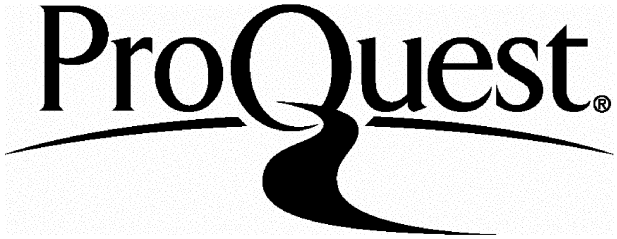
Transcription Factors Regulated by the Adenovirus E1A Gene

Patricia Alexandra Eccleston

A thesis submitted for the degree of
Doctor of Philosophy
January, 1993

University College London
Gower Street
London

Imperial Cancer Research Fund
Lincoln's Inn Fields
London



ABSTRACT

In this thesis, the yeast *S. cerevisiae* has been used to characterise the interaction of transcription factors with other proteins and with DNA. This involved three studies. Firstly, the mechanism of transcriptional activation by the adenovirus E1a gene was investigated. The initial goal was to reproduce E1a activation in yeast and subsequently take advantage of yeast genetics to analyse the mechanism in detail. The transcriptional process in yeast and mammalian cells is highly conserved. Despite this, I showed that expression of E1a in yeast was unable to modulate transcription mediated by yeast factors. One mammalian transcription factor shown to mediate E1a transactivation is ATF2. It had been proposed that a direct physical interaction occurs between the two proteins, so I investigated this possibility using a yeast genetic system. No such interaction was seen, although as a control, E1a was seen to interact with the retinoblastoma gene product.

An ATF binding activity is present in yeast itself. The second study used a yeast screen with an activation domain tagged cDNA library in order to isolate the yeast ATF transcription factor. Clones were isolated that specifically activated an ATF site. However, they may be involved in the regulation of this site, rather than encode the ATF transcription factor itself.

E2F is another mammalian transcription factor that has been shown to mediate E1a transactivation. The interaction of E2F with a number of cellular proteins may modulate its activity in a cell cycle dependent manner. One result of adenovirus infection is the dissociation of these complexes by E1a and the interaction of E2F with a viral E4 gene product. The third study involved an attempt to isolate an E2F cDNA. A yeast genetic screen was set up to isolate cDNA clones that could either bind directly to an E2F DNA binding site or bind to the adenovirus E4 protein. Screens for direct binding to the E2F site did not yield any positive clones, possibly because such a screen would depend upon E2F binding as a homodimer. A screen using the E4 protein yielded clones whose products could interact with E4 in yeast. Preliminary analysis of these clones has been carried out in *in vitro* systems and in mammalian cells.

TO MY PARENTS

ACKNOWLEDGEMENTS

Many thanks to everyone who has helped me during my work for this thesis. My supervisor Nic Jones deserves much acknowledgement for his ideas and support that was forthcoming even when times were desperate. Others include all the members of the Jones laboratory, past and present, and many members of the Triesman, Bentley, Goodbourne, Fried and Land labs. Especial thanks to Steve Dalton and Gos Micklem. In addition, I appreciated the company of all those in The George, and I must also acknowledge the loyal friends who always remembered to ring me on Friday afternoons!

TABLE OF CONTENTS

ABBREVIATIONS

APS	Ammonium persulphate
ARS	Autonomously replicating sequence
ATP	Adenosine triphosphate
bp	Base pair
BSA	Bovine serum albumin
cAMP	3', 5' cyclic adenosine monophosphate
CAT	Chloramphenicol acetyl transferase
CEN	Centromere
Ci	Curie
cpm	Counts per minute
dATP	3' deoxyadenosine 5' triphosphate
dCTP	3' deoxycytosine 5' triphosphate
dGTP	3' deoxyguanosine 5' triphosphate
dNTP	3' deoxynucleotide 5' triphosphate
dTTP	3' deoxythymidine 5' triphosphate
ddATP	2', 3' dideoxyadenosine 5' triphosphate
ddCTP	2', 3' dideoxycytosine 5' triphosphate
ddGTP	2', 3' dideoxyguanosine 5' triphosphate
ddNTP	2', 3' dideoxynucleotide 5' triphosphate
ddTTP	2', 3' dideoxythymidine 5' triphosphate
DNA	Deoxy ribonucleic acid
DNAse	Deoxy ribonucleic acid endonuclease
DTT	Dithiothreitol
EDTA	Ethylenediamine tetra acetic acid
Fig	Figure
g	Gram
G,A,T,C	Guanine, Adenine, Thymine, Cytosine
GTP	Guanosine triphosphate
HEPES	N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulphonic acid]
hr	Hours
K	Thousand
kb	Kilobase pair
kg	Kilogram
M	Molar
mA	Milli ampere
mg	Milligram
min	Minutes
ml	Milli litre

mM	Millimolar
mRNA	messenger RNA
MW	Molecular weight
ng	Nanogram
OD	Optical density
oligo	Synthetic oligonucleotide
ONPG	O-nitrophenol-b-D-galactopyranoside
ORF	Open reading frame
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PFU	Plaque forming units
PMSF	Phenylmethyl sulphonyl flouride
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT	Room temperature
S	Sedimentation coefficient
SDS	Sodium dodecylsulphate
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	Tris (hydroxymethyl) aminomethane
tRNA	Transfer RNA
UAS	Upstream activator sequence
V	Volt
W	Watt
WT	Wild type
µ	micro
Δ	Deletion

Single letter amino acid code:

A	Alanine
C	Cysteine
D	Aspartic acid
E	Glutamic acid
F	Phenylalanine
G	Glycine
I	Isoleucine
K	Lysine
L	Leucine
M	Methionine
N	Asparagine
P	Proline

Q	Glutamine
R	Arginine
S	Serine
T	Threonine
V	Valine
Y	Tyrosine

CHAPTER 1: INTRODUCTION

The regulation of gene expression is fundamental to eukaryotic development, and the primary control point of this regulation is the initiation of RNA transcription. The DNA tumour viruses have proved to be a very productive and convenient model system for studying the regulation of gene expression, and in particular, transcriptional initiation. The viral genes themselves are highly regulated, being expressed at specific times during the infective process. Since the virus depends almost entirely upon host cell machinery, studying the mechanism by which this regulation is accomplished has provided valuable information relevent to the regulation of cellular genes. The human adenoviruses have been studied intensely because their simple genome facilitates genetic analysis. The early region 1a (E1a) gene of adenovirus was found to have a key role in gene regulation and transformation by adenovirus. Work for this thesis has centred around the involvement of E1a in transcriptional activation.

In this introduction I will describe transcription regulation in general; I will then introduce the budding yeast S.cerevisiae, which has been used as a system for studying transcription because of the genetic and molecular manipulations it enables, and the evolutionary conservation of the transcription mechanism. The regulation of transcription by E1a will then be discussed, and why studying this regulation has led to the detailed analysis of two transcription factors, ATF and E2F. Relevant background for the experimental work will be discussed in this chapter, and more recent developments that have a direct bearing on my work will be reviewed further in the discussion section.

1.1 General features of transcriptional regulation

The regulation of gene expression at the level of the initiation of RNA transcription is controlled by promoter and enhancer elements, located mainly upstream of the transcriptional start site, occasionally downstream. Transcription initiation from an RNA polymerase II (polII) transcribed gene is dependent upon the core elements of the promoter. The most common of these are the TATA motif, and the initiator element located about 30 nucleotides downstream of it. Polypeptides which bind in a complex around the TATA box region and are required for basal levels of transcription have been defined as general transcription factors. Upstream of the TATA box region are proximal and distal cis-acting DNA elements that bind specific transcription factors that either increase or decrease transcription initiation. Sometimes, these control elements are in an enhancer region that can function in both orientations and at variable distances from the transcription start site. (See Dynan, 1989 for review). The combination of factors binding to a particular promoter determines the pattern of gene expression. To understand the regulation of gene expression, it is necessary to know the detailed mechanism of the interaction of transcription factors with DNA and with the initiation machinery.

1.1.1 The initiation complex

In addition to polII, several general transcription factors have been characterised that are important in transcription initiation, including TF IIA, IIB, IID, IIE, IIF, IIH and IIJ. Human cDNA clones for all but IIA and IIJ are available. These factors assemble in an ordered manner to establish a functional preinitiation complex. TFIID, the only one of these factors with an intrinsic specific DNA binding ability, binds to the TATA element, followed by or facilitated by TFIIA. IIA may subsequently leave the complex. This complex is recognised with or without TFIIA, by IIB which is thought to bridge IID and pol II. PolII is escorted by IIF, and TFIIE, H and J are then required in that order to complete the complex, which will initiate RNA synthesis upon addition of rNTPs and ATP. Phosphorylation of polII is likely to be the trigger that drives the complex into the elongation stage. (For review see Zawel and Reinberg, 1992).

TFIID occurs as a multisubunit protein complex with a molecular mass of greater than 100kD. The only TFIID polypeptide that has been cloned is the TATA binding protein (TBP) which is 38kD in humans (eg, Kao et al., 1990). The other TBP-associated factors (TAFs) that form the TFIID complex are not required for basal levels of transcription at least *in vitro*, but are necessary for the stimulation of transcription by the transcriptional activators that bind to the proximal and distal elements.

1.1.2 Modular structure of transcription factors

It has become apparent that transcription factors are composed of distinct domains that can function independently. The DNA binding domain is normally located in a separate region to parts of the protein required for transcriptional regulation. This is evident from the behaviour of mutants that retain one of these properties but not the other, and of chimaeric proteins in which DNA binding and activation domains can be interchanged. Brent and Ptashne (1985) initially demonstrated this by swapping domains between LexA and GAL4. Some features of DNA binding and activation domains will be discussed.

1.1.3 Binding and dimerisation domains

DNA binding domains often include sequences mediating dimerisation, since binding often requires dimerisation. Different motifs involved are discussed below, though several transcription factors have domains that do not appear to belong to any of these groups, eg SRF.

The helix-turn-helix motif was first discovered in bacterial DNA binding proteins such as CRO, cI and CAP, whose three dimensional structure was shown to consist of two α helices separated by a β turn. The proteins bind as dimers using both monomers to recognise and bind to target sequences. This motif has evolved into the homeodomain of transcription factors such as the yeast factors mat α 1 and 2 that regulate mating type, and the homeotic genes of drosophila important in development (Shepherd et al., 1984). A more complex structure is found in factors such as the

octamer binding proteins Oct 1 and 2. They have a POU domain which includes a homeobox and extra POU specific region (Herr et al., 1988).

The DNA binding region of TFIIIA was found to contain nine repeats of a sequence with two pairs of cysteine and histidine residues. This folds into a finger-like structure of 20 to 30 amino acids which binds a single atom of zinc and directly contacts DNA (Miller et al., 1985). Many other examples of this motif have since been found in DNA binding proteins (Evans and Hollenberg, 1988). Members of the steroid hormone receptor family have two zinc fingers of four cysteine residues which maybe of a different family to the other. GAL4 and E1a also have this type of finger. In the case of E1a it may be associated with protein-protein interactions rather than binding to DNA.

The leucine zipper structure is important in the dimerisation of transcription factors such as C/EBP, Jun, Fos and CREB/ATF (Landschulz et al., 1988). It contains a repeat of generally four to five leucines occuring at every 7th position in an a helical structure. Leucines align on the same side of the helix to form a hydrophobic surface and in the dimer the leucines from two helices interact side by side. X-ray diffraction showed that the leucine zipper of GCN4 is a coiled coil orientated in parallel (Rasmussen et al., 1991). Other hydrophobic residues between the leucines are important in the structure, and additional residues must also account for the specificity in dimer formation. This specificity lies within the zipper region, and has been shown in some cases to be determined by interhelical electrostatic interactions between oppositely charged residues (O'Shea et al., 1992). As well as forming homodimers, zipper containing proteins selectively form heterodimers between different family members and between different families, serving to increase the diversity of promoter-bound complexes. Dimer formation results in the correct structure for DNA binding, often to dyad-symmetric target sequences, the two monomers interacting with adjacent half-sites. A basic region adjacent to the zipper actually contacts the DNA and determines the DNA binding specificity (bZIP family).

The basic DNA binding domain can also be associated with a domain where two amphipathic helices are separated by a non-helical loop (bHLH family). It may function in a similar manner to the leucine zipper to mediate dimerisation, although no structural evidence for this is yet available. Examples of such a helix-loop-helix structure include the factors involved in muscle differentiation such as MyoD, and E12 and E47 which bind to the immunoglobulin enhancer. Homodimers and heterodimers can be formed to increase the diversity of DNA bound factors (Murre et al., 1989). For both bZIP and bHLH families different interactions can lead to functional differences exhibited by the heterodimers both in terms of their ability to bind DNA and their regulation (for review, see Jones, 1990a).

1.1.4 Activation domains

There are several classes of domains capable of mediating transcription activation. They have been defined by the prevalence of certain residues, but do not show

complete sequence identity within a family, and several known activation domains do not have a recognisable motif. The yeast activators GAL4 and GCN4, and the herpes simplex virus (HSV) VP16 activator have acidic stretches of amino acids enriched in glutamate and aspartate that may form amphipathic α helices. To look at the contribution of particular residues to activation, random mutations of the GAL4 activation domain were made and tested in yeast. Mutations that increased activation, increased the acidity of the domain, although not all down mutations decreased acidity. Features in addition to acidity must also be necessary for activation (Gill and Ptashne, 1987). Ma and Ptashne(1987) showed that random fragments of bacterial DNA fused to the GAL4 DNA binding domain could weakly activate transcription in yeast if they had an overall negative charge. Many of these activating fragments could form amphipathic α helices with negative charges on one surface and hydrophobic residues on the other. A synthetic α helix of this structure will weakly activate transcription in yeast, though not if the same residues are in a scrambled order, indicating again that the structure of the activating region as well as its overall charge is important for activation (Giniger and Ptashne, 1987). However, although acidic activators can activate transcription well in mammallian cells, not many mammalian transcription factors contain acidic activation domains. Examples that do are Jun and perhaps the glucocorticoid receptor. Other classes of activation domain are less well characterised. A glutamine-rich region is found in activation domains such as that of Sp1 and Oct2 (Courey and Tjian, 1988), and a proline rich activation domain in factors such as CTF/NF1 and AP2 (Mermod et al., 1989).

1.1.5 Mechanism of activation.

Models for how transcription factors stimulate the general machinery at the TATA box revolve around the theory that activation domains directly or indirectly contact one of the basal factors, looping out the intervening DNA and thereby stabilising or activating the target factor on the promoter. There are at least three lines of evidence for this. Firstly, the footprint of TFIID on the TATA box is altered *in vitro* by upstream activators such as GAL4 and ATF (Horikoshi et al, 1988a and b). Secondly, specific interactions have been shown between VP16 and TBP and TFIIB (see below), and between GCN4 and polII *in vitro* (Brandl and Struhl, 1989). Finally, high level expression of an activator can inhibit normal transcription, presumably by sequestering a target factor away from the promoter. An example of this was shown by Gill and Ptashne (1988) where derivatives of GAL4 expressed at high levels inhibited transcription from certain genes. The inhibition was stronger with GAL4 mutations that were stronger activators and the DNA binding domain of GAL4 was not required. The actual target of activation domains remains under speculation. There is recent evidence for the general factors being the direct target, such as TBP or TFIIB, or additional mediator or adaptor proteins. Alternatively, the target could be a component of chromatin.

Evidence for a direct interaction between activation domains and components of the basic transcriptional machinery include experiments using affinity chromatography.

Both TBP and TFIIB have been shown to bind to a column containing the VP16 activation domain. The binding of both is sensitive to a transcriptionally inactive form of VP16 that has a single amino acid mutation. (Stringer et al., 1990; Ingles et al., 1991; Lin et al., 1991). However, the interactions could reflect non-specific ionic interactions rather than have any functional relevence. Lin and Green (1991) showed that the interaction of TFIIB with an acidic activator seemed to be a stronger than the interaction with TBP, and that the recruitment of TFIIB by the activator into the pre-initiation complex was the rate-limiting step.

There is evidence that additional factors may be targeted by activation domains apart from the general factors. Berger et al (1990) showed in a yeast *in vitro* system that a GAL4/VP16 fusion inhibited activation by the dA:dT activator, but did not affect basal transcription. Kelleher et al (1990) showed similar inhibition both *in vitro* and *in vivo*, and partially purified a novel yeast component that served as a mediator to relieve the inhibition. Additional evidence for the existence of such mediators is the fact that purified basal transcription factors are unable to support activated transcription *in vitro* unless cellular extracts are added. These targets of activation domains that are additional to the general factors have been called adaptors and may be equivalent to some of the TAFs seen in mammalian cells in the TFIID complex. They are postulated to mediate interactions between upstream activators and the basal machinery, either as a bridge or alternatively they may be required for subsequent steps in activation. No clones for such adaptors have yet been obtained from mammalian cells, although highly purified activities exist. Some functions of TAFs are clear, and may be different to adaptors. TAFs are likely to be important in promoter specificity, for example a different set of TAFs is found in the TFIID complex at a polymerase I transcribed promoter from a polII transcribed one (Comai et al., 1992). Therefore TAFs may dictate the choice of RNA polymerase. They may also have an important role in transcription at TATA-less promoters. Some TAFs are negative regulators of transcription. For example, a recently cloned activity is the Dr1 protein, which is thought to inhibit the binding of TFIIA and/or IIB with TBP and therefore repress both basal and activated transcription (Inostroza et al., 1992) Identification of the TAFs and adaptors and resolving how they interact with upstream transcription factors is vital in understanding the mechanism of activation.

Further experiments have lead to the proposal that different types of activation domain see different targets, or see different adaptors which may interact with a common target. For example, a TBP associated fraction was able to restore activated transcription *in vitro* by a proline rich or glutamine rich activation domain, but not by an acidic one (Tanese et al., 1991). Also, by overexpressing an activating region that can inhibit the activity of another one (as described above), the targets of these various regions can be compared. Tasset et al (1990) suggested that the two activation domains of the human oestrogen receptor (hER) were different from each other and from VP16 based on their inability to transcriptionally interfere with each other. The targets of the

hER activation domains seemed to be required by VP16, but the target of VP16 was not required for the transcriptional activity of hER, possibly implying more than one target interacting in series.

A further level of complexity is added to transcription regulation by the chromatin structure of DNA *in vivo*. Models for transcription factor action must incorporate this, and an important role of transcription factors may be to free the promoter regions of histones. Chromatin has a strong repressive effect on transcriptional initiation (see Felsenfeld, 1992 for review). Transcription activation *in vitro* is seen to a much larger extent when the DNA is in a nucleosome structure. Both the binding and activation domain of GAL4 is required to alleviate the repression of transcription by nucleosomes, suggesting that the acidic domain can enhance the ability of the basal complex to compete with the nuclesomes in binding to the promoter (Workman et al., 1991).

Transcriptional activation by factors bound to the promoter is synergistic in that multiple binding sites activate 10 to 100 times more than a single site (reviewed in Ptashne, 1988). The mechanism by which synergy occurs is unknown, and may involve multiple features. It can occur between unrelated activators and can depend upon both the number of proteins bound to the promoter and the number of activation regions. Models suggested include that multiple activators simultaneously contact one or more components of the basal transcriptional machinery, leading to a cooperative increase in complex assembly. Alternatively, synergy could be mediated by the DNA binding domain having a direct role in the displacement of histones, or by itself interacting with the transcriptional machinery or, in some cases, by cooperative DNA binding. (Oliviero and Struhl, 1991; Emami and Carey, 1992).

1.2 Transcriptional regulation in *S.cerevisiae*

Many essential elements of transcriptional regulation have been conserved in eukaryotic evolution and identification of these should help elucidate central mechanistic features (see Guarente and Birmingham-McDonogh, 1992 for review). Similarities can be seen to many aspects of the mammalian transcription apparatus in the fission yeast *S. pombe* and the budding yeast, *S. cerevisiae*. Since *S. cerevisiae* allows greater molecular and genetic manipulations, this was the yeast chosen to use in these studies. Features of transcriptional regulation specific to *S. cerevisiae* will be described, with some examples of the way these features have been conserved.

The main elements of the mammalian promoter have been conserved in *S. cerevisiae*. A sequence resembling the mammalian TATA box is found in most yeast promoters near the transcription initiation site. The yeast upstream activation site (UAS) has equivalent characteristics to the mammalian enhancer. However, there are some differences in initiation. Yeast factors cannot activate transcription if bound downstream of the initiation site, although they can in mammalian cells. Also, the

distance between the TATA box and the initiation site is large and variable in *S. cerevisiae* promoters, rather than shorter and of fixed distance as in mammalian and *S. pombe* ones. Purified *S. cerevisiae* TFIID initiates transcription at the correct distance for mammalian promoters in a mammalian *in vitro* transcription system (Buratowski et al., 1988), so factors other than TFIID, or in addition to it, must select the start site in yeast. An example of such a factor is encoded by the SUA7 gene (see below).

1.2.1 General transcription factors

Many of the mammalian general factors that form a complex on the TATA element have conserved equivalents in yeast. Conservation is seen in several subunits of the RNA polymerases. The gene for the large subunit of RNA polymerase II has been cloned from yeast and mammals and was found to have a homologous primary sequence. Furthermore, in *S. cerevisiae*, a deletion of the C-terminal domain, which consists of heptapeptide repeats, was lethal, but viability was restored with the equivalent hamster domain (Allison et al., 1988). Mutations in the SPT15 gene of *S. cerevisiae* were isolated as suppressors of mutations in the 5' regions of genes which affected their transcription. SPT15 was found to encode yTBP, and is required for normal transcription initiation (Eisenmann et al., 1989; Hahn et al., 1989b). The sequence of this gene enabled cloning of homologs from *S. pombe*, Drosophila, Arabidopsis and humans due to the high degree of evolutionary conservation of the sequence (eg Fikes et al., 1990; Hoffman et al., 1990). The C-terminal core region of about 180 amino acids, which is sufficient for DNA binding and basal transcription, is at least 80% identical between all these species. Yeast TBP lacks the extended N-terminus present in the human and drosophila clones. It appears as a single polypeptide: if it does associate with other factors in the same way that hTBP associates with TAFs, then the interactions must be weaker. Both purified yTFIID and yTFIIA will complement in basal transcription in a mammalian *in vitro* transcription reaction lacking these factors, indicating their functional conservation (Buratowski et al., 1988; Hahn et al., 1989a). However, hTBP cannot replace the yeast factor *in vivo*, implying that there are differences for activated transcription. The yeast SUA7 gene is required for normal start site selection and has 35% identity to human TFIIB (Pinto et al., 1992). Whether this conservation extends to the functional level remains to be determined.

