopped enzyme assay, shown by the lack of increase in the counts per minute (Table 3.3). This result was confirmed using the enzyme-linked

assay. Gal3p does not cause a decrease in absorbance of the reaction solution, and therefore does not have detectable galactokinase activity (data not shown).

To measure the K_m for Gal3p+SA with respect to galactose, three sets of varying substrate concentration reactions were set up, each set including a different Gal3p+SA concentration. The absorbance of these reactions was measured over time, and the results used to calculate the rate of galactokinase activity. These rates were used to plot a Lineweaver-Burk graph (Fig. 3.18). The K_m values for each Gal3p+SA concentration were calculated and are shown in Table 3.6. The average K_m value was 0.203 ± 0.067 mM. This data was also plotted as a Woolf graph (Fig. 3.19). The K_m values calculated for each Gal3p+SA concentration are shown in Table 3.6. The average K_m value obtained for Gal3p+SA with respect to galactose was 0.426 ± 0.19 mM. The V_{max} values were also calculated for each enzyme concentration, and these are detailed in Table 3.6. These values show an expected increase in the V_{max} as the concentration of enzyme was increased.

The calculated K_m values for Gal3p+SA with respect to galactose are consistent with the K_m values obtained for Gal1p with respect to galactose. The Lineweaver-Burk method produces a similar value, albeit with less variation for Gal1p, while the Woolf plot produces higher values, but in each case almost all the separate values lie within the same range. These values are also the same order of magnitude to the K_m observed by Schell and Wilson (1977) for Gal1p with respect to galactose, and to those observed in this study using the stopped enzyme assay. The similarity indicates that Gal3p+SA and Gal1p both require approximately the same concentration of galactose to work at half their maximal rate.

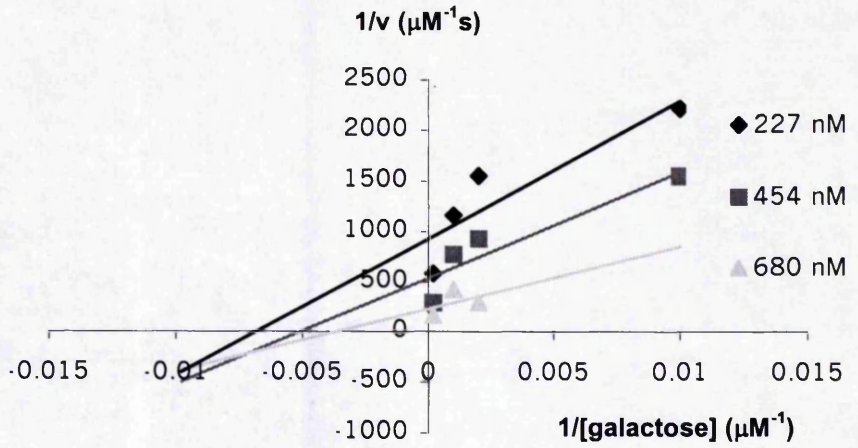


Figure 3.18 Lineweaver-Burk plot of Gal3p+SA galactokinase activity, varying substrate galactose, measured using the enzyme linked assay. This Lineweaver-Burk plot was obtained by measuring Gal3p+SA galactokinase rate at varying galactose concentrations. These rates are used to plot 1/rate (1/v) against 1/substrate concentration (1/[galactose]). Each series on the graph is obtained from one Gal3p+SA concentration, as indicated. The intercept with the x-axis is equal to $-1/K_m$, and the intercept with the y-axis equal to $1/V_{max}$.

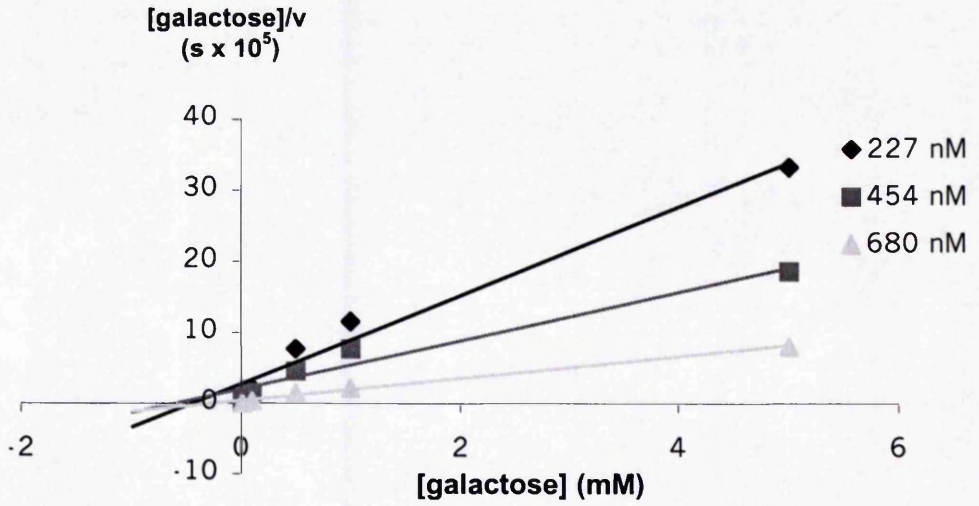


Figure 3.19 Woolfe plot of Gal3p+SA activity, varying substrate galactose, measured using the enzyme linked assay. This Woolfe, or Hanes plot was obtained by measuring Gal3p+SA galactokinase rate at varying galactose concentrations. These rates are used to plot substrate concentration/rate ([galactose]/v) against substrate concentration ([galactose]). Each series in the graph was obtained from one Gal3p+SA concentration, as indicated. The intercept with the x-axis is equal to $-K_m$, and the intercept with the y-axis is equal to K_m/V_{max}.

Enzyme Concentration (nM)	Lineweaver-Burk		Woolf	
	K_m (mM)	V_{max} (μMs^{-1})	K_m (mM)	V_{max} (μMs^{-1})
227	0.146	0.0008	0.449	0.0005
454	0.196	0.0020	0.604	0.0011
680	0.277	0.0030	0.226	0.0060
Average value (K_m only)	0.203 ± 0.067		0.426 ± 0.190	

Table 3.6 Michaelis constant (K_m) and V_{max} of Gal3p+SA with respect to galactose measured using the enzyme linked assay

This table shows the catalytic values for each Gal3p+SA concentration measured, shown in the first column. The values were calculated from the Lineweaver-Burk plot shown in Figure 3.17, and from the Woolf or Hanes plot shown in Figure 3.18. The bottom line shows the average K_m values obtained from each method with their standard deviation. V_{max} is dependent on enzyme concentration, therefore average values are not relevant.

The K_m for Gal3p+SA with respect to ATP was measured in a similar way to that of galactose. Three sets of varying substrate concentration reactions were set up, this time the ATP concentration was varied. Each set contained a different concentration of Gal3p+SA.The absorbance was measured over time, and rates were calculated. This data was used to plot a Lineweaver-Burk graph (Fig. 3.20) and a corresponding Woolf graph (Fig. 3.21). The K_m values were obtained for each enzyme concentration, and these are shown in Table 3.7. The average figures were then calculated. By the Lineweaver-Burk method, the K_m for Gal3p+SA with respect to ATP is 0.406 ± 0.31 mM. For the Woolf method, it is 0.175 ± 0.025 mM. It is noticeable that these values vary from those obtained for the corresponding experiment on Gal1p. Gal1p displayed average K_m values with respect to ATP of 0.021 mM (Lineweaver-Burk method), and 0.028 mM (Woolf method). The K_m of Gal3p+SA was observed to be a much higher average value, the Lineweaver-Burk average K_m is almost forty times as high, but with a large degree of variation between different enzyme concentrations. The range in which all the values appear is significantly higher than those observed for Gal1p, indicating that Gal3p+SA requires a higher concentration of ATP for the enzyme activity to be half maximum rate. The V_{max} values for each enzyme concentration were also calculated and these are shown in Table 3.7. Both sets of V_{max} values increase as the concentration of enzyme increases, although the values for 567 nM and 680 nM are the same, calculated by both the Lineweaver-Burk and Woolf methods. This may indicate that the enzyme is becoming saturated.

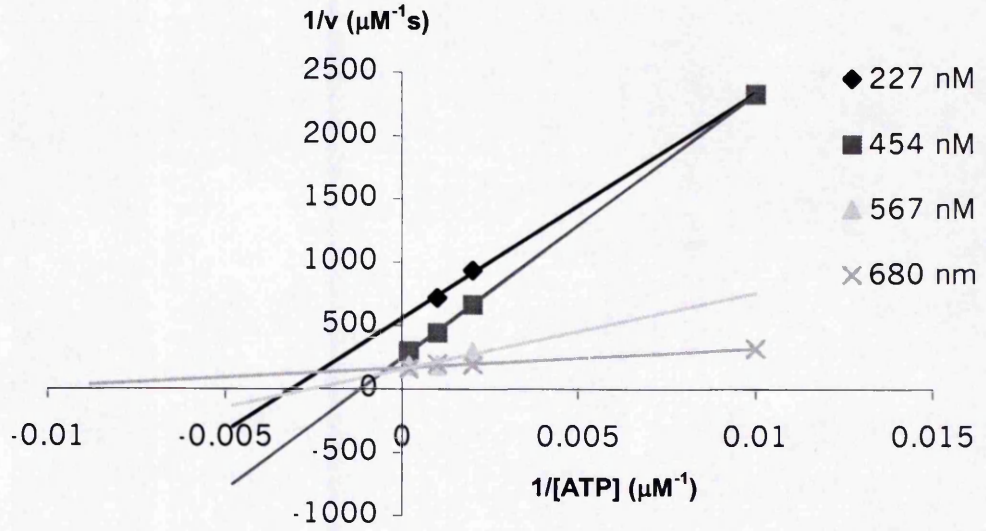


Figure 3.20 Lineweaver-Burk plot of Gal3p+SA activity, varying substrate ATP, measured using the enzyme linked assay. This Lineweaver-Burk plot was obtained by measuring Gal3p+SA galctokinase rate at varying ATP concentrations. These rates are used to plot 1/rate (1/v) against 1/substrate concentration (1/[ATP]). Each series on the graph was obtained from one Gal3p+SA concentration, as indicated. The intercept with the x-axis is equal to $-1/K_m$, and the intercept with the y-axis equal to $1/V_{max}$.

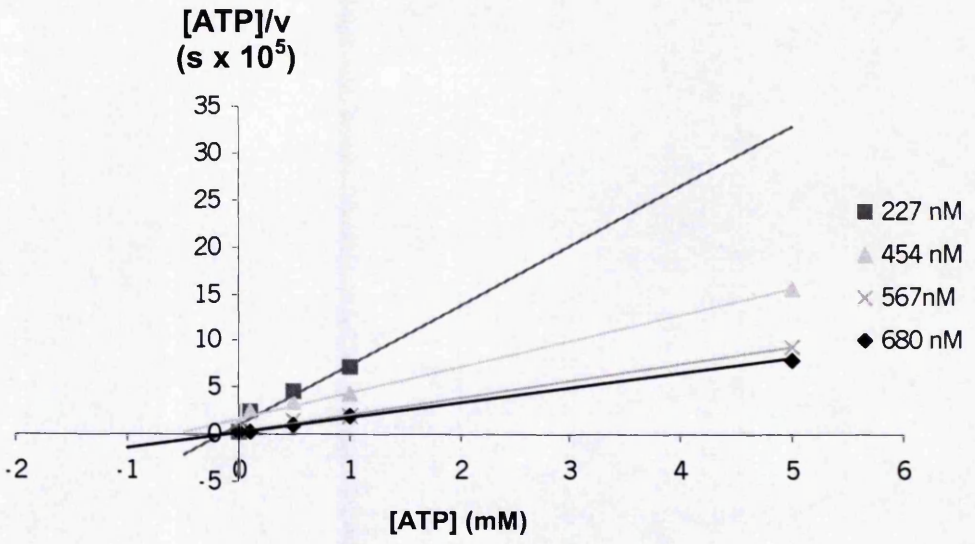


Figure 3.21 Woolfe plot of Gal3p+SA activity, varying substrate ATP, measured using the enzyme linked assay. This Woolfe, or Hanes plot was obtained by measuring Gal3p+SA galctokinase rate at varying ATP concentrations. These rates are used to plot substrate concentration/rate ([ATP]/v) against substrate concentration ([ATP]). Each series on the graph was obtained from one Gal3p+SA concentration, as indicated. The intercept with the x-axis is equal to $-K_m$, and the intercept with the y-axis is equal to K_m/V_{max}.

Enzyme Concentration (nM)	Lineweaver-Burk		Woolf	
	K_m (mM)	V_{max} (μMs^{-1})	K_m (mM)	V_{max} (μMs^{-1})
227	0.318	0.002	0.169	0.0006
454	0.83	0.003	0.196	0.001
567	0.378	0.005	0.197	0.005
680	0.088	0.005	0.144	0.005
Average value (K_m only)	0.406 ± 0.31		0.175 ± 0.025	

Table 3.7 Michaelis constant (K_m) and V_{max} of Gal3p+SA with respect to ATP measured using the enzyme linked assay. This table shows the catalytic values for each Gal3p+SA concentration measured, shown in the first column. The values are calculated from the Lineweaver-Burk plot shown in Figure 3.19, and from the Woolf or Hanes plot shown in Figure 3.20. The bottom line shows the average K_m values obtained from each method with their standard deviation. V_{max} is dependent on the concentration of enzyme.

The turnover number was calculated for Gal3p+SA. K_{cat} values were calculated for each reaction and the average value found to be and found to be $0.0055 \pm 0.0028\ s^{-1}$, indicating that an average of 0.0055 moles of product were produced by one mole of enzyme every second.

The results from the studies on Gal3p+SA confirm that the protein does contain intrinsic galactokinase activity. The results showed some differences between the enzyme constants calculated for Gal1p and those for Gal3p+SA. The k_{cat} values showed that Gal3p+SA is a much less efficient galactokinase than Gal1p; fewer molecules are converted by Gal3p+SA than Gal1p every second. This is perhaps unsurprising, as there may be more factors contributing to the activity of Gal1p that the amino acid sequence of homology region III. The K_m values also show differences between Gal1p and Gal3p+SA. The K_m values with respect to galactose appear to be relatively equal, indicating that both proteins require the same amount of galactose to work at half maximum rate. This suggests that both proteins use galactose at the same rate. The K_m values with respect to ATP display a large difference between the two proteins. The K_m of Gal1p is much lower that that of Gal3p+SA, indicating that Gal3p+SA required more ATP than Gal1p for the enzyme activity to be half maximum rate. This indicates that while the two proteins use galactose at the same rate, Gal3p + SA is unable to use ATP as efficiently as Gal1p. These experiments have demonstrated that the addition of just two amino acids to the sequence of Gal3p gives a gain of function mutation to the molecule. The galactokinase activity gained, while not equal to that of the related protein Gal1p, is sufficient to compensate for a deletion of Gal1p, allowing yeast to grow on galactose.

3.8 Discussion

Gal3p with the insertion of amino acids from Gal1p can still function as the inducer of the *GAL* genes. The insertions were centred on homology region III (Fig. 1.5), as it had been observed that there was a difference between the conserved sequence of galactokinases, including Gal1p, and the sequence of Gal3p, the inducer of the *GAL* genes. The chimeric proteins, when transformed into a Δ*gal1* yeast strain, permitted growth in galactose. This suggests that homology region III, when inserted into Gal3p, confers galactokinase activity onto this protein. Transformation of Gal3p that contains an insertion of just two amino acids, a serine and an alanine, at position 164, into a Δ*gal3* yeast strain, permits the cells to grow on galactose as the sole carbon source. This indicates that these two amino acids are, at least in part, responsible for galactokinase activity in Gal1p. Transformation of Gal3p including an insertion of alanine at position 164 allowed yeast to complement a *GAL3* deletion, but did not permit growth on galactose in a strain deleted for *GAL1*. Mutant forms of Gal3p containing either two alanine residues or two serine residues, when transformed into Δ*gal3* did not permit growth on galactose, indicating that the proteins did not retain their Gal3p activity, perhaps due to incorrect folding *in vivo*. This indicates that the insertion is amino acid specific. Transformation of the Gal3p+SA mutant protein into *Δgal1, Δgal3* yeast allowed the yeast to grow on galactose as the sole carbon source, indicating that this protein had both gained galactokinase activity and retained the inducer function of Gal3p.

Initially, the stopped enzyme assay was used to attempt to quantify the enzymatic constants of Gal1p and Gal3p+SA. This assay had been used previously to calculate the Michaelis constants (K_m) of Gal1p with respect to either substrate, ATP or galactose.

The assay was initially successful, the K_m of Gal1p was determined to be 0.86 ± 0.12 mM by the Lineweaver-Burk method and 0.64 ± 0.37 mM by the Woolf method. As discussed above, the Lineweaver-Burk method has a tendency to emphasise errors in rate at low substrate methods, this is because the graph is plotted by taking reciprocal values of both the rate and the substrate concentration. When the results are assessed, the points plotted at low x values have a higher margin of error. The Woolf graph uses substrate concentration on the both axis. This means that any errors in substrate concentration during the setting up of the assay will be emphasised. Both these methods are used extensively in the literature to work out enzyme constants. In these experiments, both methods were used on each set of data. In each case, the average values obtained for each set of experiments lie within the total range of values produced by both methods. The stopped enzyme assay was used to test the kinetics of Gal1p with respect to ATP. These experiments showed that the assay was saturated with substrate. Concurrent *in vitro* work on Gal3p+SA had showed that the stopped enzyme assay was unsuitable for measuring the kinetics of this protein's galactokinase activity. The stopped enzyme assay was abandoned, and an enzyme assay adapted from (Ali *et al.*, 1993) was developed to measure galactokinase activity.

The enzyme-linked assay was used to measure the K_m, V_{max}, and k_{cat} values of Gal1p and Gal3p+SA. The data was plotted as both Lineweaver-Burk and Woolf graphs to obtain values by both methods. The constants are summarised in Table 3.8. K_m values for Gal1p were published previously at 0.15 mM for ATP and 0.6 mM for galactose (Schell and Wilson, 1977). This study used conventional purification, the protein did not possess a poly-histidine tag. Table 3.8, shows the K_m values for galactose being approximately 0.20 mM ± 0.16 mM (Lineweaver-Burk method) and

	Gal1p		Gal3p+SA	
Method	Lineweaver Burk	Woolf	Lineweaver Burk	Woolf
K_m galactose (mM)	0.20 ± 0.16	0.34 ± 0.19	0.203 ± 0.067	0.426 ± 0.19
K_m ATP (mM)	0.021± 0.003	0.028 ± 0.023	0.406 ± 0.31	0.175 ± 0.025
k_{cat} (s^{-1})	3.1 ± 0.77		0.0055 ± 0.0028	

Table 3.8 Summary of the catalytic constants of galactokinase

The K_m values of Gal1p and Gal3p+SA with respect to galactose and ATP are shown above. The k_{cat} values, also known as the turnover number, are shown for each enzyme.

0.34 ± 0.19 mM (Woolf method) for Gal1p, compared to 0.6 mM for previous results (Schell and Wilson, 1977). The K_m value for ATP (Table 3.8) are 0.021 ± 0.003 mM (Lineweaver-Burk method) and 0.028 ± 0.023 mM (Woolf method) for Gal1p, with the previously published figure being 0.15 mM (Schell and Wilson, 1977). Gal3p+SA shows K_m values of 0.203 ± 0.067 mM (Lineweaver-Burk) and 0.426 ± 0.19 mM (Woolf method)) for galactose, and 0.406 ± 0.31 mM (Lineweaver-Burk method) and 0.175 ± 0.025 mM (Woolf method) for ATP.

Insertion of the serine and alanine may promote galactokinase action by a number of mechanisms:

1. Increased efficiency or rate of ATP binding.
2. As part of the catalytic mechanism.
3. Changing the conformation of the active site.

The binding affinity of Gal3p for ATP is currently unknown, therefore it could not be determined if the affinity of Gal3p+SA for ATP was increased by the addition of the two amino acids. An increased affinity for ATP *in vivo*, caused by the insertion of the two amino acids into Gal3p may explain why the catalytic process take place.

The serine and alanine residues may play an important role in the catalytic mechanism. This change would be more likely due to the serine with its –OH side chain rather than the methyl group of the alanine. These two residues may allow Gal3p+SA to perform the catalytic step after galactose and ATP have bound. The spacing question was addressed by a construction of a Gal3p with an insertion of AA at position 164, but the protein was not expressed, or degraded, so the effect of spacing remains unknown.

Comparison of the two proteins in Table 3.8 shows that Gal3p+SA is deficient as a galactokinase when compared to Gal1p. This would suggest other factors than the two "missing" amino acids also help to maximise the galactokinase activity of Gal1p. It is certainly possible that the lesser binding of ATP plays a part in producing a lower activation, although this may not be the only cause. The reaction does take place in Gal3p+SA, so all the catalytic components are present, but there may be other differences *e.g.* spacing differences in the tertiary structure, or differences in the identity of crucial amino acids, perhaps having the same charge as those in Gal1p, but not identical size. A conformational change may play a part in galactokinase activity. Gal1p may undergo a change from inactive to active complex during catalysis, Gal3p may be unable to undergo this change, but the addition of the SA motif may allow it to happen. A large change is unlikely however, because of the rate of galactose phosphorylation the enzyme displays. It also has to be taken into account that Gal3p has retained inducer ability, (Fig. 3.2, lines 11 and 12). It is not clear from this experiment whether this is at Gal1p (slow) or Gal3p (fast) levels, as the protein is overexpressed within the yeast, which would mask any difference. An *in vitro* transcription assay showed that Gal3p+SA induced transcription at the same efficiency as Gal3p (Platt *et al.*, 2000). This implies that Gal80p binding is not involved with homology region III. Gal80p binding to Gal3p is stabilised by galactose and ATP (Yano and Fukasawa, 1997). Therefore, if these small molecules bind to or interact with homology region III, their action on the protein allows it to bind to Gal80p. This may imply a change in Gal3p from a non-inducing conformation into an inducing one. Alternatively, or in addition, Gal3p binding to Gal80p with galactose and ATP may produce a conformational change in Gal80p, therefore allowing Gal4p to activate transcription.

Chapter 4

Affinity labelling and crosslinking studies on the *GAL* genetic switch *in vitro*

4.1 Introduction

Gal3p and Gal1p share a high degree of sequence identity, and the results in Chapter 3 have pointed towards a small sequence difference being responsible, at least in part, for the contrasting functions of the two proteins. Gal1p contains the conserved homology region III, while Gal3p has an incomplete version of this conserved motif (Fig. 1.5). The addition of just two amino acids, a serine and an alanine, at position 164 converted Gal3p into a galactokinase. A study of the kinetics of the Gal3p +SA chimera showed a similar Michaelis constant (K_m) to Gal1p with respect to galactose, but showed a different K_m value with respect to ATP. Gal3p+SA is an impaired galactokinase when compared to Gal1p, and the differences in K_m values with respect to ATP indicate that the defective component was ATP binding or hydrolysis. Gal3p+SA requires a greater concentration of ATP for the enzyme activity to be half the maximum rate than Gal1p. Homology region III, the site of the insertion of the two amino acids, may form part of the ATP binding site of Gal1p. This sequence is also involved in the active site of the enzyme, as indicated by the change in the k_{cat} value. K_m values do not give a direct measure of binding, but rather how much substrate is required for enzyme activity to be half the maximum rate (1/2 V_{max}). It has not been possible to measure binding constants for ATP or galactose to Gal1p and Gal3p by a variety of methods including equilibrium dialysis and non-equilibrium gel filtration (David Timson, personal communication), perhaps because the binding reaction is in rapid exchange where the ligands are associated, released, and reassociated rapidly.

There are a number of conserved sequences identified between Gal3p, Gal1p and other galactokinases. Some of these sequences may play roles in the interactions involved in the control of *GAL* gene expression, as well as enzyme interactions. The *GAL* genetic

switch comprises of three proteins, Gal4p, Gal80p and Gal3p, although the role of Gal3p can be taken by Gal1p. There is currently some discussion regarding the interplay of these proteins during the alleviation of repression. An *in vitro* transcription assay study showed that Gal3p activated transcription in the presence of galactose and ATP and that Gal3p and Gal80p remained bound to form a tripartite complex (Platt and Reece, 1998). Another study has suggested that Gal3p does not enter the nucleus of the cell, and exerts its effect on Gal80p by preventing entry to or facilitating exit from the nucleus, therefore allowing Gal4p to activate transcription with no interference from Gal80p (Peng and Hopper, 2000). Further investigation into the *in vitro* interactions between the three proteins may resolve some of this ambiguity.

4.2 Investigation into the location of the ATP binding site of Gal1p

A number of ATP binding sites have been identified on different proteins using labelled ATP analogues, which can be bound irreversibly to the ATP binding site. After digestion of the protein, the labelled sequence can be identified and sequenced. Many ATP binding sites contain a lysine residue, and ATP binding site labels can react with this group to covalently bind to the protein sequence. Homology region III does not contain a conserved lysine residue, however both homology regions II and IV contain a conserved basic residue, which is lysine in both Gal1p and Gal3p (Fig. 1.5). Identification of a lysine residue that binds ATP but is outside homology region III would give insight into the three dimensional architecture of the protein. Experiments to identify the ATP binding site were initially carried out on Gal1p, as it was possible to confirm that Gal1p bound to ATP analogues using the enzyme-linked assay. One of the aims of these experiments was to subsequently identify the ATP binding site on Gal3p,

to compare ligand binding between the two proteins and assess the effect on *GAL* gene activation.

4.3 2'3'-didealdehyde ATP (oATP) – an ATP binding site label

The molecule 2'3'-didealdehyde ATP (oATP) has been used to covalently label ATP binding sites. Easterbrook-Smith *et al* (1976) synthesised this molecule to label the Mg-ATP binding site of pyruvate carboxylase. In the presence of sodium borohydride ($NaBH_4$), a strong reducing agent, oATP was shown to bind covalently and specifically with a lysine side chain in the protein. Most proteins that bind ATP also bind magnesium as a co-factor. Easterbrook-Smith *et al*(1976) found that Mg-oATP bound to pyruvate carboxylase. Magnesium was included in the labelling reactions undertaken here, to allow for possible need for the ion as a co-factor. It has subsequently been shown that Gal1p and Gal3p do require Mg-ATP for function (Timson *et al.*, 2002).

oATP is an ATP-binding site label that can be synthesised from unlabelled or radioactive ATP and requires simple conditions for covalent binding. It was therefore a good candidate molecule to be used in a first attempt to label the ATP binding site of Gal1p.

4.3.1 Preparation of oATP

oATP was prepared from ATP by periodate oxidation as described by Easterbrook-Smith *et al* (1976). The concentration was determined by absorbance and the purity was assessed by thin layer chromatography. [^{32}P]-oATP was synthesised by the same method, as detailed in Materials and Methods (Section2.10.1). A thin layer chromatogram of the purification process is shown in Fig. 4.1. [^{32}P]-ATP migrates some

distance up the chromatogram, displaying an R_F value of 0.24 as expected from the work of (Easterbrook-Smith *et al.*, 1976) (Fig 4.1, fraction 1). In the subsequent column fractions, [^{32}P]-oATP was observed that does not migrate, or migrates very slightly along the chromatogram, displaying an R_F value of 0.02. (Fig. 4.1, fractions 2-5) (Easterbrook-Smith *et al*, 1976). There was a small amount of [^{32}P]-ATP present in the solution. Fractions 3 and 4 were pooled to use in labelling reactions. Unlabelled oATP was synthesised from ATP by the same method, and the thin layer chromatogram was visualised using a short wave UV lamp held at a 45° angle, data not shown.

For confirmation that oATP was being produced, rather than the hydrolysation of the phosphate groups of ATP resulting in a decreased R_F value, ATP, ADP and AMP were exposed to TLC. ATP migrated furthest, displaying the expected R_F value of approximately 0.25. ADP displayed an R_F value of 0.1, and AMP displayed an R_F value of 0.06, data not shown. The separation of the molecules was clear, and all had migrated some distance from their spotted position. Therefore, in the experiment shown in Fig 4.1, it was confirmed that oATP was being produced and the change in the R_F value of the nucleotide in the fractions observed was not due to the removal of phosphate groups from ATP. This result could not be shown because of the difficulty in photographing a TLC plate at 45° to the UV light source.

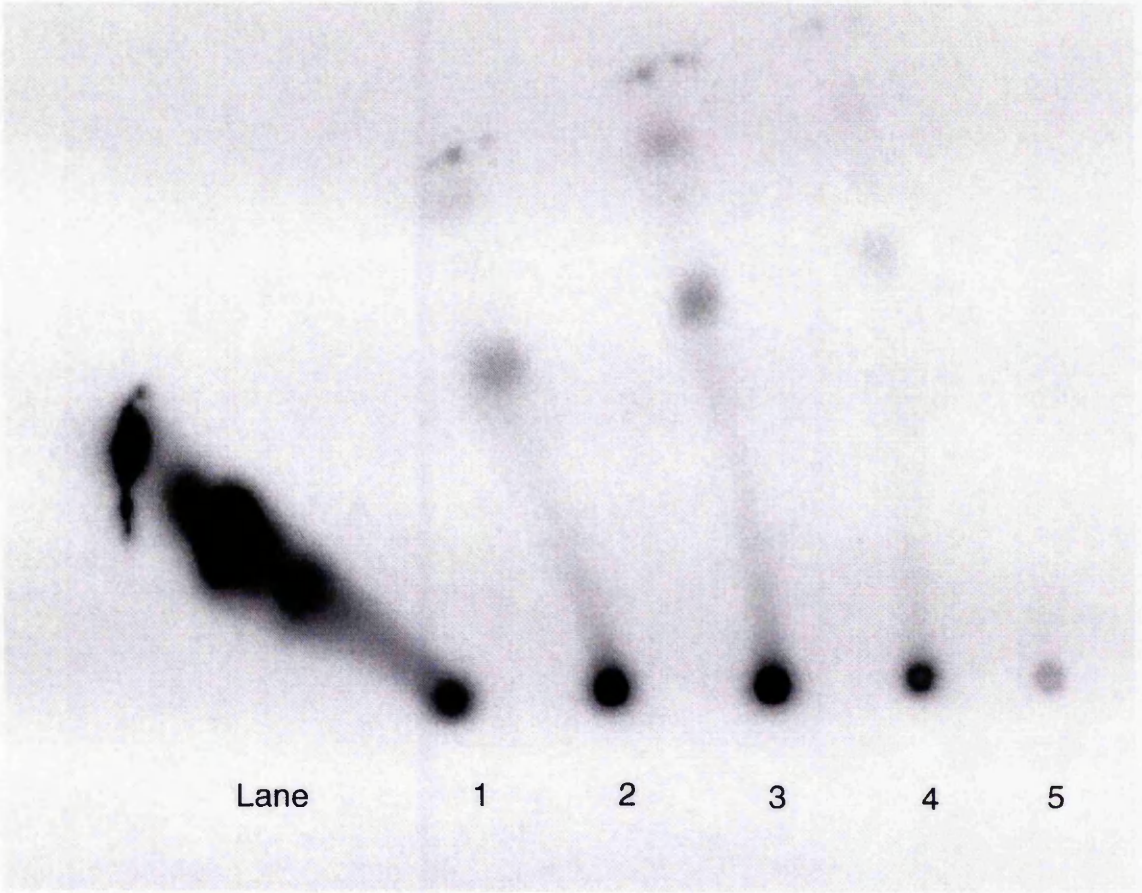


Figure 4.1 Chromatogram showing production of ^{32}P-oATP

After purification through a Sepharose column, collected fractions were exposed to thin layer chromatography, A 20 µl aliquot of each fraction was spotted onto the chromatogram. Lane 1 shows ^{32}P-ATP, and lanes 2-5 show the fractions collected from the column. ATP has an R_F value of 0.24, oATP's R_F value is 0.02. There is a small amount of ^{32}P-ATP present in the fractions shown in lanes 2,3 and 4, but the majority of ^{32}P-ATP has been converted to ^{32}P-oATP.