1.2.2 Upstream DNA binding proteins

Some examples follow of yeast transcription factors which have similar DNA binding specificities as factors in mammalian cells. In some cases the degree of conservation also extends to the level of the protein-protein interactions of these factors.

GCN4 and yAP1

GCN4 is a positive regulator of amino acid biosynthesis. GCN4 and a related but distinct yeast protein yAP1, which is also an activator, bind to a promoter element with a similar consensus sequence to the mammalian TPA-responsive element (TRE). The TRE is the binding site for members of the mammalian AP1 family, which includes cJun and cFos. The DNA binding domains of GCN4, yAP1 and Jun are homologous and they also share a leucine zipper structure (Moye-Rowley et al., 1989). Jun will substitute for GCN4 in yeast to activate the HIS3 gene to a similar extent as GCN4 (Struhl, 1988), and the acidic activation domain of GCN4 will function in mammalian cells as a transactivator, although it does not have the oncogenic properties of Jun (Oliviero et al., 1992).

yATF

In mammalian cells, the consensus sequence of the TRE differs by only a base pair from the CRE, and this change distinguishes the family members able to bind to these sites. This specificity is conserved in yeast, since there is a factor with binding activity identical to mammalian ATF and distinct from GCN4 or yAP1. The ATF/CREB binding site can act as an efficient UAS in yeast, indicating that yATF may be an activator. A point mutation in the binding site will decrease the efficiency of the site to act as a UAS, and also abolish the gel retardation activity of a yeast extract seen on a probe. Like mammalian CREB1, the yATF complex can be phosphorylated by protein kinase A, but whether this reflects evolutionary conservation of an aspect of its regulation remains to be shown (Jones and Jones, 1989; Lin and Green, 1989). The cloning of a factor that binds to the ATF site in yeast will help to elucidate its regulation.

CCAAT box binding proteins

Another example of the conservation of DNA binding specificity and addditionally of protein-protein interactions is shown by the CCAAT box binding proteins. In yeast, a heteromeric complex of HAP2,3 and 4 binds this element to regulate genes involved in the biogenesis of mitochondria. In mammals there are several classes of CCAAT box binding proteins, but the complex of CP1 and 2 binds the element with a sequence most similar to the yeast one. The subunits CP1A and CP1B are required for DNA binding and can be exchanged with HAP3 and HAP2 repectively (Chodosh et al., 1988). A conserved region of HAP2 that has 60% homology with its human homolog, contains the domains required for DNA binding and interaction with HAP3. However, there is evolutionary divergence in the regulation of these complexes since yeast contains an additional subunit HAP4, which responds to regulatory signals and provides an acidic activation domain. The mammalian HAP2 homolog has its own activation domain.

MCM1

A yeast factor encoded by the MCM1 gene has a DNA binding specificity related to the serum response factor (SRF). Both MCM1 and SRF appear to be activators and they share extensive homology over a region that defines DNA binding, dimerisation and formation of a ternary complex. SRF responds to serum induction, while MCM1 responds to mating pheremones and both mediate their response through a multiprotein complex. SRF can form a ternary complex on DNA with p62, which may be SRF accessory protein 1(SAP1) or Elk-1. This complex is probably essential for the induction of the c-fos gene. MCM1 associates with different polypeptides, dependent upon the promoter to which it is bound, to regulate cell type specific transcription. Activation of an α-specific UAS such as that of the STE3 gene, requires the association of MCM1 with the protein α1, whereas it can repress a specific genes via a complex with α2. The ternary complex it forms with STE12 to activate the a specific gene STE2 most resembles the SRF complex. MCM1 can interact with SAP1 or Elk1 using the same sub domain as SRF, and by replacing four SRF residues with the corresponding MCM1 residues, SRF becomes able to interact with STE12 (Hipskind et al., 1991; Mueller and Nordheim, 1991; Dalton and Treisman, 1992).

1.2.3 Regulation via GAL4

The yeast GAL4 protein is required for transcription of the galactose metabolising genes that are activated during growth in galactose medium and repressed during growth in glucose. It binds as a dimer to the UAS_G upstream of genes such as GAL1 and GAL10 and its binding site will confer galactose inducibility on heterologous promoters. The GAL4 binding and dimerisation domain is localised within amino acids 1 to 147, which will not activate transcription. The C-terminus contains two activation domains which will work when fused to a heterologous DNA binding domain and will also activate transcription in mammalian cells (Kakidani and Ptashne, 1988). This domain has a high proportion of acidic residues. Under uninduced conditions, GAL4 occupies its binding site, but the activation domains are thought to be blocked by an interaction with the negative regulator GAL80. Upon induction in galactose medium, the model currently thought to operate is that some unknown signal changes the conformation of the GAL4-GAL80 complex, so that the GAL4 activation domain is exposed (Leuther and Johnston, 1992). Several mechanisms contribute to the repression of galactose metabolising genes by glucose; one involves GAL4 being prevented from binding to its UAS (Giniger et al., 1985).

1.2.4 Activation domains

The type of activation domain found in most yeast activators is thought to belong to a class of acidic domains that may function by a mechanism that is universally conserved. Yeast activators such as GAL4 and GCN4, which contain acidic activation domains, will stimulate transcription in mammalian cells (Kakidani and Ptashne,

1988). There are several examples of mammalian activation domains that will function in yeast when tethered to the promoter, some of which have been shown to contain acidic residues, but have no obvious sequence homology beyond this. c-Fos or Jun, when fused to the DNA binding domain of the bacterial repressor LexA, will activate in yeast via the LexA operator (Lech et al., 1988; Struhl, 1988) and hormone receptors such as the oestrogen and glucocorticoid receptors will activate transcription from their respective binding sites in response to their ligands (Metzger et al., 1988). p53 has also been shown to be a potent transactivator in yeast (Fields and Jang, 1990), as has the herpes virus activator, VP16 (Cousens et al., 1989). Many non-acidic mammalian activation domains that are, for example, glutamine or proline rich, have not been shown to work in yeast. This could be explained by the differences in the structure of some general factors, such as TBP, and by the existence of species-specific adaptors for each type of activation domain, with adaptors for acidic activation domains being either conserved, or not required at all. However, the situation is perhaps more complicated, since many of the mammalian activation domains that work in yeast do not seem to belong to the same class, and may interact with different adaptors. For example, transcription repression experiments have shown that the target for the activation domains of VP16 and GAL4 is also required by Jun, but probably not required by the glucocorticoid receptor, and not required at all by the oestrogen receptor. As mentioned, all these activators work in yeast. (Oehler and Angel, 1992).

Two yeast factors that may be possible examples of transcriptional adaptors have recently been cloned by the use of genetics. GAL11 (or SPT13) is a gene required for full expression of the galactose inducible genes in yeast and is also thought to be required for the transcription of additional genes. A null mutant exhibits pleiotrophic effects of defective transcription. The GAL11 product may form a complex with GAL4, and is itself a transcriptional activator when fused to a DNA binding domain (Himmelfarb et al., 1990; Long et al., 1991). The second example is ADA2, also a non-essential gene, which is required for the full activity of certain acidic activation domains such as VP16 and GCN4. Mutations in ADA2 do not affect basal levels of transcription (Berger et al., 1992). Further analysis of genes such as these should help to reveal the mechanisms of action of eukaryotic transcription factors.

1.3 The adenovirus E1a gene

The adenovirus genome encodes products necessary for a viral infection. The proteins synthesised at late times are required for virion assembly, while the early proteins are required for viral replication. Regulation of the transcriptional initiation of these early and late genes is an important way their expression is controlled. The product of the adenovirus E1a gene has key multifunctional roles: it is the first gene to be expressed upon viral infection of cells and it activates the transcription of the other

viral early genes during a productive infection. Co-operating with the E1b gene or other nuclear oncogenes, E1a can transform primary cells.

Early after an adenovirus infection of cells, there are two major transcripts produced from the E1a gene, the 13S and 12S mRNAs generating products of sizes 289 and 243 amino acids respectively. Upon comparing the amino acid sequence of the 13S E1a products from different adenovirus serotypes, three evolutionarily conserved regions can be seen, designated conserved regions (CR) 1 to 3. The conservation of these domains indicates their functional importance, as does the behaviour of mutations within these domains. The 13S E1a product has a unique region of 49 amino acids comprising CR3 that is almost entirely deleted in the 12S product. This unique domain imparts upon the 13S protein a transactivation ability that is not as apparent in the 12S version. Mutations within CR3 abolish the ability of E1a to activate early viral transcription. Both 12S and 13S products have overlapping functions apart from those shown by CR3. (See Moran and Matthews, 1987 for review). They are highly phosphorylated, but no functional consequence of this modification has been shown (Tremblay et al., 1989), nor are any other modifications known of. The proteins are localised in the nucleus.

The functions of CR1 and 2 are both critical for the transforming ability of E1a. They bind to a number of cellular proteins which can be immunoprecipitated with E1a in extracts from infected cells (Harlow et al., 1986). The identity of many of these proteins is becoming clear, and some examples follow, many of which are involved in cell cycle regulation. The105kD product of the retinoblastoma gene (Rb) is a negative regulator of cell growth and loss of both alleles is associated with some tumours (Whyte et al., 1988). p107 is structurally related to Rb in its E1a-interacting domain and has been recently cloned (Ewen et al., 1991), but its function is not yet understood. It can associate with a polypeptide p60, which is the product of the cyclin A gene, and with cyclin E. A 300kD protein interacts with the N-terminus of E1a, another region critical in its transformation ability. (See Dyson and Harlow, 1992 for review). Other DNA tumour viruses also target these proteins, such as the polyoma and SV40 T antigens and the papillomavirus E7 gene, which bind to Rb and p107. This stresses further that the physical interactions between viral and cellular proteins may be important in the transformation process. The transactivating ability of E1a imparted by CR3 may also be dependent upon physical interactions with cellular proteins. Once the cellular targets have been identified, understanding the consequences of the interaction of E1a with such proteins should elucidate the important cellular pathways that it interferes with. The transactivating activities of E1a will be discussed in greater detail below.

1.3.1 General features of E1a transactivation

13S E1a stimulates transcriptional activity from all the adenoviral early promoters, as well as from two cellular genes: the human heat shock gene and b-tubulin (see Nevins, 1989 for review). More examples of activation of cellular genes are apparent using transfection assays; perhaps the chromatin structure of genomic DNA

prevents more general activation by E1a. Both RNA polymerase II and III dependent transcription can be affected by E1a. The diverse promoters that can be activated by E1a do not share any one common sequence element, and each promoter has a variety of different sequences implicated in E1a activation. Maximal basal transcription and E1a activation both require the presence of all the elements of the adenovirus promoters (see fig1.1). A study of a series of artificial promoters showed no correlation between a specific binding element and E1a activation (Taylor and Kingston, 1990). In addition, E1a has not been shown to have any sequence-specific DNA binding ability. To explain these properties, activation must occur through a mechanism which modulates the activity of the basic transcriptional machinery or of a wide variety of transcription factors. The individual elements that can mediate an E1a response have been studied in detail with the hope that once a model example has been established with one element, a similar general mechanism will operate for each case. Elements implicated in E1a inducibility because of their role in the adenovirus promoters and their ability to function in heterologous promoters, have included the ATF/CREB site and the E2F site, which have been chosen for further study here.

1.3.2 E1a activation via ATF/CREB

The E2A, E3 and E4 early promoters all contain an element with the core sequence CGTCA which can bind activating transcription factor (ATF) (Lee et al., 1987). In the E4 promoter, the ATF sites are critical for basal and E1a induced transcription, and will also confer E1a inducibility on a heterologous promoter (Pei and Berk, 1989). The ATF site can be identical to the cAMP response element (CRE) which binds the closely related CRE binding protein (CREB). The CRE can mediate activation by cAMP and Ca^{2+}, and transcription can be activated synergistically by E1a and cAMP. At least ten different mammalian cDNAs have so far been cloned that bind to the ATF/CREB site (Hai et al., 1989). They share a homologous basic DNA binding region and contain a leucine zipper dimerisation domain, but most are unrelated outside these domains. As well as forming homodimers, several of them can form heterodimers able to bind to the ATF site, thus creating additional functional diversity. Heterodimerisation also occurs between ATF/CREB proteins and members of the AP1 family (Benbrook and Jones, 1990; Hai and Curran, 1991).

Characterisation of the different family members has begun. Due to endogenous CREB proteins in mammalian cells, studies have often been carried out using fusions to a heterologous DNA binding domain, such as that of the yeast GAL4 protein, since no mammalian proteins bind to its recognition site. GAL4 fusions were made of three of the family members, CREB1, CREB2 (or CRE-BP1, ATF2) and ATF1, and their activation by E1a measured on a reporter with GAL4 binding sites. It was found that only CREB2 could be activated strongly by E1a, CREB1 was activated weakly and ATF1 not at all. This contrasted with the response to cAMP, where CREB1 was strongly activated, ATF1 weakly and CREB2 not at all (Flint and Jones, 1991). Deletion studies of the

Promoters stimulated by Adenovirus E1a

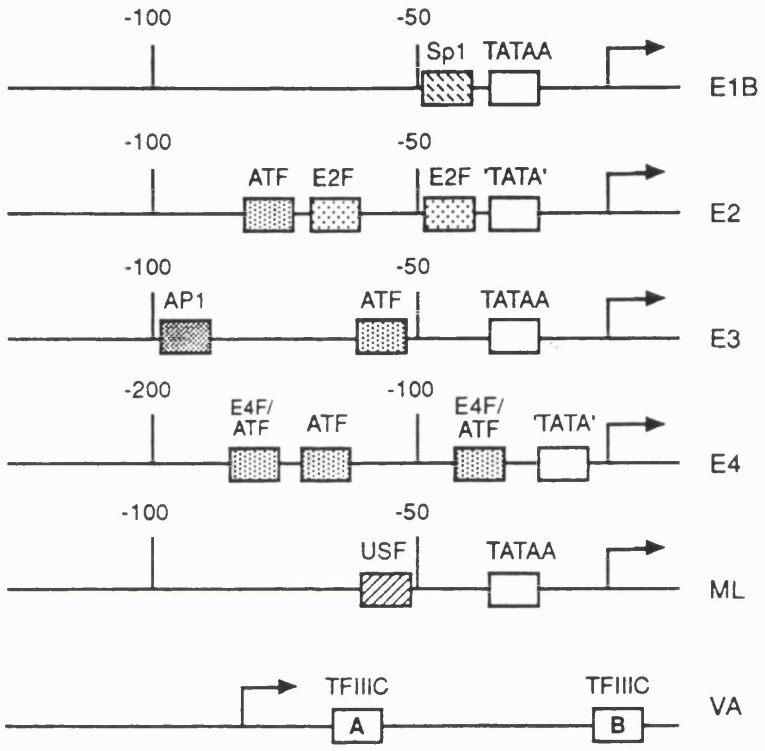


Fig 1.1

Adenovirus promoters stimulated by E1a showing the cis-acting elements and DNA binding factors important for transcription. Promoter elements are indicated by the boxes and arrows indicate the transcriptional start sites. All the promoters are transcribed via RNA polymersase II, except the VA gene which is transcribed via RNA polymerase III and contains promoter elements downstream of the transcriptional start sites.

GAL4/CREB2 fusion showed that the N-terminal region of CREB2 between residues 19 and 112 was essential and sufficient for the E1a response. Within this region lies a potential metal binding finger (C2H2 structure) which may be involved in a protein-protein interaction. This structure is critical since mutations of the cysteine residues decrease activation. Maekewa et al (1991) came to a similar conclusion, but suggested that sequences in addition to the N-terminus of CREB2 were required. Their experiments were performed with CREB2 fused to the heterologous DNA binding domain of cMyb, which, unlike GAL4, can interact with DNA as a monomer. Perhaps an interaction of CREB2 with E1a only occurs with CREB2 as a dimer, therefore in the cMyb fusion, the dimerisation domain of CREB2 is required. Further studies are necessary to elucidate this discrepancy. There may be other members of the ATF/CREB family that can be activated by E1a, but the response by CREB2 can serve as a model to study the mechanism by which E1a activation takes place.

Mechanism

E1a is unlikely to increase the concentration of any factor, since some E1a activation can be seen *in vitro* and in the absence of protein synthesis (Green et al., 1988). There is no evidence that E1a is a kinase, nor that it changes the level of binding of transcription factors when a promoter is activated, with some possible exceptions where phosphorylation is involved, as described later.

The CR3 of E1a (residues 139 to 189) can function as a minimal activation domain on the adenovirus early gene promoters (Lillie et al., 1987). When fused to the DNA binding domain of GAL4 it will stimulate both a reporter containing GAL4 binding sites, and one lacking them but containing instead the E4 promoter. Deletions of CR3 as GAL4 fusions were analysed with these reporters in mammalian cells (Lillie and Green, 1989). A deletion of the N-terminus of CR3 to residue 150 abolished activation on both reporters, but activation could be restored by fusing the deletion to the activation domain of the herpes simplex virus VP16 protein. A C-terminal deletion of CR3 beyond residue 179 abolished activation on the E4 reporter even as a VP16 fusion, but could still activate the GAL4 binding sites. A point mutation of residue 180 had the same specificity of activation as this deletion. It was speculated that the N-terminal region of CR3 may be interacting with the general transcriptional machinery. Some acidic residues are present in this domain and it could be functioning like an acidic activation domain such as that of VP16. The C-terminal region of CR3 may target E1a to the promoter, since it was not required when the promoter contained GAL4 sites. CR3 also contains a zinc finger (Culp et al., 1988), the integrity of which was not tested in these experiments. Subsequent experiments have indicated its importance and will be discussed later.

E1a is not thought to bind DNA directly, therefore a target for the C-terminus of CR3 could be a DNA-bound factor, such as CREB2. One experiment to confirm this idea has been previously mentioned, where a GAL4/CREB2 fusion could be activated by E1a.

Additionally, a GAL4/E1a fusion was made where the E1a was transcriptionally silent due to a deletion of its activating region N-terminal of residue 150. This could be activated on a reporter with GAL4 sites by CREB2 fused to the activation domain of VP16 (Liu and Green, 1990). This indicates that there could be a physical association between E1a and CREB2, but does not make clear if such an interaction is direct - other proteins may be required to mediate or stabilise the interaction.

Therefore, these experiments indicate that E1a can function as part of the complex at the promoter via an interaction with a cellular transcription factor bound upstream, and it may contribute an activation domain to interact with the general machinery. This could be analogous to the mechanism of action of the herpes simplex virus (HSV) activating protein, VP16 (or Vmw65), which also contains two distinct functional domains. One domain is necessary for the formation of a complex with the cellular transcription factor Oct1, other host cell factors and DNA (Kristie and Sharpe, 1990). Like E1a, VP16 does not have a high specificity of interaction with DNA by itself. The other domain is a highly acidic region which comprises a potent activation domain that has been shown to interact the with general factors TBP and TFIIB. However, VP16 will only activate through the distinct cis-acting elements to which Oct1 binds, and does not have the ability to activate such a diverse set of promoters as E1a does.

1.3.3 Other factors involved in 13S-mediated transactivation.

AP1 - The cellular AP1 activity binds to the site TGA(C/G)TCA that is present in the adenovirus E3 promoter. Under certain conditions, these sites can mediate E1a inducibility, which can be synergistic with activation by protein kinases C or A. An AP1 site is also one element which can also mediate repression by E1a, but this function is not dependent upon CR3, and the mechanism by which it occurs remains obscure. (Rochette-Elgy et al., 1990) showed a general repressive activity of E1a upon different enhancer elements.

TFIID - mutational analysis of the E1B and hsp70 promoters showed that in some cases E1a could activate through the TATA element alone. Suprisingly, activation through the hsp70 promoter TATA element was abolished by mutations in the N-terminus of E1a, thus defining a new region involved in activation separate from CR2 or 3 (Kraus et al., 1992). However, not all TATA elements are inducible by E1a (Pei and Berk, 1989). E1a may not recognise all TFIID complexes. With the cloning of the TATA-binding protein, the possibility of it being a direct target for E1a can be investigated.

TFIIIC - E1a can stimulate polymerase III transcription, such as of the adenovirus VA RNA gene, and *in vitro* this stimulation is mediated by a TFIIIC-containing fraction. In this case, TFIIIC appears to be activated by phosphorylation (Hoeffler et al., 1988). Activation by E1a via phosphorylation of a promoter-bound factor has also been shown to be of some importance for the E2F and E4F factors (Raychaudhuri et al., 1989). E4F binds to sites on the E4 promoter. Both of these factors show an increase in

binding activity in association with E1a-mediated phosphorylation. The importance of CR3 in E1a-mediated phosphorylation has not been established, nor whether the effect is directly or indirectly due to E1a.

Additional elements that have been shown to mediate E1a inducibility include the CCAAT, USF, Sp1 and NFkB binding sites.

1.3.4 Role of E1a in E2F-mediated transcription

The E2 gene encodes viral DNA replication proteins. Its promoter contains two copies of the binding site for the cellular transcription factor E2F, which have an inverted arrangement that forms a dyad repeat separated by a spacer. Both sites are important for basal and E1a-induced transcription, and can confer E1a inducibility upon a heterologous promoter (Pei and Berk, 1989). Other elements in the E2A promoter, such as the ATF site, may be activated by E1a perhaps via a mechanism as outlined above. A sequence homologous to the E2F binding site is also found in the promoter region of E1a and of several cellular genes that are required for the S phase of the cell cycle, such as dihydrofolate reductase (DHFR), thymidine kinase, DNA polymerase α, cdc2, c-myc and c-myb (see fig1.2). The presence of the E2F site is required for the activity of some of these genes, and at least the DHFR and c-myc genes can be activated by E1a. The E2F sites in these genes are a single overlappping motif, rather than the arrangement of two sites found in the E2A promoter (Hiebert et al., 1991).

As mentioned above, phosphorylation may play a role in E2F binding, and E1a may regulate this. E2F binding activity in HeLa cell extracts can be stimulated by a fraction from adenovirus infected cells in an ATP-dependent manner. This does not occur with a fraction infected by a virus in which the E1a gene has been deleted. Phosphatase can inactivate this E2F binding activity from infected cells, and cAMP-dependent protein kinase can restore this binding activity to the phosphatase-treated extracts (Bagchi et al., 1989).

A separate mechanism of activation of the E2A gene in infected cells via the E2F sites involves another viral protein as well as E1a, a 19kD product of the E4 gene. E4 encodes a complex set of differentially spliced mRNAs with 7 open reading frames (ORFs). A fusion by mRNA splicing between ORF 6 and 7 encodes the 19kD protein, which forms a complex with E2F and stimulates its stable, co-operative binding to the two sites on the E2A promoter. E2F can bind to the E2A promoter in the absence of the E4 complex, but only weakly. The complex between E4 and E2F can also form on one E2F binding site, or in solution, as shown by the fact that antibodies to E4 can deplete an extract of E2F binding activity (Huang and Hearing, 1989). The mechanism of how E4 participates in the promoter-bound complex is unclear. Its role may be solely to induce stable binding of E2F, or additionally it may contribute directly to activation. The stoichiometry of the complex is also unclear. E4 may be bridging two molecules of E2F, or it may itself form a dimer, each monomer interacting with a molecule of E2F. The arrangement of E2F sites on the E2A promoter appears to be unique, and the stimulation

Conserved E2F binding site motifs

E2	TTTCGCG-(18bp)-CGCGAAA
E1A	TTTCGCGGGAAA
DHFR	TTTCGCGCCAAA
c-myc	CTTGGCGGGAAA

Fig 1.2

E2F binding sites in the adenovirus type 5 E2 and E1A promoters, and in the human DHFR and c-myc promoters.

of cooperative binding by E4 does not appear to occur on E2F sites lacking this arrangement. Since at present E4 has not been found to interact with any other cellular transcription factor, it would seem that E4 is a specific activator for the E2A gene.

E1a has an additional role in the formation of the E2F-E4 complex apart from activating the transcription of the E4 gene. In many cells, E4 is prevented from interacting with E2F since E2F is complexed with several cellular proteins. E1a is able to dissociate these complexes and thus allow E4 to bind to E2F and activate the E2A promoter (Bagchi et al., 1990). This activity of E1a is dependent on CR1 and 2 (Raychaudhuri et al., 1991). These conserved domains have already been demonstrated to bind to several known cellular proteins, such as the Rb and p107 proteins. These proteins were found to be part of the complexes seen with E2F, explaining their dissociation by E1a (eg, Chellappan et al, 1991). In agreement with this, Rb was shown to be complexed with cellular proteins that had DNA binding specificity for the E2F site (Chittenden et al., 1991). It is thought that Rb inhibits the transcriptional activity of E2F when bound to it, so the release of free E2F by E1a could promote activation even independently of E4. S phase specific genes with E2F sites in their promoters, such as the ones mentioned above, could be a target for activation by E1a. Rb could be repressing these genes by interacting with E2F. Rb function in the cell cycle is still unclear. Only the underphosphorylated form appears to bind E2F, which is the state Rb is found in during G1. In binding to E2F, Rb may repress expression of the S phase genes in G1. Since Rb has been found in the DNA binding complexes with E2F, this repression may be active, rather than simply a sequestering of E2F away from the promoter. With the phosphorylation of Rb during late G1, E2F should be free to activate S phase specific genes. Therefore the interaction of E1a with Rb may have a similar effect as the phosphorylation of Rb that naturally occurs in the cell cyle.

The situation in the cell cycle is further complicated by the appearance of complexes of E2F with p107. The function of E1a in dissociating the interaction between E2F and p107 is less clear than with Rb, since the activity of the E2F-p107 complex itself is unclear. The E2F-p107 complex is associated with two different cyclins depending upon the stage of the cell cycle. An association with cyclin E is seen during during the G1 phase of the cell cycle. Towards S phase, this complex with cyclin E disappears, and a complex with cyclin A accumulates. cdk2 kinase is associated with cyclins A or E as part of these complexes at the appropriate stage of the cell cycle. It would appear that an active kinase complex is forming at the promoter via E2F, but its target remains under speculation. It may be phosphorylating components of the complex, including E2F and may function to regulate E2F activity at different times of the cell cycle. (Devoto et al., 1992; Lees et al., 1992).

At the time I began these experiments, E2F was defined solely by its binding activity; a cDNA clone for E2F should help to elucidate the importance and regulation of these complexes, and the function of E1a. The purification of E2F has yielded a protein of

apparent molecular weight of 54kd that can bind DNA, but more than one polypeptide may comprise this E2F binding activity (Yee et al., 1989).

1.4 Studying protein-protein interactions in yeast.

The modular nature of transcription factors enables binding and activation domains to function independently and when fused to heterologous proteins. Several transcriptional activating complexes contain binding and activation domains on different polypeptides, where the activation domain is recruited to the promoter solely by protein-protein interaction with the DNA binding factor, and does not bind DNA alone. For example, the HAP4 activator is recruited by HAP2 and 3; and the Gal80 repressor is recruited by GAL4. In other complexes, the activator may make contact with both protein and DNA, such as the viral VP16 activator.

Since an activation domain can function when brought indirectly to a promoter, an assay can be used in yeast or mammalian cells to look at protein-protein interactions. In yeast, if the repressor Gal80 is fused to the VP16 activation domain, it is converted into an activator due to its interaction with GAL4 on the promoter (Ma and Ptashne, 1988). Fields and Song (1989) showed that any two interacting proteins could be assayed in this way if one is fused to a DNA binding domain and the other to an activation domain that works in yeast. They also showed how a yeast genetic screen could be used to identify new interacting proteins by using a yeast expression library of cDNA tagged with an activation domain. The target protein is fused to a DNA binding domain and must have low intrinsic activation ability when expressed in yeast with a suitable reporter. Clones of interest are identified from the library by their ability to activate the reporter via interaction with the target. (Chien et al., 1991). Using the lacZ reporter, yeast cells expressing fusion proteins that activate transcription will have a blue phenotype on the X-gal substrate. The plasmids in these cells can subsequently be recovered and characterised, and the cDNA generating the phenotype is thus immediately available. Obviously such a library could also be used to identify clones that interacted directly with the promoter region of the reporter, rather than activating via a protein-protein interaction (see fig1.3). The libraries used in the screens described here are not expressed constitutively, but are expressed under the inducible promoter from the GAL1-10 gene. This promoter is repressed when cells are grown in glucose medium and activated in galactose medium. Expression of the cDNA fusions is induced by changing the growth medium from glucose to galactose once the colonies have reached a certain size after transformation. This is useful because any clones that are toxic in yeast may still be expressed if they are succesfully transformed into the cells.

Dalton and Triesman (1992) succesfully used such a screening stategy to isolate SAP1, a protein recruited by SRF to the SRE to form a ternary complex. SRF bound to the SRE was used as the target, and was transcriptionally silent in yeast. The isolation of SAP1 corresponded to outcome iii) of fig1.3. In addition, SRF itself

Fig1.3 Strategy for protein interaction screening of cDNA libraries in yeast.

A) A yeast expression library of cDNA tagged with an activation domain that will function in yeast (eg VP16).

B) The target protein for the screen fused to a DNA binding domain must have only a low intrinsic activation ability on the appropriate lacz reporter in the absence of the library. This activation ability is measured by a lacz colorimetric assay and should generate a white phenotype. The cDNA library is introduced into a yeast strain containing both this target protein and the reporter.

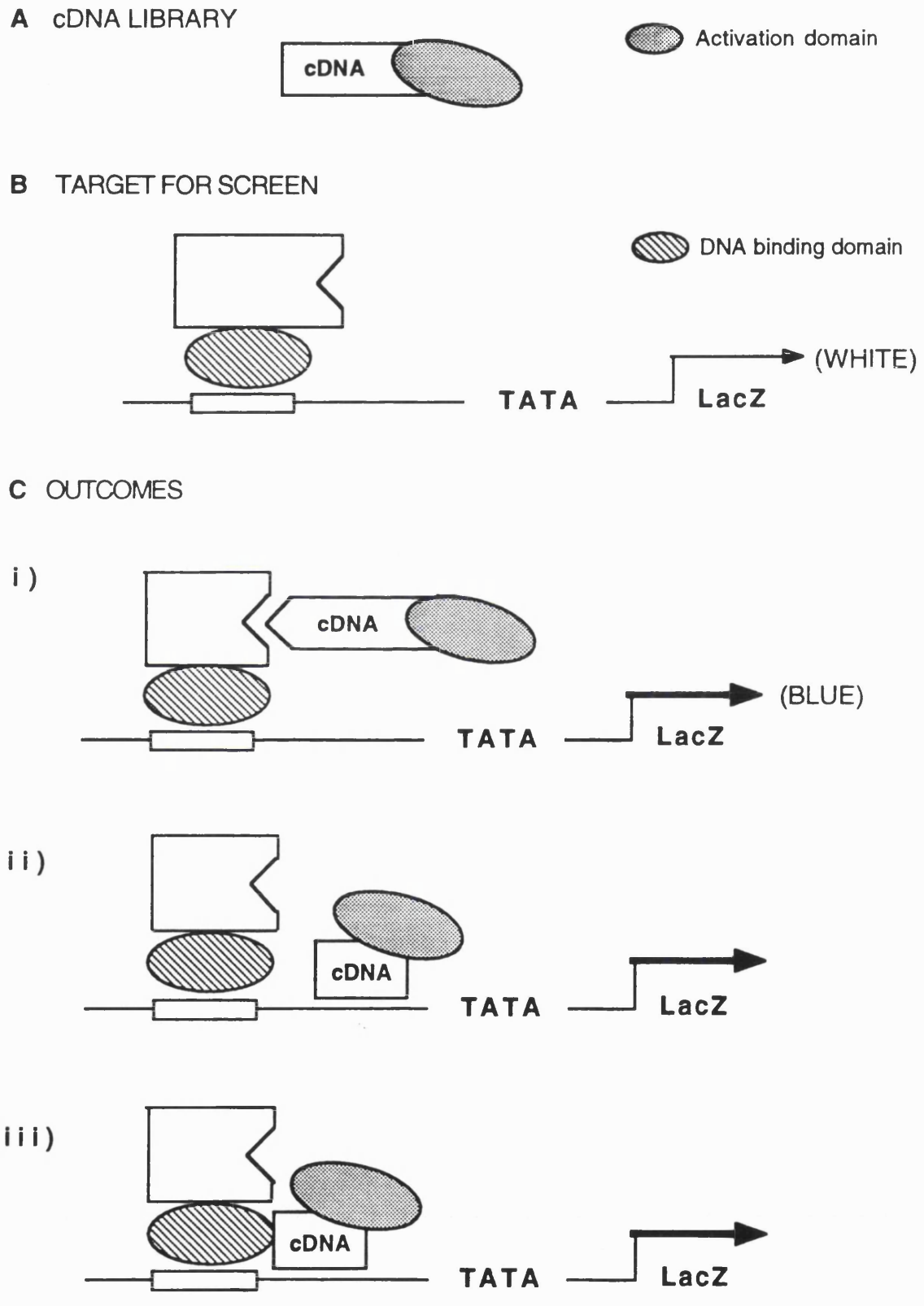
C) Possible outcomes that would generate a positive signal (blue phenotype) upon expression of both the target protein and the library:

the cDNA product **i)** interacts with the target protein

ii) binds directly to the promoter region of the reporter

iii) interacts with the DNA binding domain that is fused to the target protein

Fig1.3 Strategy for protein interaction screening of cDNA libraries in yeast



was isolated, corresponding to outcome ii). Human and yeast tagged libraries were used in experiments described in this thesis. Two different types of screen were carried out, as defined by fig1.3. One type of screen was for a DNA binding protein in which it was hoped to isolate a cDNA that would generate outcome ii). The other strategy was to isolate a cDNA whose product interacted with a target protein, which would correspond to outcome i).

The colony colour assay in yeast using the b-galactosidase reporter enables large numbers of cDNAs to be screened and positive colonies (blue) to be picked out easily above a large background of negative ones (white). It follows that once a protein-protein or protein-DNA interaction can be visualised in yeast via a colorimetric assay, then the use of yeast genetics becomes a powerful tool to analyse random mutations. The mutants that affect activity can be picked out among those that do not, enabling a functional assay in the initial screen, rather than having to laboriously check clones with no mutation or one which does not affect activity.

Experiments in this thesis exploit manipulations such as these which are possible in yeast. Chapter 2 describes experiments using a protein-protein interaction system in yeast to clarify the action of E1a at the promoter and its possible interaction with yeast or mammalian transcription factors. In chapter 3, a yeast library tagged with an activation domain was used to attempt to isolate a yeast cDNA with binding specificity for the ATF site. Finally, a mammalian library tagged with an activation domain was used in attempts to isolate the E2F transcription factor either via its ability to bind DNA or via a protein-protein interaction with the adenovirus E4 protein.

CHAPTER 2: RESULTS

A YEAST SYSTEM TO STUDY E1A TRANSACTIVATION

The aim of these studies has been to explore the possibility that E1a can modulate transcription in yeast - either endogenous transcription, or that mediated by mammalian factors introduced into yeast. An underlying rationale is the high degree of conservation of the transcription mechanism between yeast and mammalian cells, and the fact that some viral transactivators such as the HSV VP16 and human papilloma virus E2 proteins can stimulate the yeast basic transcription machinery (Cousens et al., 1989; Lambert et al., 1989). If E1a transactivation could be shown in yeast, then a more detailed analysis of the mechanism involved would be possible by using the yeast genetic system.

2.1 The effect of E1a on yeast transcription factors.

INTRODUCTION.

I first sought to determine whether E1a might be able to stimulate transcription in yeast via endogenous yeast factors, because some of the cis-acting elements that can mediate E1a inducibility in mammalian cells may have conserved functions in yeast. Artificial promoters containing either ATF, AP1 or E2F sites and a TATA-box can mediate E1a inducibility in mammalian cells (Pei and Berk, 1989). For example, the ATF sites present in the E2A, E3 and E4 adenovirus promoters are important for mediating the E1a induction of these promoters, and the ATF site can also function in this way on a heterologous promoter.Similarly, the AP1 site on the adenovirus E3 promoter, and the E2F sites on the E2A promoter also mediate E1a activation and can function on heterologous promoters. (Reviewed in Nevins, 1989). Promoters containing ATF or AP1 sites have been found to be active in *S. cerevisiae*. Factors that bind to these sites have been identified and shown to have identical binding specificities to their mammalian counterparts. In mammalian cells, the ATF/CREB binding site can be activated by both E1a and cAMP. This is explained in part by the finding that there is a large family of proteins that bind to this site, with individual members of the family having different regulatory properties. For example, CREB2 can be activated by E1a, while CREB1 is only weakly activated. Conversely, CREB1, but not CREB2, can be activated by the cAMP pathway in response to protein kinase A phosphorylation. (Flint and Jones, 1991). In yeast, yATF appears to be phosphorylated, which may contribute to its efficient binding to DNA. Phosphatase treatment decreases the binding efficiency of yATF to its site. Interestingly, the binding activity of the phosphatase-treated yATF can be restored *in vitro* by the catalytic subunit of kinase A (Jones and Jones, 1989). This suggests that yATF is phosphorylated by yeast kinase A, although an *in vivo* role of such phosphorylation has not yet been shown. Nevertheless, this lead to the possibility that in addition to cAMP, E1a might also affect yATF activity. At the start of this study, the activity of the E2F site had only been tested in *S. pombe*. Here, the E2A promoter is active, and this activity is affected by the same mutations in the E2F sites that decrease

activity in mammalian cells (R. Jones, unpublished). I examined the activity of the E2F sites in *S. cerevisiae*, together with the possibility that they could mediate activation by E1a.

Therefore in this chapter, I describe the activity in *S. cerevisiae* of reporter genes containing promoters having ATF, AP1 or E2F binding sites upstream of a TATA box. Coexpression of an E1a-containing plasmid with these reporters allowed their sensitivity to this transactivator to be examined.

RESULTS

The relative activity of several reporter constructs was measured in *S. cerevisiae*. The reporters consisted of the CYC1 promoter linked to β-galactosidase, in which the upstream activating sequences of CYC1 were replaced with various transcription factor binding sites. They were on multicopy (2μ-based) plasmids which were transformed into the yeast strain Y700. Yeast transformants on selective medium were assayed for lacz activity. No exogenous genes encoding binding proteins were present and therefore the activity of the reporters was entirely due to endogenous yeast factors binding to their promoter regions. All the promoters were active (Fig 2.1). Similar lacz activity of the reporters was given by either three copies of the ATF or AP1 sites, or by the E2F sites as arranged on the E2A promoter. In the same manner as in mammalian cells, three copies of the ATF or AP1 sites were more active than one copy. This was the first time that the E2F sites had been tested in *S. cerevisiae*. The specificity of their activity will be discussed further in Chapter 4.

To test whether E1a could modulate the activity of these reporter constructs, the plasmid pTWE1a was used (Abe et al., 1988). This allowed expression of the 13S mRNA-encoded protein of E1a under control of the yeast GAL7 gene promoter (Fig 2.2A). pTWE1a could be cotransformed with the reporter constructs into yeast since it had a different selectable marker from the reporters. The reporters contained the URA3 selectable marker, while pTWE1a had the HIS3 marker, therefore, growing cells in the absence of uracil and histidine allowed selection for both plasmids. The GAL7 promoter allows tight control of E1a expression by changing the carbon source of the growth medium. In glucose medium the promoter is repressed, and in galactose it is induced. To check that the synthesis of E1a was indeed being induced by this promoter, and to test the effectiveness of the carbon source regulation, western blots of yeast extracts prepared from these cells were performed. During yeast growth in glucose, E1a protein was not detectable by a Western blot. Expression of the E1a plasmid was induced by growing the yeast in glucose-free, galactose medium for 8 hours. E1a protein could be detected in extracts made from these cells (See Fig 2.2B). The antibody used for the Western was a monoclonal antibody specific for the 13S protein (Harlow et al., 1985).

The mobility of the E1a protein on an SDS-polyacrylamide gel was compared between E1a expressed in yeast and that produced from insect cells using the

ATF, AP1 and E2F binding sites can each act as a UAS in yeast.

A

CONSTRUCT	UAS	β- GALACTOSIDASE ACTIVITY IN Y700
pB3E4	3xATF	710
pB1E4	1xATF	160
pB3AP1	3xAP1	900
pB1AP1	1xAP1	180
p4-1	E2F	800

B

ATF	TACGTCAT
AP1	TTAGTCAG
E2F	TTTCGCG^{18bp}-------CGCGAAA

Fig2.1

A) The yeast strain Y700 was transformed with each reporter and quantitative lacz assays were performed.

B) Core sequences of the binding sites used upstream of the lacz gene. Reporter plasmids have the URA3 marker, and are 2 μbased.

Fig 2.2 Expression of E1a in yeast

A) pTWE1a construct (Abe et al, 1988). The full length coding sequence of E1a from residues 1 to 289 is flanked by the GAL7 promoter and termination sequences. The plasmid pTW lacks E1a coding sequences.

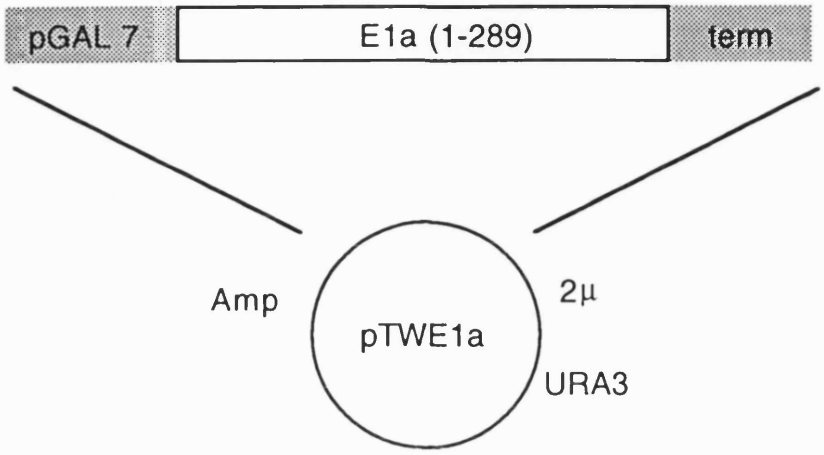
B) Western blot showing induction of expression of E1a protein in yeast upon change of growth medium from glucose to galactose. Y700 containing either pTW or pTWE1a was grown overnight in either glucose or galactose. The medium lacked histidine to select for the plasmid. Extracts were made from the cells and proteins resolved on a 10% SDS-polyacrylamide gel. After transfer to a nitrocellulose filter, proteins were visualised with the M73 anti-E1a antibody (Harlow et al, 1985).