4.3.2 oATP binds to Gal1p at the ATP binding site

For labelling reactions to be successful, oATP must bind to the ATP binding site of Gal1p. This experiment used the enzyme-linked assay for galactokinase activity. If oATP binds to the ATP binding site it will interfere with ATP binding and affect the rate at which the galactokinase produces galactose-1-phosphate. This could be assessed by measuring the rate of the galactokinase reaction with varying concentrations of oATP added to the enzyme-linked assay. Sets of galactokinase assays were set up, each with a range of ATP concentrations (0.1 mM – 10 mM). To each set of assays an amount of oATP was added (0 mM - 0.2 mM). The rate of each reaction was used to plot a Lineweaver-Burk plot, each line on the graph representing an oATP concentration (Fig 4.2). The data was also used to produce a Woolf plot (Fig. 4.3). The kinetic constants (K_m and V_{max}) were calculated for each oATP concentration (Table 4.1). These figures show two important trends. The V_{max} values for these reactions are relatively similar, with all the values lying between 0.9 $\mu Mmin^{-1}$ and 1.5 $\mu Mmin^{-1}$. The lack of change in V_{max} is indicative of competitive inhibition. These data infer that if sufficient substrate were added, it would be possible to overcome the inhibition and reach maximal activity. The K_m values increased as the concentration of oATP in the reaction was increased. This indicates that oATP was interfering with ATP-protein binding and as the concentration of oATP increases, a greater concentration of ATP is required by the enzyme to work at half the maximum rate. These data suggest that oATP is a competitive inhibitor of Gal1p with respect to ATP, and this inhibition is reversible allowing the enzyme to subsequently produce galactose-1-phosphate.

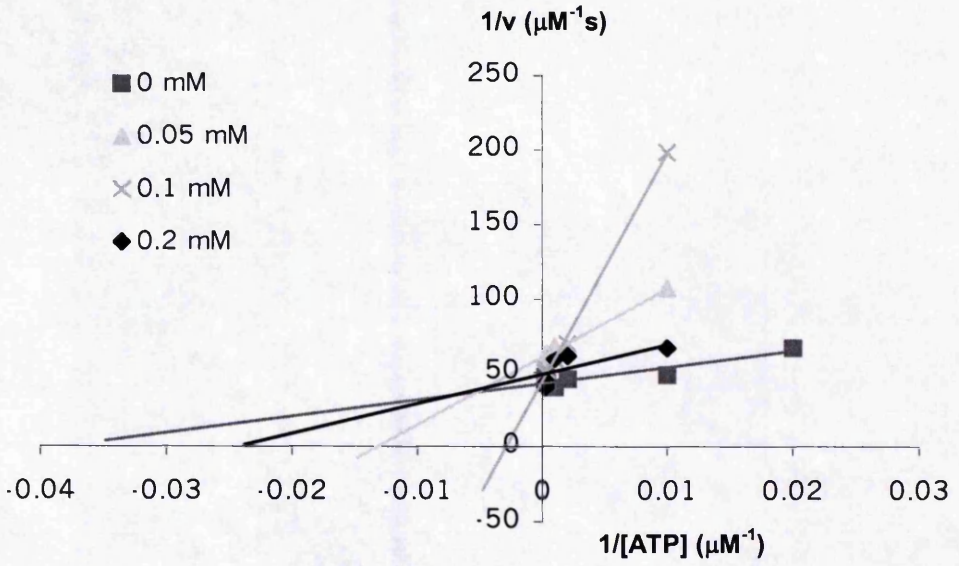


Figure 4.2 Lineweaver-Burk plot of oATP inhibition of Gal1p

oATP was included in a set of galactokinase reactions, each containing 7.2 nM Gal1p, 5 mM galactose and 0.01-5 mM ATP. The rates of these reactions were calculated and used to plot a Lineweaver-Burk graph. Each series in the graph corresponds to a different oATP concentration, as indicated on the left of the graph. The intercept on the x-axis is equal to $-1/K_m$, and the intercept on the y-axis is equal to $1/V_{max}$.

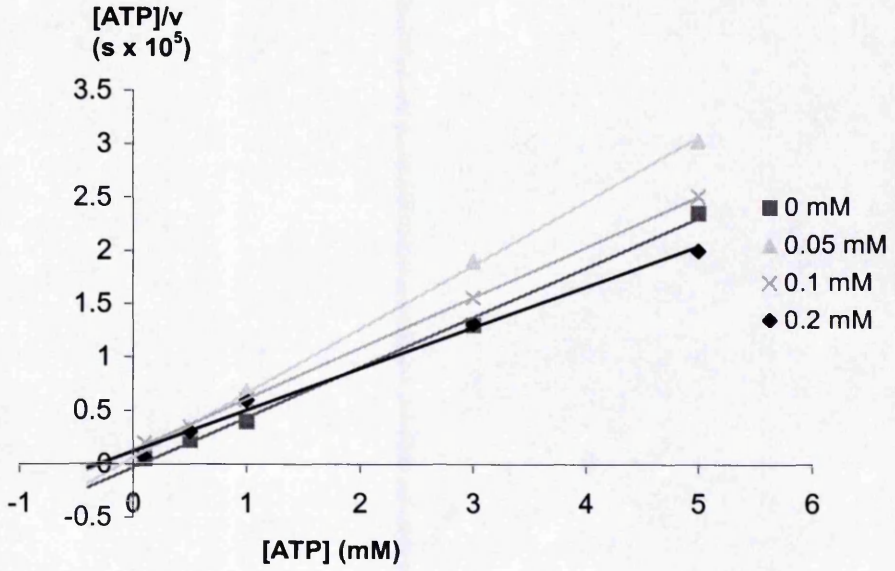


Figure 4.3 Woolfe plot of oATP inhibition of Gal1p

oATP was included in a number of sets of galactokinase reactions, each containing 7.2 nM Gal1p, 5 mM galactose and 0.01-5 mM ATP. The rates of these reactions were used to plot the Woolfe plot shown above. Each series in the graph represents one oATP concentration, as indicated on the right. The intercept of each line with the y-axis is equal to K_m/V_{max}, and the intercept with the x-axis is equal to $-K_m$.

oATP concentration	Lineweaver-Burke		Woolf	
	K_m (mM)	V_{max} (μM/min)	K_m (mM)	V_{max} (μM/min)
0	0.025	1.4	0.11	1.5
0.05	0.077	1.0	0.09	0.9
0.1	0.343	1.3	0.27	1.2
0.2	0.041	1.2	0.30	1.5

Table 4.1 K_m and V_{max} values of Gal1p with respect to ATP inhibited by oATP

The table shows the K_m values and V_{max} values for Gal1p with respect to ATP at a range of oATP concentrations. The enzyme-linked assay for galactokinase was set up with 7.2 nM Gal1p and a range of ATP concentrations. Included in each set of ATP concentration reactions was a concentration of oATP, the first column of the table. The K_m and V_{max} values for each of these concentrations are shown above.

4.3.3 Initial experiments showed no covalent modification of Gal1p

The competitive nature of oATP inhibition of Gal1p enzyme activity suggests that oATP bound to Gal1p at the same site as ATP, and was therefore a good candidate molecule to use in affinity labelling of Gal1p. The labelling of the ATP binding site of Gal1p was undertaken using [^{32}P]-oATP. Pyruvate carboxylase, the enzyme labelled by Easterbrook-Smith *et al*, was included in labelling reactions as a control. It may be that Gal1p requires galactose before adopting a conformation that is able to bind ATP. Therefore, duplicate reactions were set up, one containing galactose, and one without, before addition of $NaBH_4$ to produce a reducing environment allowing oATP to bind covalently to the protein *via* the ε-amino group on lysine residues, often found in ATP binding sites. Gal1p may metabolise a small amount of the galactose present due to contaminating [^{32}P]-ATP from the oATP solution. Galactose was added in excess to minimise any effect this may have on labelling. Labelling reactions were carried out as described in Materials and Methods (Section 2.10.3), and based on the protocol of Easterbrook Smith *et al*, (1976). Initial experiments resulted in no detectable radioactive signal associated with either protein (data not shown). There were a number of important parameters that were adjusted to affect the outcome of the experiment.

The initial period of incubation of the proteins with oATP was tested to assess whether this had an effect on experimental outcome. This could affect the labelling in a number of ways. If the interaction between Gal1p and [^{32}P]-oATP was slow, it will take more time for the two molecules to interact. This is unlikely however, as oATP shows interference in the galactokinase assay rapidly, suggesting a rapid reaction time. The galactokinase reaction did, however, contain a higher concentration of oATP. A range of initial incubation periods was tested, from 5 min to 2 hr. All other conditions

remained as standard. No covalent modification of either protein was seen (data not shown).

Unsuccessful labelling in this experiment may have been due to the relatively low concentration of the component parts of the labelling reaction compared to [^{32}P]-oATP. To counteract this, the protein concentration was increased. If a small percentage of protein interacts with [^{32}P]-oATP at any given time, then increasing the protein concentration may lead to an increase of labelled protein molecules loaded on the gel, and therefore increase the radioactive signal. Increasing the Gal1p concentration to 5 μM, did not affect the outcome of the experiment, no labelling of either protein was observed (data not shown). It was not possible to increase the concentration of [^{32}P]-oATP due to the concentration of [^{32}P]-ATP commercially available.

It was not possible to label Gal1p or pyruvate carboxylase with oATP. Factors that may be responsible for this will be discussed below. Affinity labelling of Gal1p with oATP was abandoned in favour of another ATP analogue.

4.4 Another photoreactive analogue of ATP, 8-azido ATP

8-azido-ATP is a photoaffinity label that was first synthesised and used by Czarnecki *et al* (1979). These co-workers synthesised the compound and the radioactive counterpart. 8-azido-ATP is currently commercially available in both radioactive and non-radioactive forms, and was used in experiments designed to label the ATP binding site of Gal1p. Specific labelling of proteins by [^{32}P]-8-azido-ATP occurs when the ligand is bound to the protein at the ATP binding site and subsequent photoactivation covalently binds the [^{32}P]-8-azido-ATP to the protein. [^{32}P]-8-azido-ATP was used by Abraham *et*

al to label terminal deoxynucleotidytransferase (TdT), identifying the subunit that binds to deoxynucleoside triphosphates (Abraham *et al.*, 1983). TdT is a protein composed of two subunits with molecular weights of 26 kDa and 10 kDa. The 26 kDa subunit was identified as the site of ATP binding. TdT was therefore used in these experiments as a control and the method used for the labelling of Gal1p was based on the protocol of Abraham *et al.* (1983).

4.4.1 8-Azido ATP binds at the ATP binding site of Gal1p

Before labelling reactions were undertaken, it had to be established whether 8-azido-ATP binds to Gal1p at the ATP binding site. This experiment was carried out using the enzyme-linked assay in a similar manner to oATP. The assay included range of ATP concentrations and a standard excess concentration of galactose. To each set of reactions a different concentration of 8-azido-ATP was added. The rate of the galactokinase reaction was then calculated by following the absorbance of the reaction mix at 340 nm over time. These rates were used to construct Lineweaver-Burk and Woolf graphs (Fig. 4.4 and 4.5). These plots allow the kinetic constants to be calculated (Table 4.2). The kinetic data from this experiment showed some variation, but due to the cost of 8-azido-ATP were used to determine the kinetic constants of the reaction. The Lineweaver-Burk method produced values showing some variation; the K_m appeared to decrease at high levels of 8-azido-ATP. The Woolf method, however, produced values typical of a competitive reversible inhibitor binding to the ATP binding site of the enzyme, indicating that 8-azido-ATP was binding at the ATP binding site of the enzyme. 8-azido-ATP was therefore used as an affinity label to attempt to label the ATP binding site of Gal1p.

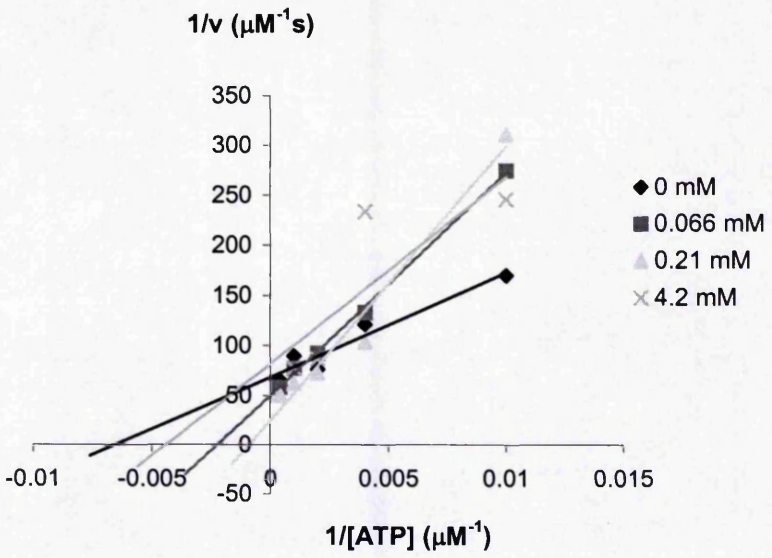


Figure 4.4 Lineweaver-Burk plot of 8-azido-ATP inhibition of Gal1p

8-azido-ATP was included in a set of galactokinase reactions, each containing 7.2 nM Gal1p, 5 mM galactose and 0.01-5 mM ATP. The rates of these reactions were calculated and used to plot a Lineweaver-Burk graph. Each series in the plot corresponds to a different 8-azido-ATP concentration, as indicated on the right of the graph. The intercept on the x-axis is equal to $-1/K_m$, and the intercept on the y-axis is equal to $1/V_{max}$.

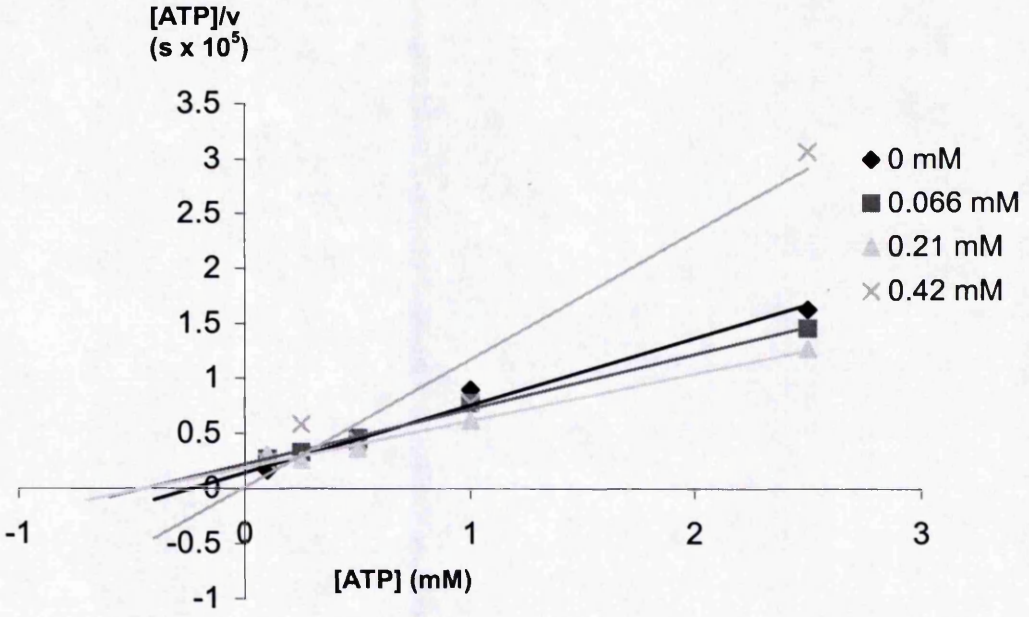


Figure 4.5 Woolfe plot of 8-azido-ATP inhibition of Gal1p

8-azido-ATP was included in a number of sets of galactokinase reactions, each containing 7.2 nM Gal1p, 5 mM galactose and 0.01-5 mM ATP. The rates of these reactions were used to plot the Woolfe plot shown above. Each series in the graph represents one 8-azido-ATP concentration, as indicated on the right. The intercept of each line with the y-axis is equal to K_m/V_{max}, and the intercept with the x-axis is equal to $-K_m$.

8-azido-ATP concentration (mM)	Lineweaver-Burk		Woolf	
	K_m (mM)	V_{max} (μM/min)	K_m (mM)	V_{max} (μM/min)
0	0.174	1.0	0.24	1.0
0.066	0.44	1.2	0.45	1.1
0.21	1.12	2.4	0.45	1.3
0.42	0.31	1.0	0.61	1.2

Table 4.2 K_m and V_{max} values of Gal1p with respect to ATP inhibited by 8-azido-ATP

The table shows the K_m values and V_{max} values for Gal1p with respect to ATP at a range of 8-azido-ATP concentrations. The enzyme-linked assay for galactokinase was set up with 7.2 nM Gal1p and a range of ATP concentrations. Included in each set of ATP concentration reactions was a concentration of 8-azido-ATP, shown in column 1 of the table.

4.4.2 [^{32}P]-8-azido-ATP appears to label TdT and Gal1p

The kinetic data provided evidence that there is a likelihood of 8-azido-ATP binding at the ATP binding site of Gal1p. Therefore labelling reactions were undertaken to identify the site interaction of [^{32}P]-8-azido-ATP with the ATP binding site of Gal1p. Labelling reactions of Gal1p and TdT were set up as described in Materials and Methods (Section 2.10.4). Initially labelling of the two proteins was seen after separation by polyacrylamide gel (Fig. 4.6). There are two reactions on this gel, lane 1 contained TdT, lane 2 Gal1p. One band can be seen in each lane, indicating labelling by the radioactive compound. There is a protein present in the TdT preparation of approximately 60 kDa that has been unexpectedly labelled. The 60 kDa band was visible after Coomassie staining (data not shown). Unfortunately, the 26 kDa subunit was not visible in Fig. 4.6. It was outside the resolution of this polyacrylamide gel. The work done by Abraham *et al.* (1983) showed that it was the 26 kDa subunit of TdT that bound to [^{32}P]-8-azido-ATP. In Fig. 4.6, both proteins, the 60 kDa protein in lane 1 and Gal1p in lane 2 appeared to be labelled equally, as judged by visual assessment of the radioactive signal. There could be a number of explanations for this result; the band seen in Fig.4.6 lane 1 may represent an ATP binding protein, and therefore [^{32}P]-8-azido-ATP was labelling this protein. There was also a possibility that the protein observed in Lane 1 does not bind ATP and non-specific labelling by [^{32}P]-8-azido-ATP was occurring in the reaction. If non-specific binding were occurring, then [^{32}P]-8-azido-ATP would not be a suitable molecule for labelling of Gal1p and further identification of the ATP binding site. To test the possibility of non-specific labelling, the experiment was repeated with a second control protein.

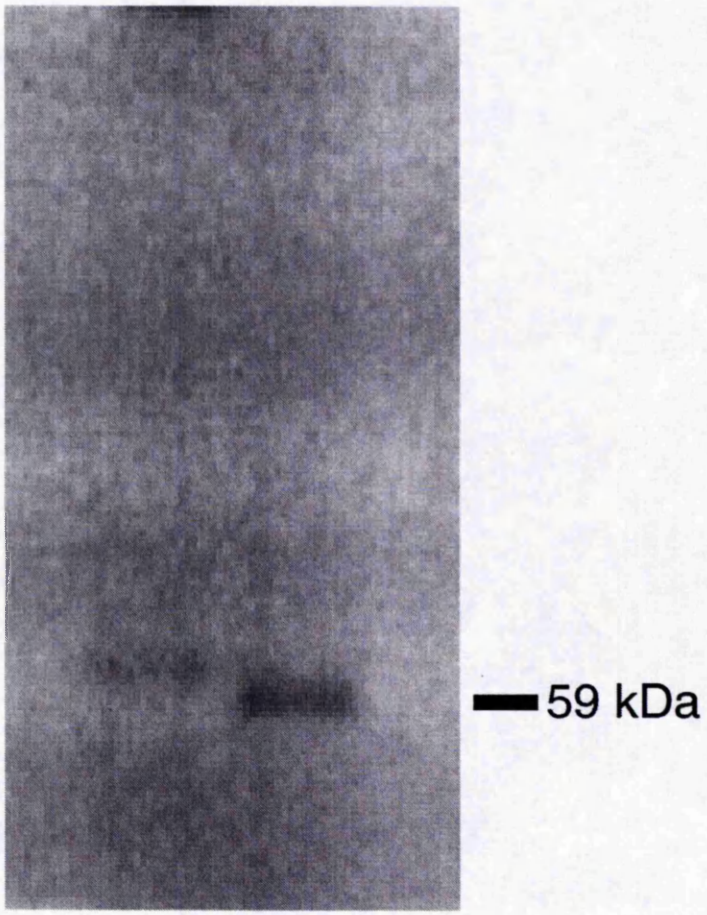


Figure 4.6 ^{32}P-8-azido ATP labels TdT and Gal1p

TdT and Gal1p were labelled with [^{32}P]-8-azido ATP. The proteins were separated by SDS-PAGE and radioactive signals detected by phosphorimage analysis. A band of approximately 60 kDa was labelled which did not correspond to a TdT subunit. Gal1p (59 kDa) was labelled by ^{32}P-8-azido ATP,

4.4.3 [^{32}P]-8-azido-ATP labels TdT, Gal1p and Gal4p (1-93+768-881) in a non-specific manner

To test the possibility of non-specific binding, three labelling reactions were set up. The labelling of TdT and Gal1p was repeated, and a third reaction contained a construct of Gal4p (1-93+768-881), which contains the Gal4p DNA binding and activation domains. This protein is not known to contain an ATP binding site. Three labelling reactions were set up containing Gal1p, TdT and Gal4p (1-93+768-881), as described in the Materials and Methods (Section 2.10.4). [^{32}P]-8-azido-ATP labelled each of these proteins (Fig. 4.7). The 26 kDa subunit of TdT is visible on this gel, but appeared to be no more labelled than the larger 60 kDa band (Fig. 4.7 lane 1). Gal1p was again labelled by [^{32}P]-8-azido-ATP (Fig. 4.7, lane 2). Gal4p (1-93+768-881) was also labelled by [^{32}P]-8-azido-ATP (Fig. 4.7, lane 3). The signal observed in the Gal4p reaction appears strong, although this was probably due to the increased concentration of the protein included in the reaction compared to that of the other reactions. Two major bands can be seen, the lower representing the monomer of Gal4p (1-93+768-881), the higher and larger band is approximately the correct size for the dimer of Gal4p (1-93+768-881). Covalent dimerisation may be due to crosslinks caused by UV irradiation. In subsequent experiments, Gal4p (1-93+768-881) was included at lower concentrations to minimise crosslinking between monomers.

If [^{32}P]-8-azido-ATP was labelling TdT and Gal1p specifically at the ATP binding site then it should not bind to Gal4p (1-93+768-881). The labelling seen in Figure 4.7 may be a background level of labelling, and the correct conditions for specific labelling of the ATP binding site had not been met. If these conditions had been obtained, the 26 kDa subunit of TdT should have been labelled to a much higher degree than the other polypeptides exposed to labelling. The nature of the labelling that occurred under

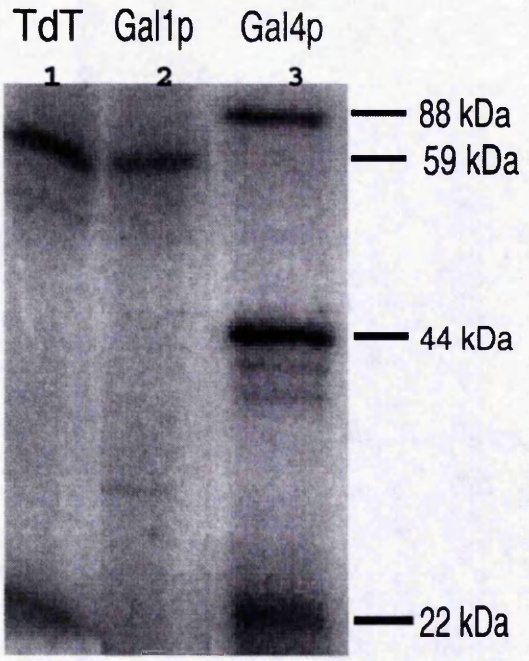


Figure 4.7 ^{32}P-8-azido-ATP labels Gal1p and Gal4p (1-93+768-881)

Three ^{32}P-8-azido-ATP labeling reactions were carried out containing TdT, Gal1p and Gal4p. Proteins were separated by SDS-PAGE and radioactivity detected by phosphorimager analysis. Lane 1 shows the reaction containing comercial TdT preparation (5 µl), lane 2 Gal1p (1.5 µM), and lane 3 Gal4p (1-93+768-881) (2.5 µM).

The 23 kDa subunit of TdT has been labelled, in addition to the unknown protein at 60 kDa. In lane 3, Gal4p (1-93+768-881) is present as both a monomer (22 kDa). Bands are present at approximately the correct positions to indicate a dimer (44 kDa), and a tetramer (88 kDa) of Gal4p (1-93+768-881).

the current experimental conditions was investigated.

4.4.4 [^{32}P]-8-azido-ATP labelling of Gal1p is affected by ATP

Specific labelling of proteins by [^{32}P]-8-azido-ATP occurs when the ligand is bound to the protein at the correct site and subsequent exposure to UV irradiation covalently binds the two molecules together. This site-specific labelling is dependent on the [^{32}P]-8-azido-ATP being bound to the correct site when the solution is exposed to UV. It may be possible to compete with this binding by addition of ATP to the reaction. ATP bound to the protein would result in a decrease in the number of molecules of [^{32}P]-8-azido-ATP interacting at the ATP binding site. The observed level of specific binding of [^{32}P]-8-azido-ATP would therefore decrease. However, if the covalent binding of [^{32}P]-8-azido-ATP was non-specific, the level of [^{32}P]-8-azido-ATP binding would be unaffected by addition of ATP to the labelling the reaction.

[^{32}P]-8-azido-ATP, in the presence and absence of ATP labelled Gal1p to different extents (Fig. 4.8, lanes 1 and 2). The presence of ATP reduced the overall level of protein labelling. No difference could be observed in the labelling of Gal4p (1-93+768-881) when ATP was added to the labelling reaction (Fig. 4.8, lanes 3 and 4). These results indicate that there may be a mixture of specific binding and non-specific binding of [^{32}P]-8-azido-ATP to Gal1p, whereas binding to Gal4p (1-93+768-881) is wholly non-specific.

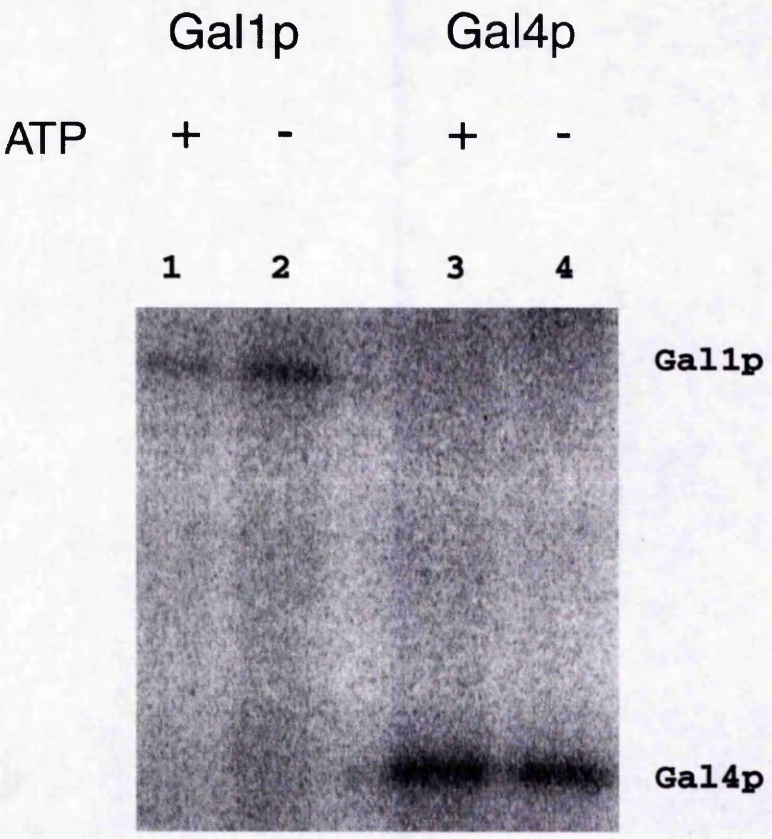


Figure 4.8 ATP reduces the level of ^{32}P-8-azido-ATP binding to Gal1p but not to Gal4p

Labelling reactions of Gal1p (1.5 μM) and Gal4p (1-93+768-881) (1.5 μM) were carried out in the presence and absence of 10 mM ATP. Proteins were separated by SDS-PAGE and radioactivity was detected by phosphorimage analysis.

4.4.5 [^{32}P]-8-azido-ATP binding to Gal1p is affected by exposure to UV

ATP affects the binding of [^{32}P]-8-azido-ATP to Gal1p, but not to Gal4p (1-93+768-881). Further experiments were designed to investigate the role that UV irradiation was playing in this labelling reaction was subsequently investigated. Gal1p is labelled to a different extent by [^{32}P]-8-azido-ATP when it is exposed to UV irradiation compared to a reaction kept in the dark (Fig. 4.9, lanes 1 and 2). Gal4p (1-93+768-881) is labelled by the molecule equally, whether the solution is exposed to UV or not (Fig. 4.9, lanes 3 and 4). This may indicate a different kind of interaction of the [^{32}P]-8-azido-ATP with Gal4p (1-93+768-881), perhaps a general interaction with the protein rather than a specific interaction at ATP binding sites. Gal1p labelling shows variation, suggesting that both specific and non-specific may be occurring.

4.4.6 [^{32}P]-8-azido-ATP does not label Gal1p at a specific ATP binding site at a detectable level

Experiments were designed to investigate whether the site of specific labelling to Gal1p observed in Fig. 4.8 could be identified. Gal1p was labelled with [^{32}P]-8-azido-ATP and the reaction digested with the proteolytic enzyme chymotrypsin. The fragments were separated by electrophoresis, but no signal could be detected from the protein fragments (Fig. 4.10). For a protein fragment to be identified the proportion of specific binding to a certain site on the protein must be high to allow the labelled fragment to be detected against the background of non-specific labelling that occurred throughout the protein. This indicated that the level of specific binding was not at a sufficient level to detect and identify the sequence that binds [^{32}P]-8-azido-ATP.

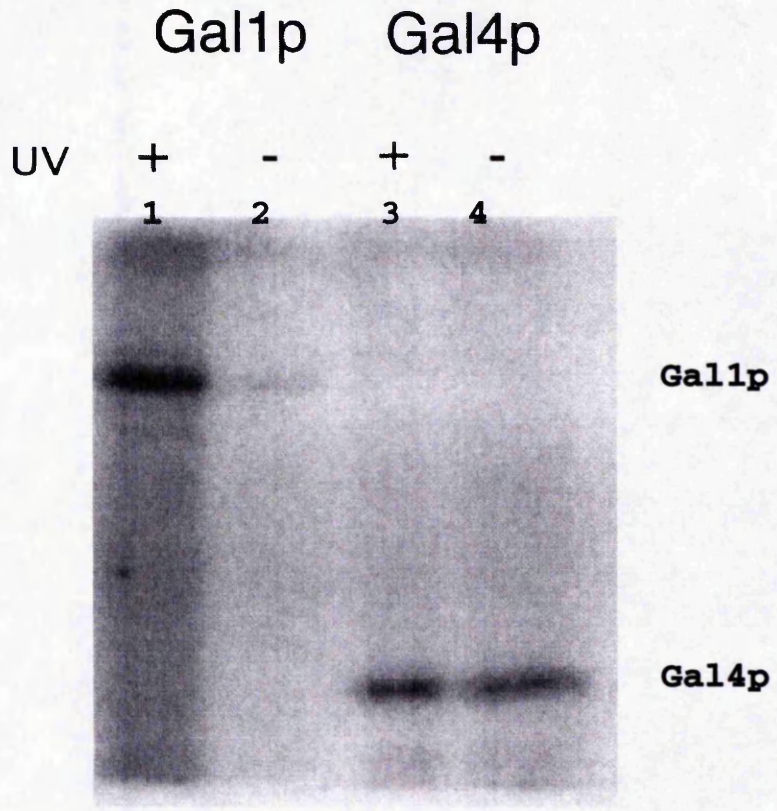


Figure 4.9 ^{32}P-8-azido-ATP labels Gal1p and Gal4p in the presence and absence of UV irradiation to different extents

Four labelling reactions were set up, two containing Gal1p, two Gal4p (1-93+768-881). One reaction from each set was exposed to UV irradiation. The proteins were separated by SDS-PAGE and radioactivity was detected by exposure to phosphorimage analysis.

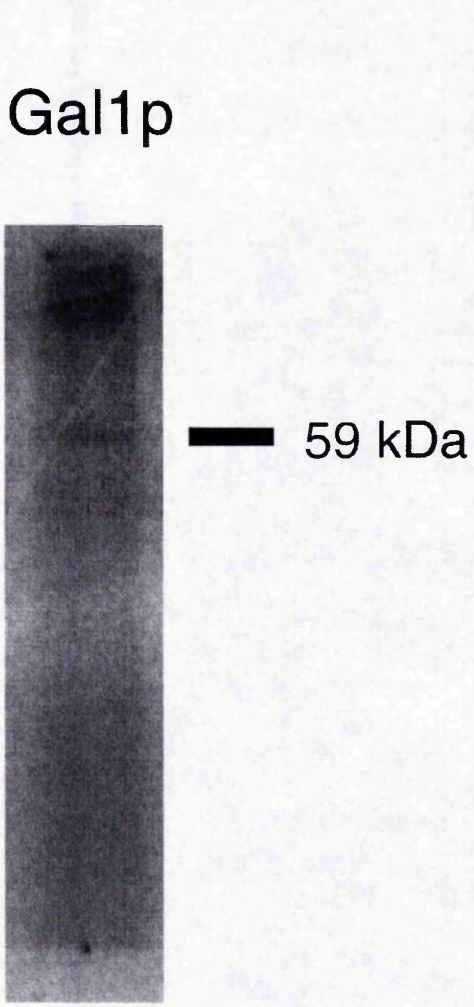


Figure 4.10 Digested labelled Gal1p does not show a specific labelled digestion product

Gal1p was labelled with ^{32}P-8-azido-ATP, followed by digestion by chymotrypsin. The fragments were separated by SDS-PAGE and radioactivity was detected by phosphorimage analysis. A smear of radioactivity was seen throughout the lane, with a faint band seen at the expected position of Gal1p, as judged by comparison with coomassie stained polyacrylamide gel of digested Gal1p.

Despite some specific binding to Gal1p occurring, as seen in Section 4.4.4, the level of this binding was not sufficient to detect an increased signal from the ATP binding polypeptide. It was therefore not possible to determine the polypeptide sequence responsible for binding ATP. *In vitro* investigations into the interaction of ATP with Gal1p and Gal3p remained unsuccessful. Subsequent published data on the structure of homoserine kinase has shown that this related protein contains a novel ATP binding sequence that does not utilise the side chain of lysine observed in classic ATP binding (Krishna *et al.*, 2001). This may explain the inability to label Gal1p with ATP derivatives that covalently label proteins *via* the active lysine side chain. This will be discussed further below. Further *in vitro* experiments were designed to investigate the interactions between the proteins of the *GAL* genetic switch.

4.5 Investigation of the interactions between proteins of the *GAL* genetic switch

Experiments were designed to investigate the interactions between the proteins of the *GAL* genetic switch *in vitro* using covalent crosslinking reagents. In addition to studying the interactions between the three wild-type *GAL* proteins, a mutant of Gal80p was studied, Gal80^{S-2}p. This is the Gal80 protein containing a single amino acid change Glu^{351}Lys. It has been shown to repress the *GAL* genes, and this repression is not alleviated by the addition of excess galactose (Nogi and Fukasawa, 1984).

4.5.1 Crosslinking of Gal4p (1-93+768-881), Gal80p and Gal80^{S-2}p

Protein-protein crosslinkers covalently join molecules that are close to each other in solution. It is a powerful method for investigating interactions between proteins. Experiments were designed to examine the interactions between Gal4p (1-93+768-881) and Gal80p or Gal4p (1-93+768-881) and Gal80^{S-2}p *in vitro*. There is a wide range of

crosslinkers available, and suitability of crosslinker is dependent on both the reactive groups and the length of the crosslinker molecule.

Bis(sulphosuccinimidyl) suberate (BS_3) is a crosslinker which contains NHS ester groups that react with primary amine groups to form covalent bonds. These groups occur at the N-terminus and in the side chain of the amino acid lysine. It will span a space between proteins of 1.1 nm (Partis *et al*, 1983). BS_3 crosslinked Gal4p (1-93+768-881) to Gal80p, and Gal4p (1-93+768-881) to Gal80^{S-2}p (Fig. 4.11, Lanes 6 and 7). The extent of crosslinking was not significantly different when the two proteins were compared suggesting similar affinities for Gal4p (1-93+768-881). Interestingly, when the two forms of Gal80p were exposed to crosslinker in the absence of Gal4p (1-93+768-881), Gal80p, but not Gal80^{S-2}p, crosslinks to run as a doublet on the gel (Fig. 4.12, lanes 3 and 4). This may be the result of internal crosslink and suggests that, although there is only a single amino acid changed between wild type and mutant protein, there may be significant structural alterations between Gal80p and Gal80^{S-2}p.

4.5.2 Crosslinking of Gal80p and Gal3p

Gal80p and Gal3p are known to interact in the presence of galactose and ATP (Platt and Reece, 1998), and this interaction alleviates the repression of the *GAL* genes. Crosslinking experiments were set up to observe the interaction between these two proteins *in vitro*, before subsequent attempts to crosslink all three proteins. It was not possible to crosslink Gal3p with Gal80p using BS_3 (data not shown). The data presented here shows crosslinking experiments carried out on the two proteins using the chemical

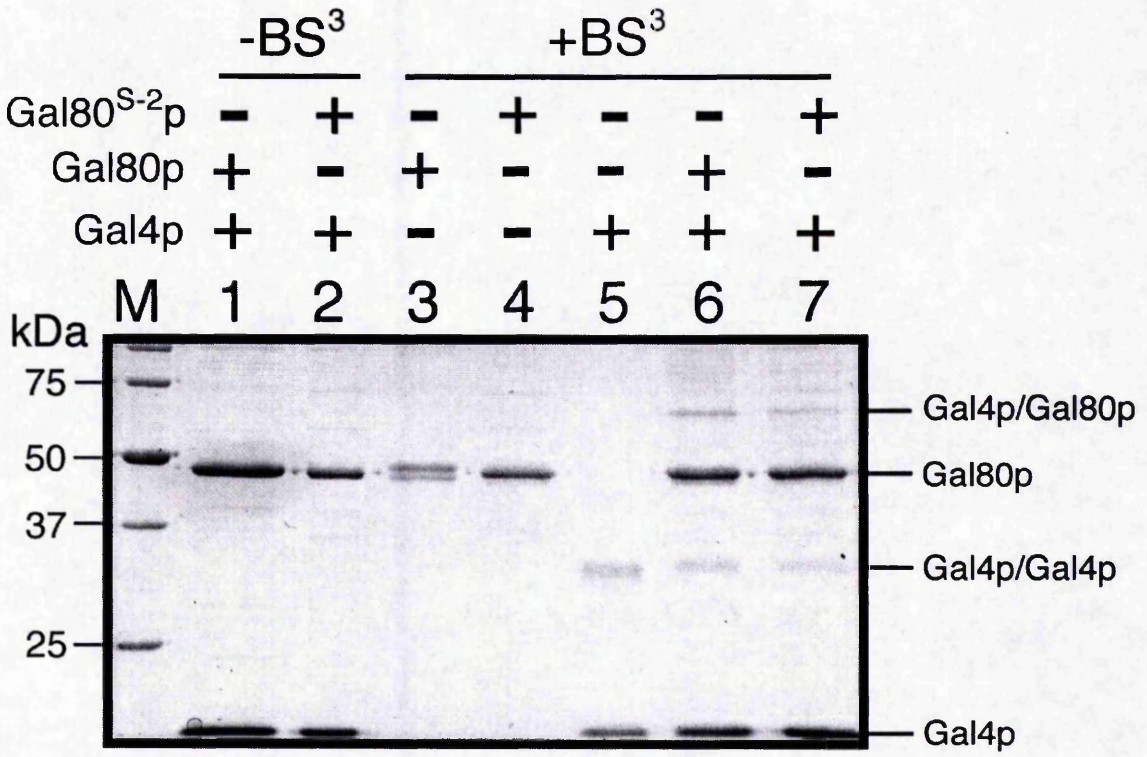


Figure 4.11 Gal4p binds Gal80p and Gal80p^{S-2}p.

Gal4p, Gal80p and Gal80^{S-2}p were mixed as indicated in the presence and absence of the crosslinker BS_3. Products were resolved by SDS-PAGE and visualised by coomassie staining. Gal4p monomers resolve at 16 kDa, dimers at around 35 kDa, Gal80p and Gal80^{S-2}p are both 48 kDa.

crosslinker glutaraldehyde. Glutaraldehyde is an amide-amide crosslinker commonly used as a protein-protein crosslinker (Hopwood, 1969).

The initial aim of the experiment to crosslink Gal3p to Ga80p was not met, however the experiment showed a number of other interesting results. Gal80p, in the presence of crosslinker formed dimers, and the doublet observed in Fig. 4.11 was observed again (Fig. 4.12, lane 3). Gal3p and Gal80p, in the presence of crosslinker, ATP and galactose, did not crosslink, but a large complex of protein was observed at in the well of the gel (Fig. 4.13, lanes 4 to 7). This may be non-specific aggregation of protein. No bands were observed at the expected size positions that would indicate a specific reaction between Gal3p and Gal80p. When Gal3p plus one of the ligands, galactose or ATP, was present in the reaction, the quantity of the upper component of the Gal80p doublet appeared to increase. This may indicate further that Gal80p changes its conformation on contact with either Gal4p or Gal3p. When this is compared to the gel shown in Fig. 4.12, it can be seen that when Gal4p (1-93+768-881) and Gal80p are in the presence of crosslinker, the Gal80p doublet appears to run at the mid-point again. This may suggest that Gal4p induces a change in the structure of Gal80p, allowing an internal crosslink to form.

Specific interactions between Gal80p and Gal3p could not be seen when visualised by SDS-PAGE followed by staining of the proteins, but there may be a small proportion of the two proteins interacting. Coomassie staining may not be sensitive enough to detect this low level of protein. The experiments were repeated, with western blotting used to identify the proteins involved in each band on the gel. A blot using Gal3p antibodies failed to show Gal3p at higher positions on the blot when in the presence of

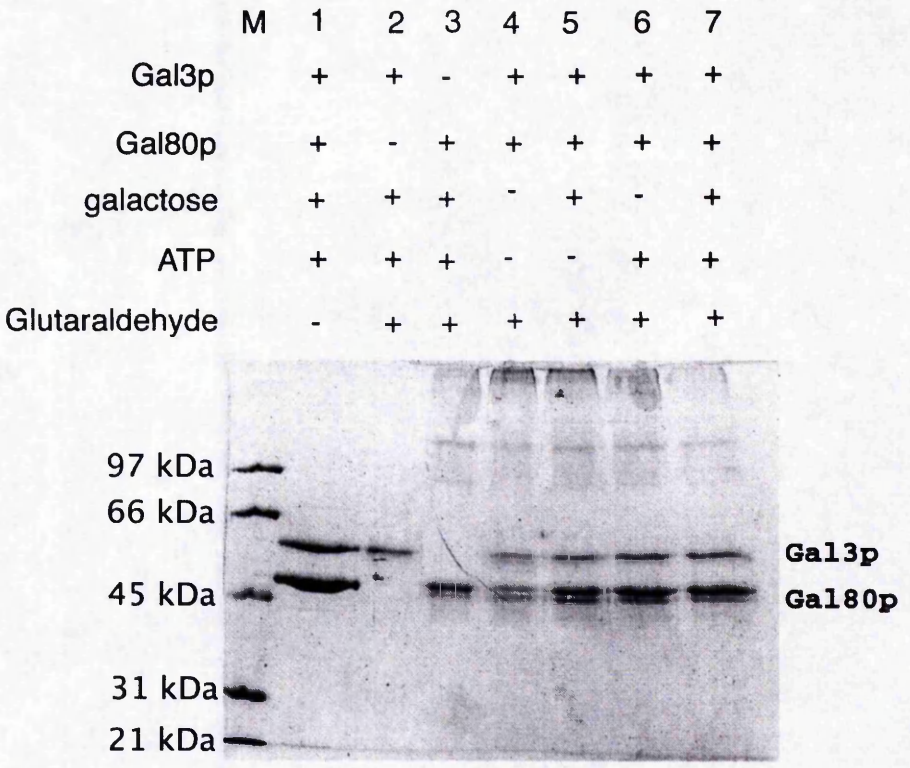


Figure 4.12 Gal80p forms multimeric complexes

Gal80p and Gal3p were mixed together with galactose and ATP as indicated in the presence and absence of the crosslinker glutaraldehyde. The products were resolved by SDS-PAGE and visualised using coomassie stain. When Gal80p (48 kDa) is exposed to crosslinker, bands at the expected size position for Gal80p dimers are present at approximately 100 kDa.

Gal80p and crosslinker (Fig. 4.14, lanes 1 and 2). Gal80p antibodies detected a large quantity of Gal80p at higher positions on the gel in the presence of crosslinker whether Gal3p was included in the reaction or not (Fig. 4.14, lanes 3 and 4). These complexes were Gal80p dimers and multimers, which have been found to be exist in free solution (Melcher and Xu, 2001). It can be seen in Fig. 4.14, lanes 1 and 2, that the quantity of the Gal3p appeared to decrease when the proteins were exposed to crosslinking agent in the presence of Gal80p. This may indicate that Gal3p and Gal80p associated with each other, but formed large aggregations of proteins, which were unable to be resolved by polyacrylamide gel electrophoresis and therefore were not subsequently exposed to western blotting.

4.6 Discussion

Experiments were undertaken to identify the ATP binding site of the galactokinase Gal1p, with the intention of localising the ATP binding site of Gal3p, the inducer of the *GAL* genes. Two derivatives of ATP were used in these experiments, oATP and 8-azido ATP. [^{32}P]-oATP was synthesised in the laboratory from [^{32}P]-ATP, and the resulting solution was judged to be >90% [^{32}P]-oATP (Fig. 4.1). The enzyme-linked assay for galactokinase activity was used to confirm that oATP binds to Gal1p at the ATP binding site. When oATP was added to the enzyme-linked assay, a pattern of competitive inhibition was seen with respect to ATP. This indicated that oATP was interfering with the binding of ATP (Fig. 4.2). Experiments to label Gal1p covalently were designed. These were based on the work of Easterbrook-Smith *et al.* (1977). Experiments to label the ATP binding site of Gal1p were unsuccessful. The experimental parameters were adjusted to increase the concentrations of the component factors.

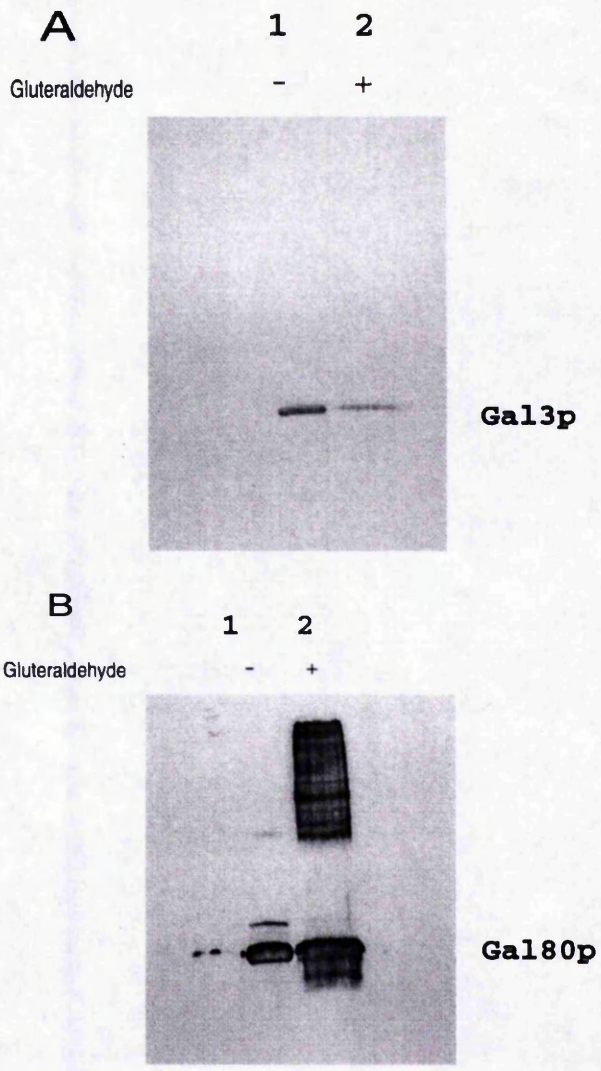


Figure 4.13 Gal3p is not present in crosslinked complexes that contain Gal80p

Gal3p, Gal80p, ATP and galactose were mixed together in the presence and absence of gluteraldehyde as indicated, and exposed to western blotting. A shows the western of the experiments exposed to Gal3p antibodies (monoclonal), and B shows identical reactions exposed to Gal80p antibodies (polyclonal).

of the reaction. The concentration of each component was limited by a variety of factors. Gal1p concentration was limited by the purified solution, and it has been found that at high concentrations Gal1p will precipitate out of solution (S. Hankin, personal communication). The concentration of [^{32}P]-oATP was limited by the supplied concentration of [^{32}P]-ATP. The final concentration of [^{32}P]-oATP was 2% of the starting material, due to necessary dilutions during the conversion reaction. The concentration of the reducing agent $NaBH_4$ could not be increased as at higher concentrations the integrity of proteins was compromised and could not be observed after electrophoresis and coomassie staining (data not shown). After adjustment of the experimental parameters, including incubation times and component concentrations, no covalent modification of Gal1p was observed.

The ATP derivative 8-azido-ATP was used in labelling experiments to attempt to label the ATP binding site of Gal1p. [^{32}P]-8-azido-ATP was obtained commercially and the labelling method based on that of Abraham *et al.* (1984). Evidence suggested that 8-azido-ATP bound to the ATP binding site using the enzyme-linked assay for galactokinase activity. Analysis of the kinetic data by the Woolf method supported this interaction, while the Lineweaver-Burk analysis showed some ambiguity. In the labelling reactions the enzyme TdT was used as a control, but a labelled protein band at approximately 60 kDa was observed (Fig. 4.3). [^{32}P]-8-azido-ATP labelled Gal1p at the expected position of 59 kDa. Both bands could be observed after coomassie staining of a polyacrylamide gel (data not shown) and there were no major contaminating bands in the Gal1p solution. The labelling of the 60 kDa protein in the TdT preparation raised the possibility of non-specific labelling of proteins by [^{32}P]-8-azido-ATP. To test for the presence of non-specific binding the experiment was repeated, and Gal4p (1-93+768-881) was included in a labelling reaction. Gal4p is not known to contain ATP binding

sites. [^{32}P]-8-azido-ATP labelled Gal4p (1-93+768-881), (Fig. 4.7). The nature of the binding was then investigated. ATP competition experiments were carried out on Gal1p and Gal4p (1-93+768-881) (Fig. 4.8). If the labelling were non-specific, then no difference in labelling would be observed when ATP was present in the reaction. This was the case in the labelling of Gal4p (1-93+768-881), no difference was observed between two labelling reactions, one containing ATP, the other without. However, when this experiment was carried out with Gal1p, a difference in labelling intensity was seen. The ATP-containing reaction showed a lower level of labelling than the reaction that did not contain ATP. This indicated that a proportion of Gal1p labelling was at the ATP binding site. The presence of a signal when ATP was included in the reaction indicated that there was still a proportion of labelling of Gal1p that is non-specific. The effect of UV on the labelling reaction was investigated. Different levels of labelling were observed in Gal1p and Gal4p (1-93+768-881), dependent on whether or not the samples were exposed to UV (Fig. 4.9). This may indicate a different method of labelling of each protein by [^{32}P]-8-azido-ATP. These experiments indicated that while [^{32}P]-8-azido-ATP was labelling the proteins non-specifically, there may be an element of specific labelling occurring at the ATP binding site of Gal1p. It was attempted to identify this site by digestion of labelled Gal1p. The labelled protein was digested with chymotrypsin. When the digested protein fragments were separated by SDS-PAGE and stained with Coomassie, an evenly spaced pattern of digestion was observed (data not shown). After labelling and digestion of Gal1p, no label associated with a specific band of digested protein was observed (Fig. 4.10). The smear of radioactivity seen throughout the lane suggests that [^{32}P]-8-azido-ATP is associating throughout the length of the protein, and is not specifically bound to one location on the protein. It was therefore not possible to identify the location of the ATP binding site using [^{32}P]-8-azido-ATP.

Experiments to investigate the interactions between the proteins of the *GAL* genetic switch *in vitro* were designed. Crosslinkers were used to examine the interactions between these proteins. The crosslinker BS_3 was used to demonstrate that a truncated version of Gal4p binds to Gal80p *in vitro*. It was not possible to crosslink Gal3p to Gal80p using the crosslinker BS_3. BS_3 crosslinks amine groups, causing covalent links between proteins. It crosslinks proteins together at a distance of 1.1 nm. It is possible that the backbones of Gal3p and Gal80p do not contain amino groups at a suitable distance apart, precluding covalent modification of the proteins. BS_3 was used to investigate the interaction of Gal4p (1-93+768-881) and the super-repressor Gal80p^{S-2}, and this was compared to the interaction with Gal80p. Gal80p^{S-2} binds to Gal4p (1-93+768-881), and when the two proteins are crosslinked, the resulting SDS-PAGE gel displays a band equivalent to a dimer of Gal4p (1-93+768-881) bound to a monomer of Gal80p^{S-2}. The Gal80p crosslinking reaction, when compared to Gal80p^{S-2} showed an interesting difference. When crosslinker was added to Gal80p alone, the protein was resolved on the polyacrylamide gel as two proteins bands, situated very closely together. This doublet was not observed when Gal80p^{S-2} was exposed to crosslinker. When Gal4p (1-93+768-881) was present in the crosslinking reaction, the presence of this doublet decreased, and Gal80p that had not bound to Gal4p appeared to be held mostly in one conformation. This suggests that Gal80p is present in two conformations in solution, and that Gal80p^{S-2} is present in one. The absence of the doublet in the crosslinking reaction containing Gal4p (1-93+768-881) and Gal80p may suggest that Gal4p alters the conformation of Gal80p. These experiments may explain the reasons why Gal80p^{S-2} is a super-repressor. Gal80p, while binding Gal4p in one conformation, may bind Gal3p in the other. If Gal80p^{S-2} occurs only in one conformation, the Gal4p binding form, and is unable to switch conformations, then it may not be able to bind to

the inducer Gal3p. This would result in an inability to react to galactose in the media, and the *GAL* genes would remain repressed.

The doublet observed in crosslinked Gal80p is present when the protein is included in a reaction containing the crosslinker glutaraldehyde. When Gal3p is included in this reaction, the proportion of Gal80p in each conformation changes. Instead of a doublet of equal intensity, one of the components of the doublet is increased, while the other decreased. This also suggests that the conformation of Gal80p when bound to Gal4p changes when activation takes place. It was not possible to see Gal80p crosslinked to Gal3p. Western blotting, which can identify the components of each band with a higher sensitivity, did not identify Gal3p at a higher position on the gel. This could mean that Gal3p is not forming crosslinks with Gal80p. Another possibility is that crosslinks are being formed, but this interaction covers the binding site for the monoclonal antibody, resulting in a lack of detection of Gal3p in complex with Gal80p. When the crosslinking reaction was exposed to polyclonal Gal80p antibodies, large complexes were identified as containing Gal80p. It has been shown that Gal80p forms multimers, and that this may be a regulatory feature (Melcher and Xu, 2001).

These experiments have investigated *in vitro* interactions between the proteins and small molecules of the *GAL* genetic switch. Experiments to identify the binding site of ATP to Gal1p were unsuccessful, due to the probable lack of a lysine residue in the active site of the galactokinase enzyme. Galactokinase is a member of the GHMP group of proteins, all of which share conserved motifs and are involved in the phosphorylation of sugars (Bork *et al.*, 1993). Recent structural studies on related proteins, including homoserine kinase, mevalonate kinase and phosphomevalonate kinase have shown that these proteins do not bind ATP *via* an active lysine residue (Fu *et al.*, 2002; Krishna *et*

al., 2001; Romanowski *et al.*, 2002). The structure of homoserine kinase bound to ADP has been solved, and the ATP/ADP binding motif found to be the consensus sequence PX_3GSSAA, found in GHMP kinases. This sequence formed a novel phosphate-binding loop, which did not utilise a lysine residue to bind ADP (Krishna *et al.*, 2001). This suggests that galactokinase may also bind ATP *via* the GLSSA sequence present in homology region III, and may not require a lysine residue for ATP binding. This would explain why the labelling reactions using ATP derivates, designed to bind covalently to lysine residues, were unsuccessful in binding to Gal1p.

It was not possible to confirm that the proteins of the *GAL* genetic switch form a tripartite complex on induction of the *GAL* genes. *In vitro* interactions between Gal4p and Gal80p were observed, as were interactions between monomers of Gal80p. Subsequent work has crosslinked Gal3p to Gal80p with the crosslinker EDC, but it has not been possible to crosslink a tripartite complex (D. Timson, personal communication). Recent work has identified the cellular locale of Gal3p as being both within and outside the nucleus, whereas Gal80p is located within the nucleus (Kumar *et al.*, 2002). This research disagrees with the theory that Gal3p is located outside the nucleus, and controls the localisation of Gal80p within the cell (Peng and Hopper, 2000). More evidence has been collected to support the theory that Gal80p changes conformation on induction of the *GAL* genes. The super-repressor $Gal80p^{S-2}$ cannot change conformation, and this may explain why the repression of the *GAL* genes cannot be alleviated in a $gal80p^{S-2}$ strain. Gal80p is bound to Gal4p in one conformation, and when Gal3p, along with galactose or ATP, is present, the conformation of Gal80p changes, allowing alleviation of repression.

Chapter Five

Investigation into the induction of the *GAL* genes *in vivo*

5.1 Introduction

The transcription of the *GAL* genes is known to be under control of three proteins, Gal4p, Gal80p and Gal3p; the *GAL* genetic switch. The inducer, Gal3p interacts with Gal80p, in the presence of galactose and ATP, to alleviate repression of transcription, allowing Gal4p to activate gene expression (Suzuki-Fujimoto *et al.*, 1996) Gal1p can also affect control on the *GAL* genes. In the absence of Gal3p, Gal1p will induce the *GAL* genes, although the induction occurs 2-3 days after addition of galactose to the growth media (Bhat and Hopper, 1992), the LTA phenotype (Rotman and Spiegelman, 1953). Transcriptional control of the *GAL* genes has been observed directly using an *in vitro* transcription assay, where it was found that Gal1p activates the *GAL* genes at a lower efficiency than Gal3p (Platt and Reece, 1998). *In vivo*, the activation of the *GAL* genes was observed by growth of yeast on galactose as the sole carbon source, as shown in Chapter 3. This method has some drawbacks when measuring the transcription of the *GAL* genes in yeast containing mutant sequences. Yeast that does not contain an active galactokinase (normally encoded by the *GAL1* gene) cannot grow on galactose as the sole carbon source (Chapter 3). This presents a problem when measuring the effect of Gal1p on transcription *in vivo*, as the growth of yeast cannot be used to assess the activation of the *GAL* genes by Gal3p. Reporter genes would allow the detection of protein, but the assay described here allows transcription to be observed independently of translation. To measure the rate of transcription *in vivo*, an RT-PCR assay was developed. This assay quantifies the mRNA produced from activation of the *GAL10* gene, which is under control of the *GAL* genetic switch. The *GAL1* gene is often used as a paradigm of the *GAL* genes, but some of the experiments described below take place in a strain deleted for *GAL1* or introduce the *GAL1* gene on a plasmid, therefore examining the transcription of this gene would be inappropriate. The transcription of the

ACT1 gene was also measured, as this should be unaffected by the addition of galactose to the media. By plotting the ratio of the amount of these gene products, it was possible to follow the level of activation of the *GAL10* gene. Yeast cells were grown in media containing galactose and ethanol. Ethanol can be used as a carbon source by yeast, but should not be preferentially chosen over galactose. Yeast cells grown in galactose and ethanol will result in activation of the *GAL* genes and in yeast strains without a galactokinase the yeast will continue to grow, using ethanol as a carbon source. The *in vivo* activation of the *GAL* genes was followed by observing the mRNA levels of the *GAL10* gene, and this can be observed in a Δ*gal1* yeast strain.

Experiments were designed to follow the activation of the *GAL* genes *in vivo* in a variety of genetic backgrounds. These data may give further insights into the individual roles played *in vivo* by different members of the *GAL* genetic switch. Further experiments were designed to follow transcription of the *GAL* genes controlled by plasmid-borne genes encoding either Gal1p or Gal3p, in a Δ*gal3* background. The *GAL1* and *GAL3* genes were expressed on the plasmid under the control of either the *GAL1* or the *GAL3* promoter to examine the effect that the inducer gene has on the transcription of the *GAL* genes. This was investigated in Chapter 3, but due to the lack of sensitivity of the growth assay, no firm conclusions could be reached. These experiments will give insight into the role of the promoters in the pattern of expression seen in both wild type expression and in the LTA phenotype, resulting in greater understanding in the role each protein plays in the activation of the *GAL* genes *in vivo*.