Lane 1: 5μl insect cell lysate from cells overproducing E1a

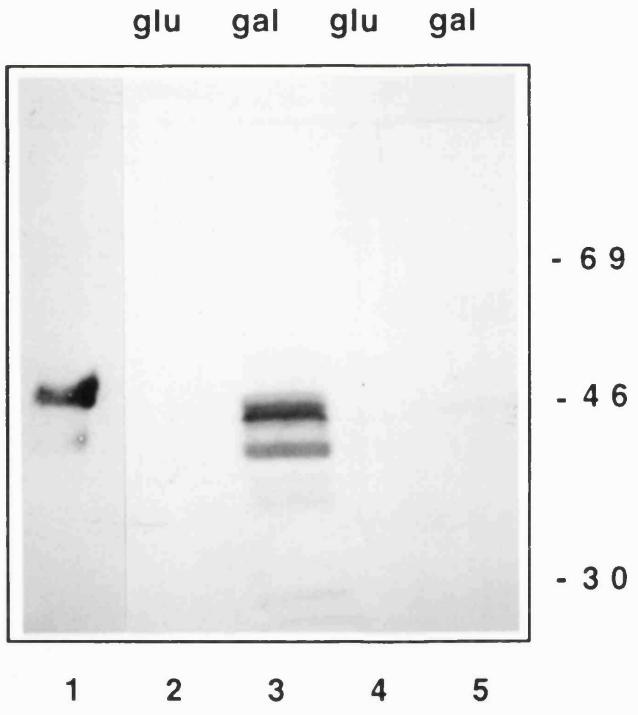
Lanes 2 and 3: Extract from Y700 containing pTWE1a.

Lanes 4 and 5: Extract from Y700 containing the pTW vector lacking E1a coding sequences.

A


pGAL 7
E1a (1-289)
term
Amp
pTWE1a
2μ
URA3

B


glu
gal
glu
gal
- 69
- 46
- 30
1
2
3
4
5

baculovirus system (provided by G. Patel). Multiple phosphorylation of the E1a protein is known to occur in mammalian cells giving rise to two predominant forms on a one-dimensional denaturing gel, and multiple forms on a 2D-gel. Insect cell-produced E1a protein is known to migrate with identical mobility to that produced in mammalian cells (Patel and Jones, 1990), although only one form predominated in Fig2.2; the faster migrating form can only be seen faintly. The relative ratio of the two forms differed between extracts. E1a expressed in yeast also migrated as two predominant forms, but both forms had slightly greater mobility than the equivalent insect cell-produced proteins. This discrepancy could result from a slightly different pattern of post-translational modifications such as phosphorylation. In yeast, it has been shown by 2-D gel analysis that all the E1a proteins produced are phosphorylated (Handa et al., 1987), but the pattern of phosphorylation may not be identical. However, it should be stressed that there is no evidence that any post-translational modifications, including phosphorylation, are thought to be critical for E1a function. For example, E1a expressed in yeast or bacteria is able to bind to proteins such as Rb and TBP (see later). Additionally, mutation of the known major phosphorylation sites does not significantly affect the various known E1a functions (eg Richter et al, 1988; Tremblay et al., 1989).

Having confirmed the expression of pTWE1a, it was cotransformed with each reporter plasmid into the yeast strain Y700. β-galactosidase activities were measured in non-induced and induced conditions for E1a expression and with and without the E1a plasmid; E1a did not alter the level of LacZ activity of any of the reporter plasmids. Their activity remained the same as in Fig2.1. Expression of E1a in this yeast system did not seem to cause any other detectable phenotype; there was no apparent change in the transformants' growth rate or size.

There are a number of possible reasons why no transactivation by E1a was seen. Two of the most prominent ones will be addressed in the next sections:

i) If E1a functions by a direct interaction with factors on the promoter, then it may not be able to functionally interact with the general transcription machinery found in yeast. The activity of the E1a activation domain was therefore tested in yeast.

ii) Unlike their mammalian counterparts, the yeast upstream factors may not be able to mediate an E1a effect, for example they may not be able to interact with E1a.

Function of the activation domain of E1a in yeast

CR3 of E1a has been shown to function as an activation domain in mammalian cells when fused to a heterologous DNA binding domain such as GAL4. These experiments used a reporter with GAL4 binding sites. Deletion of the N-terminal region of CR3 abolishes the activation (Lillie and Green, 1989). There are some acidic residues in this region of CR3, which could potentially function as a typical acidic activation domain. As discussed earlier, acidic activation domains are thought to be able to function in yeast, thus it was thought possibile that the activation domain of E1a could also do so. As a negative control for any E1a activation seen, we had available a mutation in CR3 which

dramatically decreases activation in mammalian cells. The mutation consisted of a proline to leucine residue change at position 150 (P150L).

To see if the activation domain of E1a can function in yeast, it was tethered to a yeast promoter via a fusion with a heterologous DNA binding domain. In order to directly compare results to the study in mammalian cells described above, that of the yeast GAL4 protein was used. The activities of two GAL4/E1a fusions were measured. One consisted of wild type E1a sequences including CR2 and 3, the other had the mutation of P150L in CR3, which has a dramatically decreased effect on activation in mammalian cells. The plasmid pMA424 expresses the DNA binding domain of GAL4 under a constitutive promoter, and was used to construct the E1a fusions. They were transformed into the yeast strain YT6::171 which contains an integrated reporter consisting of the GAL4 binding site upstream of the Lac z gene. In this yeast strain, the genes GAL4 and GAL80, whose products normally bind to the GAL4 site, have been deleted, ensuring that the activity directed by this site was due only to the GAL4/E1a fusion. YT6::171 transformants were assayed for lacz activity. The expression of wild type GAL4/E1a gave only a low basal level of Lacz activity. This was in comparison with a mammalian activator that works strongly in yeast (CBF), expressed as a fusion in the same way. The mutation in CR3 (in plasmid GAL4/P150L) had the same low Lacz activity as the wild type E1a fusion. (Fig 2.3A)

Therefore, the results show that the activation domain of E1a can only interact with the yeast general transcriptional machinery very weakly. The acidic residues present in CR3 could be giving this low activation activity, but clearly this level would not be sufficient to see activation of the ATF, AP1 or E2F reporter genes tested.

Since they had such low activity, it was necessary to check that these GAL4/E1a fusions were expressed correctly in yeast. I took advantage of the observation that E1a binds strongly to several proteins in mammalian cells, of which the best characterised is the 105kD product of the retinoblastoma gene (Rb). The interaction between E1a and Rb is known to be direct, implying that if Rb and E1a are co-expressed in yeast, they should be able to interact. The GAL4/E1a fusions should contain sequences sufficient for the interaction of Rb. They contain all of CR2, and although CR1 is missing, it has been shown that a peptide of CR2 alone is sufficient for Rb binding, although optimal binding requires CR1 (Dyson and Harlow, 1992). Interaction between the GAL4/E1a fusion and Rb should result in the recruitment of Rb to the promoter. In order to measure such recruitment, the Rb was tagged with an activation domain that can work in yeast. The acidic activation domain of (HSV)VP16 was used, and a construct made consisting of a VP16 fusion of Rb driven by the CUP1 promoter. The CUP1 promoter is inducible by copper, which was added to the medium in the form of CuSO4. The two GAL4/E1a plasmids were co-transformed with pVP16/Rb and the activity of the reporter in YT6::171 measured. The expected activation by VP16/Rb above the GAL4/E1a background did occur, confirming that the GAL4 fusions were correctly expressed and that they could bind DNA (Fig 2.3B). Therefore, incorrect expression of the GAL4

Fig 2.3 Activity of the E1a activation domain in yeast

Constructs were transformed into the yeast strain YT6::171, which contains an integrated lacz reporter driven by the GAL4 binding site. Quantitative lacz assays were performed.

A) GAL4(1-147) alone, or as a fusion to either E1a, E1aP150L or CBF was transformed into yeast under selection for His3.

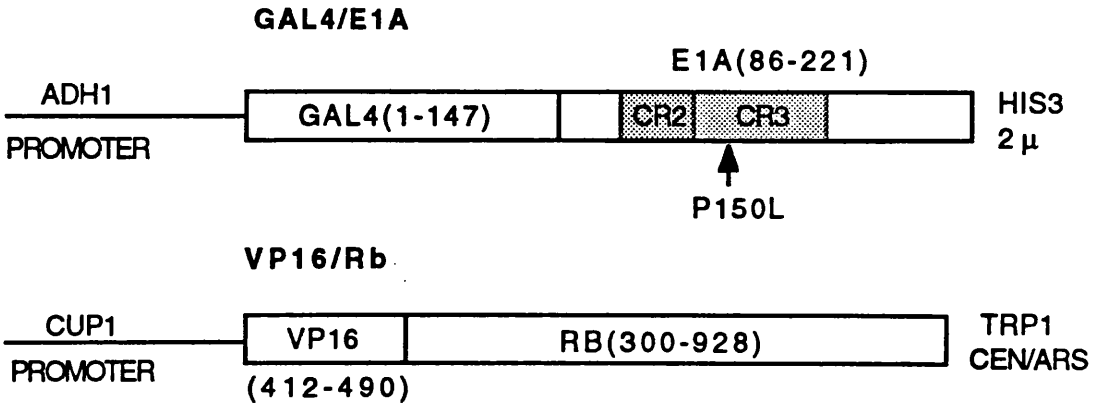
B) The GAL4/E1a fusions were cotransformed with either VP16 alone or VP16/Rb under His3 and Trp1 selection. The plasmids under control of the CUP1 promoter were induced with 1mM CuSO4 in the medium.

C) Diagram of GAL4/E1a and VP16/Rb constructs.The conserved regions of E1a are indicated.

Fig 2.3 Activity of the E1a activation domain in yeast

	CONSRUCTS	β- GALACTOSIDASE ACTIVITY IN YT6::171
A	GAL4	<1
	GAL4/E1a	21
	GAL4/E1aP150L	25
	GAL4/CBF	284
B	GAL4/E1a + VP16	21
	GAL4/E1a + VP16/Rb	143
	GAL4/E1aP150L + VP16/Rb	100

C



fusions was not an explanation for the low activity of the E1a activation domain seen in yeast.

These results therefore suggested a possible explanation for the inability of E1a to activate any of the previously described reporters in yeast. The activation domain of E1a is unable to interact with the yeast basic machinery, and therefore even if E1a could be targeted to specific promoters in yeast, activation was unlikely to ensue.

To circumvent this problem, E1a was tagged with an activation domain that would work in yeast. The VP16 activation domain was again chosen. This enabled an investigation to be continued into the targetting of E1a to upstream sequences. It would overcome the problem of the low activity of the E1a activation domain in yeast, but still allow the effect of E1a on the reporters to be tested. However, the use of this E1a fusion no longer addressed the interaction of E1a with the yeast basic machinery. Also, now the assay depended upon E1a operating by interacting with factors on the promoter, rather than activating indirectly, such as via some modification mechanism.

A fusion of VP16 to the N-terminus of E1a was made, and its expression in yeast determined by western analysis of cell extracts with an E1a antibody. Its functional activity was also assessed, again by use of the association of E1a with Rb. On this occasion, Rb was fused to a heterologous DNA binding domain to target it to the promoter of the appropriate reporter. The binding domain used was that of the bacterial LexA protein, which binds efficiently to its LexA operator (LexAop) in yeast. Since there are no endogenous proteins that can bind to this site, it is the only protein to do so. LexA/Rb was co-expressed with the VP16/E1a plasmid in the yeast strain Y35 which contains an integrated LacZ reporter under the control of the LexAop. Both the Rb and E1a constructs were inducible by galactose. Neither the LexAop alone nor together with the LexA/Rb fusion gave any activity in yeast. However, activation was seen in the presence of VP16/E1a. This activation was E1a-dependent since no activation was seen with VP16 alone (Fig2.4).

This experiment showed that VP16/E1a could activate the yeast general transcription machinery when tethered to a yeast promoter via a protein - protein interaction with Rb. To see if such an interaction could occur with yeast transcription factors, VP16/E1a was expressed in Y700 containing the ATF-, AP1-, and E2F-Lacz reporters. The basal level of all the reporters remained unaltered (ie identical to the activities in Fig2.1). This implied that no detectable interaction was occuring between E1a and the yeast transcription factors.

One problem with using such reporters could be that they are already maximally induced in yeast, so the expression of E1a cannot increase activation further. However, there is a more likely explanation of the negative result obtained. If E1a transactivates in mammalian cells by directly binding to transcription factors such as CREB2, then the yeast factors binding to the reporters may have a divergent structure from such mammalian factors, and be unable to interact with E1a. In mammalian cells there are

The VP16/E1a construct is expressed in yeast

A

CONSRUCTS	β-GALACTOSIDASE ACTIVITY IN Y35
LexA/Rb + VP16	<1
LexA/Rb + VP16/E1a	20

B

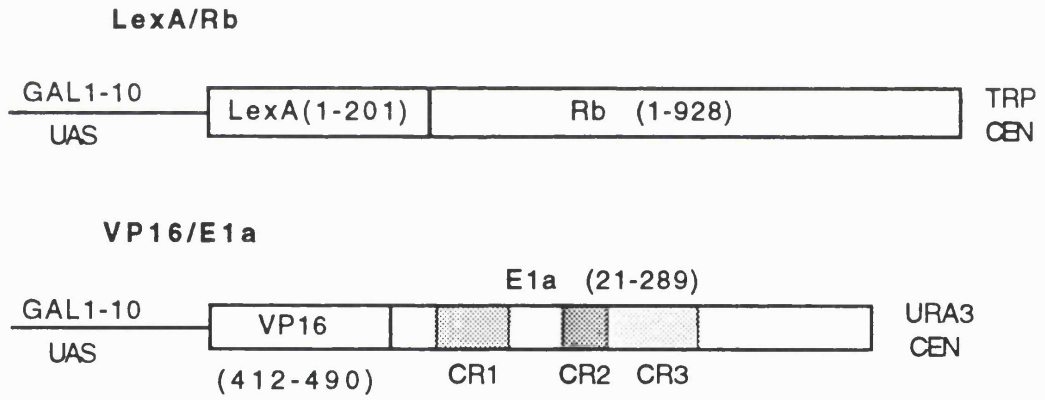


Fig 2.4

A) The LexA/Rb construct was cotransformed with VP16 or VP16 fused to E1a, into the yeast strain Y35 which contains an integrated lacz reporter driven by the LexA operator. Quantitative lacz assays were performed.

B) Diagram of LexA/Rb and VP16/E1a constructs. The conserved regions of E1a are indicated.

multiple ATF/CREB proteins, only a subset of which respond to E1a. Until a cDNA for yATF is obtained and its structure compared to the mammalian proteins, it is not possible to predict what features, if any, of the regulation of the ATF/CREB site in mammalian cells, are conserved in yeast. The next step was to see if E1a is interacting directly with a mammalian factor sensitive to its action. The most direct way to test this possibility was to introduce mammalian factors into yeast.

2.2 Expression of mammalian CREB in yeast

INTRODUCTION

A problem arises from introducing mammalian factors into yeast, since yeast contains an endogenous background of such factors, which have similar DNA binding specificities to the mammalian ones. An ideal approach to look at E1a action would be to introduce mammalian factors into mutant yeast strains lacking these factors. However, in most cases such strains are not available.

An alternative is to fuse the mammalian factor to a heterologous DNA binding domain. It should be feasible to use this system in yeast to study E1a activation, since work in mammalian cells using CREB 1 and 2 has shown that both can be activated by E1a as GAL4 fusions. CREB2 is activated to a much greater extent (Flint and Jones, 1991).

The yeast system should demonstrate if such E1a transactivation occurs by a mechanism of direct interaction with a mammalian promoter-bound factor. If E1a does transactivate in yeast via such a mammalian factor, then a model of direct interaction is most likely. Alternative models could be that a bridging protein between E1a and the mammalian factor may be conserved in yeast, or, less likely, E1a may be acting through an enzymatic pathway conserved in yeast, though not able to operate via the yeast transcription factors.

In yeast, CREB2 bound to the promoter via a heterologous DNA binding domain should recruit E1a to the promoter if a direct interaction occurs between them. If E1a is fused to VP16, then activation of the yeast general transcription machinery will result. This activation would only be seen if the basal activity of CREB2 in yeast is low. The assay should also work if the heterologous fusions of E1a and CREB2 were switched. In this way, E1a would be expressed as a GAL4 fusion (as tested previously), and CREB2 as a VP16 fusion.

A major impetus for showing the interaction between E1a and CREB2 using the yeast system was that if such an interaction could be demonstrated, then the regions of the two proteins required for the interaction could be defined using random mutagenesis. Mutants could be readily selected in yeast. This would complement studies in mammalian cells using mutated CREB2 proteins that have already indicated the importance of the N-terminal region of CREB2 in its E1a inducibility. Also, if mutants of one protein suppressed a corresponding mutation in the other, this would be strong genetic evidence

for a direct interaction between the two proteins and negate the presence of any yeast bridging protein. Such sets of mutant proteins could also be tested for their ability to interact in mammalian cells.

RESULTS

Full length CREB1 and 2 cDNAs were fused to a heterologous DNA binding domain. That of SRF was used; it can efficiently bind to the SRE in yeast, without causing activation. In addition, the particular SRE used had been selected to be transcriptionally silent in yeast, since it does not bind the yeast SRF-like protein, MCM1, yet is still able to bind SRF efficiently (Dalton and Treisman, 1992). The activity of the CREB/SRF fusions was assayed in the yeast strain S62L which contains an integrated lacz reporter under the control of the SRE. No detectable Lacz activity was observed (Fig2.5A), indicating that the activation domains of CREB1 and 2 do not function in yeast. This might be expected since neither clone shows evidence of having an acidic activation domain.

Since no activity was detected, the correct expression of the CREB fusions had to be checked. This was done by co-transforming each into yeast together with a plasmid containing a fusion between the VP16 activation domain and SAP1. SAP1 has been shown to form a ternary complex with SRF on the SRE with strong avidity. Upon binding to SRF, SAP1 should bring VP16 to the promoter and thereby activate transcription. When the CREB fusions and VP16/SAP1 were co-expressed in S62L, such activation did occur (Fig2.5B). This implied that SRF was correctly expressed and able to bind to DNA. Therefore the CREB fusions must also be expressed in the correct reading frame, since they were fused N-terminally to SRF.

The CREB/SRF fusions were now co-expressed with pVP16/E1a. No activation at all was seen (Fig2.5C). One explanation of this result is that an additional protein is required to mediate any interaction between CREB and E1a.

To further substantiate the result that an interaction cannot be seen between E1a and CREB2 in yeast, other constructs were introduced into yeast, combining different sets of fusions of the two proteins. The construct VLCR2 was made (L.Clark). This comprised of a fusion between the VP16 activation domain, the LexA DNA binding domain and the full length CREB2 protein. The ORF was checked by sequencing and also functionally in yeast in two ways:

i) VLCR2 was transformed into the yeast strain Y35 which contains an integrated LexA operator. Upon inducing the expression of VLCR2 with galactose, induction of the lacz reporter was also seen, generating a blue phenotype in the colorimetric assay (Fig2.6A). This resulted from the interaction of the LexA DNA binding domain with its site, thus bringing VP16 to the promoter to activate transcription. Therefore both the LexA and VP16 parts of the fusion protein must be functional.

ii) The ability of the CREB2 part of the fusion to bind to an ATF site was also tested. VLCR2 was transformed into a yeast strain containing an integrated Lacz reporter driven by ATF binding sites. Binding of CREB2 to the ATF site should result in the VP16

Fig 2.5 Activity of mammalian CREB in yeast

Constructs were transformed into the yeast strain S62L and quantitative lacz assays performed. CREB denotes both CREB1 and CREB2.

A) SRF alone or fused to either CREB1 or 2 was transformed into yeast under selection for Trp.

B) SRF fused to either CREB1 or 2 was cotransformed with VP16 or VP16/SAP1 under selection for Trp and Ura.

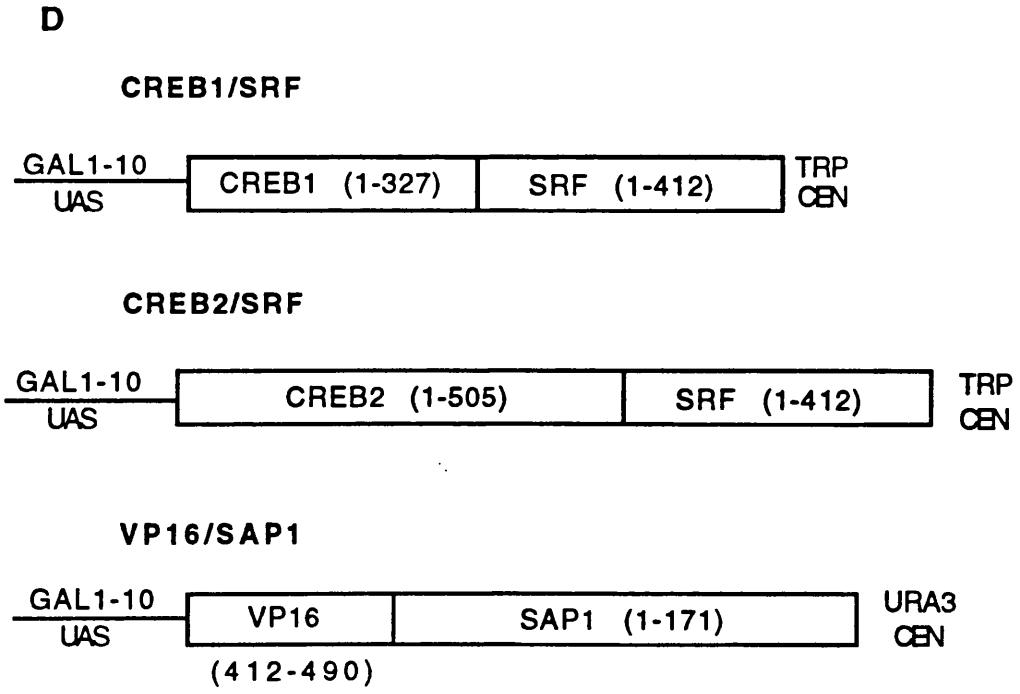
C) SRF fused to either CREB1 or 2 was cotransformed with VP16/E1a

D) Diagram of CREB/SRF and VP16/SAP1 constructs.

Fig2.5 Activity of mammalian CREB in yeast

	CONSRUCTS	β- GALACTOSIDASE ACTIVITY IN S62L
A	SRF	<1
	CREB1/SRF	<1
	CREB2/SRF	<1
B	CREB/SRF + VP16	<1
	CREB/SRF + VP16/SAP1	150
C	CREB/SRF+ VP16/E1a	<1

D



Summary of yeast transformations:
E1a and CREB2 do not interact in yeast.