5.2 Transcription of the *GAL* genes in yeast

The *ACT1* gene was chosen as a control for these experiments. The *ACT1* gene produces actin, a protein involved in the molecular structure of the cell. The transcription of the *ACT1* gene is not under the control of the *GAL* genetic switch, and is not under nutrient control (Kinoshita *et al.*, 1992). The transcriptional activation of the *GAL10* gene was measured against the level of the *ACT1* gene transcription. The RT-PCR method was optimised until it was possible to perform RT-PCR on both *GAL10* and *ACT1* in the same reaction. This eliminates experimental inconsistencies that may have arisen had the *ACT1* reaction been carried out in a separate PCR reaction. Both reactions were in the exponential phase of reaction, as measured by testing a range of cycle conditions. The *ACT1* PCR reaction was not at saturation levels that would result in an equal level of PCR product in each lane. Positioning of the samples in the PCR machine did not affect the outcome of the experiment, tested by identical reactions carried out in each reaction holder of the PCR machine (data not shown).

The transcription of the *GAL* genes was followed in the yeast strain JPY5. This strain is wild-type for all the *GAL* genes. The RT-PCR method was used to assay the level of *GAL10* mRNA in the cell, as described in Materials and Methods (Sections 2.13.1 and 2.13.2) (Fig. 5.1). The *GAL* genes were activated quickly, within one hour after the addition of galactose to the media. The zero sample was taken from the cells when growing in media containing raffinose as the sole carbon source, before inoculation into media containing galactose and ethanol. The products observed on the agarose gel show that *GAL10* transcription remained switched on throughout the 72 hours of the experiment (Fig. 5.1). The plot of expression of the ratio of *GAL10*:*ACT1* indicates that transcription of the *GAL* genes was switched on,

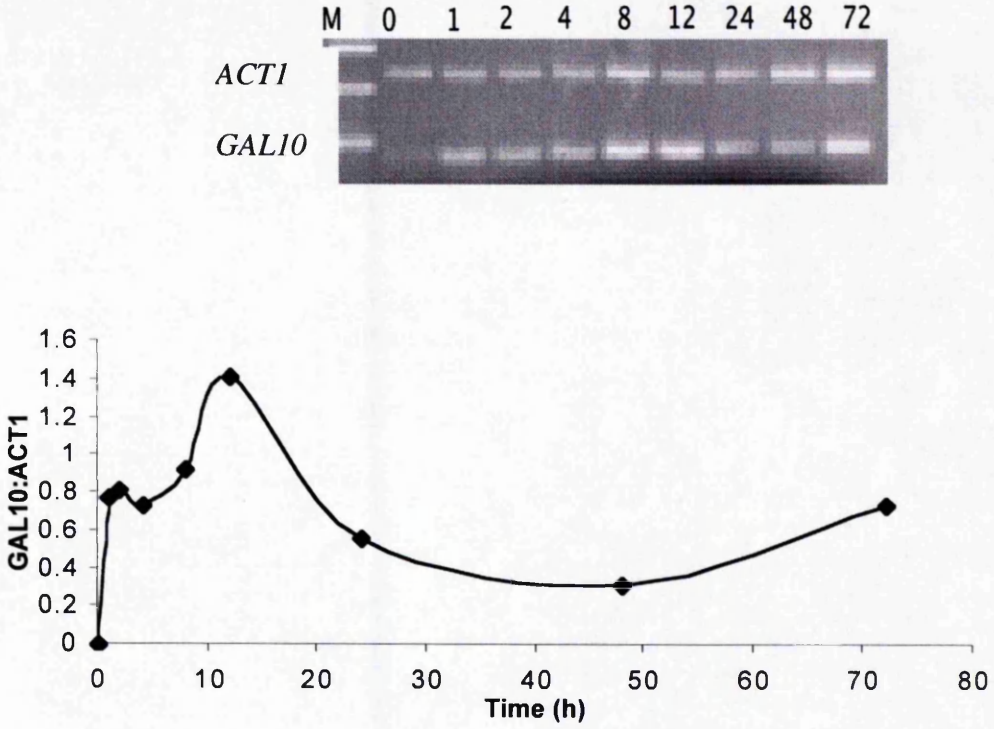


Figure 5.1 Expression of the *GAL* genes in wild-type yeast

RT-PCR assays were carried out on mRNA samples taken over time from wild type yeast. The RT-PCR products were separated by agarose gel eletrophoresis Labelled lanes refer to the time (h) after galactose was added to the media. The *ACT1* gene product is 340 bp in length, and the *GAL10* product 190 bp. The ratio of the intensity of the RT-PCR products in each reaction was plotted against time, showing the relative mRNA levels of the *GAL10* gene product.

followed by a steadying of the rate between 1 and 4 hours, followed by another rise to a high level before decreasing to just above the switched off state. This indicates the *GAL* genes are switched on quickly, and after the initial production of the required genes in the first 12 hours, the transcription stabilises at a lower level.

5.2.1 *GAL* gene expression in yeast deleted for *GAL1*

The effect of deleting this protein on the *in vivo* transcription of the *GAL* genes was investigated using the RT-PCR method detailed above (Fig. 5.2). Transcription should be controlled by Gal3p. The agarose gel indicates that the *GAL10* gene is activated rapidly after addition of galactose to the media. The graph of the levels of mRNA also indicates that transcription becomes activated rapidly after the addition of galactose to the growth media, and the activation remains high throughout the 72 hour period.

5.2.2 *GAL* gene expression in yeast deleted for *GAL3*

Gal1p is the galactokinase of the Leloir pathway and is known to induce *GAL* gene expression in the absence of Gal3p (Bhat and Hopper, 1992). Yeast, deleted for *GAL3* and containing over-expressed Gal1p shows the same rate of growth as yeast containing Gal3p (Chapter 3, Fig 3.3, lines 2 and 3). The RT-PCR method was used to measure *GAL* gene expression in a Δ*gal3* yeast strain (Fig. 5.3). In the absence of the *GAL3*, the *GAL* genes were not induced when compared to the level of expression of the *ACT1* genes, but remained at low levels throughout. The level of mRNA remained below that of *ACT1* throughout the experiment. This was unexpected, as the LTA phenotype, where growth of yeast occurs after 2-3 days is mediated by *GAL1* and should have been observed in this experiment. This will be discussed in greater detail below.

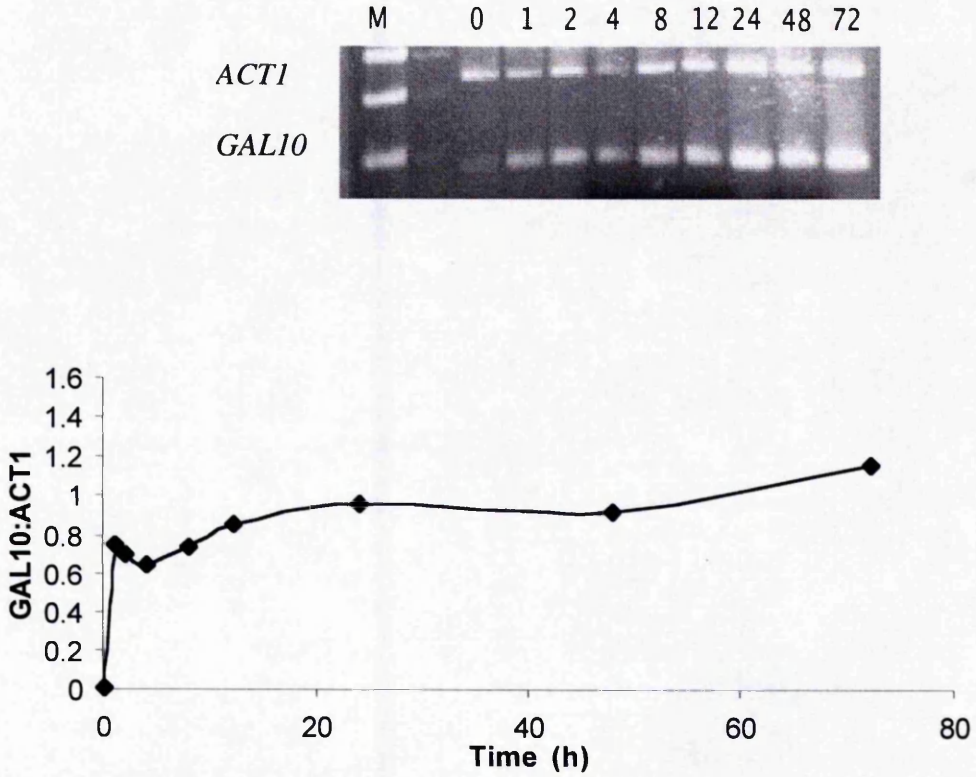
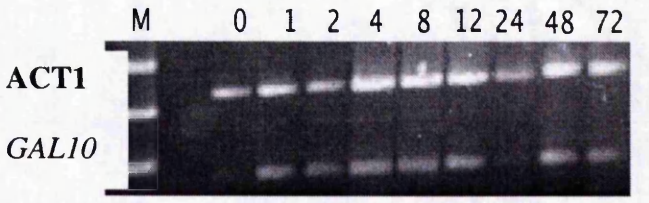


Figure 5.2 Expression of the *GAL* genes in Δ*gal1* yeast

The RT-PCR assay was carried out on mRNA samples taken over time from yeast deleted for the *GAL1* gene. The RT-PCR products were separated by agarose gel; labelled lanes refer to the time (h) after galactose was added to the media. The *ACT1* gene product is approx. 340 bp in length, and the *GAL10* product 190 bp. The ratio of the intensity of the RT-PCR products in each reaction was plotted against time. This shows the mRNA level of the *GAL10* gene product.



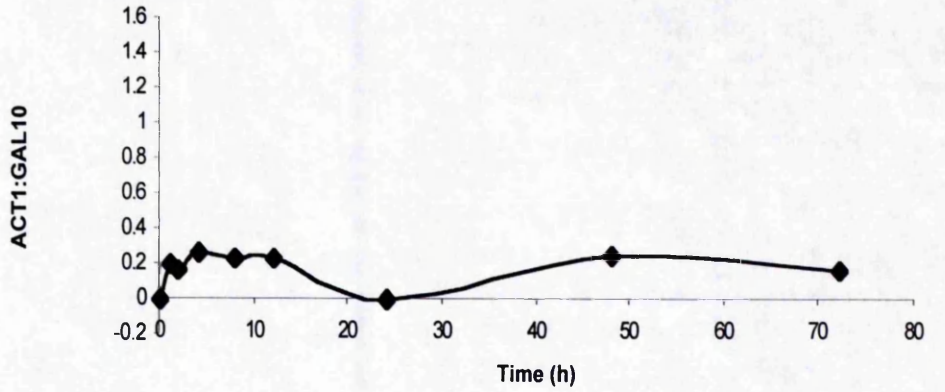


Figure 5.3 Expression of the *GAL* genes in Δ*gal3* yeast

The RT-PCR assay was carried out on mRNA samples taken over time from yeast deleted for the *GAL3* gene. The RT-PCR products were separated by agarose gel, labelled lanes refer to the time (h) after galactose was added to the media. The *ACT1* gene product is approx. 340 bp in length, and the *GAL10* product 190 bp. The ratio of the intensity of the RT-PCR products in each reaction was plotted against time. This shows the mRNA level of the *GAL10* gene product.

5.2.3 Expression of the *GAL* genes in yeast deleted for *GAL1* and *GAL3*.

It has been shown that the *GAL* genes can be activated *in vitro* by both Gal3p and Gal1p (Platt and Reece, 1998). When these genes are deleted, yeast cannot grown on galactose as the sole carbon source (Chapter 3, Fig. 3.3, line 1). The expression of the *GAL* genes in a Δ*gal1*,Δ*gal3* yeast strain was observed using the RT-PCR method described above (Fig. 5.4). There was a low level of *GAL10* mRNA present in the samples taken during the first 8 hours of the experiment. This may be a background level of transcription and not dependent on galactose, as indicated by a faint band in the zero time lane of the gel, although it did rise after addition of galactose to the media. The plot of the *GAL10* mRNA levels shows no transcription activation of the *GAL10* gene when compared to the levels of *ACT1* mRNA. Again, the level of *GAL10* mRNA remained below that of the *ACT1* gene throughout the experiment. In this yeast, Gal80p was repressing the transcription of the *GAL* genes, and this repression was not alleviated by the addition of galactose as there was no inducer present in the cell.

5.2.4 Expression of the *GAL* genes in a yeast strain deleted for *GAL1, GAL3* and *GAL80.*

In Δ*gal3*,Δ*gal1*,Δ*gal80* yeast it would be expected that the *GAL* genes would be under no galactose regulated control, and the activator Gal4p would activate transcription the *GAL* genes, as long as there were no other negative regulators affecting the transcription. The RT-PCR assay, described above, was used to follow the transcription of the *GAL10* gene in this yeast strain (Fig. 5.5). The transcription of the *GAL* genes was active when no galactose was present in the media. It can be seen from the gel that at each time point, including the zero time lane when no galactose was present that the *GAL10* gene is transcribed at an activated level. The transcription of the *GAL* genes

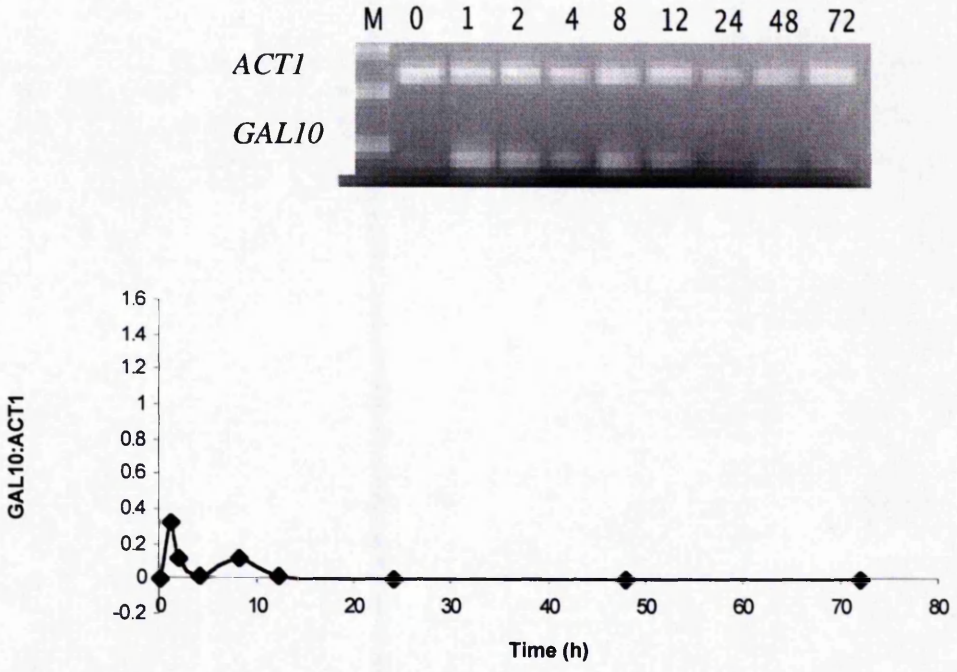


Figure 5.4 Expression of the *GAL* genes in Δ*gal1*, Δ*gal3* yeast

RT-PCR assays were carried out on mRNA samples taken over time from yeast deleted for both the *GAL1* gene and the *GAL3* gene. The RT-PCR products were separated by agarose gel electrophoresis; labelled lanes refer to the time (h) after galactose was added to the media. The *ACT1* gene product is 340 bp in length, and the *GAL10* product 190 bp. The ratio of the intensity of the RT-PCR products in each reaction was plotted against time.

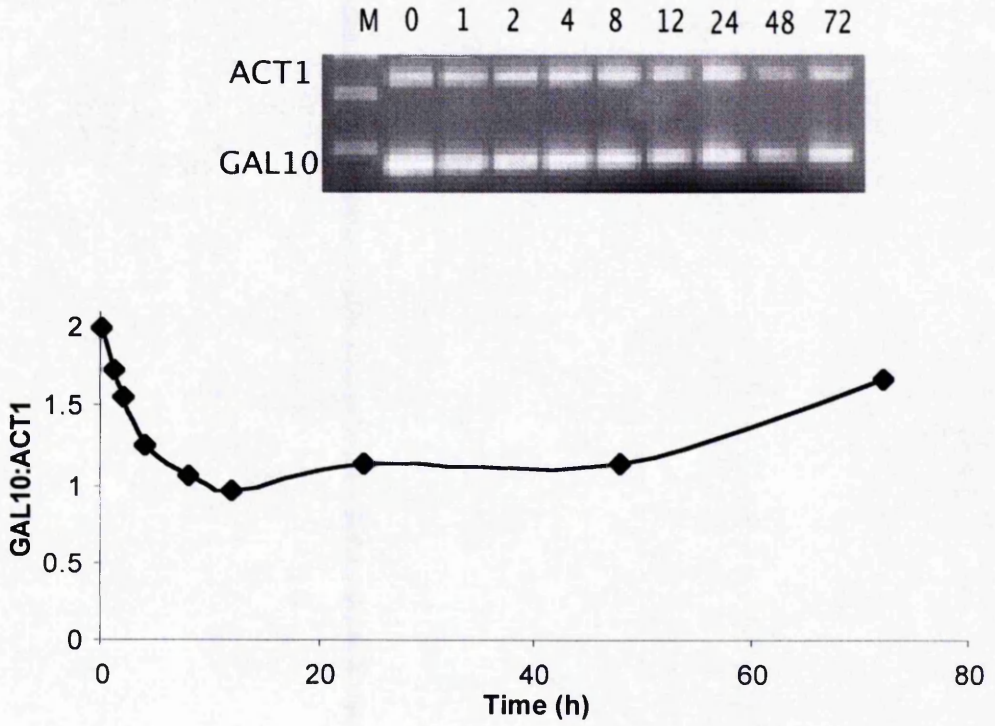


Figure 5.5 Expression of the *GAL* genes in Δ*gal1*, Δ*gal3*, Δ*gal80* yeast

The RT-PCR assay was carried out on mRNA samples taken over time from yeast deleted for the *GAL1* gene, the *GAL3* gene and the *GAL80* gene. The RT-PCR products were separated by agarose gel electrophoresis; labelled lanes refer to the time (h) after galactose was added to the media. The *ACT1* gene product is 340 bp in length, and the *GAL10* product 190 bp. The ratio of the intensity of the RT-PCR products in each reaction was plotted against time.

is not under nutrient control.

The expression of the *GAL* genes measured by the RT-PCR assay confirmed previous work on the control of the *GAL* genes in yeast. This method could therefore be used to investigate the role that the promoters of the two inducers, Gal3p and Gal1p play in the activation of these genes, resulting in different phenotypes in Δ*gal1* and Δ*gal3* yeast.

5.3 The control of *GAL* gene expression is dependent on both the inducer gene and promoter

GAL gene expression can be controlled by both Gal3p and Gal1p. Yeast cells without a genomic copy of the inducer *GAL3*, that are transformed with a multicopy plasmid expressing either *GAL1* or *GAL3* grow on galactose within twenty hours (Chapter 3, Fig 3.2, lines 2-5), indicating that the transcription of the *GAL* genes is switched on rapidly. To investigate differences between the expression of the *GAL* genes controlled by either plasmid-borne *GAL1* or *GAL3*, the RT-PCR method described above was used. This experiment was carried out in a strain deleted for *GAL3*. The results above indicated that there is minimal transcription occurring in this 72 hour period by *GAL1* and this should not interfere with the transcription driven by the plasmid-borne gene.

5.3.1 Expression of the *GAL* genes induced by plasmid-borne Gal3p

The Δ*gal3* yeast cells were transformed with pAP25, a plasmid that contains the *GAL3* gene expressed under the control of the *GAL3* promoter. Transcription of the *GAL* genes was measured over time using the RT-PCR method described above. Transcription of *GAL10* was induced rapidly, although does tend to decrease at later time points (Fig. 5.6). This would indicate that the *GAL* genes were switched on rapidly and the

induction was mediated by Gal3p. The decrease observed at later time points may be due to depletion of galactose in the media, as a rise in expression levels was observed after 24 hours when the culture was inoculated into fresh media. This experiment was carried out over 48 hours.

5.3.2 Expression of the *GAL* genes induced by Gal3p under the control of the *GAL1* promoter

The Δ*gal3* yeast cells were transformed with pAP101, which expressed the *GAL3* gene controlled by the promoter of the *GAL1* gene. The expression levels of the *GAL10* gene after addition of galactose to the growth media were followed by the RT-PCR method. The *GAL10* gene was switched on after approximately 4 hours, and fully activated levels were observed after 12 hours (Fig. 5.7). The activated level of transcription was sustained throughout the 72 hour period of the experiment.

5.3.3 Expression of the *GAL* genes induced by plasmid-borne Gal1p

The transcription of the *GAL* genes in a Δ*gal3* strain transformed with pAP102, which contains the *GAL1* gene under the control of the *GAL1* promoter, was assessed using the RT-PCR method, described above. This plasmid was present at a high copy number in the cell, providing many copies of this gene. The *GAL* genes were switched on gradually, with fully activated transcription observed after 12 hr (Fig. 5.8). This activated level of expression remained high throughout the remainder of the 72 hour period.

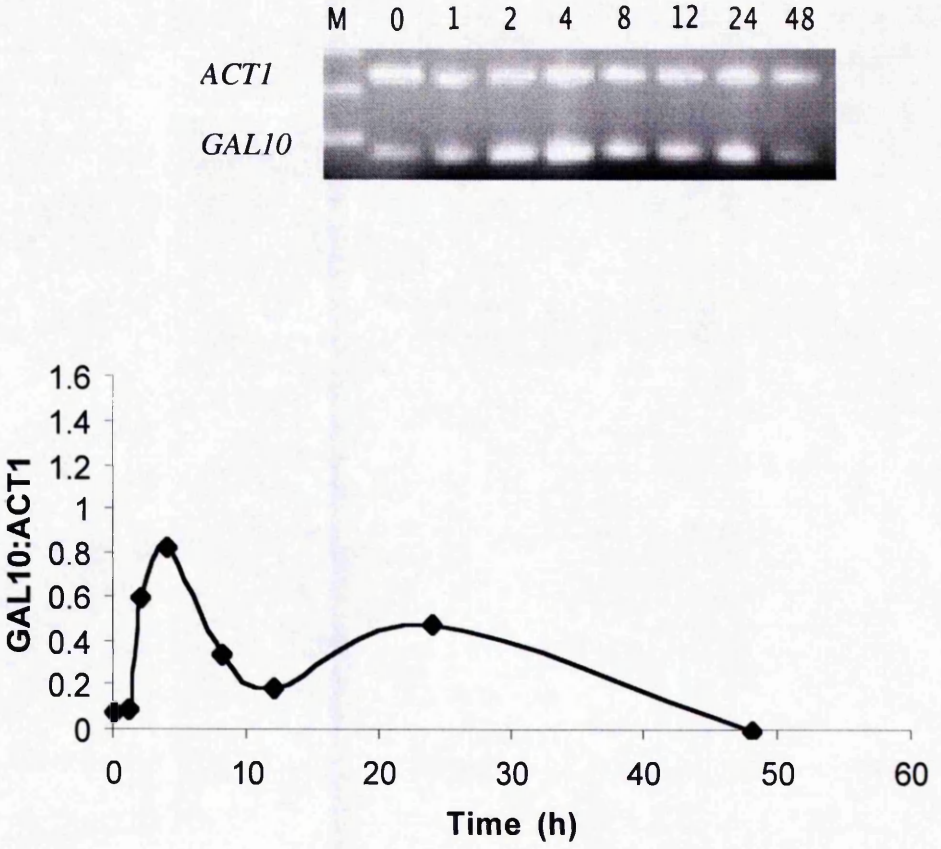


Figure 5.6 *GAL* gene expression controlled by Gal3p expressed from the *GAL3* promoter

The RT-PCR assay was carried out on mRNA samples taken over time from Δ*gal3* yeast transformed with a plasmid expressing *GAL3* from the *GAL3* promoter. The RT-PCR products were separated by agarose gel; labelled lanes refer to the time (h) after galactose was added to the media. The *ACT1* gene product is approx. 340 bp in length, and the *GAL10* product 190 bp. The ratio of the intensity of the RT-PCR products in each reaction was plotted against time.

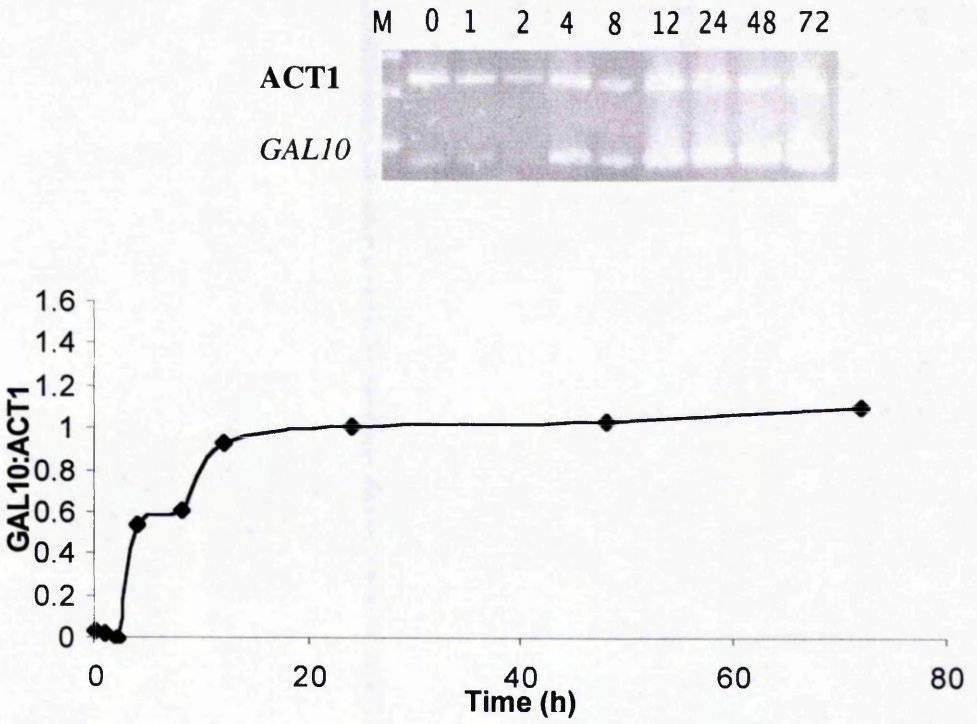


Figure 5.7 *GAL* gene expression controlled by Gal3p expressed from the *GAL1* promoter

The RT-PCR assay was carried out on mRNA samples taken over time from Δ*gal3* yeast transformed with a plasmid expressing *GAL3* from the *GAL1*promoter. The RT-PCR products were separated by agarose gel; labelled lanes refer to the time (h) after galactose was added to the media. The *ACT1* gene product is approx. 340 bp in length, and the *GAL10* product 190 bp. The ratio of the intensity of the RT-PCR products in each reaction was plotted against time. This shows the mRNA level of the *GAL10* gene product.

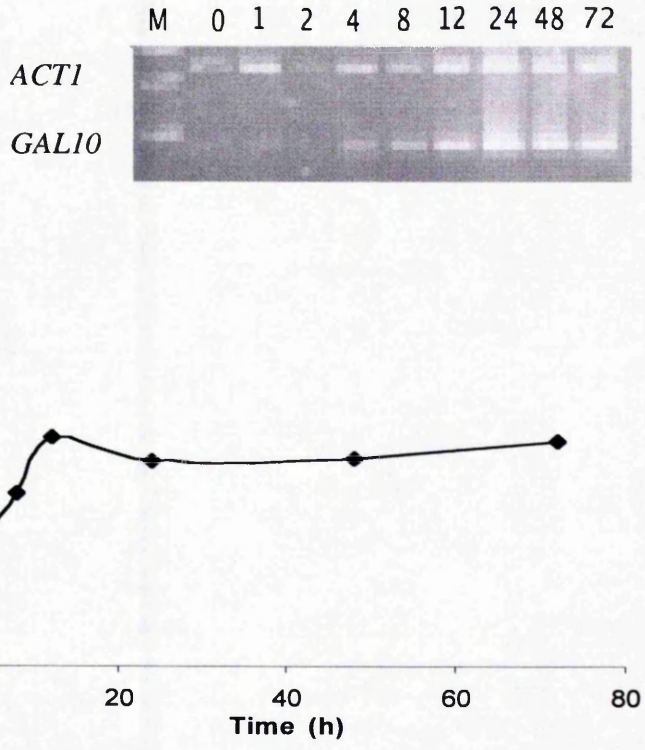


Figure 5.8 *GAL* gene expression controlled by Gal1p expressed from the *GAL1* promoter

The RT-PCR assay was carried out on mRNA samples taken over time from Δ*gal3* yeast transformed with a plasmid expressing *GAL1* from the *GAL1* promoter. The RT-PCR products were separated by agarose gel; labelled lanes refer to the time (h) after galactose was added to the media. The *ACT1* gene product is approx. 340 bp in length, and the *GAL10* product 190 bp. The ratio of the intensity of the RT-PCR products in each reaction was plotted against time. This shows the mRNA level of the *GAL10* gene product.

5.3.4 Expression of the *GAL* genes induced by Gal1p under the control of the *GAL3* promoter

The Δ*gal3* yeast cells were transformed with pRJR374 that contains a copy of the *GAL1* gene under the control of the *GAL3* promoter. Again, the expression of the *GAL10* gene after addition of galactose to the growth media was followed using the RT-PCR method, described above. Transcription of the *GAL* genes was switched on rapidly, after 1 hour there was a very strong signal from the *GAL10* position on the agarose gel. The graph of expression levels also displays a rapid induction, followed by a decrease in the relative level of transcription after 24 hr (Fig 5.9).

5.4 Discussion

The RT-PCR method is a powerful tool for following the *in vivo* transcription of the *GAL* genes. In Chapter 3 *GAL* gene activation was assessed by growth on galactose as the sole carbon source. The RT-PCR method described in this chapter allowed the levels of transcription of the *GAL* genes to be observed in a Δ*gal1* strain. The yeast cells were grown in raffinose, and then introduced into media containing galactose and ethanol. Raffinose, like ethanol, is a non-inducing non-repressing carbon source, and experiments where the yeast cells were grown on a mixture of raffinose and galactose were attempted, but transcription of the *GAL* genes did not occur in a wild type strain for at least twelve hours (data not shown). Therefore the yeast cells were inoculated into fresh media that contained galactose and ethanol. The ethanol allowed growth in a yeast strain deleted for *GAL1* by providing a carbon source that could be metabolised by the cell.