	CONSRUCTS	REPORTER	β– GALACTOSIDASE ACTIVITY
A	VLCR2	LexAop	BLUE
	VLCR2	ATF	BLUE
	VCR2	ATF	BLUE
B	LCR2 + VP16/E1a	LexAop	WHITE
		ATF	WHITE
C	LexA/E1a + VCR2	LexAop	PALE BLUE
	+ VP16	LexAop	PALE BLUE

D

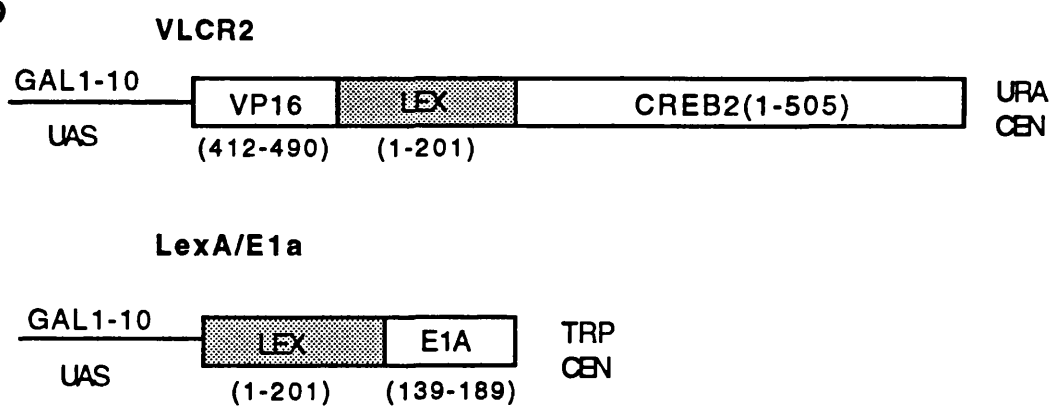


Fig2.6

Constructs were transformed into yeast strains Y35 (LexAop reporter) or MA7310 (ATF reporter) and qualitative lacz assays performed.

A) Construct VLCR2 was checked by transforming it into both Y35 and MA7310. VCR2 was checked by transforming it into MA7310.

B) LexA/CR2 was cotransformed with VP16/E1a into both Y35 and MA7310.

C) LexA/E1a was cotransformed with either VP16/CR2 or VP16 into Y35.

D) Diagram of VLCR2 and LexA/E1a constructs.

part of the fusion being brought to the promoter and activating transcription. Since the ATF site in yeast is already very active, and might mask activation by VLCR2, a mutant yeast strain MA7310 was used. This has a decreased activity of the reporter driven by ATF sites. Activation via another site such as the AP1 site is not affected in this strain, indicating that is not defective in general transcription, but more specifically in transcription via yATF. (For further details of the generation of this strain, see Chapter 3). Expression of VLCR2 in MA7310 led to an activation of the reporter driven by ATF sites, indicating that VLCR2 could bind to an ATF site and activate transcription via VP16 (Fig2.6A).

The VLCR2 construct was now used in two ways: i) by deleting the VP16 portion, it was used as a LexA/CREB2 fusion to see if this could be activated by VP16/E1a via either an ATF site or the LexAop, and ii) by deleting the LexA portion it was used as a VP16/CREB2 fusion to see if this could activate E1a tethered to the promoter via the LexA binding domain.

i) LexA/CREB2 was cotransformed with VP16/E1a into the yeast strains Y35 and MA7310. In Y35 the activity of the integrated LexAop reporter was tested, and in MA7310, the activity of the reporter driven by ATF sites was tested. In neither case was activation seen, indicating that no interaction was occuring between E1a and CREB2, whether CREB2 was bound to DNA via its own DNA binding domain, or via a heterologous one. Thus the yeast cells retained a white phenotype in the lacZ colorimetric assays (Fig2.6B).

ii) The VP16/CREB2 construct could activate the ATF reporter in strain MA7310, confirming its correct expression (Fig2.6A). It was then cotransformed with a LexA/E1a fusion into the yeast strain Y35. The LexA/E1A fusion consisted of CR3 of E1a inserted downstream of the LexA DNA binding domain. This construct could not be checked by cotransforming it with VP16/Rb because it lacks the sequences required for Rb interaction. It was checked by sequencing, and it was also found that its low transcriptional activity in yeast (pale blue phenotype) correlated with the activity expected, since the GAL4/E1a fusions expressed previously also had this phenotype. When LexA/E1a and VP16/CR2 were expressed in combination, no activation of the reporter was seen above that expected from the basal activity of LexA/E1a (Fig2.6C).

All these experiments would indicate that E1a and CREB2 cannot be seen to interact in such a yeast system. This may lead the idea that a mediator protein is required. However, there are other possible explanations why no transactivation by E1a was seen. E1a may transactivate through a biochemical pathway not present in yeast, or the proteins may not be post-translationally modified correctly in yeast. These possibilities will be expanded upon in the Discussion section.

CHAPTER 3: RESULTS

A YEAST SCREEN FOR THE YEAST ATF TRANSCRIPTION FACTOR

Introduction

This chapter will describe further characterisation of the ATF DNA binding activity in yeast. The goal was to isolate a cDNA clone for yATF. A genetic screen was developed in which a yeast cDNA library was introduced into yeast, and colonies isolated with increased transcription of a lacz gene driven by ATF sites. Because of the already high endogenous activity of the ATF site in yeast, a mutant strain was used for the screen with decreased activity of the ATF binding site.

It has been shown in Chapter 2 and previously that reporters containing ATF sites are active in yeast (Fig2.1). Furthermore, this yATF DNA binding activity has the same specificity as mammalian ATF, as determined by gel retardation analysis upon wild type and mutant binding sites. Since the ATF site can act as an efficient upstream activator *in vivo*, this suggests that yATF is a transcriptional activator. (Lin and Green, 1989; Jones and Jones, 1989).

The study of yATF may elucidate conserved mechanisms that would help to understand how mammalian ATF is regulated. The ATF/CREB site plays an important role in the control of gene expression in mammalian cells. A large family of factors can bind to it, and several have been shown to have their individual mode of regulation by different signal transduction pathways. The ATF site may similarly be of central importance in yeast biology. It would be interesting if as well as conservation at the level of DNA binding specificity, regulation of the activity of factors binding to the ATF site in yeast also showed conservation. One mechanism of regulation via the ATF/CREB site in mammalian cells is via phosphorylation by protein kinase A. CREB1 is activated by such a mechanism (Gonzalez and Montminy, 1989). In yeast it has been shown that yATF can also be phosphorylated by protein kinase A (Jones and Jones, 1989). Such phosphorylation has not been proven to be an *in vivo* method of regulation, but implies that the relationship between yATF and the mammalian factors may extend beyond the conservation of DNA binding specificity. The availablity of yeast genetics provides some advantages in studying a yeast factor over the study of a mammalian one. Once its sequence is known, the gene for the yeast factor can be deleted. Unlike the situation in mammalian cells, genetic knockout of a yeast gene is comparatively easy, and yeast can be propagated as haploids. Analysis of such knockout strains is invaluable in characterising the function of the gene in question. In addition, genetic screens involving the suppression of such mutant strains often provides useful information, for example the identity of downstream targets. Finally, the study of a yeast factor may be simpler since it may not be a member of such a large family as is the mammalian CREB/ATF family.

Previous efforts to isolate a cDNA encoding yATF have consisted of more conventional approaches than the alternative strategy that I adopted. Lin and Green (1989) purified yATF as a 66kD polypeptide. During experiments in our lab to purify

and clone yATF, DNA binding activity was lost, although parallel experiments to purify yAP1 were successful. A screen with the ATF binding site on a λgt11 library also did not yield a clone (Jones, 1990b).

Strategy of the yATF screen

The yeast strain Y700b has an integrated reporter consisting of 3 ATF sites upstream of the LacZ gene. A postdoctoral fellow in the laboratory, Dr. S. Kuge, had isolated mutant strains of Y700b that had significantly lower levels of expression of the lacz reporter gene driven by the ATF sites. These mutant strains were crucial for the strategy I adopted in attempting to isolate the yATF gene. I will describe first the method of isolation of these mutant strains.

The ATF sites confer a fairly high transcriptional activity upon the lacz reporter in Y700b. The mutant strains were isolated by screening for spontaneous Y700b- mutants that had lowered LacZ activity. However, a second citerion was additionally used, since relying solely on the Lacz assay would have involved not only assaying a vast number of colonies, but also detecting down mutants (white in the lacz colourimetric assay) amongst large numbers of wild type colonies (blue). It would also have been difficult to assess the specificity of these white colonies. Therefore an initial selection was carried out with a separate reporter, where only down mutants survived because high transcriptional activity induced cell death. High numbers of colonies (10^6) could then be assayed on one 9cm plate. The few survivors could subsequently be assayed for decrease in Lacz production. Therefore, as well as reducing the number of colonies assayed for lacz, using two assays increased the specificity of the clones finally selected.

The initial screen used the CAN1 gene as a reporter. This encodes a permease specific for the uptake of arginine, and its toxic analogue, L-canavanine. Y700b is can1$^-$ and therefore lacks this permease. As a result, it will survive on medium containing canavanine. When Y700b was transformed with a plasmid containing the CAN1 gene under its wild type promoter, as expected, the cells were inviable on canavanine. Replacing the CAN1 promoter with three ATF binding sites meant that cell survival depended on the efficiency of a promoter that was regulated by yATF. Y700b containing 3xATF-CAN1 was inviable on medium containing 10μg/ml canvanine. This concentration of canavanine was picked because it did allow mutants of Y700b with decreased ATF activity to survive. These surviving clones were then assayed for Lacz activity and those with a lowered activity were isolated. They seemed to be specifically mutated in the ATF pathway, since the activity of the AP1 binding site was not affected. The AP1 site differs by only one base pair from the ATF site.

Several down mutants were isolated by this strategy. The exact nature of the mutation in these strains is unknown, as is the number of different complementation groups that these mutants represent. Possible targets for a mutation must be either the ATF transcription factor itself or part of its regulatory pathway. To investigate the mutant strains further, a gel retardation assay was used. The complexes in yeast

extracts seen to bind to an ATF site probe appeared to be the same in both wild type and mutant strains. Therefore the DNA binding activity of yATF appeared to be unaffected in the mutant strains.

Strain MA7310 was chosen for the screen I set up to isolate yATF. It had an activity of the 3xATF-Lacz reporter which was decreased 50-fold over wild type (Fig 3.1A, data from S.Kuge). This strain could have been used for a complementation screen in yeast to isolate the gene that has been mutated. However, since the nature of the mutation in any of the mutant strains was unknown, and since there was no evidence that the mutation was in yATF itself, obtaining the complementing gene may not have led to the yATF gene. An alternative approach was thought more likely to yield a cDNA with DNA binding activity. This involved introducing a yeast library tagged with the VP16 activation domain into the mutant strain and screening for increased Lacz activity via the ATF sites. cDNA clones that interact with the promoter of the lacz reporter will activate transcription via the VP16 activation domain, and yeast cells expressing them will be blue in the colony colour assay. Positive cDNA clones obtained can then be checked for their specificity of interaction with the ATF site.

The wild type Y700b strain could possibly have been used for this screen, since a cDNA product that can bind DNA and is fused to the strong activation domain of VP16, might be expected to have a high transcriptional activity. However, use of the mutant strain enabled a more sensitive library screen than that possible in a wild type strain because of the lower basal activity of the lacz reporter.

Results.

As described above, a reporter containing the AP1 site driving the lacz gene had been introduced into the mutant yeast strain MA7310, and its transcription was activated to the same extent as in a wild type strain. This showed that the mutant strain was still capable of supporting normal levels of transcription, and therefore the mutation did not result in a decrease in transcription in general. As an additional control that the mutant strain was still capable of mediating activation via its ATF binding sites, the mammalian transcription factor CREB2 was introduced into the strain. The activation domain of CREB2 does not stimulate transcription in yeast (see chapter 2), so in order to circumvent this problem, it was fused to the activation domain of VP16. VP16/CR2 was transformed into MA7310 and lacz levels measured. As expected, a strong activation of the Lacz reporter was seen (Fig 3.1B). This clearly demonstrated that the ATF binding site could still mediate transcriptional activation in this mutant strain, and supported the possibility of cloning a yeast factor with ATF DNA binding affinity, using a VP16-tagged library.

Therefore, strain MA7310 was transformed with a randomly primed yeast cDNA library, fused C-terminal to the VP16 activation domain, in a galactose inducible vector (G. Micklem. See fig3.2A). 800,000 transformants were grown on nylon filters placed

Activity of the MA7310 yeast strain

	STRAIN	CONSRUCTS	β- GALACTOSIDASE ACTIVITY
A	Y700b	-	100
	MA7310	-	2.1
	MA208	-	2.5
B	MA7310	VP16	2.1
		VP16/CR2	170

Fig3.1

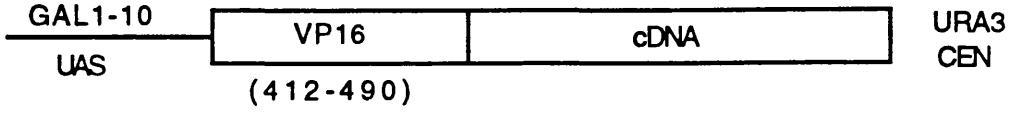
Quantitative lacz assays were performed.

A) The activity of the 3xATF-lacz reporter in the MA7310 and MA208 mutant strains was compared to the wild type strain Y700b

B) VP16 alone and a VP16/CREB2 fusion were transformed into MA7310 and their lacz values compared.

Three yeast cDNA clones activate transcription in MA7310

A



B

CONSRUCTS	β-GALACTOSIDASE ACTIVITY IN MA7310	APPROXIMATE SIZE OF cDNA (Kb)
VP16	2	-
A8	41	0.7
A18	66	1.0
A20	62	0.5

Fig3.2

A) Diagram of VP16/cDNA library construct used for the screen.

B) The three yeast VP16/cDNA clones and VP16 alone were transformed into MA7310, and quantitative lacz assays performed. The approximate size of each cDNA was estimated by restriction digest analysis on agarose gels.

on glucose agar medium. Growing the colonies on filters enabled them to be easily transferred to galactose medium to induce expression of the library. Growth on glucose was continued for about 48 hours and then the filters switched to galactose for a further 16 hours. Upon assaying the colonies on the filters for lacz activity, five colonies of blue (positive) phenotype were isolated. They were purified from any contaminating colonies of negative phenotype by replating dilutions to give single colonies, repeating the lacz assay and picking single blue colonies. The cDNA-containing plasmid was cured from these colonies by growing under non-selective conditions (YPD medium). Three colonies had a cDNA-dependent positive phenotype since they were no longer blue when cured of the Ura plasmid. The plasmids containing these cDNAs were prepared from the yeast and purified via E.coli. Upon retransforming into MA7310 the blue phenotype was restored, showing that the cDNAs could transmit the positive phenotype. The lacz activity of MA7310 expressing each of these cDNA clones was measured quantitatively (Fig 3.2B).

The cDNA clones were also transformed into several other yeast strains to determine their specificity. When transformed into the wild type strain Y700b, they were able to activate further the ATF reporter, despite its high endogenous activity. They were also able to activate the ATF-lacz reporter of another mutant strain obtained in the same screen, MA208. To test the specificity of the clones' ability to activate the ATF site, they were introduced into strains containing other reporters. They were tested on an AP1 site, and two mutant ATF sites, also on the SRE and the E2F site. They failed to activate any reporters containing these sites. For the two mutant ATF sites tested, there was no difference in the sequence of the reporter apart from base changes within the ATF site that decreased binding activity and the ability of the site to act as a UAS, in both yeast and mammalian cells. Since the cDNA products could no longer activate transcription via these mutated sites, this implied that their ability to activate was specific for the ATF site. (For summary, see fig3.3).

The three cDNA clones were sequenced from their N-terminus using a primer to VP16 (fig3.4). They were of a different sequence to each other and were not homologous to any sequence present in the data bank. There are motifs that might be expected of a DNA-binding factor. All the members of the mammalian ATF/CREB family share a homologous region consisting of a basic DNA binding domain. In addition they all have a leucine zipper motif, but outside these regions, most of them have completely divergent sequences. A basic DNA binding domain and a leucine zipper might be an expected feature of yATF. Other yeast transcription factors have this feature, for example, yAP1. The lack of such a motif in the sequences of these cDNA clones could either indicate that they are not DNA binding proteins, or that they must bind DNA using a different motif. However, it is important to note that the sequencing of the clones is not complete. Their approximate sizes are indicated in Fig3.2B.

The most likely explanation for the specific increase in activation via the ATF site in the presence of these clones, is that the products they encode either bind the ATF site directly, or form part of the complex binding to that site. This issue was initially

Three yeast cDNA clones specifically activate an ATF site

A

YEAST STRAIN	REPORTER	INCREASE IN LACZ ACTIVITY
Y700B	ATF	YES
MA208	ATF	YES
Y700	mATF1/2	NO
	AP1	NO
S62L	SRE	NO
Y4-1	E2F	NO

B

ATF	TACGTCA
mATF-1	TACTCCA
mATF-2	TACGCCA
AP1	TTAGTCA

Fig3.3

A) The three yeast cDNA clones were transformed into the yeast strains as indicated and Lacz assays performed. It was noted if they activated a reporter above its endogenous level of activity.

B) Core sequence of binding sites on reporter

Fig 3.4 DNA sequences of yeast cDNA clones A8, 18 and 20.

Sequencing is from the 5' end of each cDNA using a primer to the VP16 part of the vector. The predicted amino acid sequences in one open reading frame are shown.

Fig 3.4 DNA sequence from 5' end of cDNA A8.

```
   D  V  T  P  A  S  P  I  L  T  S  S  Q  T  P  H  Y  S  N  S
AAGACGTTACGCCCGCATCACCAATATTAACAAGTAGTCAAACGCCGCATTACTCAAACT
        10         20         30         40         50         60

   L  Y  N  A  P  F  A  V  S  S  P  P  D  P  L  P  N  L  F  T
CGCTTTATAACGCACCTTTTGCTGTTTCCTCTCCACCAGATCCTTTACCAAACCTTTTTA
        70         80         90        100        110        120

   T  T  S  E  K  V  F  P  K  I  N  V  L  I  V  E  D  N  V  I
CCACCACAAGTGAAAAAGTTTTCCCCAAAATTAATGTTTTAATAGTTGAAGACAACGTCA
       130        140        150        160        170        180

   N  Q  A  I  L  G  S  F  L  R  K  H  K  I  S  Y  K  L  A  K
TCAACCAAGCTATCTTAGGTTCCTTTCTGAGGAAACACAAAATCTCATATAAACTGGCTA
       190        200        210        220        230        240

   N  G  Q  E  A  V  N  I  W  K  E  G  G  L  H  L  I  F  M  D
AAAATGGTCAAGAAGCTGTTAATATTTGGAAGGAAGGCGGTCTTCATTTAATATTTATGG
       250        260        270        280        290        300

   L  Q  L  P  V  L  S  G  I  E  A  A  K  Q  I  R  D  F  E  K
ATTTACAGCTGCCTGTCTTGTCTGGTATAGAAGCTGCCAAGCAGATTAGGGACTTCGAAA
       310        320        330        340        350        360

   Q  N  G  I  G  I  Q  K  S  L  N  N  S  H  S  N  L  E  K  G
AACAAAATGGCATTGGCATTCAAAAAAGTCTCAATAACTCACACTCCAATCTTGAAAAAG
       370        380        390        400        410        420

   T  S  K  R  F  S  Q  A  P  V  I  I  V  A  L  T  A  S  N  S
GTACTTCAAAGAGATTCTCTCAGGCGCCCGTGATTATTGTAGCATTGACCGCATCTAACT
       430        440        450        460        470        480

   Q  M  D  K  R  K  A  L  L  S  G
CTCAGATGGATAAAAGAAAAGCACTTCTTTCTGGT
       490        500        510
```

Fig 3.4 DNA sequence from 5' end of cDNA A18.

```
   F  T  Y  T  N  E  M  Y  A  H  V  V  N  M  F  K  I  N  L  F
GGTTCACTTACACGAATGAGATGTACGCTCATGTGGTGAACATGTTCAAAATCAATCTGT
       10        20        30        40        50        60

   R  P  I  P  P  P  V  N  P  V  G  D  I  Y  D  P  D  E  D  E
TTAGACCTATTCCACCACCAGTAAATCCAGTTGGTGACATTTATGACCCAGATGAAGATG
       70        80        90       100       110       120

   P  V  N  E  L  A  W  P  H  M  Q  A  V  Y  E  F  F  L  R  F
AACCTGTTAACGAACTAGCCTGGCCTCATATGCAAGCTGTTTACGAATTCTTTTTAAGGT
      130       140       150       160       170       180

   V  E  S  P  D  F  N  H  Q  I  A  K  Q  Y  I  D  Q  D  F  I
TTGTGGAAAGTCCTGATTTCAATCATCAGATTGCTAAACAATATATTGATCAGGACTTTA
      190       200       210       220       230       240

   L  K  L  L  E  L  F  D  S  E  D  I  R  E  R  D  C  L
TTTTAAAGTTACTGGAATTATTTGATAGCGAAGATATCAGAGAAAGAGACTGTTT
      250       260       270       280       290
```

Fig 3.4 DNA sequence from 5' end of cDNA A20.

```
   Y  A  S  V  Q  S  I  L  N  D  S  F  D  E  R  A  E  T  L  H
CTTATGCCAGCGTGCAAAGCATACTAAATGATTCATTCGATGAGAGAGCAGAGACCCTAC
       10        20        30        40        50        60

   C  A  L  S  C  Q  S  E  K  Q  D  D  T  E  F  S  R  S  E  S
ATTGTGCATTAAGCTGCCAATCTGAAAAACAAGATGACACCGAGTTTTCCAGAAGTGAAA
       70        80        90       100       110       120

   S  E  Y  I  F  M  T  E  E  D  R  N  L  R  G  S  W  I  G  E
GCTCGGAATACATATTTATGACGGAAGAAGACAGAAACCTACGGGGCAGTTGGATCGGTG
      130       140       150       160       170       180

   P  K  E  C  F  T  F
AGCCAAAAGAGTGTTTTACCTT
      190       200
```

addressed using a gel retardation assay, which might detect such interactions. As mentioned above, the ATF specific complex in a gel retardation assay of yeast extracts appears to be the same in both Y700b and MA7310. Yeast extracts were made from the Y700b yeast strain containing either the three VP16/cDNA clones, pSD06, or VP16/CR2. The yeast were harvested after inducing the expression of the constructs they contained by growing in galactose overnight. Upon incubating yeast extract with a labelled ATF probe, the endogenous ATF binding activity could be seen by gel retardation (Fig3.5, complex I). As well as complex I, an additional complex was seen on both probes, of greater mobility on the 3xATF probe, and slower mobility on the ATF(FN) probe. None of these complexes were altered by the expression of VP16 from pSD06 (not shown). When VP16/CREB2 was expressed (lane 5), a more slowly migrating complex was seen (II) as well as the endogenous complexes. However, no such band was seen with extracts prepared from yeast expressing any of the cDNA clones, nor did the endogenous ATF complexes appear to be altered on either probe (lanes 2 to 4). Both the endogenous ATF and VP16/CREB2 complexes were able to be competed by excess cold probe (lane 6). The AP1 site was used as a mutant ATF site competitor in lane 7. It was not able to compete with the VP16/CREB2 complex binding to either ATF site, confirming the specificity of this complex. The AP1 site competitor was able compete with the residual endogenous ATF binding activity on the 3xATF site, but at lower levels of competitor this effect is not seen (Jones and Jones, 1989).

In conjunction with this experiment, the cDNA clones were transferred as VP16 fusions into a T7 vector, transcribed into RNA and translated in rabbit reticulocyte lysate (fig3.6). Upon addition of these *in vitro* translated products to a gel retardation reaction with an ATF probe, the endogenous ATF binding activity of the reticulocyte lysate could be seen, but no other complexes could be seen with the expression of any of the cDNA clones.

These gel retardation results in both yeast and *in vitro* systems would suggest that the cDNA clones are neither binding directly to DNA, nor are part of the ATF site binding complex. However, interactions directly with DNA could be weaker than those of the CREB2 used as a control if only a partial clone is being expressed. The lacz activity of VP16/CREB2 on the ATF sites is higher than any of the cDNA clones. The cDNA products could also be part of the complex on the ATF site, but not be binding DNA directly themselves. In this case, the protein-protein interactions involved may be too weak to visualise in a gel retardation assay.

As a final experiment to test whether these clones could interact directly with an ATF site, they were transferred as VP16 fusions into the mammalian expression vector, MLV. Transient transfections were performed in HeLa cells with a reporter consisting of 3xATF sites upstream of the CAT gene. This has a low basal level which could be activated strongly by the E1a gene, but the three VP16/cDNA clones failed to activate it (fig3.7).

From these experiments, it would seem that although these yeast cDNAs have a very specific phenotype in yeast, they may not be binding DNA directly themselves,

Fig 3.5 Gel retardation of yATF binding activity in Y700b.

Y700b yeast extract was incubated in a binding reaction with two different labeled probes: A: 3xATF site, B: ATF(FN) site
The reaction mixtures were loaded onto a 6% polyacrylamide gel, electrophoresed to separate DNA-protein complexes from free probe, and autoradiographed.
Yeast had been transformed with the following plasmids:
Lane1: no plasmid,
Lanes2, 3 and 4: VP16/cDNA clones A8, 18 and 20 respectively
Lanes5-7: VP16/CREB2.
Lanes 6 and 7 contained 100ng of competitor DNA as follows:
Lane6: cold probe (3xATF or ATF(FN)),
Lane7: AP1 site.
Complex(I) is due to endogenous yATF activity,
complex(II) is due to VP16/CR2 binding to ATF site.

Gel Retardation of Yeast ATF Binding Activity

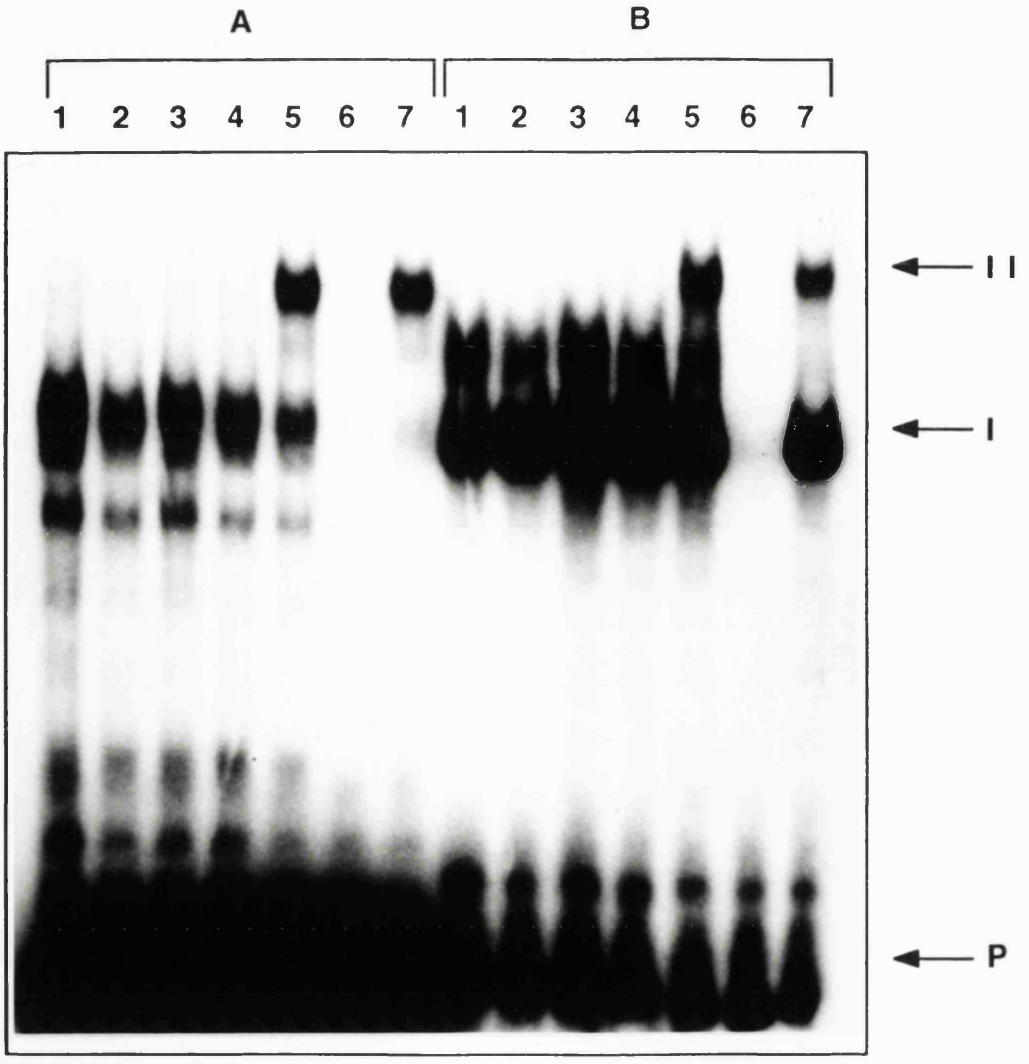
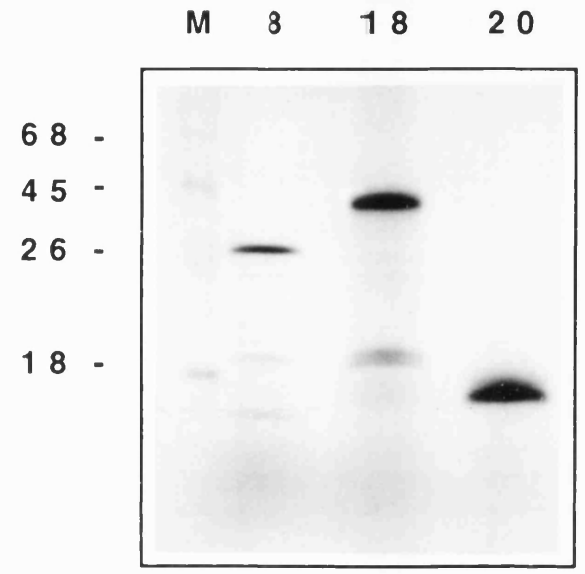


Fig 3.6 In vitro translation of cDNA clones from yATF screen

5µl reticulocyte lysate programmed with RNA from cDNA clones A8, 18 or 20 was loaded onto a 10% SDS-polyacrylamide gel with loading buffer. Markers (M) were also run.

In vitro translation of cDNA clones from yATF screen



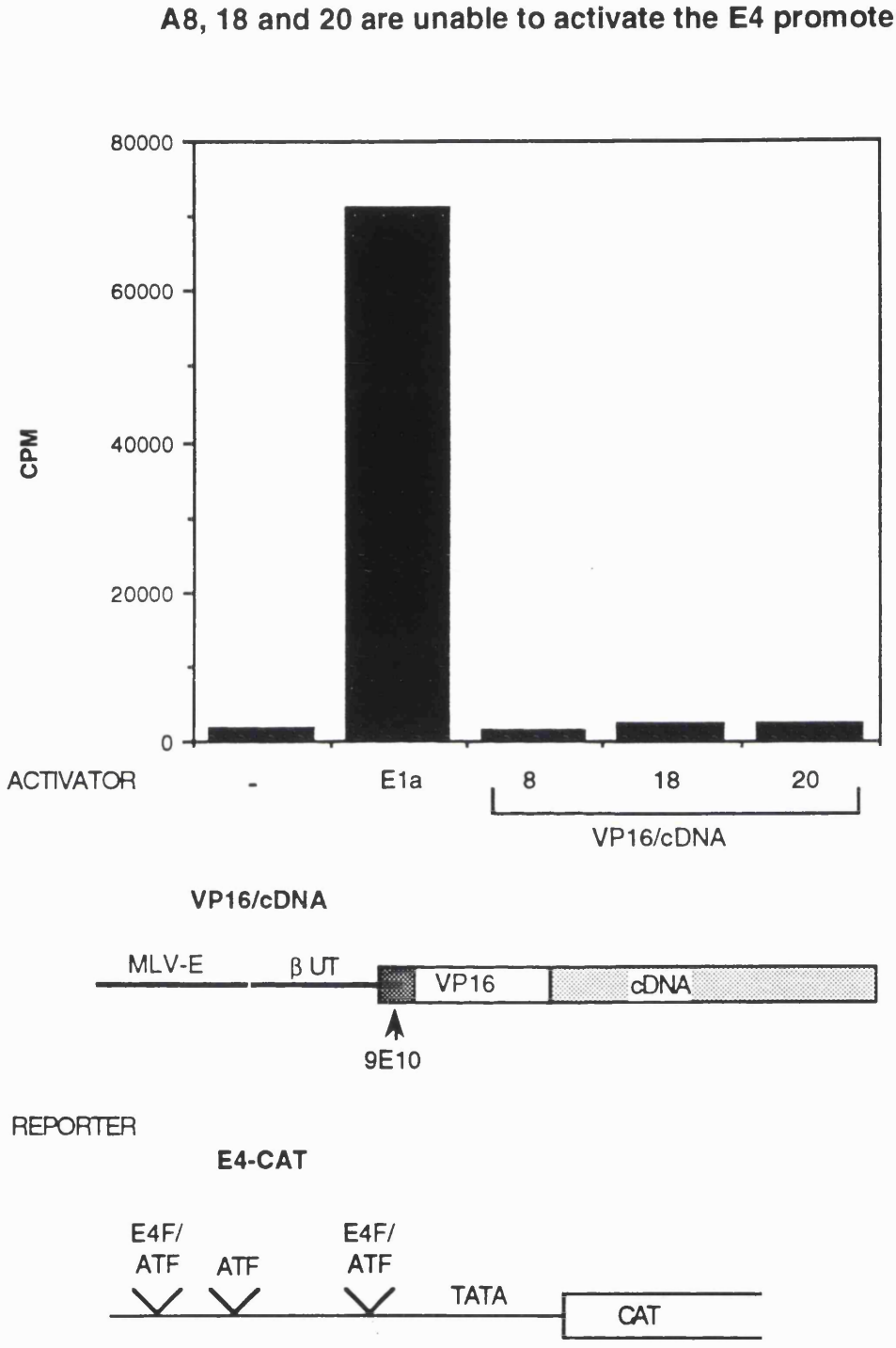


Fig3.7

Constructs were transfected into HeLa cells. 2μg of a reporter consisting of the E4 promoter fused to the CAT gene, was cotransfected with either 2μg of E1a (pCE plasmid) or 5μg of clones A8, 18 or 20 in the MLV vector. Diagrams are shown of the MLV expression vector containing the VP16/cDNA, and of the CAT reporter construct.

and may be interacting with a yeast factor in order to activate the ATF site. At present, any such interaction cannot be detected by the methods used. It would be of interest to delete the genomic copy of these genes and determine whether the deletion was lethal, or of its affect on the activity of the ATF site.

CHAPTER 4: RESULTS

YEAST SCREENS FOR THE MAMMALIAN E2F TRANSCRIPTION FACTOR

INTRODUCTION.

The E2F factor has been defined as a binding activity that plays a key role in the transcriptional regulation of the adenovirus E2A gene (Loeken and Brady, 1989). Sequences homologous to the E2F binding site have also been found upstream of several cellular genes whose products are involved in DNA replication. As described in the Introduction, the activity of E2F can be modified by both viral and cellular proteins. Purification of this factor has yielded a polypeptide of 54K molecular weight capable of binding to the E2F site with the correct specificity. During this purification, it was found that active E2F molecules are of relatively low abundance in the cell compared to other transcription factors such as Sp1 and AP1, which may have hampered searches for a cDNA. (Yee et al., 1989). As well as its DNA binding specificity, other properties known of E2F that may suggest strategies to isolate a cDNA, are that it can form a complex with the product of the ORF6/7 of the adenovirus E4 gene, and with cellular proteins such as Rb and p107. This chapter will describe two strategies to obtain a cDNA for E2F; one exploiting the DNA binding activity of E2F and the other its interaction with the E4 protein.

For both strategies, use was made of a library of randomly primed HeLa cell cDNA, inserted C-terminal to the (HSV)VP16 activation domain, in a yeast vector that allows galactose inducible expression of the fusion protein. This library had already been used successfully in a yeast genetic screen to isolate SAP1, a protein recruited by SRF to the SRE to form a ternary complex (Dalton and Treisman, 1992). In the screen for SAP1, a partial cDNA of SRF itself was also isolated that bound directly to the reporter.

To isolate cDNA clones that can bind to the E2F sites, the library was transformed into a yeast strain containing E2F sites upstream of a lacz reporter. VP16/cDNA clones that interact with the promoter will activate transcription of the reporter via the VP16 activation domain, and yeast cells expressing them will therefore have a blue colour in the lacZ colony colour assay. Any such positives can be checked whether they encode a protein that specifically binds to the E2F sites by re-transforming recovered cDNA-containing plasmids into different reporter strains which are available. For example, clones specific for the E2F site should not activate a reporter containing a mutated E2F site.

The second approach was to attempt to clone a cDNA encoding a protein that could interact with the E4 protein. E4 forms a stable complex with E2F on the 2 E2F sites of the E2A promoter. Additionally, E4 antibodies can immunoprecipitate E2F binding activity showing that the complex can form independently of DNA and in solution (Huang and Hearing, 1989). The E4 complex with E2F can be seen in a gel retardation experiment on either a single E2F site or a double one as found in the E2A promoter. These properties indicate that the physical association between E4 and E2F is fairly strong. To use E4 as a target for a yeast genetic screen, E4ORF6/7 was expressed as a

fusion to a heterologous DNA binding domain. This fusion gave a low basal level of transcriptional activation in a yeast strain with an appropriate reporter. The strain served as a recipient for the HeLa cDNA library fused to the VP16 activation domain. Yeast cells scoring positive in the Lacz assay should express proteins that interact with either the E4 or the SRF part of the fusion, or that interact directly with the promoter region of the reporter. Such specificity can subsequently be checked after ensuring that the positive phenotype can be transmitted by the cDNA clone to another yeast strain, and is therefore cDNA-specific.

4.1 Screening for a factor binding directly to an E2F site

4.1.1 Activity of E2F sites in Yeast

Chapter 2 showed that a part of the E2A promoter containing the two E2F sites placed upstream of the Lacz gene, gave a fairly high transcriptional activity with the multi-copy (2μ-based) plasmid p4-1. This is presumably due to a yeast activity that binds to this part of the E2A promoter, and may be specific for the E2F sites. Other variations of this reporter were also tested (See Fig4.1A). p3-1 contains the same two sites, but arranged in tandem rather than in inverted orientation. p2-1 has one E2F site, and p1-1 has a mutant version of the E2F sites in the same orientation as p4-1. The activity of these reporters in yeast reflects their activity in mammalian cells: the mutant E2F sites have a very low activity, and the single site and tandemly orientated sites have a lower activity than the wild type site. In mammalian infected cells, the wild type sites (4-1) are much more active than the tandem arrangement (3-1) since the wild type arrangement allows the E4 protein to stimulate the cooperative binding of E2F to these sites. It is unclear why the orientation of the two sites also causes this difference in yeast; cooperative binding of factors is perhaps possible if the sites are arranged in this way. Alternatively the flanking regions around the core may play some role: reversing the orientation of the proximal site may, for example, create a new DNA binding site since the core will be flanked by different 5' and 3' sequences.

The reporter 4-1, with E2F sites as arranged on the E2A promoter upstream of the lacz gene, was integrated into the yeast genome to give yeast strain Y4-1. This was accomplished by cloning the lacz gene and the UAS into the middle of the Ura3 gene, and then integrating by homologous recombination into the genomic Ura3 gene. Now that each yeast cell only carried one copy of the reporter (checked by Southern analysis), it had a much lower Lacz activity (fig4.1B), and enabled an approach to clone a mammalian cDNA that might bind to the E2F sites.

Lacz activity of E2F site in yeast

A

LAC Z REPORTER	β- GALACTOSIDASE ACTIVITY IN Y700
E2F(4-1)	800
E2F(3-1)	250
E2F(2-1)	250
E2F(1-1)	10

B

YEAST STRAIN	β- GALACTOSIDASE ACTIVITY
Y4-1	5
Y700	<1

C

p4-1 TTTTCGCG-(18bp)-CGCGAAA

p3-1 TTTTCGCG----------TTTTCGCG

p2-1 TTTTCGCG

p1-1 TTTTGGGG----------CCCCAAA

Fig4.1

A) Reporters on 2μ-based plasmids were transformed into the yeast strain Y700 and quantitative lacz assays were performed.

B) The lacz activity was compared between strain Y4-1, which contains an integrated p4-1 reporter, and Y700.

C) Core sequences of the E2F binding sites upstream of the lacz gene.

4.1.2 Yeast screen.

The Y4-1 strain was transformed with the VP16-tagged library, plated out on nylon filters placed on top of the agar plates, and grown on selective glucose medium for 24 to 36 hours. The filters were transferred to galactose medium for 12 to 16 hours to induce expression of the cDNA library, then a lacZ assay was performed on the colonies. Positive blue colonies were picked. The endogenous activity of the E2F sites appeared as a faint blue colour after about two hours, but this still enabled stronger positives to be seen above this background. Approximately $3x10^6$ colonies were screened in this way, and 41 dark blue colonies picked. After purification of the positive colonies from any contaminating negative ones, they were cured of the cDNA-containing plasmid (URA marker) by growing under non-selective conditions (YPD medium). 7 of the darkest blue colonies failed to cure of their plasmid, indicating that an integration event had taken place. This integration was not investigated further, but since the library and the reporter were derived from the same CYC1 promoter sequences, it was presumed that homologous recombination had occurred so that the GAL promoter from the library was driving the integrated Lacz gene and thus giving a very dark blue inducible phenotype. The other colonies appeared to cure of their plasmid, and upon retesting for lacz activity, were no longer of positive phenotype. This indicated that they scored positive as a result of the presence of the VP16/cDNA. A population of plasmids from each yeast colony were then recovered via E. coli, re-transformed into Y4-1 and re-screened for lacZ activity. If more than one plasmid were present per yeast colony, the one generating the positive phenotype should not be missed by this method. However, upon re-screening, the positive phenotype was not recovered.

It is difficult to speculate why the positive phenotype was not seen upon re-transformation, and why the colonies were initially showing cDNA-dependent blue phenotypes. A mutation event could have occured in the colonies picked that increased their transcriptional activity, and somehow this only occured in a VP16-dependent manner. However, since the phenotype could not be transmitted by the cDNA-containing plasmid alone, further exploration was not undertaken.

4.2 Screening for a cDNA interacting with E4

4.2.1 Expression of E4 in yeast.

In order to use the E4 protein as a target for a screen with the VP16-tagged library, the gene was fused to a heterologous DNA binding domain to bring it to the promoter region of a reporter. The DNA binding domain of SRF was used, in conjunction with a reporter with an SRE upstream of lacz. This reporter is integrated to make the S62L strain (Dalton and Triesman, 1992).

The ORF6/7 of E4 was cloned by PCR into the SRF-containing plasmid pSD09, to create an E4/SRF fusion. The E4 part of the fusion, plus its junction with the SRF gene,

was sequenced to check for polymerase errors and to confirm that the ORF of the junction was correct. The construct was transformed into the yeast strain S62L, containing the integrated SRE-LacZ reporter, and its LacZ activity compared to that of SRF alone. SRF bound to this reporter does not activate the lacz gene, so the yeast are white in a colony colour assay, while the E4/SRF fusion gave a very pale blue phenotype in this strain (Fig4.2A). This background was lower than that of the E2F sites driving LacZ used for the previous screen and would therefore enable a more sensitive screen with the VP16/cDNA library.