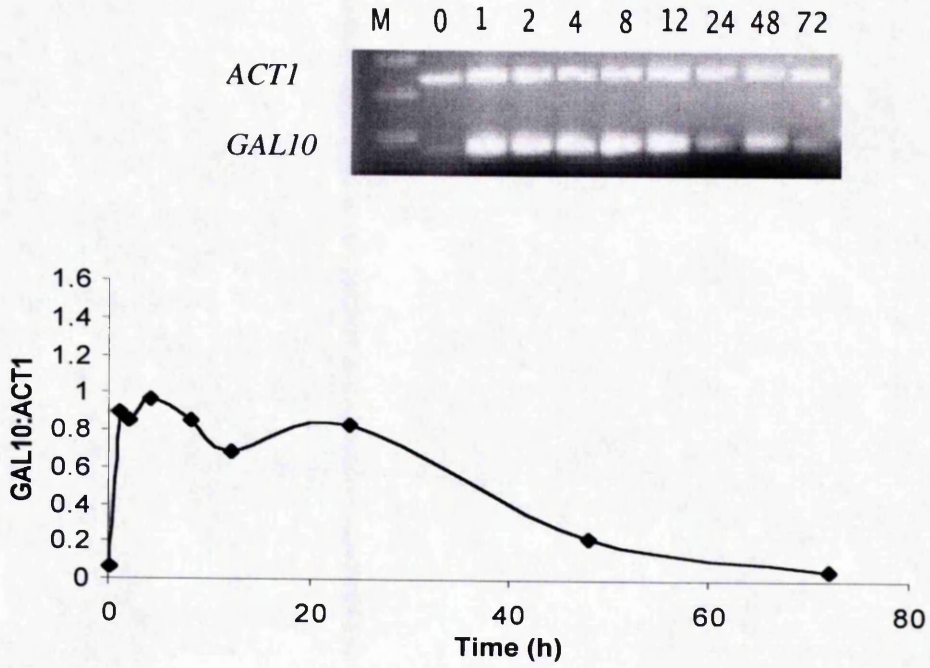


Figure 5.9 *GAL* gene expression controlled by Gal1p expressed from the *GAL3* promoter

The RT-PCR assay was carried out on mRNA samples taken over time from Δ*gal3* yeast transformed with a plasmid expressing *GAL1* from the *GAL3* promoter. The RT-PCR products were separated by agarose gel; labelled lanes refer to the time (h) after galactose was added to the media. The *ACT1* gene product is approx. 340 bp in length, and the *GAL10* product 190 bp. The ratio of the intensity of the RT-PCR products in each reaction was plotted against time. This shows the mRNA level of the *GAL10* gene product.

Transcription of the *GAL* genes was tested by the RT-PCR method in a variety of genetic backgrounds. These experiments were designed to validate the RT-PCR method before studying the effect of the promoter on gene activation. The results of these control experiments will be only briefly summarised here, as all results were what would be expected from past studies. In a wild-type yeast strain rapid activation of the *GAL* genes was observed, which stayed at activated levels for approximately 12 hours, followed by a decrease in the amount of *GAL10* mRNA present in the cell. This decrease may because the growth media was being depleted of galactose, and as the concentration of galactose in the environment decreased, the transcription of the *GAL* genes was reduced. The cells were re-inoculated in fresh media after 24 hours, and a slight rise in the level of transcription can be seen at this point. The transcription of the *GAL* genes in a *GAL1* delete strain displayed an rapid increase to activated levels. The pattern of expression varies from wild type expression in that it subsequently remains at activated levels throughout the 72 hour period. In a Δ*gal1* strain, the yeast cannot metabolise galactose, therefore galactose within the cell was not depleted. This would result in constant induction of the *GAL* genes, as the galactose signal for transcription remains present.

In a yeast strain deleted for the inducer *GAL3* the level of transcription was at a low level throughout, faint bands were seen in most lanes of the agarose gel. This may indicate that in the LTA phenotype, the growth observed after 2-3 days is not due to a rapid switching on of the *GAL* genes by Gal1p once a threshold has been met. It may be due to a gradual increase in the levels of the enzymes of the Leloir pathway until they are sufficient for growth to occur. This would cause a feedback loop at the Gal1p promoter. In these experiments the yeast are already growing, utilising ethanol as a carbon source, so a sudden increase in growth due to the metabolism of galactose would

not be observed. It may be, however, that this experiment did not continue for sufficient time for a rapid increase in transcription, due to Gal1p, to be observed.

Expression of the *GAL10* gene was assessed in strain that were deleted for both *GAL1* and *GAL3*. The *GAL* genes were repressed by Gal80p, and the addition of galactose to the media did not alleviate this repression because neither of the possible inducer proteins was present. In a yeast strain deleted for *GAL1, GAL3,* and *GAL80,* the transcription of the *GAL10* gene is not under the control of galactose. The levels of transcription remains at activated levels throughout the experiment. Gal4p is activating the *GAL* genes independently of galactose in the environment.

Following these results, which showed that RT-PCR was a reliable method for following the *in vivo* activation of the *GAL* genes, the induction of the *GAL* genes controlled by plasmid-borne genes was measured in a yeast strain deleted for *GAL3*. These experiments aimed to identify if the promoters of *GAL1* and *GAL3* are involved in the regulation of the *GAL* genes. Yeast was transformed with a plasmid encoding *GAL3*, expressed by the *GAL3* promoter sequence. The *GAL10* gene was expressed rapidly, but activated expression was followed by a drop in the levels of transcription, possibly caused by the decreasing concentration of galactose in the media. The pattern of expression is similar to the pattern observed in wild-type yeast. The corresponding experiment was carried out with the plasmid expressing the *GAL3* gene, this time controlled by the *GAL1* promoter. The expression of the *GAL10* gene is induced slowly over the first 12 hours, but again the expression of the gene remains high throughout the 72 hour timeframe of the experiment. The *GAL1* promoter causes the *GAL* genes to be switched on more slowly, even though the protein Gal3p usually induces the *GAL* genes rapidly.

The experiment was repeated with a plasmid containing the *GAL1* gene, controlled by the *GAL1* promoter. *GAL* gene expression was again induced slowly. This time no decrease was observed over the time frame of the experiment. This was not expected, as the cell should be able to metabolise galactose due to the galactokinase activity of Gal1p.

The corresponding experiment with the transformed plasmid expressing the *GAL1* gene controlled by the *GAL3* promoter was carried out. The expression of the *GAL10* gene was switched on rapidly, indicating that the *GAL3* promoter must cause this rapid induction. Transcription decreased after 24 hr, probably due to depletion of galactose in the growth media.

These experiments highlight a difference in transcription of the *GAL* genes when controlled by different inducer proteins *in vivo*. These results are summarised in Table 5.1. Initial experiments to investigate whether the promoter had an effect showed no difference whether the inducer protein was driven from the *GAL1* or the *GAL3* promoter (Chapter 3, Section 3.2). In these experiments it has been shown that the *GAL1* promoter sequence, when driving an inducer gene, leads to slower overall activation of the *GAL* genes than the *GAL3* promoter. It can therefore be proposed that it is the promoter of the *GAL1* gene that causes the LTA phenotype, and not an inability of the protein to activate the transcription of the *GAL* genes. However, another interesting point is raised by the pattern of gene expression when the inducer protein is controlled by the *GAL1* promoter, the transcription of the *GAL* genes remains high throughout the 72 hr period of the experiment.

Protein	Promoter	Induction	Sustained for 72 hr?
Gal3p	*GAL3*	Rapid	No
Gal3p	*GAL1*	Gradual	Yes
Gal1p	*GAL1*	Gradual	Yes
Gal1p	*GAL3*	Rapid	No

Table 5.1 Summary of the results of transcriptional control experiments

The results from Sections 5.3.1 to 5.3.4 are summarised, showing the effect that the promoter of the inducer gene (encoding either Gal3p or Gal1p) had on the transcriptional control of the *GAL* genes.

This was not due to galactose being present, as transcription remained high even when Gal1p was present. The galactokinase activity of Gal1p will metabolise galactose, depleting it from the cell. This may be due to the different number of Gal4p binding sites on the *GAL3* and the *GAL1* promoter. The *GAL1* promoter contains more Gal4p binding sites than the *GAL3* promoter. A further mechanism that allows transcription to be activated has been identified. Gal4p S699 is phosphorylated, allowing transcription at low galactose concentration (Rohde *et al.*, 2000). This residue is phosphorylated by Srb10p of the Mediator subcomplex, and occurs in parallel with gene activation (Hirst *et al.*, 1999). This phosphorylation of Gal4p may allow the *GAL1* promoter to stay active even at low galactose concentration, causing Gal1p to be produced, which continues to activate the *GAL* genes. The *GAL3* promoter, with fewer Gal4p binding sites, may not activate the transcription at low galactose concentrations.

Chapter 6

Final Overview

6.1 Final overview

The experiments described in this thesis have been used to investigate the molecular mechanism of the *GAL* genetic switch in *S. cerevisiae*. The protein inducer of the *GAL* genes, Gal3p, was converted into a galactokinase by the insertion of just two amino acids, a serine and an alanine into position 164 (Chapter 3). The site of the insertion lies in a region of Gal3p that shares a high degree of similarity with many galactokinase enzymes, including Gal1p in *S. cerevisiae*, which includes the consensus sequence PX_3GGSSA found in members of the GHMP kinase superfamily, which includes galactokinase (Bork *et al.*, 1993). The coupled enzyme assay described by Norby (1971) was adapted to measure galactokinase activity. This enzyme-linked assay was used to investigate the kinetics of purified Gal1p and Gal3p+SA (Chapter 3). Gal3p+SA was found to complement a *GAL1* deletion *in vivo*, but displayed a lower galactokinase activity compared to Gal1p. The calculation of Michaelis constants, K_m, of the two enzymes revealed that while Gal1p and Gal3p+SA had comparable values with respect to galactose, there were significant differences observed between the two enzymes with respect to ATP. Gal3p+SA required a greater concentration of ATP to work at half the maximum rate, and has a lower catalytic constant (k_{cat}) value, indicating that it turns over galactose and ATP at a lower rate. This suggests that the insertion produced a catalytic site defective in ATP binding or hydrolysis. This may suggest that these residues play a role in ATP interaction in Gal1p. Further experiments to investigate the location of the ATP binding site of Gal1p were unsuccessful (Chapter 4). Recently published work has shown that a GHMP kinase protein, homoserine kinase, binds ATP *via* the PX_3GSSAA motif, forming a novel ATP binding loop which does not utilise the side chain of the amino acid lysine in ATP binding (Krishna *et al.*, 2001). This explains the lack of success using ATP analogues in the experiments presented here (Chapter 4).

The ATP analogues used both interact with ATP binding sites *via* lysine side chains (Abraham *et al.*, 1983; Czarnecki *et al.*, 1979). If Gal1p binds ATP/ADP in a similar manner to homoserine kinase using the GLSSA motif, then this would preclude covalent labelling of Gal1p with oATP or 8-azido-ATP. The homoserine kinase structure may also explain why the addition of two amino acids to Gal3p confers galactokinase activity on this protein. Without these two amino acids, the ATP-binding loop would be truncated, perhaps not allowing either the correct positioning of the phosphate groups or the prolonged ATP binding required for enzyme action. Many fungi, for example *K. lactis*, contain just one protein that induces the genes for galactose metabolism and also has galactokinase activity. The two proteins found in *S. cerevisiae* are probably the result of a duplication of the genome that occurred 100 million years ago (Wolfe and Shields, 1997). This would suggest that the Gal3p lost galactokinase activity, while retaining inducer activity. Gal1p is an impaired inducer compared to Gal3p *in vitro* (Platt and Reece, 1998). This suggests that after duplication, the two genes evolved separate functions; one as a transcriptional inducer, the other as a galactokinase.

In vitro experiments to examine the interactions between the proteins of the *GAL* genetic switch, while unsuccessful in identifying interactions between Gal80p and Gal3p, did produce an interesting result (Chapter 4). Crosslinks were observed between Gal80p and Gal4p, a well documented interaction (Ma and Ptashne, 1987a). The results suggested that the repressor Gal80p appeared to be in a different conformation in solution when compared to the super-repressor Gal80^{S-2}p, which contains the single point mutation Glu351Lys. This result raises the possibility of Gal80p existing in two conformations, one binding Gal4p and the other binding Gal3p. This may explain why yeast cells expressing Gal80p^{S-2}p are unable to activate transcription of the *GAL* genes;

the protein may not be able to form the Gal3p-binding conformation. Gal3p and Gal80p were not found to crosslink *in vitro*, although subsequently these proteins have been crosslinked using another chemical crosslinker (Timson *et al.*, 2002). The original aims of these experiments included identifying the whether the three proteins of the *GAL* genetic switch formed a three-way complex, as observed previously by Platt and Reece (1998) using band shift assays. Recent work using ChIP assays, which identify proteins bound at a DNA sequence, have found that when yeast are grown on galactose, the amount of Gal80p present at the promoter decreases. Gal3p is not seen at the promoter (Peng and Hopper, 2002). These workers suggest that Gal3p is located outside the nucleus, and on addition of galactose to the media, Gal3p sequesters Gal80p and prevents access to the nucleus and repression of the *GAL* genes. This model of activation is shown in Figure 6.1. The localisation of Gal4p may also play a role in the control of the *GAL* genes, although localisation experiments would need to be carried out to investigate this possibility.

Experiments were designed to follow the transcription of the *GAL* genes *in vivo* and for the most part results agreed with the published data (Chapter 5). These experiments did not use a standard reporter gene, but instead measured the expression of *GAL10*. The data produced from an experiment on yeast cells deleted for the inducer *GAL3* showed a low level transcription after galactose was added to the growth medium. Galactose was added to the media after the 0 hour time point, the 24 hour time point and the 48 hour time point. The level of transcription was substantially lower than activated transcription observed in wild-type yeast, but appeared higher than the levels of transcription observed when neither Gal1p nor Gal3p were present. These data suggest that the LTA phenotype is not caused by a sudden switching on of the *GAL* genes, but

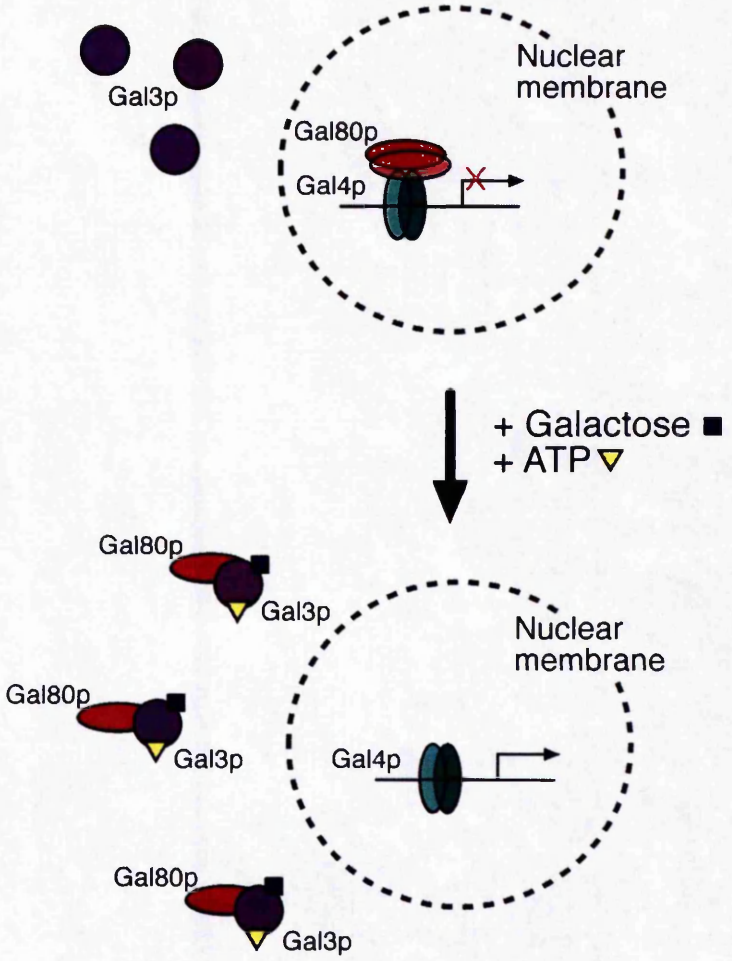


Figure 6.1 Diagrammatical representation of the induction of the *GAL* genes

Current evidence points towards the induction of the *GAL* genes being due to Gal3p, in the presence of galactose and ATP, retaining the repressor Gal80p outside the nucleus. This allows Gal4p to contact the transcriptional machinery and activate gene expression.

rather a post-transcriptional mechanism that occurs after the protein levels have reached a sufficient concentration. Gal1p would need to overcome this barrier, as a greater concentration of the protein is required to activate the *GAL* genes. This would result in a feedback loop at the *GAL1* promoter, and an increase in transcription would be expected after a time-lag. These results may be an artefact of this experiment, or they may indicate that a hitherto unknown interaction involved in transcription of the *GAL* genes in the absence of Gal3p mediated this activity. The increase in growth may be due to a gradual increase of transcription of the *GAL* genes rather than a sudden rise. Further investigation into Gal1p-mediated transcription using plasmid-borne *GAL1* tested the role of the promoter of the inducer gene in the transcription of the *GAL10* gene. When the transcription of the inducer protein was controlled by the *GAL1* promoter, the level of transcription remained high, even after galactose concentrations had presumably dropped due to conversion to galactose-1-phosphate. As discussed in Chapter 5, this may be due to the phosphorylation of Gal4p, which provides a second mechanism for active transcription of the *GAL* genes in low galactose concentrations (Rohde *et al.*, 2000). As the *GAL1* promoter contains more Gal4p binding sites than the *GAL3* promoter (Johnston, 1987), this phosphorylation mechanism may form a greater degree of control of the *GAL1* gene than the *GAL3* gene. The *GAL3* promoter appeared to be much more responsive to galactose; transcription of *GAL10* was activated quickly, and dropped at lower galactose levels. This was independent of the identity of the inducer gene, either *GAL1* or *GAL3*. The differences between the two promoters agree with the function of their respective genes. Gal3p is required to activate the *GAL* genes rapidly with sensitivity to the galactose concentration in the cell. Conversely, Gal1p is required to be transcribed at high levels in response to galactose to allow the cell to metabolise the sugar. Interestingly, when galactose levels drop in wild type yeast, the transcription

of the *GAL10* gene also displays a lower rate; Gal1p is not activating transcription when Gal3p is present.

In a wider context, recent work on the *GAL* system has concentrated on two areas. The first, the mechanism of induction, is discussed above. The second area of study involves looking at the interactions of the protein of the *GAL* genetic switch. Recently it has been discovered that SAGA, the histone acetyltransferase complex, functions as a co-activator of transcription by Gal4p at the *GAL1* promoter (Bhaumik and Green, 2001; Larschan and Winston, 2001). These experiments were carried out using ChIP assays, and found that both the Gal4p activation domain and induction by galactose was required for SAGA recruitment. If SAGA was not recruited, transcription did not occur, indicating the requirement for this protein complex for transcription of the *GAL1* gene. Further work has shown that the repressor Gal80p blocks the interaction between SAGA and Gal4p, but not the interaction of another activator Gcn4p with SAGA, indicating that the repression is specific to the *GAL* genetic switch. Gal80p could specifically replace SAGA at the *GAL1* promoter on a nucleosome array (Carrozza *et al.*, 2002).

6.2 Future Directions

Work presented here has shown that the homology region III of Gal1p is involved with ATP binding and catalysis. The insertion of two amino acids into this region in Gal3p converts the inducer into a galactokinase. Structural data of the protein would confirm if Gal1p binds to ATP at this region, as homoserine kinase and mevalonate kinase have been found to do (Fu *et al.*, 2002; Krishna *et al.*, 2001). Initial studies have found that Gal1p tends to aggregate at high concentration, which presents a problem for formation

of crystals (S. Hankin, personal communication). Structural data on galactokinase from another organism may lead to identification of the ATP binding site and active site of the enzyme, which could be compared to the sequence data on Gal1p and Gal3p. Non-structural work on the conserved motifs other that homology region III may lead to putative functions being proposed for these sequences, *e.g.* galactose binding. Gal3p+SA has the same K_m value as Gal1p, therefore to identify the interacting sequence would involve creating mutations to produce a partial loss-of-function phenotype, a difficult process. Structural studies on galactokinase, with either ligand bound to the protein would obviate the need for these experiments.

Identification of the mechanism for the activation of transcription has been hampered by various factors. While it has been shown that Gal80p interacts with Gal3p, and this interaction causes alleviation of the *GAL* gene repression, the site of this interaction is not yet clear. The most recent published work, using ChIP assays to show that when the *GAL* genes are activated, Gal80p is no longer bound at the DNA, provides the most compelling evidence that Gal3p causes Gal80p to cease interaction and repression of Gal4p (Peng and Hopper, 2002). It has been proposed that this interaction occurs outside the nucleus, and that Gal3p remains cytoplasmic, but the mechanism by which this happens is unclear. Transcription of the *GAL* genes occurs quickly, within 1-2 hours of addition of galactose, as seen in Chapter 5. If Gal3p does not enter the nucleus to sequester Gal80p, then Gal80p must be in rapid equilibrium between the nucleus and the cytoplasm to allow the rapid induction to take place. *In vivo* experiments using specific Gal3p antibodies, that do not recognise Gal1p, must be carried out, and the localisation of Gal3p and Gal80p in induced and uninduced cells, and at wild type levels, must be determined. Understanding of the location of each protein in the induced

as well as the uninduced state would allow a clearer understanding of the mechanism of action of Gal3p.

In vivo experiments to follow the transcription of the *GAL* genes have shown the RT-PCR method to be a reliable way of following activation. However, the experiments were carried out using yeast grown in batch culture in flasks, and this may have led to interference in the pattern of transcription from the growth state of the yeast. In these experiments, this was minimised by using an internal control. In flask culture it is impossible to keep the galactose levels constant, or to control the growth state of the yeast. Carrying out these experiments using chemostat cultures, synchronising the yeast cells and keeping the concentration of nutrients steady, would likely yield more consistent results. Studies on the phosphorylation state of Gal4p at different stages of these experiments may also lead to insights into the levels of phosphorylation observed at the latter stages of the experiments. The transcription remained high when the inducer protein was under the control of the *GAL1* promoter. Knowledge of the phosphorylation state of Gal4p bound to this promoter compared to Gal4p bound to the Gal3p promoter in the corresponding experiments would allow assignation of the mechanism for sustained transcription, even at low galactose levels.

Recently there has been a large increase in the knowledge of transcription, gained form a variety of *in vivo* and *in vitro* experiments in a host of different eukaryotes. However, mechanisms of transcription remain imperfectly understood. The work presented here helps in the understanding of the paradigm eukaryotic transcription system in a model eukaryote. The techniques and models developed here may prove useful to the investigation of other genetic switches in eukaryotes.

References

Abraham, K.I., Haley, B. and Modak, M.J. (1983) Biochemistry of terminal deoxynucleotidyltransferase: characterization and properties of photoaffinity labeling with 8-azidoadenosine 5'-triphosphate. *Biochemistry*, **22**, 4197-4203.

Alexiadis, V., Varga-Weisz, P.D., Bonte, E., Becker, P.B. and Gruss, C. (1998) *In vitro* chromatin remodelling by chromatin accessibility complex (CHRAC) at the SV40 origin of DNA replication. *EMBO J.*, **17**, 3428-3438.

Ali, J.A., Jackson, A.P., Howells, A.J. and Maxwell, A. (1993) The 43-kilodalton N-terminal fragment of the DNA gyrase B protein hydrolyzes ATP and binds coumarin drugs. *Biochemistry*, **32**, 2717-2724.

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.

Auble, D.T., Wang, D., Post, K.W. and Hahn, S. (1997) Molecular analysis of the SNF2/SWI2 protein family member MOT1, an ATP-driven enzyme that dissociates TATA-binding protein from DNA. *Mol. Cell. Biol.*, **17**, 4842-4851.

Bajwa, W., Torchia, T. and Hopper, J.E. (1988) Yeast regulatory gene *GAL3*: Carbon regulation; UAS_{Gal} elements in common with *GAL1, GAL2, GAL7, GAL10, GAL80*, and *MEL1*; encoded protein strikingly similar to yeast and *Escherichia coli* galactokinases. *Mol. Cell. Biol.*, **8**, 3439-3447.

Balciunas, D., Galman, C., Ronne, H. and Bjorklund, S. (1999) The Med1 subunit of the yeast mediator complex is involved in both transcriptional activation and repression. *Proc. Natl. Acad. Sci. USA*, **96**, 376-381.

Barbaric, S., Münsterkötter, M., Svaren, J. and Hörz, W. (1996) The homeodomain protein Pho2 and the basic-helix-loop-helix protein Pho4 bind DNA cooperatively at the yeast *PHO5* promoter. *Nucleic Acids Res.*, **24**, 4479-4486.

Barberis, A., Pearlberg, J., Simkovich, N., Farrell, S., Reinagel, P., Bamdad, C., Sigal, G. and Ptashne, M. (1995) Contact with a component of the polymerase II holoenzyme suffices for gene activation. *Cell*, **81**, 359-368.

Berben, G., Legrain, M. and Hilger, F. (1988) Studies on the structure, expression and function of the yeast regulatory gene *PHO2*. *Gene*, **66**, 307-312.

Bhat, P.J. and Hopper, J.E. (1992) Overproduction of the Gal1 or Gal3 protein causes galactose-independent activation of the Gal4 protein: evidence for a new model of induction for the yeast *GAL/MEL* regulon. *Mol. Cell. Biol.*, **12**, 2701-2707.

Bhat, P.J., Oh, D. and Hopper, J.E. (1990) Analysis of the *GAL3* signal transduction pathway activating *GAL4* protein-dependent transcription in *Saccharomyces cerevisiae*. *Genetics*, **125**, 281-291.

Bhaumik, S.R. and Green, M.R. (2001) SAGA is an essential *in vivo* target of the yeast acidic activator Gal4p. *Genes Dev.*, **15**, 1935-1945.

Bird, A. (2001) Molecular biology. Methylation talk between histones and DNA. *Science*, **294**, 2113-2115.

Bork, P., Sander, C. and Valencia, A. (1993) Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. *Protein Sci.*, **2**, 31-40.

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

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

Brehm, A., Langst, G., Kehle, J., Clapier, C.R., Imhof, A., Eberharter, A., Muller, J. and Becker, P.B. (2000) dMi-2 and ISWI chromatin remodelling factors have distinct nucleosome binding and mobilization properties. *EMBO J.*, **19**, 4332-4341.

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

Broach, J.R. (1979) Galactose regulation in *Saccharomyces cerevisiae*. The enzymes encoded by the *GAL7, 10, 1* cluster are co-ordinately controlled and separately translated. *J. Mol. Biol.*, **131**, 41-53.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y. and Allis, C.D. (1996) Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*, **84**, 843-851.

Bullock, W.O., Fernandez, J.M. and Short, J.M. (1987) Xl1-Blue - a high-efficiency plamid transforming *RecA Escherachia coli* strain with beta-galactosidase selection. *Biotechniques*, **5**, 376-379.

Buratowski, S., Hahn, S., Guarente, L. and Sharp, P.A. (1989) Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell*, **56**, 37-42.

Burke, T.W. and Kadonaga, J.T. (1996) *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev.*, **10**, 711-724.

Cairns, B.R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B. and Kornberg, R.D. (1996) RSC, an essential, abundant chromatin-remodeling complex. *Cell*, **87**, 1249-1260.

Candau, R., Moore, P.A., Wang, L., Barlev, N., Ying, C.Y., Rosen, C.A. and Berger, S.L. (1996) Identification of human proteins functionally conserved with the yeast putative adaptors *ADA2* and *GCN5*. *Mol. Cell. Biol.*, **16**, 593-602.

Candau, R., Zhou, J.X., Allis, C.D. and Berger, S.L. (1997) Histone acetyltransferase activity and interaction with ADA2 are critical for GCN5 function in vivo. *EMBO J.*, **16**, 555-565.

Carey, M., Kakidani, H., Leatherwood, J., Motashari, F. and Ptashne, M. (1989) An amino-terminal fragment of *GAL4* binds DNA as a dimer. *J. Mol. Biol.*, **209**, 423-432.

Carlson, M. (1997) Genetics of transcriptional regulation in yeast: connections to the RNA polymerase II CTD. *Annual Rev. Cell Dev. Biol.*, **13**, 1-23.

Carrozza, M.J., John, S., Sil, A.K., Hopper, J.E. and Workman, J.L. (2002) Gal80 confers specificity on HAT complex interactions with activators. *J. Biol. Chem.*, **277**, 24648-24652.

Carruthers, L.M. and Hansen, J.C. (2000) The core histone N termini function independently of linker histones during chromatin condensation. *J. Biol. Chem.*, **275**, 37285-37290.

Chasman, D.I., Flaherty, K.M., Sharp, P.A. and Kornberg, R.D. (1993) Crystal structure of yeast TATA-binding protein and model for interaction with DNA. *Proc. Natl. Acad. Sci. USA*, **90**, 8174-8178.

Chaves, R.S., Herrero, P. and Moreno, F. (1999) Med8, a subunit of the mediator CTD complex of RNA polymerase II, directly binds to regulatory elements of *SUC2* and *HXK2* genes. *Biochem. Biophys. Res. Commun.*, **254**, 345-350.

Chen, H., Lin, R.J., Schiltz, R.L., Chakravarti, D., Nash, A., Nagy, L., Privalsky, M.L., Nakatani, Y. and Evans, R.M. (1997) Nuclear receptor coactivator ACTR is a

novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell*, **90**, 569-580.

Clarke, A.S., Lowell, J.E., Jacobson, S.J. and Pillus, L. (1999) Esa1p is an essential histone acetyltransferase required for cell cycle progression. *Mol. Cell. Biol.*, **19**, 2515-2526.

Coin, F. and Egly, J.M. (1998) Ten years of TFIIH. *Cold Spring Harb. Symp. Quant. Biol.*, **63**, 105-110.

Conaway, R.C. and Conaway, J.W. (1993) General initiation factors for RNA polymerase II. *Annu. Rev. Biochem.*, **62**, 161-190.

Corden, J.L. (1990) Tails of RNA polymerase II. *Trends Biochem. Sci.*, **15**, 383-387.

Cormack, B.P. and Struhl, K. (1992) The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. *Cell*, **69**, 685-696.

Cornish-Bowden, A. (1995) *Fundamentals of enzyme kinetics*. Portland Press, London, UK.

Cosma, M.P., Tanaka, T. and Nasmyth, K. (1999) Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell*, **97**, 299-311.

Coulombe, B. (1999) DNA wrapping in transcription initiation by RNA polymerase II. *Biochem. Cell. Bio.*, **77**, 257-264.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R. and Kornberg, R.D. (2000) Architecture of RNA polymerase II and implications for the transcription mechanism. *Science*, **288**, 640-649.

Cramer, P., Bushnell, D.A. and Kornberg, R.D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 Å resolution. *Science*, **292**, 1863-1876.

Cui, Y. and Bustamante, C. (2000) Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure. *Proc. Natl. Acad. Sci. USA*, **97**, 127-132.

Czarnecki, J., Geahlen, R. and Haley, B. (1979) Synthesis and use of azido photoaffinity analogs of adenine and guanine nucleotides. *Methods Enzymol*, **56**, 642-653.

Dahmus, M.E. (1996) Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J. Biol. Chem.*, **271**, 19009-19012.

Darst, S.A., Kubalek, E.W. and Kornberg, R.D. (1989) Three-dimensional structure of *Escherichia coli* RNA polymerase holoenzyme determined by electron crystallography. *Nature*, **340**, 730-732.

De Vit, M.J., Waddle, J.A. and Johnston, M. (1997) Regulated nuclear translocation of the Mig1 glucose repressor. *Mol. Biol. Cell.*, **8**, 1603-1618.

Durrin, L.K., Mann, R.K. and Grunstein, M. (1992) Nucleosome loss activates *CUP1* and *HIS3* promoters to fully induced levels in the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **12**, 1621-1629.

Easterbrook-Smith, S.B., Wallace, J.C. and Keech, D.B. (1976) Pyruvate carboxylase: Affinity labelling of the magnesium adenosine triphosphate binding site. *Eur. J. Biochem.*, **62**, 125-130.

Eisen, J.A., Sweder, K.S. and Hanawalt, P.C. (1995) Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res.*, **23**, 2715-2723.