The integrity of the fusion protein also was checked in 2 other ways:

i) The first method involved a functional assay in yeast. SRF bound to the SRE can recognise the protein SAP1 which is recruited to the promoter to form a ternary complex. VP16/SAP1 was co-transformed with E4/SRF into the yeast strain S62L. The interaction between SRF and SAP1 brought VP16 to the promoter, thereby activating its transcription strongly (Fig4.2B). This indicated that SRF was being translated in the correct reading frame, and strongly suggested the correct translation of the E4 ORF since it is located N-terminal to SRF.

ii) Secondly, an *in vitro* assay was used to check that the E4/SRF fusion could interact with E2F. The ORF of the fusion construct was cloned into a T7 vector enabling transcription by T7 RNA polymerase and translation of this RNA in rabbit reticulocyte lysate. As a positive control, the unfused E4ORF6/7 in a T7 vector was also transcribed and translated. The interaction of these E4 proteins with E2F was analysed by a gel retardation assay (Fig4.3). DNA-bound E2F can be found in cells either free of complex formation with other proteins, or else complexed with cellular proteins. In Ad-infected cells, E2F is found in a complex with E4 rather than with cellular proteins. Using a labelled probe containing the two E2F sites as arranged on the E2A promoter and incubating this with HeLa cell nuclear extract, E2F-specific complexes are detected (lane1). In HeLa cells, the majority of the E2F-specific DNA binding activity is free of complex formation (I) with cellular proteins, although some complexes remain which can be seen as bands of slower mobility. With extracts prepared from Ad-infected cells, the complex (I) due to free E2F is reduced, and a new, more slowly migrating large complex is seen, which contains the E4-ORF6/7 protein (lane2). All the specific complexes are competed with unlabelled excess probe (lane5). Using the same probe incubated with uninfected HeLa cell lysate and adding unprogrammed reticulocyte lysate, an increased E2F complex (I) is seen due to E2F binding activity in the reticulocyte lysate. Adding RL programmed with E4 RNA to the HeLa nuclear extract gives a complex (II) with a similar mobility as the infection-specific form of E2F. When increasing amounts of E4/SRF protein are added to the HeLa nuclear extract, the complex due to free E2F is seen to disappear, although the slower-migrating complex seen with E4 and infected HeLa cells is not apparent (lanes6 to 8). The disappearance

Activity of adenovirus E4 (ORF6/7) in yeast

	CONSRUCTS	β-GALACTOSIDASE ACTIVITY IN S62L
A	SRF	<1
	E4/SRF	3.0
B	E4/SRF + VP16	3.0
	+ VP16/SAP1	225

C

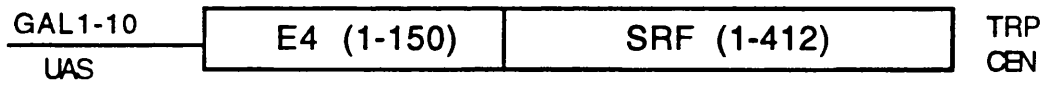


Fig4.2

A) E4/SRF and SRF (pSD09) were transformed into S62L and their activity compared by quantitative lacz assays.

B) E4/SRF was cotransformed with either VP16 alone or VP16/SAP1 and quantitative lacz assays were performed.

C) Diagram of E4/SRF construct

Fig 4.3 Gel retardation assay of E2F and E4 complexes.

A labelled probe consisting of the E2F sites from the E2A promoter (4-1 sequence) was incubated with the following extracts.

Lane1: HeLa cell nuclear extract

Lane2: Adenovirus-infected HeLa cell extract

Lanes3-9: HeLa cell nuclear extract as in lane1, plus the following:

Lane3: 4μl reticulocyte lysate (RL)

Lane4: 4μl E4 protein made in RL

Lane5: as lane3, but with 100ng unlabelled probe as competitor

Lane6: 3.5μl RL plus 0.5μl E4/SRF protein (made in RL)

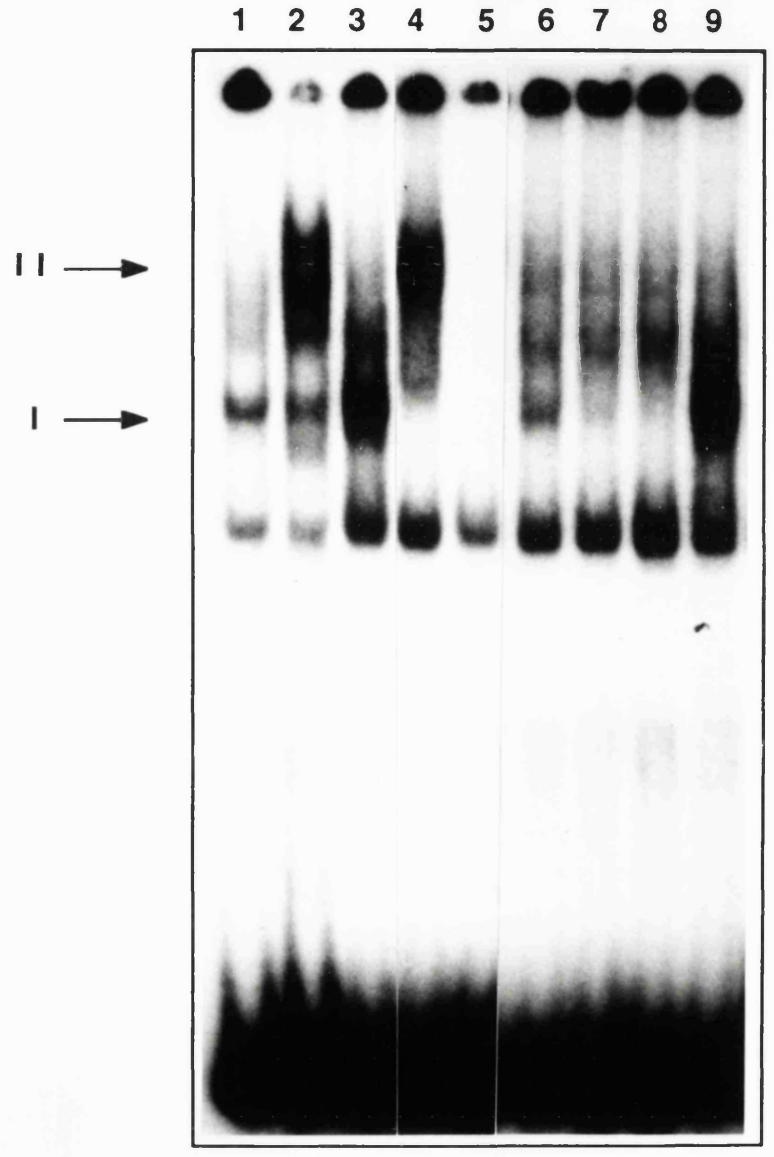
Lane7: 2μl RL plus 2μl E4/SRF protein (RL)

Lane8: 4μl E4/SRF protein (RL)

Lane9: 4μl RL plus SRF protein made in insect cells

The reaction mixtures were loaded onto a 6% polyacrylamide gel, electrophoresed at 4°C to separate DNA-protein complexes from free probe, and autoradiographed.

Gel Retardation of E2F and E4 Complexes



of the E2F complex is dependent upon the E4 sequences of the fusion protein, since addition of an equivalent amount of SRF protein (from insect cells) does not alter the E2F complex (lane9). This would indicate that the E4/SRF protein is able to bind to free E2F, although in doing so, prevents it from binding DNA. Therefore the property of E4 required for this screen - to be able to bind to E2F - is functional in the E4/SRF fusion.

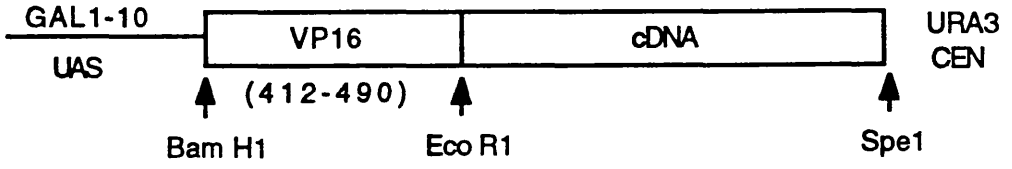
4.2.2 Screening with E4/SRF

The S62L strain containing the E4/SRF plasmid was transformed with the VP16-tagged library (fig4.4A). Approximately 10^6 colonies were obtained and transferred to galactose plates to induce expression of both the E4/SRF protein and the library products. Upon developing for lacz activity, 50 positives (blue phenotype) were picked from the filters and purified from any contaminating negative ones. In order to test whether these yeast had a cDNA-dependent positive phenotype, they were cured of the library (Ura marker) and their lacZ activity on galactose retested. A cDNA-dependent blue colony should now return to basal levels of lacz activity due only to the E4/SRF protein. One yeast colony could not be cured of its Ura plasmid indicating that an integration event had occured - this was the strongest positive picked. At least 30 clones could be cured to give white colonies. This indicated that their blue phenotype was cDNA-dependent. 13 of these had the strongest blue phenotype. These clones were chosen to test if the expressed cDNA was interacting with the E4/SRF fusion. The original yeast clones containing the cDNA were cured of the E4/SRF plasmid (Trp marker). They lost their lacz activity, indicating that the activity was dependent on the fusion protein. A quick method was used to see if this was determined by the SRF or E4 part of the fusion protein. The yeast cured of the E4/SRF plasmid, but still containing the VP16/cDNA were mated with a yeast strain of opposite mating type (Y35) containing either the original E4/SRF plasmid again, or one containing SRF alone (pSD09). Diploids were selected in medium lacking histidine (diploids complement each other to become HIS^+ since Y700 is $his3^-4^+$ and Y35 is $his4^-3^+$), and assayed for lacz activity. cDNA products that interacted with SRF would give a blue phenotype with both pSD09 and E4/SRF, while those interacting with E4 would be white with pSD09. The lacz assay indicated that 9 clones were dependent upon the E4 part of the fusion, and the rest were not. These phenotypes were confirmed by recovering the cDNA clones from the yeast cured of the Trp plasmid, and following transformation into and plasmid recovery from E. coli, retransforming them into S62L together with either E4/SRF or pSD09.

Therefore, unlike the previous screen, cDNA clones were isolated that could transmit their positive phenotype when retransformed into yeast. The lacz activity of the yeast colonies containing these cDNA clones was defined more accurately using a liquid b-galactosidase assay (Fig4.4). Section B shows the cDNA clones whose products interact specifically with E4, and section C those that showed a positive phenotype with both E4/SRF and SRF alone. Clones D2, 9, 17, 21, 43 and 44 are grouped together because subsequent sequence analysis showed that they were the same (see later).

Activity of cDNA clones from E4 screen

A



	CONSTRUCTS CO-TRANSFORMED WITH E4/SRF	β-GALACTOSIDASE ACTIVITY IN S62L	APPROXIMATE SIZE OF cDNA (Kb)
B	D2, 9, 17, 21, 43, 44	27	1.6
	D4	15	0.75
	D29	15	0.25
	D49	16	0.5
C	D25	92	0.6
	D26, 39	225	0.75
D	VP16	3	-

Fig4.4

A) Diagram of VP16/cDNA library construct used for the screen.

B) The E4-dependent cDNA clones were co-transformed with E4/SRF into S62L and quantitative lacz assays were performed.

C) Activity of E4-independent cDNA clones when co-transformed with E4/SRF.

D) E4/SRF was co-transformed with pSD06 and its basal activity measured in the absence of any cDNA fusion.

The approximate size of each cDNA was estimated by restriction digest analysis on agarose gels.

4.3 Analysis of E4-dependent clones.

4.3.1 Yeast assays for E2F-like properties.

A cDNA product that can interact with E4 might be expected to have some or all of the properties of the E2F transcription factor. This certainly may not be the case - the complex seen on the E2F site may be a heterogeneous one, with different subunits interacting with E4 than with DNA or cellular proteins such as Rb. Also, even if the same polypeptide exhibits more than one of these properties, the domains for each interaction may be separate, and the cDNA clone obtained may only contain the domain that interacts with E4. Finally, a polypeptide that interacts with E4 may not be related in any way to the E2F transcription factor. However, since simple yeast assays for some properties of E2F were available, the cDNA clones were tested in various yeast strains in the following ways.

i) To check for DNA binding ability on the E2F site, the cDNAs (still as VP16 fusions in the library vector) were transformed into the Y4-1 strain which contains an integrated E2F-Lacz reporter. Upon induction of the expression of the cDNA with galactose, no difference in lacz activity was seen between yeast containing any of the cDNAs, or containing VP16 alone (in pSD06).

ii) To test for binding to Rb, the cDNAs were introduced into Y35 (with its integrated LexAop-lacz reporter) together with the Rb/LexA construct. Rb/LexA had been previously shown to be functional in this strain since VP16/E1a could bind to it and thereby activate the lacz gene (see fig2.4). However, none of the cDNAs showed such a phenotype, so in this assay at least do not appear to be binding to Rb.

4.3.2 The region of E4 required for interaction with the clones.

During the course of these experiments, some additional information was published concerning the region of E4 required for its interaction with E2F. The C-terminal 70 amino acids of the protein are necessary and sufficient for full induction of E2F binding activity on the E2A promoter. Within this region lies a putative helix-loop-helix (HLH) structure which is required for the interaction with E2F, and it appeared that the very C-terminal 10 amino acids are also essential. (Neill and Nevins, 1991; O'Connor and Hearing, 1991)

To determine which part of E4 the cDNA products were interacting with, three deletions of E4 in the E4/SRF fusion were made and tested in yeast for their ability to bind to the cDNA products. In E4D the C-terminal 10 amino acids of ORF6/7 were deleted; in E4DH the deletion was extended to include one helix; and in E4D2H the deletion was extended further to include the entire HLH structure. These deletions were constructed by PCR, cloned into pSD09 and the E4 ORF checked by sequencing. Additionally, their expression was tested in yeast in the same way the full length E4/SRF had been tested: each one was co-transformed with the VP16/SAP1 construct. In each

case the SRF was able to bring SAP1 to the promoter indicating that it was being translated in the correct ORF. Therefore the E4 part of the fusion should also be translated correctly.

The behaviour of the E4 deletions was checked *in vitro* by looking at their ability to abolish the E2F bandshift seen with the E2A probe. Each E4/SRF fusion was transferred from pSD09 into pT7, and the DNA transcribed and translated. On an SDS polyacrylamide gel, equivalent amounts of 35S labelled protein were made as compared to the full length E4/SRF fusion (Fig4.5). In a gel retardation assay with labelled E2A probe, as expected, all three deletions failed to abolish the E2F bandshift, ie failed to interact with E2F *in vitro*, although the full length E4 fused to SRF could interact with E2F as seen previously. (See fig4.6 for summary).

These deletions were then tested in yeast by co-transforming each one in pSD09 together with each VP16/cDNA. E4D could still interact with all the cDNA products, giving a Lacz activation of the same strength as full length E4. The other deletions E4DH and E4D2H failed to interact with any of the cDNA products. (Fig4.6). All the deletions still retained the same background activity in yeast as full length E4/SRF.

These results show a discrepancy in behaviour of the deletion of the C-terminal 10 amino acids of E4 (E4D) in its interaction with the E2F binding activity and with the products of the cDNA clones. I showed that E4D cannot interact with E2F in a gel retardation assay as an SRF fusion. Others have also used gel retardation and transient transfection assays to show that the C-terminal 10 amino acids are crucial (Neill and Nevins, 1991; O'Connor and Hearing, 1991). It had been expected that if any of the cDNA products were binding to E4 with the same specificity as E2F, then they should not interact with E4D.

4.3.3 Sequence Analysis of E4-dependent Clones.

The cDNA clones were sequenced from their N-terminus using a primer to the C-terminus of VP16. This revealed that six of the clones were identical at their N-terminus. These six clones also yeilded restriction fragments of similar size on an agarose gel when digested with enzyme combinations such as BamH1 and Spe1, indicating that their C-terminal splice site in the yeast vector may also be the same. They will subsequently be referred to as clone D2. (For approximate sizes of each cDNA, see fig4.4). The remaining three cDNAs had a completely different sequence composition to each other and to the above six clones. The sequences of clones D2, 29 and 49 are shown in fig4.7. A full length sequence was only determined for clone D29. Of all 4 sequences, only one was identical to a sequence present in the data bank. This was clone D4 which was the gene H12.3 (Guillemot et al, 1989). The H12.3 gene had been isolated by chance in studies to map the chicken major histocompatibility complex (MHC) because it was linked to the MHC gene cluster. The chicken gene is identical at the amino acid level to the human one, but unlike the case in chickens, the H12.3 gene is unrelated to the MHC genes. It has a repeat structure similar to that found in the family of guanine nucleotide-

Fig 4.5 In vitro translation products of E4/SRF and deletions.

5µl of reticulocyte-lysate translated proteins were resolved by a 10% SDS polyacrylamide gel.
Lane1: full length E4/SRF fusion
Lane2: E4D/SRF
Lane3: E4DH/SRF
Lane4: E4D2H/SRF

In vitro translation products of E4/SRF and deletions.

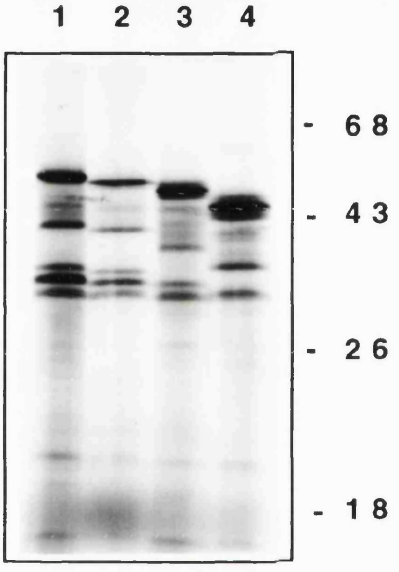
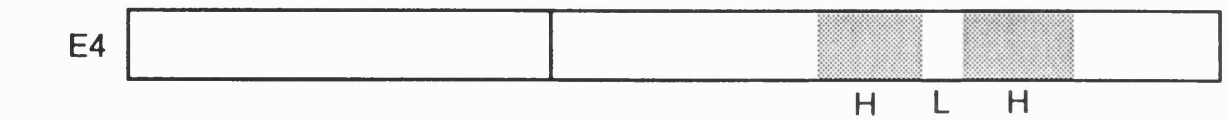


Fig 4.6 Activities of E4 deletions

E4 (ORF6 and 7) is shown as the full length protein of 150 amino acids with the helix-loop-helix domain indicated (HLH). Below are shown the three deletions of E4 with their lengths indicated. The interaction of these deletions with E2F was determined by translating each as an SRF fusion in reticulocyte lysates and performing a gel retardation assay on the E2F site. The E2F complex was abolished if the E4/SRF fusion was able to interact with E2F (see fig 4.3). To determine their interaction with the cDNA clones, each E4 deletion as a fusion to SRF was co-transformed with each cDNA into the yeast strain S62L, and qualitative lacz assays were performed.

Fig4.6 Activities of E4 deletions



	AMINO ACIDS	E2F INTERACTION (MAMMALIAN)	INTERACTION WITH cDNAS (IN YEAST)
E4	1-150	Yes	Yes
E4D	1-140	No	Yes
E4DH	1-115	No	No
E4D2H	1-96	No	No

Fig 4.7 DNA sequences of E4-dependent cDNA clones.

Sequencing is from the 5' end of each cDNA using a primer to the VP16 part of the vector. The predicted amino acid sequences in one open reading frame are shown.