Fang, S.M. and Burton, Z.F. (1996) RNA polymerase II-associated protein (RAP) 74 binds transcription factor (TF) IIB and blocks TFIIB-RAP30 binding. *J. Biol. Chem.*, **271**, 11703-11709.

Flaus, A. and Owen-Hughes, T. (2001) Mechanisms for ATP-dependent chromatin remodelling. *Curr. Opin. Genet. Dev.*, **11**, 148-154.

Flores, O., Lu, H. and Reinberg, D. (1992) Factors involved in specific transcription initiation by RNA polymerase II: identification and characterization of factor IIH. *J. Biol. Chem.*, **267**, 2786-2793.

Flynn, P.J. and Reece, R.J. (1999) Activation of transcription by metabolic intermediates of the pyrimidine biosynthetic pathway. *Mol. Cell. Biol.*, **19**, 882-888.

Frolova, E., Johnston, M. and Majors, J. (1999) Binding of the glucose-dependent Mig1p repressor to the *GAL1* and *GAL4* promoters *in vivo*: regulation by glucose and chromatin structure. *Nucleic Acids Res.*, **27**, 1350-1358.

Fry, C.J. and Peterson, C.L. (2001) Chromatin remodeling enzymes: who's on first? *Curr. Biol.*, **11**, R185-197.

Fu, Z., Wang, M., Potter, D., Miziorko, H.M. and Kim, J.J. (2002) The structure of a binary complex between a mammalian mevalonate kinase and ATP: insights into the reaction mechanism and human inherited disease. *J. Biol. Chem.*, **277**, 18134-18142.

Georgakopoulos, T. and Thireos, G. (1992) Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *EMBO J.*, **11**, 4145-4152.

Gill, G. and Ptashne, M. (1987) Mutants of Gal4 protein altered in an activation function. *Cell*, **51**, 121-126.

Grant, P.A., Duggan, L., Cote, J., Roberts, S.M., Brownell, J.E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C.D., Winston, F., Berger, S.L. and Workman, J.L. (1997) Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev.*, **11**, 1640-1650.

Grant, P.A., Sterner, D.E., Duggan, L.J., Workman, J.L. and Berger, S.L. (1998) The SAGA unfolds: convergence of transcription regulators in chromatin-modifying complexes. *Trends Cell. Biol.*, **8**, 193-197.

Greenblatt, J. (1991) RNA polymerase-associated transcription factors. *Trends Biochem. Sci.*, **16**, 408-411.

Gregory, P.D., Schmid, A., Zavari, M., Munsterkotter, M. and Horz, W. (1999) Chromatin remodelling at the PHO8 promoter requires SWI-SNF and SAGA at a step subsequent to activator binding. *Embo J*, **18**, 6407-6414.

Guschin, D., Wade, P.A., Kikyo, N. and Wolffe, A.P. (2000) ATP-Dependent histone octamer mobilization and histone deacetylation mediated by the Mi-2 chromatin remodeling complex. *Biochemistry*, **39**, 5238-5245.

Guzman, E. and Lis, J.T. (1999) Transcription factor TFIIH is required for promoter melting *in vivo*. *Mol. Cell. Biol.*, **19**, 5652-5658.

Hames, B.D. (1981) An introduction to polyacrylamide gel electrophoresis. In Rickwood, D. (ed.), *Gel electrophoresis of proteins. A practical approach*. IRL Press., Oxford and Washington DC., pp. 1-91.

Hamiche, A., Sandaltzopoulos, R., Gdula, D.A. and Wu, C. (1999) ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. *Cell*, **97**, 833-842.

Han, M. and Grunstein, M. (1988) Nucleosome loss activates yeast downstream promoters *in vivo*. *Cell*, **55**, 1137-1145.

Han, S.J., Lee, Y.C., Gim, B.S., Ryu, G.H., Park, S.J., Lane, W.S. and Kim, Y.J. (1999) Activator-specific requirement of yeast mediator proteins for RNA polymerase II transcriptional activation. *Mol. Cell. Biol.*, **19**, 979-988.

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

Hanes, C.S. (1932) Studies on plant amylases. *Biochem. J.*, **26**, 1406-1421.

Harlow, E. and Lane, D. (1988) *Antibodies - A laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Hengartner, C.J., Thompson, C.M., Zhang, J., Chao, D.M., Liao, S.M., Koleske, A.J., Okamura, S. and Young, R.A. (1995) Association of an activator with an RNA polymerase II holoenzyme. *Genes Dev.*, **9**, 897-910.

Herrera, J.E., Schiltz, R.L. and Bustin, M. (2000a) The accessibility of histone H3 tails in chromatin modulates their acetylation by P300/CBP-associated factor. *J. Biol. Chem.*, **275**, 12994-12999.

Herrera, J.E., West, K.L., Schiltz, R.L., Nakatani, Y. and Bustin, M. (2000b) Histone H1 is a specific repressor of core histone acetylation in chromatin. *Mol. Cell. Biol.*, **20**, 523-529.

Hirst, K., Fisher, F., McAndrew, P.C. and Goding, C.R. (1994) The transcription factor, the Cdk, its cyclin and their regulator: directing the transcriptional response to a nutritional signal. *EMBO J.*, **13**, 5410-5420.

Hirst, M., Kobor, M.S., Kuriakose, N., Greenblatt, J. and Sadowski, I. (1999) *GAL4* is regulated by the RNA polymerase II holoenzyme-associated cyclin- dependent protein kinase *SRB10/CDK8*. *Mol. Cell*, **3**, 673-678.

Holstege, F.C., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, E.S. and Young, R.A. (1998) Dissecting the regulatory circuitry of a eukaryotic genome. *Cell*, **95**, 717-728.

Holstege, F.C.P., Tantin, D., Carey, M., Vandervliet, P.C. and Timmers, H.T.M. (1995) The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. *EMBO J.*, **14**, 810-819.

Hopwood, D. (1969) An investigation in the cross-linking of proteins by various fixatives. *J. Anat.*, **105**, 196.

Imai, S., Armstrong, C.M., Kaeberlein, M. and Guarente, L. (2000) Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature*, **403**, 795-800.

Imbalzano, A.N., Zaret, K.S. and Kingston, R.E. (1994) Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA. *J. Biol. Chem.*, **269**, 8280-8286.

Ishihama, A. (1993) Protein-protein communication within the transcription apparatus. *J. Bacteriol.*, **175**, 2483-2489.

Ito, T., Bulger, M., Pazin, M.J., Kobayashi, R. and Kadonaga, J.T. (1997) ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell*, **90**, 145-155.

Jackson-Fisher, A.J., Chitikila, C., Mitra, M. and Pugh, B.F. (1999) A role for TBP dimerization in preventing unregulated gene expression. *Mol. Cell*, **3**, 717-727.

Jaskelioff, M., Gavin, I.M., Peterson, C.L. and Logie, C. (2000) SWI-SNF-mediated nucleosome remodeling: role of histone octamer mobility in the persistence of the remodeled state. *Mol Cell Biol*, **20**, 3058-3068.

Johnston, M. (1987) A model fungal gene regulatory mechanism: The *GAL* genes of *Saccharomyces cerevisiae*. *Microbiol. Rev.*, **51**, 458-476.

Johnston, M., Flick, J.S. and Pexton, T. (1994) Multiple mechanisms provide rapid and stringent glucose repression of *GAL* gene expression in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **14**, 3834-3841.

Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J. and Wolffe, A.P. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.*, **19**, 187-191.

Kaffman, A., Herowitz, I., Tjian, R. and O'Shea, E.K. (1994) Phosphorylation of the transcription factor *PHO4* by a cyclin-CDK complex, *PHO80-PHO85*. *Science*, **263**, 1153-1156.

Kammerer, B., Guyonvarch, A. and Hubert, J.C. (1984) Yeast regulatory gene *PPR1*. Nucleotide sequence, restriction map and codon usage. *J. Mol. Biol.*, **180**, 239-250.

Katritch, V., Bustamante, C. and Olson, W.K. (2000) Pulling chromatin fibers: computer simulations of direct physical micromanipulations. *J. Mol. Biol.*, **295**, 29-40.

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

Kim, J.L., Nikolov, D.B. and Burley, S.K. (1993) Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*, **365**, 520-527.

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.

Kimmelman, J., Kaldis, P., Hengartner, C.J., Laff, G.M., Koh, S.S., Young, R.A. and Solomon, M.J. (1999) Activating phosphorylation of the Kin28p subunit of yeast TFIIH by Cak1p. *Mol. Cell. Biol.*, **19**, 4774-4787.

Kingston, R.E. and Narlikar, G.J. (1999) ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev.*, **13**, 2339-2352.

Kinoshita, T., Imamura, J., Nagai, H. and Shimotohno, K. (1992) Quantification of gene expression over a wide range by the polymerase chain reaction. *Anal. Biochem.*, **206**, 231-235.

Knoepfler, P.S. and Eisenman, R.N. (1999) Sin meets NuRD and other tails of repression. *Cell*, **99**, 447-450.

Kobayashi, N., Boyer, T.G. and Berk, A.J. (1995) A class of activation domains interacts directly with TFIIA and stimulates TFIIA-TFIID-promoter complex assembly. *Mol. Cell. Biol.*, **15**, 6465-6473.

Kobor, M.S., Archambault, J., Lester, W., Holstege, F.C., Gileadi, O., Jansma, D.B., Jennings, E.G., Kouyoumdjian, F., Davidson, A.R., Young, R.A. and Greenblatt, J. (1999) An unusual eukaryotic protein phosphatase required for transcription by RNA polymerase II and CTD dephosphorylation in *S. cerevisiae*. *Mol. Cell*, **4**, 55-62.

Koh, S.S., Ansari, A.Z., Ptashne, M. and Young, R.A. (1998) An activator target in the RNA polymerase II holoenzyme. *Mol. Cell*, **1**, 895-904.

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

Koleske, A.J., Buratowski, S., Nonet, M. and Young, R.A. (1992) A novel transcription factor reveals a functional link between the RNA polymerase II CTD and TFIID. *Cell*, **69**, 883-894.

Komeili, A. and O'Shea, E.K. (1999) Roles of phosphorylation sites in regulating activity of the transcription factor Pho4p. *Science*, **284**, 977-980.

Kornberg, R.D. (1975) Chromatin structure: A repeating unit of histones and DNA. *Science*, **184**, 868-871.

Kornberg, R.D. and Lorch, Y. (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell*, **98**, 285-294.

Krebs, J.E., Fry, C.J., Samuels, M.L. and Peterson, C.L. (2000) Global role for chromatin remodeling enzymes in mitotic gene expression. *Cell*, **102**, 587-598.

Krebs, J.E., Kuo, M.H., Allis, C.D. and Peterson, C.L. (1999) Cell cycle-regulated histone acetylation required for expression of the yeast *HO* gene. *Genes Dev.*, **13**, 1412-1421.

Krishna, S.S., Zhou, T., Daugherty, M., Osterman, A. and Zhang, H. (2001) Structural basis for the catalysis and substrate specificity of homoserine kinase. *Biochemistry*, **40**, 10810-10818.

Kuldell, N.H. and Buratowski, S. (1997) Genetic analysis of the large subunit of yeast transcription factor IIE reveals two regions with distinct functions. *Mol. Cell. Biol.*, **17**, 5288-5298.

Kumar, A., Agarwal, S., Heyman, J.A., Matson, S., Heidtman, M., Piccirillo, S., Umansky, L., Drawid, A., Jansen, R., Liu, Y., Cheung, K.H., Miller, P., Gerstein, M., Roeder, G.S. and Snyder, M. (2002) Subcellular localization of the yeast proteome. *Genes Dev.*, **16**, 707-719.

Kuo, M.H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G., Roth, S.Y. and Allis, C.D. (1996) Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature*, **383**, 269-272.

Kuo, M.H., vom Baur, E., Struhl, K. and Allis, C.D. (2000) Gcn4 activator targets Gcn5 histone acetyltransferase to specific promoters independently of transcription. *Mol Cell*, **6**, 1309-1320.

Kuo, M.H., Zhou, J., Jambeck, P., Churchill, M.E. and Allis, C.D. (1998) Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes *in vivo*. *Genes Dev.*, **12**, 627-639.

Lamphier, M.S. and Ptashne, M. (1992) Multiple mechanisms mediate glucose repression of the yeast *GAL1* gene. *Proc. Natl. Acad. Sci. USA*, **89**, 5922-5926.

Langst, G., Bonte, E.J., Corona, D.F. and Becker, P.B. (1999) Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. *Cell*, **97**, 843-852.

Larschan, E. and Winston, F. (2001) The *S. cerevisiae* SAGA complex functions *in vivo* as a coactivator for transcriptional activation by Gal4. *Genes Dev.*, **15**, 1946-1956.

Lee, T.I., Wyrick, J.J., Koh, S.S., Jennings, E.G., Gadbois, E.L. and Young, R.A. (1998) Interplay of positive and negative regulators in transcription initiation by RNA polymerase II holoenzyme. *Mol. Cell. Biol.*, **18**, 4455-4462.

Leuther, K.K., Bushnell, D.A. and Kornberg, R.D. (1996) Two-dimensional crystallography of TFIIB- and IIE-RNA polymerase II complexes: implications for start site selection and initiation complex formation. *Cell*, **85**, 773-779.

Leuther, K.K. and Johnston, S.A. (1992) Nondissociation of *GAL4* and *GAL80 in vivo* after galactose induction. *Science*, **256**, 1333-1335.

Liao, S.M., Zhang, J., Jeffery, D.A., Koleske, A.J., Thompson, C.M., Chao, D.M., Viljoen, M., van Vuuren, H.J. and Young, R.A. (1995) A kinase-cyclin pair in the RNA polymerase II holoenzyme. *Nature*, **374**, 193-196.

Lin, Y.S., Ha, I., Maldonado, E., Reingerg, D. and Green, M.R. (1991) Binding of general transcription factor TFIIB to an acidic activating region. *Nature*, **353**, 569-571.

Lineweaver, H. and Burk, D. (1934) The determination of enzyme constants. *J. Amer. Chem. Soc.*, **56**, 658-666.

Lu, H., Zawel, L., Fisher, L., Egly, J.M. and Reinberg, D. (1992) Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature*, **358**, 641-645.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8Å resolution. *Nature*, **389**, 251-260.

Ma, J. and Ptashne, M. (1987a) The carboxy-terminal 30 amino acids of Gal4p are recognized by Gal80p. *Cell*, **50**, 137-142.

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

Ma, J. and Ptashne, M. (1988) Converting a eukaryotic transcriptional inhibitor into an activator. *Cell*, **55**, 443-446.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular cloning - A laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Marmorstein, R., Carey, M., Ptashne, M. and Harrison, S.C. (1992) DNA recognition by GAL4: Structure of a protein-DNA complex. *Nature*, **356**, 408-414.

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

Martinez, E., Zhou, Q., L'Etoile, N.D., Oelgeschlager, T., Berk, A.J. and Roeder, R.G. (1995) Core promoter-specific function of a mutant transcription factor TFIID defective in TATA-box binding. *Proc. Natl. Acad. Sci. USA*, **92**, 11864-11868.

Matsui, T., Segall, J., Weil, P.A. and Roeder, R.G. (1980) Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J. Biol. Chem.*, **255**, 11992-11996.

Melcher, K. and Xu, H.E. (2001) Gal80-Gal80 interaction on adjacent Gal4p binding sites is required for complete *GAL* gene repression. *EMBO J.*, **20**, 841-851.

Meyer, J., Walker-Jonah, A. and Hollenberg, C.P. (1991) Galactokinase encoded by *GAL1* is a bifunctional protein required for induction of the *GAL* genes in

Kluyveromyces lactis and is able to suppress the *gal3* pheontype in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **11**, 5454-5461.

Michealis, L. and Menten, M. (1913) The kinetics of invertase. *Biochem J.*, **49**, 333-369.

Mizzen, C.A., Yang, X.J., Kokubo, T., Brownell, J.E., Bannister, A.J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S.L., Kouzarides, T., Nakatani, Y. and Allis, C.D. (1996) The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell*, **87**, 1261-1270.

Myer, V.E. and Young, R.A. (1998) RNA polymerase II holoenzymes and subcomplexes. *J. Biol. Chem.*, **273**, 27757-27760.

Myers, L.C., Gustafsson, C.M., Bushnell, D.A., Lui, M., Erdjument-Bromage, H., Tempst, P. and Kornberg, R.D. (1998) The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain. *Genes Dev.*, **12**, 45-54.

Myers, L.C., Gustafsson, C.M., Hayashibara, K.C., Brown, P.O. and Kornberg, R.D. (1999) Mediator protein mutations that selectively abolish activated transcription. *Proc. Natl. Acad. Sci. USA*, **96**, 67-72.

Mylin, L.M., Bhat, P.J. and Hopper, J.E. (1989) Regulated phosphorylation and dephosphorylation of GAL4, a transcriptional activator. *Genes Dev.*, **3**, 1157-1165.

Nan, X., Campoy, F.J. and Bird, A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*, **88**, 471-481.

Niu, W., Kim, Y., Tau, G., Heyduk, T. and Ebright, R.H. (1996) Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase. *Cell*, **87**, 1123-1134.

Nogi, Y. and Fukasawa, T. (1984) Nucleotide sequence of the yeast regulatory gene *GAL80*. *Nucleic Acids Res.*, **12**, 9287-9298.

Nogi, Y. and Fukasawa, T. (1989) Functional domains of a negative regulatory protein, GAL80, of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **9**, 3009-3017.

Nonet, M.L. and Young, R.A. (1989) Intragenic and extragenic suppressors of mutations in the heptapeptide repeat domain of *Saccharomyces cerevisiae* RNA polymerase II. *Genetics*, **123**, 715-724.

Norby, J.G. (1971) Studies on a coupled enzyme assay for rate measurements of ATPase reactions. *Acta. Chem. Scand.*, **25**, 2717-2726.

Ogawa, N., Noguchi, K., Yamashita, Y., Yasuhara, T., Hayashi, N., Yoshida, K. and Oshima, Y. (1993) Promoter analysis of the *PHO81* gene encoding a 134-kDa protein bearing ankyrin repeats in the phosphatase regulon of *Sacchormoyces cerevisiae*. *Mol. Gen. Genet.*, **238**, 444-454.

Ohkuma, Y. and Roeder, R.G. (1994) Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature*, **368**, 160-163.

Otero, G., Fellows, J., Li, Y., de Bizemont, T., Dirac, A.M., Gustafsson, C.M., Erdjument-Bromage, H., Tempst, P. and Svejstrup, J.Q. (1999) Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation. *Mol. Cell*, **3**, 109-118.

Owen-Hughes, T., Utley, R.T., Steger, D.J., West, J.M., John, S., Cote, J., Havas, K.M. and Workman, J.L. (1999) Analysis of nucleosome disruption by ATP-driven chromatin remodeling complexes. *Methods Mol. Biol.*, **119**, 319-331.

Pan, G. and Greenblatt, J. (1994) Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site. *J. Biol. Chem.*, **269**, 30101-30104.

Partis, M.D., Griffiths, D.G., Roberts, G.C. and Beechey R.B. (1983) Crosslinking of proteins by ω-maleimido alkanoyl *N*-hydroxysuccinamide esters. *J. Prot. Chem.*, **2** 263-277

Parvin, J.D., Timmers, H.T. and Sharp, P.A. (1992) Promoter specificity of basal transcription factors. *Cell*, **68**, 1135-1144.

Peng, G. and Hopper, J.E. (2000) Evidence for Gal3p's cytoplasmic location and Gal80p's dual cytoplasmic-nuclear location implicates new mechanisms for controlling Gal4p activity in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **20**, 5140-5148.

Peng, G. and Hopper, J.E. (2002) Gene activation by interaction of an inhibitor with a cytoplasmic signaling protein. *Proc. Natl. Acad. Sci. USA*, **99**, 8548-8553.

Peterson, C.L. and Logie, C. (2000) Recruitment of chromatin remodeling machines. *J. Cell. Biochem.*, **78**, 179-185.

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.

Platt, A. and Reece, R.J. (1998) The yeast galactose genetic switch is mediated by the formation of a Gal4p/Gal80p/Gal3p complex. *EMBO J.*, **17**, 4086-4091.

Platt, A., Ross, H.C., Hankin, S. and Reece, R.J. (2000) The insertion of two amino acids into a transcriptional inducer converts it into a galactokinase. *Proc. Natl. Acad. Sci. USA*, **97**, 3154-3159.

Ptashne, M. and Gann, A. (1997) Transcriptional activation by recruitment. *Nature*, **386**, 569-577.

Pugh, B.F. and Tjian, R. (1990) Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell*, **61**, 1187-1197.

Read, C.M., Baldwin, J.P. and Crane-Robinson, C. (1985) Structure of subnucleosomal particles. Tetrameric (H3/H4)2 146 base pair DNA and hexameric (H3/H4)2(H2A/H2B)1 146 base pair DNA complexes. *Biochemistry*, **24**, 4435-4450.

Reece, R.J., Rickles, R.J. and Ptashne, M. (1993) Overproduction and single-step purification of GAL4 fusion proteins from *Escherichia coli*. *Gene*, **126**, 105-107.

Reid, J.L., Iyer, V.R., Brown, P.O. and Struhl, K. (2000) Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase. *Mol. Cell*, **6**, 1297-1307.

Reifsnyder, C., Lowell, J., Clarke, A. and Pillus, L. (1996) Yeast SAS silencing genes and human genes associated with AML and HIV-1 Tat interactions are homologous with acetyltransferases. *Nat. Genet.*, **14**, 42-49.

Robert, F., Douziech, M., Forget, D., Egly, J.M., Greenblatt, J., Burton, Z.F. and Coulombe, B. (1998) Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol. Cell*, **2**, 341-351.

Rohde, J.R., Trinh, J. and Sadowski, I. (2000) Multiple signals regulate *GAL* transcription in yeast. *Mol. Cell. Biol.*, **20**, 3880-3886.

Romanowski, M.J., Bonanno, J.B. and Burley, S.K. (2002) Crystal structure of the *Streptococcus pneumoniae* phosphomevalonate kinase, a member of the GHMP kinase superfamily. *Proteins*, **47**, 568-571.

Rosenberg, M. and Court, D. (1979) Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.*, **13**, 319-353.

Rotman, B. and Spiegelman, S. (1953) The conversion of negatives to positives in "slow" adapting populations of yeast. *J. Bacteriol.*, **66**, 492-497.

Roy, A., Exinger, F. and Losson, R. (1990) *cis*- and *trans*-acting regulatory elements of the yeast *URA3* promoter. *Mol. Cell. Biol.*, **10**, 5257-5270.

Roy, R., Adamczewski, J.P., Seroz, T., Vermeulen, W., Tassan, J.P., Schaeffer, L., Nigg, E.A., Hoeijmakers, J.H. and Egly, J.M. (1994) The MO15 cell cycle kinase is associated with the TFIIH transcription-DNA repair factor. *Cell*, **79**, 1093-1101.

Rundlett, S.E., Carmen, A.A., Kobayashi, R., Bavykin, S., Turner, B.M. and Grunstein, M. (1996) *HDA1* and *RPD3* are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. *Proc. Natl. Acad. Sci. USA*, **93**, 14503-14508.

Sadowski, I., Niedbala, D., Wood, K. and Ptashne, M. (1991) GAL4 is phorphorylated as a consequence of transcription activation. *Proc. Natl. Acad. Sci. USA*, **88**, 10510-10514.

Schell, M.A. and Wilson, D.B. (1977) Purification and properties of galactokinase from *Saccharomyces cerevisiase*. *J. Biol. Chem.*, **252**, 1162-1166.

Schiessel, H., Gelbart, W.M. and Bruinsma, R. (2001) DNA folding: structural and mechanical properties of the two-angle model for chromatin. *Biophys. J.*, **80**, 1940-1956.

Schjerling, P. and Holmberg, S. (1996) Comparative amino acid sequence analysis of the C_6 zinc cluster family of transcriptional regulators. *Nucleic Acids Res.*, **24**, 4599-4607.

Sendra, R., Tse, C. and Hansen, J.C. (2000) The yeast histone acetyltransferase A2 complex, but not free Gcn5p, binds stably to nucleosomal arrays. *J. Biol. Chem.*, **275**, 24928-24934.

Serizawa, H., Makela, T.P., Conaway, J.W., Conaway, R.C., Weinberg, R.A. and Young, R.A. (1995) Association of Cdk-activating kinase subunits with transcription factor TFIIH. *Nature*, **374**, 280-282.

Shen, X., Mizuguchi, G., Hamiche, A. and Wu, C. (2000) A chromatin remodelling complex involved in transcription and DNA processing. *Nature*, **406**, 541-544.

Sherman, F., Fink, G.R., and Hicks, J.B. (1986) Laboratory course manual for methods in yeast genetics. Cold Spring Harbour, New York. Cold Spring Harbour Laboratory Press.

Shiekhattar, R., Mermelstein, F., Fisher, R.P., Drapkin, R., Dynlacht, B., Wessling, H.C., Morgan, D.O. and Reinberg, D. (1995) Cdk-activating kinase complex is a component of human transcription factor TFIIH. *Nature*, **374**, 283-287.

Silver, P.A., Brent, R. and Ptashne, M. (1986) DNA binding is not sufficient for nuclear localization of regulatory proteins in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **6**, 4763-4766.

St John, T.P. and Davis, R.W. (1981) The organization and transcription of the galactose gene cluster of *Saccharomyces*. *J. Mol. Biol.*, **152**, 285-315.

Sterner, D.E., Grant, P.A., Roberts, S.M., Duggan, L.J., Belotserkovskaya, R., Pacella, L.A., Winston, F., Workman, J.L. and Berger, S.L. (1999) Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction. *Mol. Cell. Biol.*, **19**, 86-98.

Strickland, M.S., Thompson, N.E. and Burgess, R.R. (1988) Structure and function of the sigma-70 subunit of *Escherichia coli* RNA polymerase. Monoclonal antibodies: localization of epitopes by peptide mapping and effects on transcription. *Biochemistry*, **27**, 5755-5762.

Studitsky, V.M., Clark, D.J. and Felsenfeld, G. (1995) Overcoming a nucleosomal barrier to transcription. *Cell*, **83**, 19-27.

Sudarsanam, P., Iyer, V.R., Brown, P.O. and Winston, F. (2000) Whole-genome expression analysis of *snf/swi* mutants of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*, **97**, 3364-3369.

Suzuki-Fujimoto, T., Fukuma, M., Yano, K.-I., Sakurai, H., Vonika, A., Johnston, S.A. and Fukasawa, T. (1996) Analysis of the galactose signal transduction pathway in *Saccharomyces cerevisiae*: Interaction between Gal3p and Gal80p. *Mol. Cell. Biol.*, **16**, 2504-2508.

Svaren, J. and Hörz, W. (1997) Transcription factors vs nucleosomes: Regulation of the *PH05* promoter in yeast. *Trends Biochem. Sci.*, **22**, 93-97.

Svejstrup, J.Q., Feaver, W.J., LaPointe, J. and Kornberg, R.D. (1994) RNA polymerase transcription factor IIH holoenzyme from yeast. *J. Biol. Chem.*, **269**, 28044-28048.

Thompson, C.M., Koleske, A.J., Chao, D.M. and Young, R.A. (1993) A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell*, **73**, 1367-1375.

Timson, D.J., Ross, H.C. and Reece, R.J. (2002) Gal3p and Gal1p interact with the transcriptional repressor Gal80p to form a complex of 1:1 stoichiometry. *Biochem. J.*, **363**, 515-520.

Triezenberg, S.J. (1995) Structure and function of transcriptional activation domains. *Curr. Opin. Gen. Dev.*, **5**, 190-196.

Tsukiyama, T., Daniel, C., Tamkun, J. and Wu, C. (1995) ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor. *Cell*, **83**, 1021-1026.

Tyler, J.K. and Kadonaga, J.T. (1999) The "dark side" of chromatin remodeling: repressive effects on transcription. *Cell*, **99**, 443-446.

Tzamarias, D. and Struhl, K. (1994) Functional dissection of the yeast Cyc8-Tup1 transcriptional co-repressor complex. *Nature*, **369**, 758-761.

Verrijzer, C.P. and Tjian, R. (1996) TAFs mediate transcriptional activation and promoter selectivity. *Trends Biochem. Sci.*, **21**, 338-342.

Vidal, M. and Gaber, R.F. (1991) *RPD3* encodes a second factor required to achieve maximum positive and negative transcriptional states in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **11**, 6317-6327.

Vogel, K., Horz, W. and Hinnen, A. (1989) The two positively acting regulatory proteins Pho2p and Pho4p physically interact with *PHO5* upstream activation regions. *Mol. Cell. Biol.*, **9**, 2050-2057.

Wang, L., Liu, L. and Berger, S.L. (1998) Critical residues for histone acetylation by Gcn5p, functioning in Ada and SAGA complexes, are also required for transcriptional function *in vivo*. *Genes Dev.*, **12**, 640-653.

Wolfe, K.H. and Shields, D.C. (1997) Molecular evidence for an ancient duplication of the entire yeast genome. *Nature*, **387**, 708-713.

Wolffe, A.P. and Hayes, J.J. (1999) Chromatin disruption and modification. *Nucleic Acids Res.*, **27**, 711-720.

Wolffe, A.P., Urnov, F.D. and Guschin, D. (2000) Co-repressor complexes and remodelling chromatin for repression. *Biochem. Soc. Trans.*, **28**, 379-386.

Woychik, N.A. (1998) Fractions to functions: RNA polymerase II thirty years later. *Cold Spring Harb. Symp. Quant. Biol.*, **63**, 311-317.

Yano, K.I. and Fukasawa, T. (1997) Galactose-dependent reversible interaction of Gal3p with Gal80p in the induction pathway of Gal4p-activated genes of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*, **94**, 1721-1726.

Young, R.A. (1991) RNA polymerase II. *Annu. Rev. Biochem.*, **60**, 689-715.

Zhang, Y., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Bird, A. and Reinberg, D. (1999) Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.*, **13**, 1924-1935.

Zenke, F.T., Engles, R., Vollenbroich, V., Meyer, J., Hollenberg, C.P. and Breunig, K.D. (1996) Activation of Gal4p by galactose-dependant interaction of galactokinase and Gal80p. *Science* **273** (5268) 1662-5.

Appendix

The insertion of two amino acids into a transcriptional inducer converts it into a galactokinase

Adam Platt*, Helen C. Ross, Steven Hankin, and Richard J. Reece†

School of Biological Sciences, The University of Manchester, 2.205 Stopford Building, Oxford Road, Manchester M13 9PT, United Kingdom

Edited by Mark Ptashne, Memorial Sloan-Kettering Cancer Center, New York, NY, and approved January 12, 2000 (received for review November 19, 1999)

The transcriptional induction of the *GAL* genes of *Saccharomyces cerevisiae* occurs when galactose and ATP interact with Gal3p. This protein-small molecule complex associates with Gal80p to relieve its inhibitory effect on the transcriptional activator Gal4p. Gal3p shares a high degree of sequence homology to galactokinase, Gal1p, but does not itself possess galactokinase activity. By constructing chimeric proteins in which regions of the *GAL1* gene are inserted into the *GAL3* coding sequence, we have been able to impart galactokinase activity upon Gal3p as judged *in vivo* and *in vitro*. Remarkably, the insertion of just two amino acids from Gal1p into the corresponding region of Gal3p confers galactokinase activity onto the resultant protein. The chimeric protein, termed Gal3p+SA, retains its ability to efficiently induce the *GAL* genes. Kinetic analysis of Gal3p+SA reveals that the K_m for galactose is similar to that of Gal1p, but the K_m for ATP is increased. The chimeric enzyme was found to have a decreased turnover number in comparison to Gal1p. These results are discussed in terms of both the mechanism of galactokinase function and that of transcriptional induction.

The yeast *Saccharomyces cerevisiae* utilizes galactose by means of the enzymes of the Leloir pathway. When yeast are grown in the absence of galactose, the genes encoding the enzymes of the pathway (the *GAL* genes) are transcriptionally inert (reviewed in refs. 1–3). If the cells are switched to medium in which galactose is the sole carbon source, then the *GAL* genes are rapidly induced and transcribed at high levels (4). The induction of the *GAL* genes is controlled by the interplay of three proteins—a transcriptional activator, Gal4p, a repressor, Gal80p, and an inducer, Gal3p. Induction appears to occur as a result of a galactose- and ATP-dependent interaction between Gal3p and Gal80p (5–8). This association results in the formation of a transcriptionally active Gal4p-Gal80p-Gal3p complex (9). It has been suggested that the association of Gal3p with Gal80p results in the movement of Gal80p from the activation domain of Gal4p to a different part of the protein (10). The location of this second site of Gal80p interaction on Gal4p remains unclear.

The first step of the Leloir pathway is the conversion of galactose to galactose-1-phosphate by galactokinase by Gal1p (11), the product of the *GAL1* gene. Gal1p and Gal3p are highly homologous proteins (73% identity and 92% homology at the amino acid level). Unlike Gal1p, Gal3p does not possess a galactokinase activity (12). Gal1p is bifunctional in that it has galactokinase activity and is able to induce the expression of the *GAL* genes both *in vivo* (13) and *in vitro* (9), although approximately 40-fold less efficiently than Gal3p (9). Galactokinases are relatively well conserved throughout nature. For instance, Gal1p and *Escherichia coli* galactokinase, *galK*, share 50% amino acid homology. Galactokinases are, however, extremely highly conserved in five regions (Fig. 1). The functions of these regions have not been defined experimentally, but presumably they could be involved in galactose or ATP binding, or in promoting the catalytic reaction. All functional galactokinases contain an invariant GLSSSA(A/S)(F/L/I) motif within homology region III (Fig. 1). Based on sequence alignments alone, it appears that Gal3p has a truncated version of this motif, with GLSSAF being found at amino acids 161–166 of the protein.

To delineate the various functions of Gal3p (galactose binding, ATP binding, and interaction with Gal80p), we first created a series of deletion mutations of the protein and tested their ability to complement a yeast strain deleted for Gal3p function. We found that we were unable to remove more than 28 amino acids from the amino-terminal end, or 9 amino acids from the carboxyl-terminal end of the molecule without loss of function. We therefore swapped regions of Gal1p into Gal3p in an attempt to restore galactokinase function to this transcriptional inducer. Surprisingly, we found that the insertion of just two amino acids from Gal1p into Gal3p imparted galactokinase activity onto Gal3p. This chimeric protein still retains the ability to induce transcription of the *GAL* genes both *in vivo* and *in vitro*. We discuss these results both in terms of galactokinase function and the implications for transcriptional induction.

Materials and Methods

Strains and Media. *E. coli* strain DH5α was used for all DNA manipulations. The following *S. cerevisiae* strains were used; yeast nuclear extract was prepared from BJ2168 (MATα *ura3, leu2, trp1, gal2, prb1, pep4, prc1*) grown in yeast extract/peptone/dextrose media (14), proteins were overproduced in strain MC2 (*MATa trp1 ura3–52 leu2–3 prc1–407 prb1-112 pep4-3*), and genetic analysis was performed in strain JPY5 (MATα *ura3–52 his3Δ200 leu2Δ1 trp1Δ63 lys2Δ385*) (15). Disruption of *GAL1*, *GAL3*, and *GAL80* from JPY5 was achieved by using PCR-generated blaster cassettes (16).

Plasmid Construction. All plasmid manipulations were performed as described by Sambrook *et al.* (17). *GAL3*, *GAL1*, and all derivatives expressed *in vivo* from the *GAL3* promoter and terminator (−606 to −1 and +1,564 to +2,019) were carried on a yeast 2 μ plasmid with the *HIS3* auxotrophic marker (pJP139). Deletion derivatives of *GAL3* were all prepared from the parental vector pAP25. The Δ2–28, 2–40, 180–223, 198–223, 198–243, 428–520, 500–520, and 511–520 *GAL3* deletions were all constructed by PCR-mediated mutagenesis (oligonucleotide sequences are available upon request). All other deletion derivatives of the *GAL3* gene were constructed by using naturally occurring restriction sites (details available upon request). *GAL3*, *GAL1*, and all derivatives expressed *in vivo* from the *GAL1* promoter were cloned into the 2 μ, *HIS3* expression vector pYX223 (CLONTECH). All *GAL3* and *GAL1* chimeras were

This paper was submitted directly (Track II) to the PNAS office.

*Present address: Chiroscience, Cambridge Science Park, Milton Road, Cambridge, CB4 0WE, United Kingdom.

†To whom reprint requests should be addressed. E-mail: Richard.Reece@man.ac.uk.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

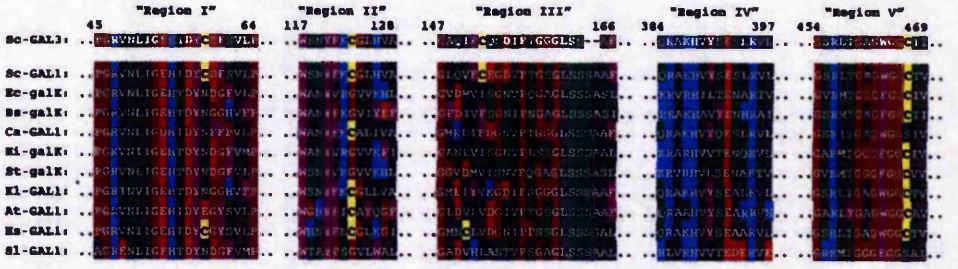


Fig. 1. Sequence comparison of galactokinase molecules. The numbers at the top refer to the Gal3p amino acid sequence. Sc, *Saccharomyces cerevisiae*; Ec, *Escherichia coli*; Bs, Bacillus subtilis; Ca, *Candida albicans*; Hi, *Haemophilus influenzae*; St, *Salmonella typhimurium*; Kl, *Kluyveromyces lactis*; At, *Arabidopsis thaliana*; Hs, *Homo sapiens*; St, *Streptomyces lividans*. Amino acids have been colored according to their properties. Blue indicates positively charge amino acids (H, K, R), red indicates negatively charged residues (D, E), green indicates polar neutral residues (S, T, N, Q), grey indicates nonpolar aliphatics (A, V, L, I, M), and purple indicates nonpolar aromatic residues (F, Y, W). Brown is used to indicate proline and glycine, whereas yellow indicates cysteine.

constructed by recombinant PCR (oligonucleotide sequences are available upon request). For high-level protein expression and purification from yeast cells, genes were cloned into pYEX-BX as described (9). All plasmids were sequenced to confirm the fidelity of the manipulations and the PCR (data not shown).

Northern and Western Blotting. Total RNA was isolated from 5 ml of yeast cells grown in the appropriate medium to an optical density (A_{600}) of 1.0 by using an RNAeasy extraction kit (Qiagen, Chatsworth, CA). RNA was then run on a formamide-agarose gel prior to alkaline transfer to Zetaprobe (Bio-Rad) according to manufacturer's instructions. Blots were hybridized sequentially with a *GAL10* probe (+1 to +2,077) and an 18S RNA probe (+279 to +1,405). Double-stranded DNA probes were prepared by using random-primed labeling as described (18).

For Western blotting, 5 ml of yeast cells were grown in the appropriate medium to an optical density (A_{600}) of 1.0. Cells were harvested, washed with 1 ml of sterile water, and the pellet placed in a boiling water bath for 1 min. The pellet was resuspended in 0.2 ml of Laemmli buffer, and 0.2 g of glass beads were added. The tube was vortexed for 90 s and boiled for 3 min twice. The supernatant was recovered, and 5 μl run on 12% SDS/PAGE. Proteins were transferred to nitrocellulose membranes that were subsequently incubated with either rabbit polyclonal Gal3p antibodies or a mouse-derived monoclonal antibody.

Galactokinase Assays. Galactokinase activity was measured by using an enzyme-linked assay system (19). Briefly, reaction mixtures (150 μl) were set up in microtiter wells containing the following: 20 mM Hepes (pH 8.0), 150 mM NaCl, 5 mM $MgCl_2$, 400 μM phosphoenol pyruvate, 1 mM DTT, 1 mM NADH, 1.1 units pyruvate kinase, 1.5 units lactic dehydrogenase (PK/LDH; Sigma). Reactions were supplemented with various concentrations of galactose (0.01–5 mM), ATP (0.005–5 mM), and Gal1p, Gal3p, or Gal3p+SA (0.72–680 nM). The plates were incubated at 30°C, and the decrease in absorbance at 340 nm was measured by using a Multiskan Ascent plate reader.

Results

Gal3p Is an Essentially Globular Protein. A series of deletion derivatives of Gal3p were constructed by using either naturally occurring restriction sites within the *GAL3* gene, or by means of PCR-based mutagenesis. Each of these proteins was produced in a yeast strain in which the *GAL3* gene had been deleted from the natural *GAL3* promoter contained on a 2 μ plasmid. We had previously determined that the kinetics of galactose induction were impaired in comparison to wild-type yeast if the *GAL3* gene was contained on an ARS-CEN plasmid (data not shown). Each of the deletion mutations was tested for its ability to grow on medium containing either raffinose or galactose as the sole carbon source (Fig. 2*A*). The phenotype of a *gal3* deletion is slow growth on galactose. Therefore, growth was scored 2 days after plating. At this time, all of the deletion plasmids permitted growth on raffinose, but yeasts containing the empty plasmid (Fig. 2*A*, line 1) showed little growth on galactose. Plasmids bearing the wild-type *GAL3* gene were able to grow well on galactose after 2 days (Fig. 2*A*, line 2). When incubated for 5 days, the growth of *GAL3* or *gal3* yeasts was indistinguishable, presumably because of leaky expression of *GAL1* (13).

Expression of deletion derivatives in which the amino-terminal 28 amino acids or the carboxyl-terminal 9 amino acids of Gal3p were removed gave rise to proteins that were able to complement the *gal3* deletion as well as the wild-type protein (Fig. 2*A*, lines 3 and 19). Deletions further into the sequence resulted in proteins that were unable to complement the *gal3* deletion (Fig. 2*A*, lines 4 and 18). We also found that a small internal deletion of Gal3p (the removal of amino acids 198–223) resulted in the formation of a partially functional protein (Fig. 2*A*, line 12). Extending this deletion either toward the amino- or carboxyl-terminal end of the molecule again resulted in a nonfunctional protein (Fig. 2*A*, lines 11 and 13). All other deletion mutations that we tested were unable to complement the *gal3* deletion.

A number of the deletion mutations were tested for their ability to induce *GAL10* mRNA expression by Northern blot analysis (Fig. 2*B*). Wild-type Gal3p induced full expression of *GAL10* within 24 hr of induction by galactose (Fig. 2*B*, line 2). An empty plasmid that failed to complement the *gal3* mutation (Fig. 2*A*, line 1) only induced *GAL10* expression to its maximal level 4 days after induction (Fig. 2*B*, line 1). We found a correlation between the ability of a protein to complement the *gal3* mutation (Fig. 2*A*) and its ability to rapidly induce *GAL10* expression (Fig. 2*B*).

To show that the deletion derivatives outlined above were produced within the cell, we performed Western blot analysis (Fig. 2*C*). The detection of Gal3p by this method is complicated by the high degree of homology between Gal3p and Gal1p. Our own polyclonal antibodies raised against Gal3p, and those described previously (6), both showed a high degree of cross-

BIOCHEMISTRY

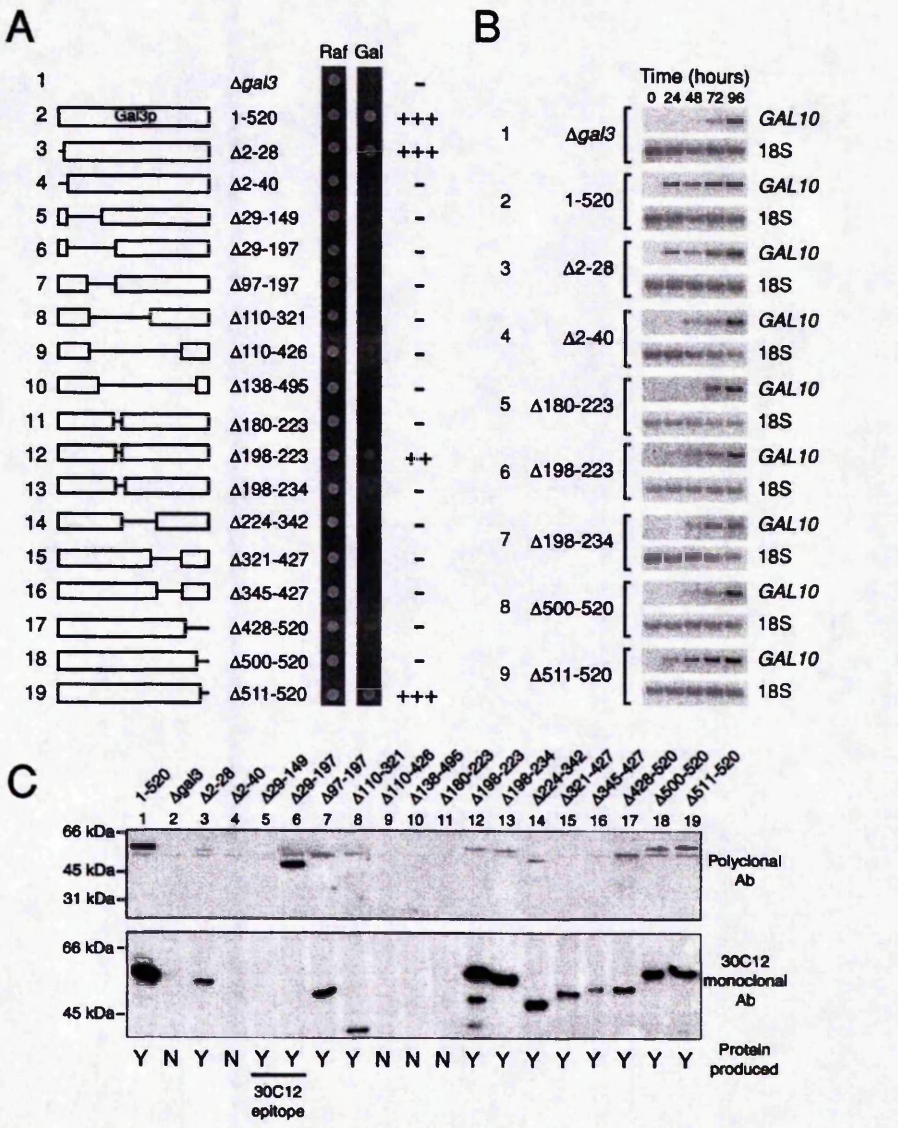


Fig. 2. The activity of Gal3p deletion mutations. (*A*) Plasmids expressing the wild-type Gal3p protein or the deletion derivative indicated were transformed into a yeast strain in which the *GAL3* gene had been disrupted. Each of the strains were then grown on either raffinose (Raf) or galactose (Gal) as the sole carbon source, and cell growth was monitored after 3 days of growth at 30°C. (*B*) The effect of various Gal3p deletion derivatives on the expression of *GAL10*. Northern blot analysis of *GAL10* and *18S* RNA expression was monitored daily after the cells were transferred to a galactose-containing medium. (*C*) Western blot analysis of Gal3p and the deletion derivatives. A yeast strain deleted for *GAL1*, *GAL3*, and *GAL80* was transformed with the appropriate Gal3p expression vector and grown on galactose for 2 days. Western blot analysis was performed on whole cell extracts using either Gal3p polyclonal antibodies or 30C12, a mouse monoclonal antibody raised against Gal3p.

reactivity with Gal1p (data not shown). We therefore analyzed the production of the Gal3p deletion derivatives in a yeast stain deleted for the *GAL1*, *GAL3*, and *GAL80* genes. The additional deletion of *GAL80* ensured constitutive induction of the *GAL3* promoter. We also performed Western blot analysis with a Gal3p monoclonal antibody (30C12) that showed no crossreactivity with Gal1p (data not shown). The production of the Δ2–40, Δ110–426, Δ138–495, and Δ180–223 Gal3p deletion derivatives were not detected in yeast cells (Fig. 2*C*, lanes 4, 9–11), and thus the inability of these proteins to complement the *gal3* mutation can be explained by a lack of protein. All of the other deletion derivatives were expressed, although their accumulation proved to be variable (Fig. 2*C*). For instance, wild-type Gal3p and Gal3pΔ2–28 both complemented the *gal3* mutation to the same extent (Fig. 2*A*, lines 2 and 3), yet the deletion derivative appears to be expressed at 10-fold lower levels than the wild-type protein (Fig. 2*C*, lanes 1 and 3). It is therefore not clear whether this difference is a true reflection of the levels of each protein or an artifact of performing Western blot analysis in a triple mutant strain. However, the inability of the majority of the deletion derivatives to complement the *gal3* mutation cannot simply be explained as a result of lack of protein. We therefore conclude that Gal3p is an essentially globular protein that is sensitive to the removal of all but the smallest amino acid stretches.

Converting Gal3p into a Galactokinase. Given the lack of success in isolating smaller, functional versions of Gal3p, we attempted to address the issue of why Gal3p, which is so homologous to Gal1p, does not possess a galactokinase activity itself. We noted that Gal3p is the only galactokinase-like molecule that deviates within the highly conserved GLSSSAA motif of homology region III (Fig. 1). We therefore concentrated on this region of the protein as a potential area to explain the differences in activity of the two proteins. We constructed a series of chimeric genes in which regions of the *GAL1* coding sequence were inserted into the *GAL3* gene. These chimeras were then tested for their ability to act as a galactokinase (by complementing a *gal1* deletion), as a transcriptional inducer (by complementing a *gal3* deletion) or as bifunctional proteins (by complementing a *gal1*, *gal3* double deletion). Proteins were either expressed from the *GAL1* promoter or from the *GAL3* promoter to eliminate the possibility that differing protein levels were interfering with the assays. The results of this analysis are shown in Fig. 3. Wild-type Gal3p, expressed either from its own promoter or from the *GAL1* promoter, was unable to complement the *gal1* deletion or the *gal1*, *gal3* double deletion (Fig. 3, lines 4 and 5). Gal3p expressed from its own promoter was better able to complement the *gal3* deletion than the same protein produced from the *GAL1* promoter (Fig. 3, lines 4 and 5). Gal1p, on the other hand, was able to complement the mutations in all three deletion strains, again showing the bifunctional nature of this protein when it is overproduced (Fig. 3, lines 2 and 3). As observed for Gal3p, Gal1p expressed from the *GAL3* promoter is better able to complement the *gal3* deletion.

We found that the insertion of homology region III from Gal1p into Gal3p was sufficient to grant the resulting chimeric protein with galactokinase activity (Fig. 3, lines 6–10). Indeed, the insertion of the two amino acids "missing" from Gal3p in homology region II (the insertion of a serine and an alanine residue at position 164 of Gal3p) allows the resultant protein to complement the *gal1*, the *gal3*, and the *gal1*, *gal3* deletions (Fig. 3, lines 11 and 12). The insertion of only an alanine residue at position 164 of Gal3p resulted in a protein that still functioned as Gal3p but was unable to complement the *gal1* mutations (Fig. 3, line 13). We also inserted either a serine residue at position 164 of Gal3p or two alanine residues. Neither of these proteins was able to complement the deletion in *gal1*, but they also did not complement the mutation in *gal3*, suggesting that these insertions may compromise the overall structure of the protein or are simply unstable (data not shown).

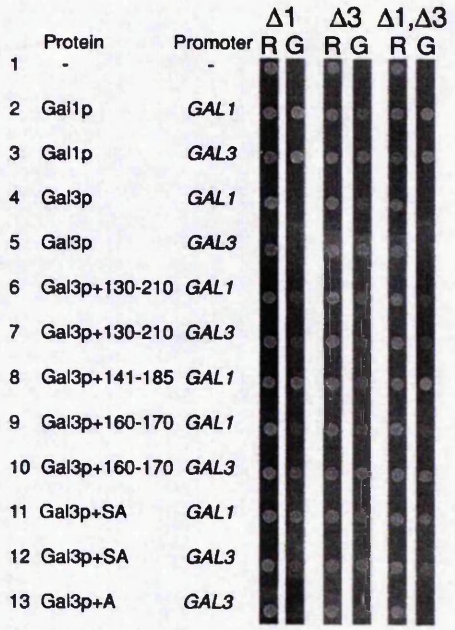


Fig. 3. The insertion of two amino acids into Gal3p converts it into a galactokinase. Plasmids expressing Gal1p, Gal3p, or the indicated chimera were transformed into yeast strains lacking *GAL1* (Δ1), *GAL3* (Δ3), or both *GAL1* and *GAL3* (Δ1,Δ3). Cells were then plated onto media containing either raffinose (R) or galactose (G) as the sole carbon source. Growth was monitored after 3 days incubation at 30°C. The chimeras are numbered to indicate the amino acids from Gal3p that have been replaced by the corresponding sequences from Gal1p.

Activity of Gal3p+SA *in Vitro*. We overproduced Gal3p+SA in yeast cells and purified it to homogeneity (data not shown). We then tested the ability of this protein to function either as a transcriptional inducer *in vitro* (Fig. 4) or as a galactokinase (Table 1). Using *in vitro* transcription assays, Gal4p (amino acids 1–93 fused to 768–881) was able to induce transcription approximately 10-fold from a template containing five Gal4p-binding sites upstream of the *E4* TATA-box (Fig. 4, lanes 1 and 2). This activation was completely inhibited by an equimolar amount of Gal80p (Fig. 4, lane 3). As observed previously (9), Gal3p or Gal1p are able to restore Gal4p-mediated transcription even in the presence of Gal80p (Fig. 4, lanes 4–6 and 10–12) with approximately 40-fold more Gal1p required to induce transcription to the same level as Gal3p. In these assays, Gal3p+SA behaved indistinguishably from wild-type Gal3p (Fig. 4, lanes 7–9), suggesting that the chimeric protein retained its ability to efficiently induce the Gal4p-Gal80p complex.

We investigated the galactokinase activity of Gal1p, Gal3p, and Gal3p+SA. Using an enzyme-linked assay, described in *Materials and Methods*, both Gal1p and Gal3p+SA displayed Michaelis–Menten kinetics (data not shown). The values we

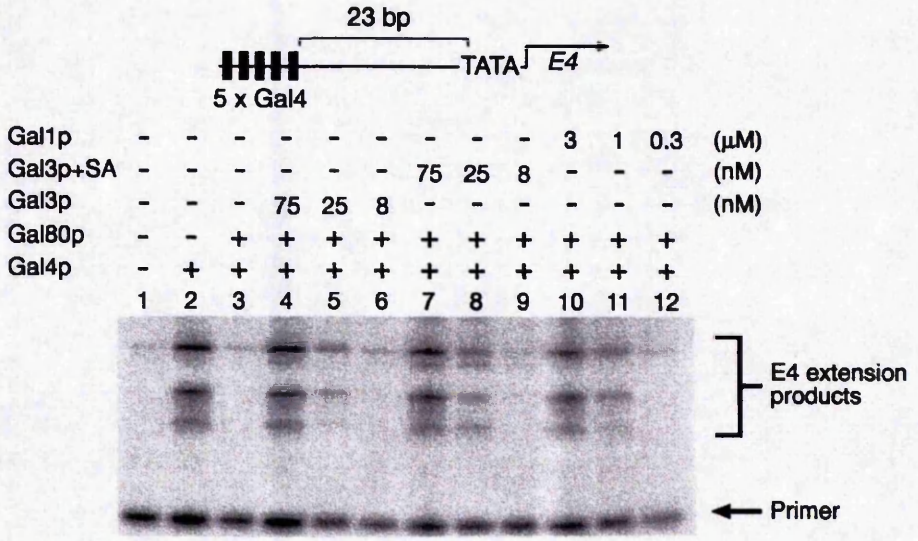


Fig. 4. The activity of Gal3p+SA *in vitro*. *In vitro* transcription and primer extension reactions were performed as described previously (9). All reactions contained 1.5 mM galactose and 1 mM ATP and, where indicated, 3 nM Gal4p (amino acids 1–93 fused to 768–881), 3 nM Gal80p and either Gal3p (75, 25, and 8 nM in lanes 4, 5, and 6, respectively), Gap3p+SA (75, 25, and 8 nM in lanes 7, 8, and 9, respectively), or Gal1p (3,000, 1,000, and 333 nM in lanes 10, 11, and 12 respectively).

obtained for the K_m of our histidine-tagged Gal1p for either galactose and ATP are similar to those reported for the wild-type enzyme (0.25 mM and 0.02 mM for galactose and ATP, respectively, compared with 0.6 mM and 0.15 mM, respectively, previously reported (11). Consistent with previous data (12), we have been unable to detect galactokinase activity of Gal3p using this or other assays (Table 1, and data not shown). The K_m for galactose of Gal3p+SA is very similar to that observed for Gal1p (Table 1). However, the K_m for ATP is greatly increased for the chimeric protein (0.17 mM compared with 0.02 mM for Gal1p; Table 1). The calculated turnover number for the chimeric protein is considerably smaller than that calculated for Gal1p (Table 1). We calculated the turnover number of Gal1p to be 20 per second, while the chimeric protein had a turnover rate some 400-fold less at 0.05 per second.

Discussion

S. cerevisiae contains two galactokinase-like molecules, Gal1p and Gal3p, which share large regions of amino acid sequence identity and similarity (Fig. 1). The function of each protein is related, but distinct. Gal1p catalyzes the formation of galactose-1-phosphate from galactose and ATP, whereas Gal3p participates in a galactose- and ATP-dependent interaction with Gal80p to induce the *GAL* genetic switch (5–7, 9). Gal1p is somewhat bifunctional in that it can, when overproduced, also induce *GAL* gene expression (13), albeit weakly (9). Gal3p, on the other hand, possesses no galactokinase activity (12). Closely related yeasts, e.g., *Kluyveromyces lactis*, contain a single galactokinase-like molecule that functions both as a galactokinase and as a transcriptional inducer (20).

Table 1. The galactokinase activity of Gal1p, Gal3p, and Gal3p+SA

Protein	K_m (mM)		Turnover number (s^{-1})
	Galactose	ATP	
Gal1p	0.25 ± 0.09	0.020 ± 0.003	20
Gal3p	ND	ND	0
Gal3p+SA	0.21 ± 0.06	0.17 ± 0.02	0.05

The galactokinase activity of each protein was measured by using an enzyme-linked assay described in *Materials and Methods*. Kinetic parameters were determined by using both double reciprocal and Hanes plots (22) and represent the average of at least four separate experiments. ND, no detectable activity.

In efforts to assign function to, and possibly identify domains of, Gal3p, we initially undertook a deletion analysis of the protein (Fig. 2). We found that all but the smallest deletions from the amino- or carboxyl-terminal ends of the protein rendered the protein nonfunctional in terms of its ability to complement a *gal3* deletion. Internal deletions were also largely nonfunctional, with the exception of the removal of amino acids 198–223 from the full-length protein, which resulted in the formation of a partially active Gal3p molecule (Fig. 2, line 12). Attempts to construct larger deletion derivatives based around the functional Δ198–223 deletion also proved unsuccessful and resulted in the formation of nonfunctional protein. We therefore conclude that Gal3p is likely to be essentially a globular protein whose structure is intolerant to deletions.

Given the high degree of homology between Gal3p and Gal1p, we utilized chimeric molecules to assign function to regions of each protein. Gal3p is not a galactokinase, but shares a high degree of homology to both Gal1p and other galactokinases (Fig. 1). Galactokinases are particularly well conserved in, what we term, homology region III. In particular, galactokinases contain a highly conserved GLSSSAA motif (amino acids 167–173 in Gal1p). Sequence alignments have shown that Gal3p, the only

molecule of this set not to have galactokinase activity, contains an apparent deletion of two amino acids in this highly conserved motif (Fig. 1). The insertion of sequence segments from homology region III of Gal1p into the corresponding amino acid sequence of Gal3p resulted in the formation of chimera with galactokinase activity (Fig. 3, lanes 6–10). That is, the chimeric proteins were able to complement a deletion of *gal1*, a deletion of *gal3*, or a double deletion. This *in vivo* assay easily identified proteins possessing galactokinase activity through complementation of the *gal1* deletion by a plasmid-borne copy of the chimeric gene. However, the isolation of the Gal3p induction phenotype by this method has proved difficult because a plasmid-borne copy of *GAL1* will complement both the *gal1* and *gal3* deletions (Fig. 3, lines 3 and 4). This effect was observed when *GAL1* was expressed from either the *GAL3* or the *GAL1* promoter and suggests that relatively modest levels of Gal1p overproduction will allow the protein to act as a transcriptional inducer.

The insertion of the missing two amino acids of homology region III from Gal1p into Gal3p (Gal3p+SA) converts Gal3p into a galactokinase. Gal3p+SA is able to complement *gal1* and *gal3* deletions *in vivo* (Fig. 3, lines 11 and 12) and is able to activate *GAL* gene expression *in vitro* indistinguishably from wild-type Gal3p (Fig. 4). During *in vitro* transcription assays, Gal3p+SA behaves like Gal3p in that a 25-fold molar excess over Gal80p is required to activate *GAL* gene expression, as seen previously (9). A much higher level of Gal1p was required to achieve the same level of transcriptional induction (Fig. 4, lanes 10–12). This shows that Gal3p+SA retains its ability to efficiently activate transcription while at the same time possessing a galactokinase activity. The insertion of a single alanine residue in place of SA (Gal3p+A) is insufficient to promote galactokinase function (Fig. 3, lane 13), but the protein is still able to complement the *gal3* deletion *in vivo*. Our attempts to analyze a single serine insertion or that of two alanines have been hampered by the fact that these proteins do not complement the *gal3* deletion, suggesting that they are misfolded or unstable.

Assays of galactokinase activity have revealed that Gal3p+SA is a relatively inefficient enzyme when compared with Gal1p (Table 1). We note, however, that the K_m for galactose is very similar for Gal1p and Gal3p+SA. The K_m for ATP is, however, greatly increased for the Gal3p+SA protein when compared with Gal1p (Table 1). This suggests that homology region III of Gal1p may be involved in ATP-binding or catalysis. Both Gal1p and Gal3p require both galactose and ATP for their mode of action. Although they perform seemingly different functions, our data suggest that Gal1p and Gal3p act through a similar mechanism. The mechanism of galactokinase action is not known, but it is reasonable to assume that binding of galactose and ATP, either as separate events or in concert, are followed by a catalysis step and then finally the release of products. Gal3p appears to be defective in its ability to perform the catalysis step of the reaction. What is the role of the serine and alanine residues in the catalysis reaction? Further kinetic and structural analyses are currently being used to address this question.

Several attempts have been made to map the interaction between Gal3p and Gal80p by searching for constitutive mutations of Gal3p (8, 21). The position of such mutations have been spread throughout the protein and have thus failed to identify a specific region of Gal3p responsible for the interaction with Gal80p. It is tempting to speculate that the binding of galactose and ATP to Gal1p and Gal3p induces a conformational change in the proteins. In the case of Gal1p, this is required for the catalysis reaction, perhaps by bringing the SA residues close to the site of catalysis, whereas in Gal3p the conformational change is required to promote association with Gal80p. Our data suggest that homology region III is required for the galactokinase reaction and that another region of the protein is responsible for the interaction with Gal80p.

We thank Cristina Merlotti and David Timson for many useful discussions and comments on the manuscript. We also thank Tony Maxwell (The University of Leciester) for advice on enzyme assays and Terri Attwood for help with the sequence alignment. This work was supported by grants from the Biotechnology and Biological Sciences Research Council and The Leverhulme Trust.

1. Johnston, M. (1987) *Microbiol. Rev.* **51**, 458–476.
2. Lohr, D., Venkov, P. & Zlantanova, J. (1995) *FASEB J.* **9**, 777–787.
3. Reece, R. J. & Platt, A. (1997) *BioEssays* **19**, 1001–1010.
4. St. John, T. P. & Davis, R. W. (1981) *J. Mol. Biol.* **152**, 285–315.
5. Zenke, F. T., Engles, R., Vollenbroich, V., Meyer, J., Hollenberg, C. P. & Breunig, K. D. (1996) *Science* **272**, 1662–1665.
6. Suzuki-Fujimoto, T., Fukuma, M., Yano, K.-I., Sakurai, H., Vonika, A., Johnston, S. A. & Fukasawa, T. (1996) *Mol. Cell. Biol.* **16**, 2504–2508.
7. Yano, K. I. & Fukasawa, T. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 1721–1726.
8. Blank, T. E., Woods, M. P., Lebo, C. M., Xin, P. & Hopper, J. E. (1997) *Mol. Cell. Biol.* **17**, 2566–2575.
9. Platt, A. & Reece, R. J. (1998) *EMBO J.* **17**, 4086–4091.
10. Sil, A. K., Alam, S., Xin, P., Ma, L., Morgan, M., Lebo, C. M., Woods, M. P. & Hopper, J. E. (1999) *Mol. Cell. Biol.* **19**, 7828–7840.
11. Schell, M. A. & Wilson, D. B. (1977) *J. Biol. Chem.* **252**, 1162–1166.
12. Bhat, P. J., Oh, D. & Hopper, J. E. (1990) *Genetics* **125**, 281–291.
13. Bhat, P. J. & Hopper, J. E. (1992) *Mol. Cell. Biol.* **12**, 2701–2707.
14. Ohashi, Y., Brickman, J. M., Furman, E., Middleton, B. & Carey, M. (1994) *Mol. Cell. Biol.* **13**, 2731–2739.
15. Wu, Y., Reece, R. J. & Ptashne, M. (1996) *EMBO J.* **15**, 3951–3963.
16. Lorenz, M. C., Muir, R. S., Lim, E., McElver, J., Weber, S. C. & Heitman, J. (1995) *Gene* **158**, 113–117.
17. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
18. Feinberg, A. P. & Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266–267.
19. Ali, J. A., Jackson, A. P., Howells, A. J. & Maxwell, A. (1993) *Biochemistry* **32**, 2717–2724.
20. Meyer, J., Walker-Jonah, A. & Hollenberg, C. P. (1991) *Mol. Cell. Biol.* **11**, 5454–5461.
21. Vollenbroich, V., Meyer, J., Engels, R., Cardinali, G., Menezes, R. A. & Hollenberg, C. P. (1999) *Mol. Gen. Genet.* **261**, 495–507.
22. Cornish-Bowden, A. (1995) *Fundamentals of Enzyme Kinetics* (Portland Press, London).

Gal3p and Gal1p interact with the transcriptional repressor Gal80p to form a complex of 1:1 stoichiometry

David J. TIMSON, Helen C. ROSS and Richard J. REECE[1]

School of Biological Sciences, The University of Manchester, 2.205 Stopford Building, Oxford Road, Manchester M13 9PT, U.K.

The genes encoding the enzymes required for galactose metabolism in *Saccharomyces cerevisiae* are controlled at the level of transcription by a genetic switch consisting of three proteins: a transcriptional activator, Gal4p; a transcriptional repressor, Gal80p; and a ligand sensor, Gal3p. The switch is turned on in the presence of two small molecule ligands, galactose and ATP. Gal3p shows a high degree of sequence identity with Gal1p, the yeast galactokinase. We have mapped the interaction between Gal80p and Gal3p, which only occurs in the presence of both ligands, using protease protection experiments and have shown that this involves amino acid residue 331 of Gal80p. Gel-filtration experiments indicate that Gal3p, or the galactokinase Gal1p, interact directly with Gal80p to form a complex with 1:1 stoichiometry.

Key words: Gal4p, galactose, gene expression, protein–protein interaction, yeast.

INTRODUCTION

Saccharomyces cerevisiae utilizes galactose via the enzymes of the Leloir pathway. When yeast cells are grown in the absence of galactose, the genes encoding the enzymes of the pathway (the *GAL* genes) are transcriptionally inert (reviewed in [1–3]). If the cells are switched to medium in which galactose is the sole carbon source, then the *GAL* genes are rapidly induced and transcribed at high levels [4]. The induction of the *GAL* genes is controlled by the interplay of three proteins: a transcriptional activator, Gal4p; a repressor, Gal80p; and a ligand sensor, Gal3p. Induction appears to occur as a result of a galactose- and ATP-dependent interaction between Gal3p and Gal80p [5–8]. The nature of the transcriptionally active complex remains unclear. It has been suggested that the association of Gal3p and Gal80p results in the formation of a transcriptionally active Gal4p–Gal80p–Gal3p complex [9]. It has also been suggested that the association of Gal3p with Gal80p results in the movement of Gal80p from the activation domain of Gal4p to a different part of the protein [10], and, in a different study, that the interaction between Gal3p and Gal80p occurs solely in the cytoplasm of yeast cells [11]. The apparent discrepancies between these data have yet to be resolved.

Gal3p is the ligand sensor of the *GAL* genetic switch [9]. The protein is highly related (73% identity and 92% similarity at the amino acid level) to Gal1p, the galactokinase of the Leloir pathway. Gal3p has no galactokinase enzymic activity itself [12], but the close relationship between Gal1p and Gal3p has been underlined by recent work showing that the insertion of just two amino acids from Gal1p into Gal3p turned the latter into a galactokinase [13].

In the present study, we set out to determine the site of interaction on Gal80p with Gal3p and the nature of the interaction between the two proteins. Our data show that both galactose and ATP are required to promote a direct physical interaction between Gal3p and Gal80p. The interaction involves amino acids in the C-terminal part of Gal80p.

MATERIALS AND METHODS

Protein expression and purification

His-tagged Gal1p, Gal3p, Gal80p and Gal80$^{\Delta 4}$p were overexpressed in, and purified from, yeast as described previously [9], except that yeast cells were broken as described in [14]. Gal4p (amino acids 1–93 fused to amino acids 768–881) was overexpressed in *Escherichia coli* and purified as described previously [15].

Protease protection experiments

Gal3p and Gal80p were mixed in buffer A [50 mM Hepes/NaOH (pH 8.0), 150 mM NaCl, 10% (v/v) glycerol and 1.4 mM 2-mercaptoethanol] at the concentrations stated in the absence or presence of ligands (galactose and ATP) as required, and incubated at 30 °C for at least 30 min. After this time, tosyllysylchloromethane ('TLCK')-treated α-chymotrypsin (Sigma) was added to a final concentration of 300 nM, and digestion allowed to proceed for 30 min at 30 °C. Products were analysed by SDS/PAGE [15% (w/v) gels].

Analytical gel filtration

Gal3p (or Gal1p) and Gal80p were mixed in buffer A in the absence or presence of ligands (200 μl total volume) and incubated at 20 °C for at least 30 min, prior to loading on to a Sephadex S75 column (bed volume, V_t, of 24 ml; Amersham Biosciences), which had been pre-equilibrated in buffer A containing the same ligand(s) at the same concentration as the protein mixture. The column was run at 0.5 ml/min and 0.5 ml fractions were collected. Proteins were precipitated from these fractions by acetone precipitation and analysed by SDS/PAGE.

The column was calibrated using a set of molecular-mass standards (Amersham Biosciences). The elution volume (V_e) for each standard protein was determined in buffer A along with the void volume of the column (V_o), which is equal to the elution volume of Blue Dextran. A standard curve was obtained by

Abbreviation used: BS^3, bis(sulphosuccinimidyl) suberate.

[1] To whom correspondence should be addressed (e-mail Richard.Reece@man.ac.uk).

plotting K_{av} against the logarithm of the molecular mass, where K_{av} is given by $(V_e - V_o)/(V_t - V_o)$.

Protein–protein cross-linking

Gal4p (2 μM) was mixed with either Gal80p (2 μM) or Gal80$^{6-8}$p (2 μM) in buffer A and incubated at 20 °C for 15 min. Bis(sulphosuccinimidyl) suberate (BS^3; Pierce) is an amine cross-linking agent with a spacer arm of 1.1 nm [16]. It was dissolved in 5 mM sodium citrate buffer (pH 5.5) and then added immediately to protein solutions at a final concentration of 0.1 mM. The reactions were then incubated at 20 °C for 30 min. To quench the reaction, Tris/HCl (pH 7.5) was added to a final concentration of 25 mM, and the reaction was incubated at 20 °C for a further 15 min. Samples were analysed by SDS/PAGE.

Analytical methods

Protein concentrations were determined using the method of Bradford [17]. SDS/PAGE was carried out as described by Laemmli [18]. Acetone precipitation of protein samples was carried out by adding an equal volume of acetone and placing the samples at −20 °C for 30 min. Samples were then centrifuged at 15000 *g* in a bench-top microcentrifuge for 10 min at 20 °C, the supernatant was removed and the pellets were dried at 94 °C. Pellets were redissolved in 15 μl of buffer A and an equal volume of SDS-loading buffer, prior to heating at 94 °C and loading on to an SDS/polyacrylamide gel. Protein sequencing was performed by Dr Linda Berry (School of Biological Sciences, University of Manchester, U.K.) using an Applied Biosystems 476A machine.

RESULTS

Gal3p protects Gal80p from proteolysis in the presence of galactose and MgATP

Protease protection experiments were used to determine the site of interaction between highly purified Gal3p and Gal80p. In the absence of Gal3p, Gal80p was digested by chymotrypsin into three principal fragments of sizes 36, 25 and 12 kDa (Figure 1A, lane 2). This digestion pattern was not affected by the presence of either MgATP or galactose alone or together (Figure 1A, lanes 3–5). A similar digestion pattern of Gal80p was observed using a variety of proteases (results not shown).

In contrast, when Gal3p was digested under identical conditions, a greater number of protein fragments were produced, although less of the starting material was consumed (Figure 1A, lane 6). This pattern was not changed by the presence of either MgATP or galactose alone or together (Figure 1A, lanes 7–9). When the two proteins were mixed together and exposed to the protease, the digestion pattern was approximately additive (Figure 1A, lane 11). This was also the case when both proteins were mixed in the presence of either MgATP or galactose alone (Figure 1A, lanes 12 and 13). However, in the presence of both ligands (Figure 1A, lane 14), the 36 and 12 kDa bands (marked with an asterisk in Figure 1) from Gal80p were reduced markedly in intensity. Protein sequencing revealed that the 36 kDa band began at amino acid 4 and the 12 kDa at residue 332 of Gal80p (results not shown).

Since the ligands themselves did not affect the digestion patterns of the proteins on their own, the most likely explanation for this change was that the binding of Gal3p to Gal80p was preventing access to the digestion site on Gal80p by the protease, resulting in protection of Gal80p. Furthermore, this experiment suggested that both ligands are required to promote the interaction between Gal3p and Gal80p. The tolerance of the system to changes in the nucleotide ligand was tested (Figure 1B). MgADP was found to substitute effectively for MgATP, but MgAMP was ineffective in promoting binding and protection from protease digestion.

A 1:1 interaction between Gal80p and Gal3p requires both ligands

Gel filtration is a powerful technique for demonstrating protein–protein interactions *in vitro*. When Gal3p alone was applied to a gel-filtration column, it was eluted as a single symmetrical peak

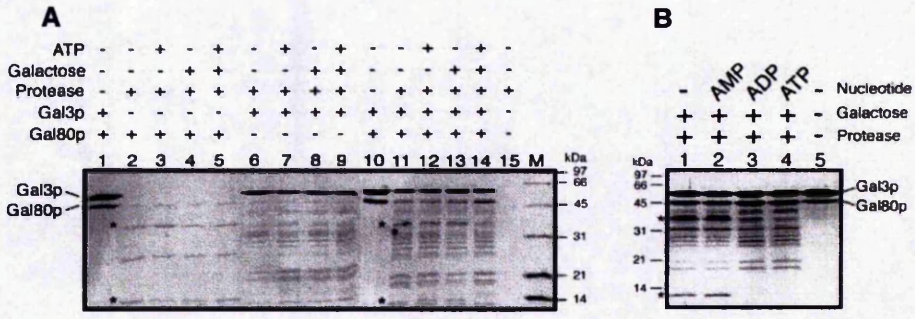


Figure 1 Binding of Gal3p to Gal80p protects the latter from limited proteolytic digestion by chymotrypsin

(**A**) Gal3p (10 μM) was mixed with Gal80p (4 μM) in the presence of galactose (2.5 mM) and/or MgATP (2.5 mM) as indicated. Reaction mixtures were digested with chymotrypsin (300 nM) for 30 min at 30 °C, and then resolved by SDS/PAGE [15% (w/v) gel]. Lane M, molecular-mass markers. (**B**) ADP can substitute for ATP in promoting binding of Gal3p and Gal80p. Reaction mixtures were set up as above, except that MgATP was substituted with MgAMP or MgADP (both at 2.5 mM) as appropriate. Molecular-mass markers (in kDa) are shown on the left.

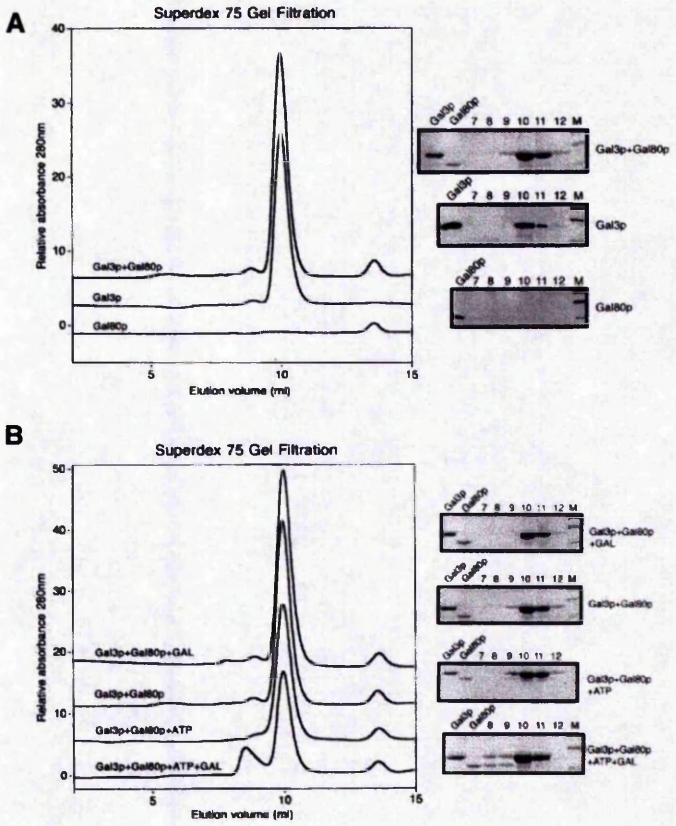


Figure 2 Gel-filtration analysis of Gal3p and Gal80p

(**A**) Gal3p (13 μM), Gal80p (3 μM), or a mixture of both, were applied to a Superdex S75 column (Amersham Biosciences). UV absorbance traces at 280 nm are shown alongside SDS/PAGE gels of acetone-precipitated samples from across the major peaks. The numbers on the gels correspond to the approximate elution volume (in ml). (**B**) A discrete Gal3p–Gal80p complex can be isolated by gel filtration in the presence of galactose (GAL) and ATP. Concentrations of proteins were as in (**A**). Ligands were added both to the sample, running buffer at a final concentration of 2.5 mM. Lane M, molecular-mass markers.

(Figure 2A), corresponding to a molecular mass of 46 kDa (actual molecular mass is 58 kDa). SDS/PAGE analysis of acetone-precipitated fractions from this peak confirmed that it corresponded to Gal3p elution. When a similar quantity of Gal80p was applied to the column, no visible peaks were observed. Gal80p was found to be present in small amounts across a wide range of elution volumes. This suggests that Gal80p does not exist as one single discrete entity in solution, and is consistent with the protein having a highly flexible conformation and/or existing in a variety of multimeric states. This result is consistent with recent observations of a variety of multimeric states of Gal80p [19]. An alternative explanation is that Gal80p is binding to the column itself, but we disfavour this explanation, since the experiments were performed at relatively high salt levels (150 mM NaCl) and adherence to the column would result in an elution profile shifted to higher volumes, rather than the lower volumes observed.

When a mixture of Gal3p and Gal80p was applied to the column, a single peak was observed at exactly the same elution volume as was seen with Gal3p alone (Figure 2A). This peak contained almost exclusively Gal3p, as shown by SDS/PAGE analysis of acetone-precipitated fractions. A similar result was

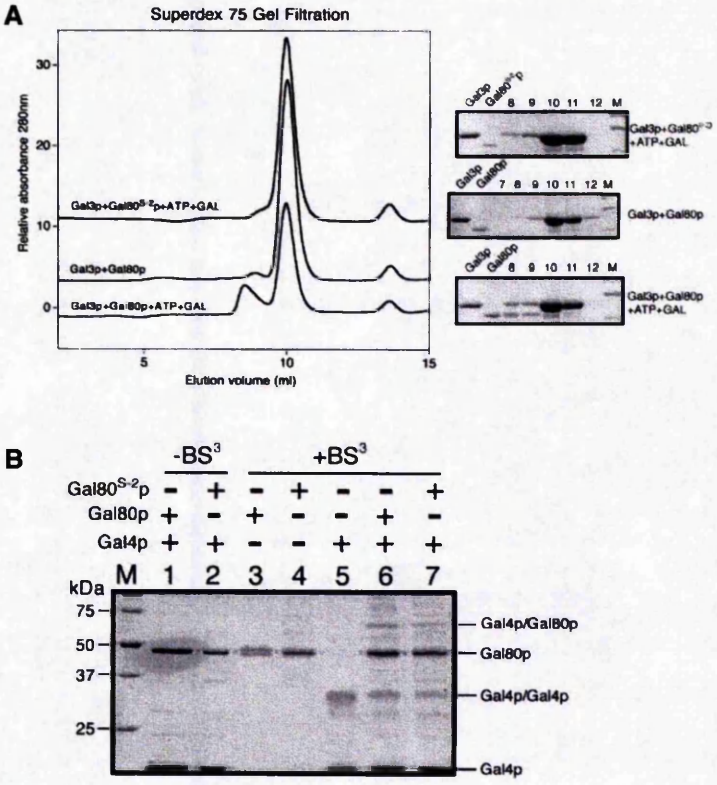


Figure 3 Gal80^{S-2}p, a super-repressor mutant, is unable to bind to Gal3p, but does bind Gal4p

(**A**) Gal80^{S-2}p (2 μM) was mixed with Gal3p (13 μM) in the presence of galactose and MgATP (both at 2.5 mM). The mixture was applied to a Superdex S75 column pre-equilibrated in buffer containing both ligands at this concentration. (**B**) Gal80p and Gal80^{S-2}p, and Gal4p (all 2 μM) were mixed as indicated in the absence or presence of the cross-linker BS^3 (0.1 mM final concentration). Products were resolved by SDS/PAGE [10% (w/v) gel]. Lane M, molecular-mass markers.

seen when a mixture of the two proteins was applied in the presence of either galactose or MgATP (Figure 2B). However, when both proteins were mixed in the presence of both ligands, a new peak appeared. This peak corresponded to a molecular mass of 91 kDa, and SDS/PAGE analysis of fractions from it revealed the presence of both Gal80p and Gal3p. Visual inspection of the bands suggested that both proteins were present in approximately equal amounts (despite an excess of Gal3p being loaded). This, taken together with the apparent molecular mass, suggests that a complex containing one molecule of Gal3p and one molecule of Gal80p had been isolated, despite Gal80p existing in free solution as a variety of multimeric states [19].

Since Gal3p protects Gal80p from proteolytic digestion at amino acid residue 331, mutations in this region might be expected to disrupt the binding. Gal80^{S-2}p is a super-repressor mutant carrying the point mutation Glu^{351} → Lys [20]. Cells carrying this mutation are unable to induce the *GAL* genes even in excess galactose. The ability of this mutant protein to bind Gal3p was tested in the gel-filtration assay and was found to be unable to interact (Figure 3A). However, both the mutant and wild-type proteins could be cross-linked to Gal4p using the amine cross-linker BS^3 (Figure 3B). The extent of cross-linking was not significantly different when the two proteins were compared, suggesting similar affinities for Gal4p. Interestingly, when the two forms of Gal80p were exposed to cross-linker in the absence of Gal4p, Gal80p, but not Gal80^{S-2}p, underwent cross-linking, which resulted in a doublet on the gel (Figure 3B, lane 3). This may suggest that Gal80p is amenable to intramolecular cross-linking, but that this ability is disrupted in the mutant protein. Additionally, although there is only a single amino acid changed

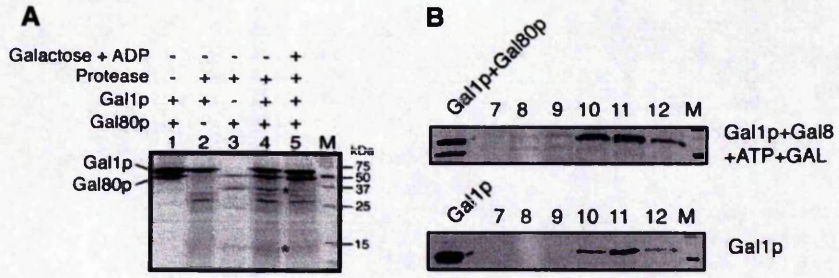


Figure 4 Interaction between Gal1p and Gal80p

(**A**) Binding of Gal1p to Gal80p protects the latter from proteolytic digestion. Gal1p (17 μM) was mixed with Gal80p (5 μM) in the presence of galactose and MgATP (both 2.5 mM) as indicated. Reaction mixtures were digested with chymotrypsin (300 nM) for 30 min at 30 °C, and then resolved by SDS/PAGE [15% (w/v) gel]. (**B**) A Gal1p–Gal80p complex can be isolated by gel filtration. A mixture of Gal1p (16 μM), Gal80p (1 μM), galactose (2.5 mM) and MgATP (2.5 mM) was applied to a Superdex S75 column equilibrated in running buffer supplemented with both ligands at 2.5 mM. Lane M, molecular-mass markers.

between the wild-type and mutant proteins, there may be significant structural alterations between Gal80p and Gal80^{s-3}p. However, any such changes only affect the ability of the protein to bind Gal3p and not its interaction with Gal4p.

Gal1p interacts with Gal80p in a similar manner to Gal3p

Gal1p was shown to bind Gal80p in both protease protection and gel-filtration experiments (Figure 4). The treatment of Gal1p with chymotrypsin yielded a large number of breakdown products (Figure 4A, lane 2). The combined digestion of Gal1p and Gal80p yielded an additive digestion pattern (Figure 4A, lane 4). The Gal80p-derived fragments (indicated by asterisks in Figure 4A, lane 4), overlapped with some of the Gal1p-derived fragments. It was still clear, however, that the protection of Gal80p from chymotryptic digestion required Gal1p, galactose and MgADP (Figure 4A, lane 5). MgATP was not tested in this experiment, as it is a substrate for the enzyme. In the presence of MgATP and galactose, a complex between Gal1p and Gal80p was isolated by gel filtration (Figure 4B). In this experiment, MgATP could be used, as the substrates were constantly replenished as the protein mixture moved down the column. Like the Gal3p–Gal80p complex, the Gal1p–Gal80p is a 1:1 stoichiometric complex, as judged by its apparent molecular mass (approx. 90 kDa) and the intensity of staining of the bands after SDS/PAGE (Figure 4B, lanes 8 and 9).

DISCUSSION

In order to understand how gene expression is switched on and off by systems such as the *GAL* genetic switch, it is necessary to know how the various protein and small molecule components of it interact and how these interactions influence each other. Both the protease protection and gel-filtration experiments show that there is a direct interaction between Gal3p and Gal80p in the presence of both ligands (ATP and galactose). It appears that there is an absolute requirement for the presence of both ligands in order to promote this interaction. Protease protection suggests that the interaction is centred at, or near, residue 331 in Gal80p. This conclusion is supported by the fact that Gal80^{s-3}p, a super-repressor mutant, is unable to bind to Gal3p, although its interaction with Gal4p is unaffected. The nature of this mutation is a single amino acid substitution at residue 351, close to the site of interaction suggested by protease protection.

Gel filtration clearly indicates that the interaction between Gal3p and Gal80p results in the formation of a 1:1 complex. Furthermore, that the complex could be isolated by this method shows that it has a relatively long half-life. Even though the equilibrium has been disturbed at the moment the mixture of complexed and uncomplexed proteins enters the gel-filtration column, sufficient remains after the 20 min required for separation so that it could be detected. A similar complex could be detected between Gal1p, the galactokinase, and Gal80p, although the yield of the complex was low. It is possible that the Gal1p–Gal80p interaction is less strong than the one between Gal3p and Gal80p. It has been noted previously that approx. 40-fold more Gal1p than Gal3p is required to activate the *GAL* genetic switch *in vitro* [9].

Gal3p acts as the ligand sensor of the *GAL* genetic switch. In the present study, we show that both its ligands, galactose and MgATP, are required for interaction with Gal80p. This interaction has a 1:1 stoichiometry and requires no other proteins or small molecules to bring it about. The high degree of similarity between the galactokinase, Gal1p, and the ligand sensor, Gal3p, suggest that much of what we learn about one will apply to the other. Although Gal3p has never been shown to be a galactokinase, Gal1p will act as the ligand sensor in the *GAL* genetic switch both *in vivo* [12] and *in vitro* ([9] and the present study). We have shown that Gal1p, when acting as a galactokinase proceeds via an ordered ternary-complex mechanism in which ATP binds first (D. J. Timson and R. J. Reece, unpublished work). It seems likely that this ordered ligand binding is preserved when Gal1p acts as a ligand sensor and also in Gal3p. We therefore conclude that *GAL* gene activation begins with the binding of ATP to Gal3p (or Gal1p). This induces a conformational change in the ligand sensor, creating a binding site for galactose. This Gal3p–ATP–galactose ternary complex is then capable of interacting with Gal80p, an event which relieves the inhibition of Gal4p and permits transcription to proceed.

We are grateful to members of the Reece laboratory for their comments on the manuscript. The work was funded by the Biotechnology and Biological Sciences Research Council and the Leverhulme Trust.

REFERENCES

1 Johnston, M. (1987) A model fungal gene regulatory mechanism: the *GAL* genes of *Saccharomyces cerevisiae*. Microbiol. Rev. **51**, 458–476
2 Lohr, D., Venkov, P. and Zlantanova, J. (1995) Transcriptional regulation in the yeast *GAL* gene family: a complex genetic network. FASEB J. **9**, 777–787
3 Reece, R. J. and Platt, A. (1997) Signaling activation and repression of RNA polymerase II transcription in yeast. BioEssays **19**, 1001–1010
4 St John, T. P. and Davis, R. W. (1981) The organization and transcription of the galactose gene cluster of *Saccharomyces*. J. Mol. Biol. **152**, 285–315
5 Zenke, F. T., Engles, R., Vollenbroich, V., Meyer, J., Hollenberg, C. P. and Breunig, K. D. (1996) Activation of Gal4p by galactose-dependent interaction of galactokinase and Gal80p. Science (Washington, D.C.) **272**, 1662–1665
6 Suzuki-Fujimoto, T., Fukuma, M., Yano, K.-I., Sakurai, H., Vonika, A., Johnston, S. A. and Fukasawa, T. (1996) Analysis of the galactose signal transduction pathway in *Saccharomyces cerevisiae*: interaction between Gal3p and Gal80p. Mol. Cell. Biol. **16**, 2504–2508
7 Yano, K. I. and Fukasawa, T. (1997) Galactose-dependent reversible interaction of Gal3p with Gal80p in the induction pathway of Gal4p-activated genes of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U.S.A. **94**, 1721–1726
8 Blank, T. E., Woods, M. P., Lebo, C. M., Xin, P. and Hopper, J. E. (1997) Novel Gal3 proteins showing altered Gal80p binding cause constitutive transcription of Gal4p-activated genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **17**, 2566–2575
9 Platt, A. and Reece, R. J. (1998) The yeast galactose genetic switch is mediated by the formation of a Gal4p/Gal80p/Gal3p complex. EMBO J. **17**, 4086–4091
10 Sil, A. K., Alam, S., Xin, P., Ma, L., Morgan, M., Lebo, C. M., Woods, M. P. and Hopper, J. E. (1999) The Gal3p-Gal80p-Gal4p transcription switch of yeast: Gal3p destabilizes the Gal80p–Gal4p complex in response to galactose and ATP. Mol. Cell. Biol. **19**, 7828–7840
11 Peng, G. and Hopper, J. E. (2000) Evidence for Gal3p's cytoplasmic location and Gal80p's dual cytoplasmic-nuclear location implicates new mechanisms for controlling Gal4p activity in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **20**, 5140–5148
12 Bhat, P. J. and Hopper, J. E. (1992) Overproduction of the GAL1 or GAL3 protein causes galactose-independent activation of the GAL4 protein: evidence for a new model of induction for the yeast *GAL/MEL* regulon. Mol. Cell. Biol. **12**, 2701–2707
13 Platt, A., Ross, H. C., Hankin, S. and Reece, R. J. (2000) The insertion of two amino acids into a transcriptional inducer converts it into a galactokinase. Proc. Natl. Acad. Sci. U.S.A. **97**, 3154–3159
14 Ansari, A. and Schwer, B. (1995) SLU7 and a novel activity, SSF, act during the PRP16-dependent step of yeast pre-mRNA splicing. EMBO J. **16**, 4001–4009
15 Reece, R. J., Rickles, R. J. and Ptashne, M. (1993) Overproduction and single-step purification of GAL4 fusion proteins from *Escherichia coli*. Gene **126**, 105–107
16 Partis, M. D., Griffiths, D. G., Roberts, G. C. and Beechey, R. B. (1983) Cross-linking of protein by ω-maleimido alkanoyl *N*-hydroxysuccinimide esters. J. Protein Chem. **2**, 263–277
17 Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**, 248–254
18 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**, 680–685
19 Melcher, K. and Xu, H. E. (2001) Gal80–Gal80 interaction on adjacent Gal4p binding sites is required for complete *GAL* gene repression. EMBO J. **20**, 841–851
20 Nogi, Y. and Fukasawa, T. (1984) Nucleotide sequence of the yeast regulatory gene *GAL80*. Nucleic Acids Res. **12**, 9287–9298

Received 1 November 2001; accepted 30 January 2002