Fig4.7 DNA sequence from 5' end of cDNA clones D2, 9, 17, 21, 43, 44.

```
P  H  A  H  I  S  E  Q  A  V  W  A  L  G  N  L  A  G  D  G
CCCCATGCTCACATCAGTGAACAAGCTGTCTGGGCTCTAGGAAACTTAGCAGGTGATGGC
        10        20        30        40        50        60

 S  V  F  R  D  L  V  I  K  Y  G  A  V  D  P  L  L  A  L  L
TCAGTGTTCCGAGACTTGGTTATTAAGTACGGTGCAGTTGACCCACTGTTGGCTCTCCTT
        70        80        90       100       110       120

 A  V  P  D  M  S  S  L  A  C  G  Y  L  R  N  L  T  W  T  L
GCAGTTCCTGATATGTCATCTTTAGCATGTGGCTACTTACGTAATCTTACCTGGACACTT
       130       140       150       160       170       180

 S  N  L  C  R  N  K  N  P  A  P  P  I  D  A  V  E  Q  I  L
TCTAATCTTTGCCGCAACAAGAATCCTGCACCCCCGATAGATGCTGTTGAGCAGATTCTT
       190       200       210       220       230       240

 P  T  L  V  R  L  L  H  H  D  D  P  E  V  L  A  D  T  C  W
CCTACCTTAGTTCGGCTCCTGCATCATGATGATCCAGAAGTGTTAGCAGATACCTGCTGG
       250       260       270       280       290       300

 A  I  S  Y  L  T  D  G  P  N  E  R  I  G  M  V  V  K  T  G
GCTATTTCCTACCTTACTGATGGTCCAAATGAACGAATTGGCATGGTGGTGAAAACAGGA
       310       320       330       340       350       360

 V  V  P  Q  L  V  K  L  L  G  A  S  E  L  P  I  V  T  P  A
GTTGTGCCCCAACTTGTGAAGCTTCTAGGAGCTTCTGAATTGCCAATTGTGACTCCTGCC
       370       380       390       400       410       420

 L  R  A  I  G  N  I  V  T  G  T  D  E  Q  T  Q  V  V  I  D
CTAAGAGCCATAGGGAATATTGTCACTGGTACAGATGAACAGACTCAGGTTGTGATTGAT
       430       440       450       460       470       480
```

```
 A  G  A  L  A  V  F  P  S  L  L  T  N  P  K  T  N  I  Q  K
GCAGGAGCACTCGCCGTCTTTCCCAGCCTGCTCACCAACCCCAAAACTAACATTCAGAAG
      490       500       510       520       530       540

 E  A  T  W  T  M  S  N  I  T  A  G  R  Q  D  Q  I  Q  Q  V
GAAGCTACGTGGACAATGTCAAACATCACAGCCGGCCGCCAGGACCAGATACAGCAAGTT
      550       560       570       580       590       600

 V  N  H  P  L  V  P  F  L  V  S  V  L  S  K  A  D  F  K  T
GTGAATCATCCATTAGTCCCATTCCTTGTCAGTGTTCTCTCTAAGGCAGATTTTAAGACA
      610       620       630       640       650       660

 Q  K  E  A  V  W  A  V  T  N  Y  T  S  G  G  T  V  E  Q  I
CAAAAGGAAGCTGTGTGGGCCGTGACCAACTATACCAGTGGTGGAACAGTTGAACAGATT
      670       680       690       700       710       720

 V  Y  L  V  H  C  G  I  I  E  P  L  M  N  L  L  T  A  K  I
GTGTACCTTGTTCACTGTGGCATAATAGAACCGTTGATGAACCTCTTAACTGCAAAGATA
      730       740       750       760       770       780

 P  R  I  I  W  Y  L  D  A  I  S
CCAAGGATTATCTGGTATCTGGATGCCATTTCAA
      790       800       810
```

Fig4.7 Complete DNA sequence of cDNA D29.

```
   G  S  Y  G  S  S  S  Q  S  S  S  Y  G  Q  P  Q  S  G  S  Y
CGGGAAGTTACGGTAGCAGTTCTCAGAGCAGCAGCTATGGGCAGCCCCAGAGTGGGAGCT
        10        20        30        40        50        60

   S  Q  Q  P  S  Y  G  G  Q  Q  Q  S  Y  G  Q  Q  Q  S  Y  N
ACAGCCAGCAGCCTAGCTATGGTGGACAGCAGCAAAGCTATGGACAGCAGCAAAGCTATA
        70        80        90       100       110       120

   P  P  Q  A  M  D  S  R  T  S  T  T  A  A  V  V  V  E  V  E
ATCCCCCTCAGGCTATGGACAGCAGAACCAGTACAACAGCAGCAGTGGTGGTGGAGGTGG
       130       140       150       160       170       180

   V  E  V  E  V  T  M  A  R  S  I  L  H  H  S  G  L  D  C  H
AGGTGGAGGTGGAGGTAACTATGGCAAGATCAATCCTCCACCACTCTGGTCTTGATTGCC
       190       200       210       220       230       240

   N  R  H  H  C
ATAACCGCCACCACTGC
       250
```

Fig 4.7 DNA sequence from 5' end of cDNA D49.

```
 V  E  E  E  D  P  G  G  L  S  E  K  D  F  L  V  G  L  L  D
GTTGAAGAAGAGGATCCGGGTGGCCTCAGTGAAAAGGATTTTCTCGTAGGCCTTCTCGAT
        10        20        30        40        50        60

 A  P  S  D  L  I  P  D  S  V  E  Q  D  V  N  E  E  G  V  A
GCACCCAGCGATCTCATCCCTGATAGTGTCGAGCAGGATGTCAATGAAGAAGGTGTAGCT
        70        80        90       100       110       120

 L  G  G  D  V  T  L  G  Q  E  H  F  V  V  A  A  L  H  Q  V
CTCGGCGGGGATGTTACCCTTGGCCAGGAACACTTTGTTGTAGCTGCCCTCCATCAGGTA
       130       140       150       160       170       180

 L  L  Q  G  H  W  V  L  D  V  D  I  G  L  Y  V  L  G  R  Q
TTGCTCCAGGGACACTGGGTGCTTGATGTAGACATTGGTCTGTATGTCCTTGGCAGGCAG
       190       200       210       220       230       240

 P  L  Q  L  R  V  E  L  S  H  R  S  G  Q  Q  E  E  E  V
CCGCTCCAACTCCGTGTGGAACTCAGCCACCGGTCTGGACAGCAGGAAGAGGAGGTGA
       250       260       270       280       290
```

binding protein β (Gβ) subunits. This repeat structure is also found in other proteins that are not thought to form part of the G-protein complex, such as the yeast TUP1 and CDC4 proteins. A recent report showed that the same gene, H12.3, had been isolated from a screen for cDNA products that could interact with Rb (W-H Lee, unpublished).

The sequences of the other three cDNA clones were also compared to sequences present in the data bank. The only homology to be seen was to D2. This cDNA was homologous to a mouse-expressed cDNA, the function of which is unknown.

4.3.4 Analysis of E4-dependent clones *In vitro.*

Assessing the relevance of the interaction of each cDNA with E4 would be aided by demonstrating that these interactions can occur in a different system from yeast. Initially, *in vitro* methods were attempted. E4 protein and the cDNAs were co-translated in reticulocyte lysate in an attempt to show an interaction between them by co-immunoprecipitation or gel retardation. Either E4 or the cDNA was expressed as a fusion with the myc epitope 9E10, to which an antibody is available (G. Evan).

E4ORF6/7 was transferred into the vectors T7 and T7Tag. T7Tag will express the 9E10 epitope at the N-terminus of the translated protein. cDNA clones D2 and 29 were cloned into the T7 and T7Tag vectors as VP16 fusions, and clones D4 and 49 were cloned in via their Eco R1 sites ensuring that the correct open reading frame was expressed. (Clones 4 and 49 have an internal Bam H1 site so cannot be excised as VP16 fusions). An example is shown in Fig4.8 of a co-translation of 35S methionine-labelled E4 and D2 with either one or both tagged with the myc epitope (lanes1 to 4). The co-translated proteins were incubated in an immunoprecipitation reaction, unbound protein removed from the reaction by washing, and the bound products visualised on an SDS polyacrylamide gel (lanes5 to 8). Tagged E4 or D2 were always brought down by the antibody, but neither tagged protein was able to interact with and therefore bring down specifically the untagged protein in the reaction (lanes 5 and 7).

Using these reticulocyte lysate translated proteins, the ability of E4 to interact with the cDNA products was also tested using a gel retardation assay on E2F sites. If the translated cDNAs interacted with E4, they might be expected to perturb the complex seen with translated E4 protein and endogenous reticuloctye lystate E2F binding activity. However, no difference in the complex of E4 with E2F was seen in the presence of the translated cDNA clones. The complex remained as in fig4.3 (lane4). One problem in looking for such interactions with reticulocyte lysate-expressed proteins is the presence of endogenous E2F binding activity in the lysate. This may block the interaction of E4 with the exogenously-expressed cDNA clones.

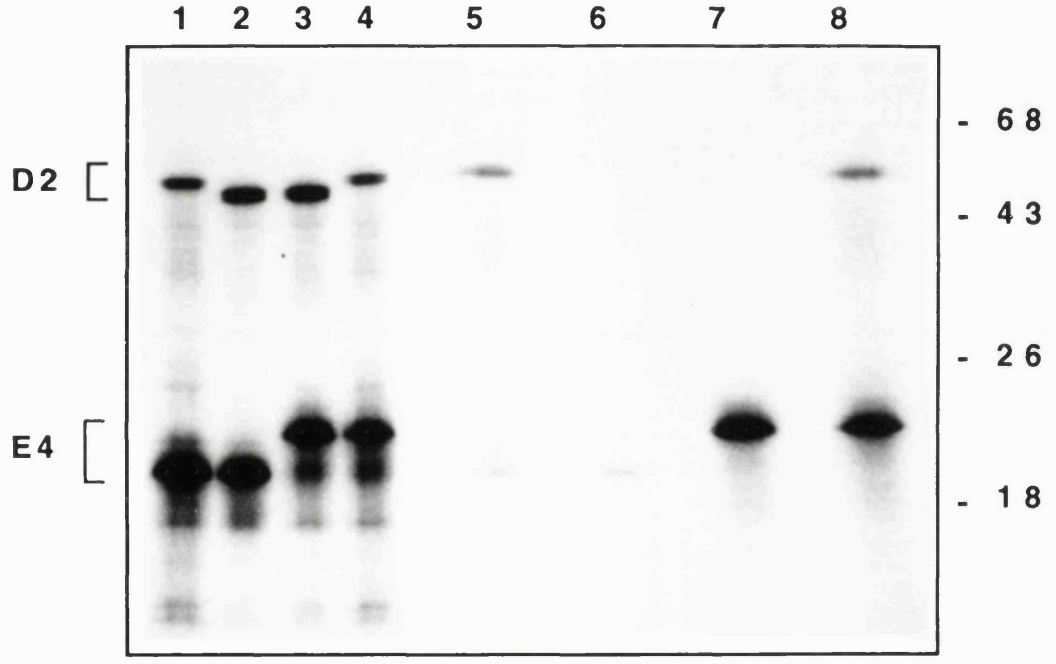
4.3.5 Analysis of E4-dependent clones In mammalian cells.

Clones D2 and 29 were transferred as VP16 fusions into the mammalian expression vector MLV. Transient transfection assays in NIH3T3 cells were performed to see if any of the cDNA products formed part of the E2F complex binding to the E2A

Fig 4.8 Co-immunoprecipitation reactions of in vitro translated E4 and D2 proteins.

Lanes 1-4: E4 and D2 were cotranslated in reticulocyte lysate and 5µl run on a 10% SDS polyacrylamide gel. The slower migrating form of each protein is a fusion to the myc 9E10 epitope, ie lanes 1 and 4 for D2 and lanes 3 and 4 for E4.
Lanes 5-8: An immunoprecipitation reaction was performed with the co-translated products of lanes 1-4, and the products bound to agarose beads were eluted by boiling in SDS sample buffer and resolved on lanes 5-8, respectively.

Co-immunoprecipitation reactions
of in vitro translated E4 and D2



promoter. As a reporter, part of the E2A promoter containing the two E2F sites fused upstream to the CAT gene was used. The reporter E2A(-97) was clearly inducible when co-transfected with a plasmid expressing the E1a protein, but the VP16/cDNA clones did not affect its activity (Fig4.9).

Another strategy in mammalian cells was to try and mimic the interaction seen in yeast of E4 with the cDNA clones. A fusion of E4ORF6/7 was made to the DNA binding domain of GAL4. This was co-transfected with a CAT reporter bearing 5 GAL4 binding sites upstream of the TATA box (G5E4CAT). A high transcriptional activity of the GAL/E4 construct was seen on the reporter, compared to the GAL4 DNA binding region alone. Typical activation activity of E4 would be about 20-fold above basal levels. No additional activation was seen when any of the VP16/cDNAs were transfected with the GAL4/E4 fusion. The high basal activity of GAL4/E4 was unexpected, and it is possible it may have masked activation by the cDNA clones.

4.4 Analysis of E4-Independent clones.

cDNA clones D25, 26 and 39 appeared to require the presence of SRF to exert their blue phenotype, but not the presence of E4. Clones 26 and 39 were specific to the presence of SRF. They activated transcription when cotransformed with SRF, or any fusion of SRF, but not with any other DNA binding domains, such as that of LexA. Partial sequence analysis from the N-terminus of D26 and 39 using a primer to the VP16 part of the fusion showed that they both encoded SAP1.

Clone D25 was not specific to the presence of SRF. It also activated the reporter with E2F sites in the Y4-1 strain, and the LexA operator with LexA bound. It did not activate the SRE alone, nor the mutant E2F site, nor the LexA operator without the presence of LexA. These results indicated that the clone could activate the reporter directly, presumably by directly binding to it, but only if another protein (whether of yeast, bacterial or mammalian origin) was also bound. A possible explanation is that this protein was necessary to perturb the chromatin structure in order to allow D25 to bind. Sequence analysis of Clone D25 (fig4.10) showed that it was almost entirely identical to a gene product termed fos-interacting protein (FIP), which was isolated from a screen for proteins that interacted with the leucine zipper of fos (Blanar and Rutter, 1992). FIP is a member of the basic, helix-loop-helix, leucine zipper (bHLH-ZIP) family of transcription factors that includes USF and cMyc. At the amino acid level it has almost 100% identity to USF in the basic-HLH domain, and 53% identity overall. The sequence of D25 terminates at residue 235 of FIP, and is missing the last leucine of the leucine zipper. The published sequence of FIP is 254 amino acids in length, but the N-terminus is not complete since the sequence does not start at a methionine. The sequence of D25 diverges completely from that of FIP at the N-terminus. It has 70bp of unique sequence still within the same open reading frame. It is interesting that at the junction where the unique D25 sequence becomes identical to the FIP sequence, the nucleotides A-G are

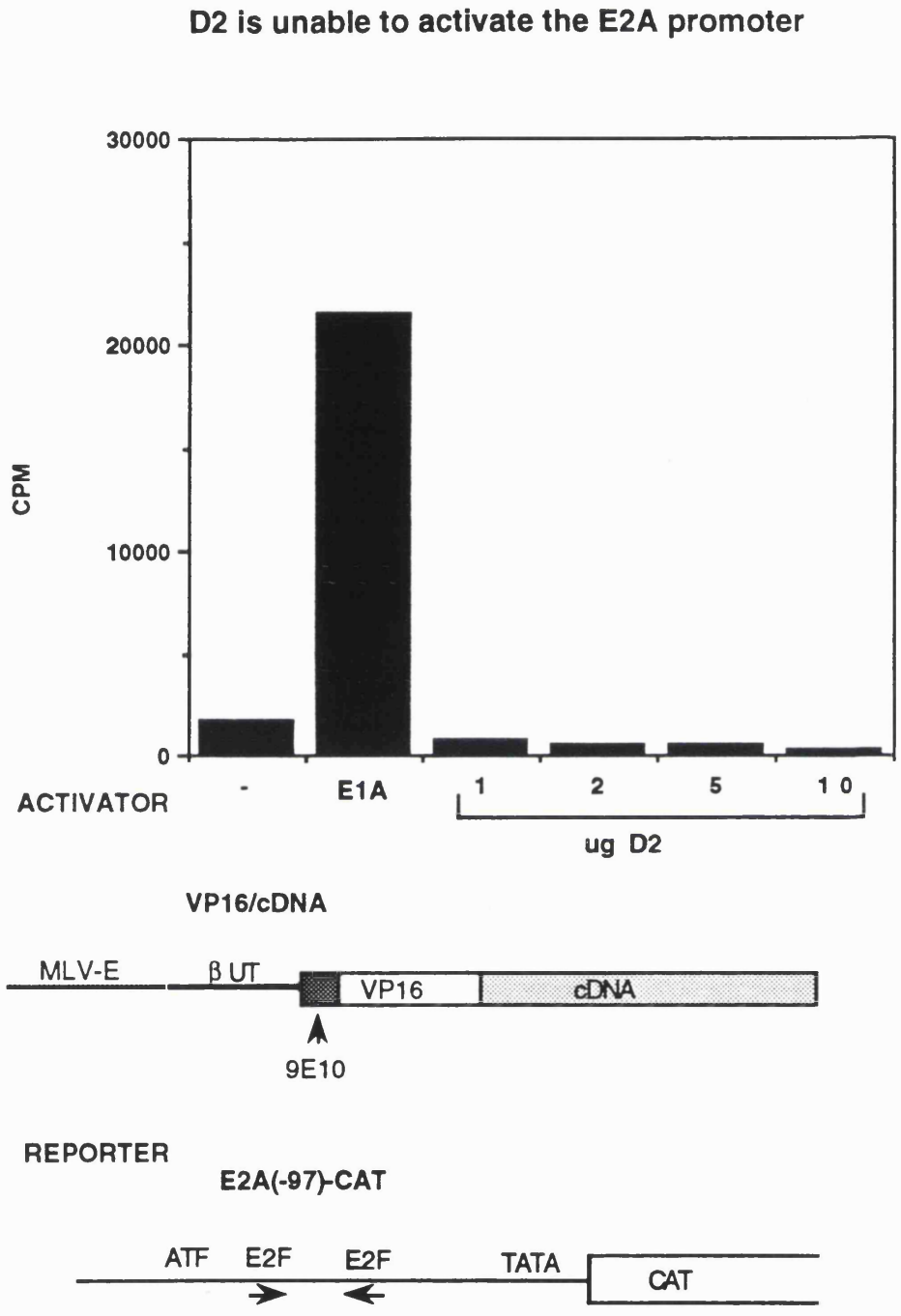
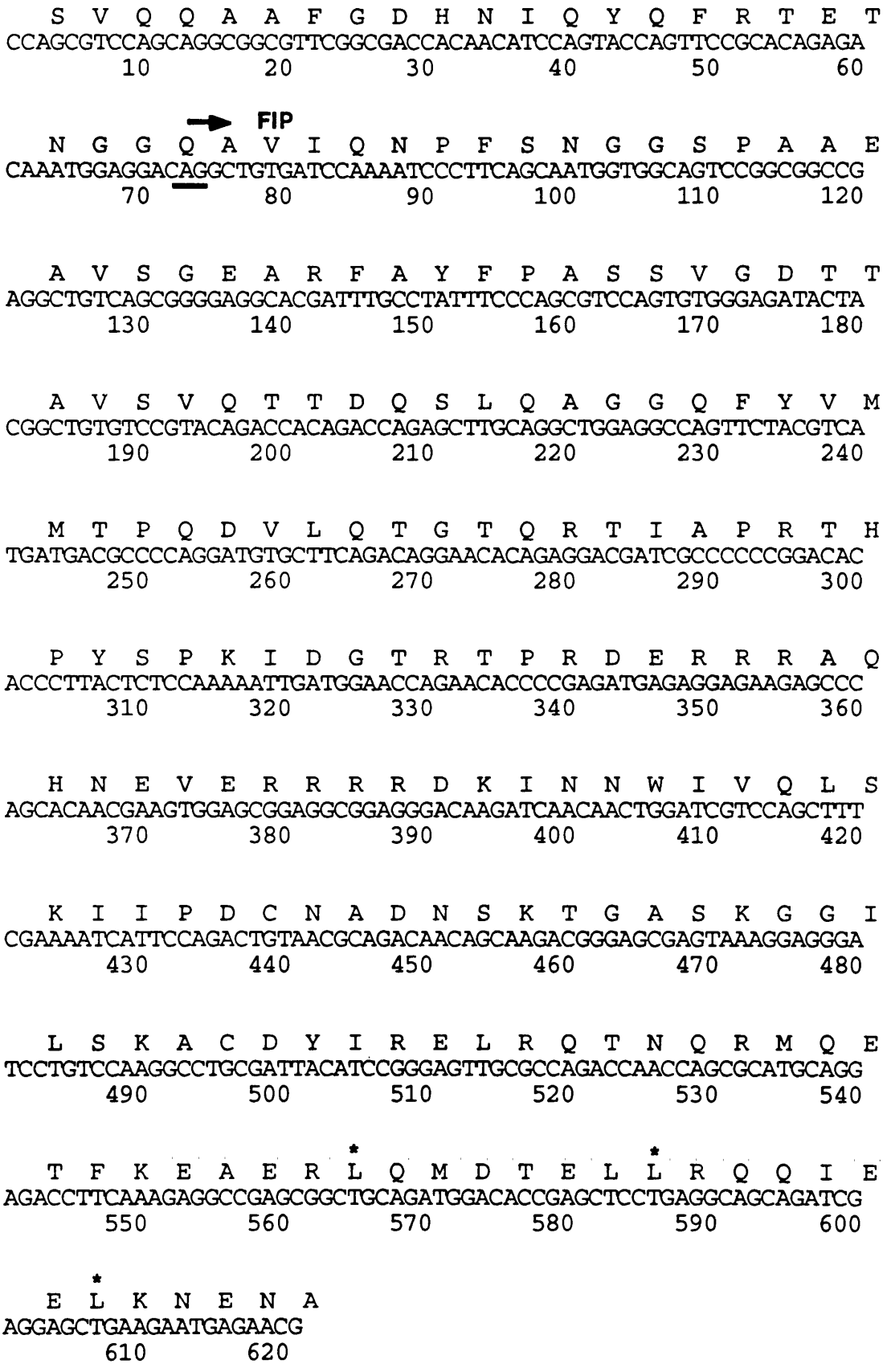


Fig4.9
NIH3T3 cells were transfected with 2μg of the E2A-CAT reporter plasmid and either 2μg of a plasmid (pCE) expressing the E1a protein, or increasing amounts of the D2 clone in an MLV expression vector. Diagrams are shown of the MLV expression vector containing the VP16/cDNA and the CAT reporter construct.

Fig 4.10 Complete DNA sequence of cDNA D25.

Sequencing is from the 5' end of the cDNA using a primer to the VP16 part of the vector. The predicted amino acid sequence in one open reading frame is shown. An arrow marks where the open reading frame becomes identical to that of FIP. * indicates the leucine residues that form a putative leucine zipper domain. The AG sequence that might form a splice acceptor site is underlined.

Fig 4.10 Complete DNA sequence of cDNA D25.



found. It is possible that they could indicate an intron-exon boundary. The divergent sequence at the N-terminus of D25 could be an alternatively spliced mRNA of FIP.

Since D25 had such high degree of homology to USF, its activity was tested on a USF binding site in mammalian cells. The D25 cDNA was transferred to the mammalian expression vector MLV as a VP16 fusion. In transfection experiments in NIH3T3 cells with a reporter consisting of the Ad major late promoter (MLP) fused to the CAT gene, MLV25 gave approximately 3-fold activation. It did not alter the activity of reporters without a USF site such as the E4 promoter (fig4.11). Activation of the MLP was not seen with the constructs MLV2 or 29.

It is possible that D25 may have been isolated from the yeast screen by its ability to bind to a USF-like element in the reporter. Within the promoter sequence of the CYC1-based reporter, there is the sequence CACATG that could form a binding site for the USF-myc family of transcription factors. It fits the consensus sequence for this family of CANNTG, and is similar to the E box sequence CACGTG.

4.5 Analysis of RBP3 properties

There have been two recent reports of a cDNA encoding a protein that can directly bind to Rb and that has E2F-like properties (Helin et al., 1992; Kaelin et al., 1992). The same cDNA (RBP3 or RBAP-1) was isolated by both groups by its ability to interact with Rb. Experiments in these reports indicated that the cDNA could activate transcription either from the E2A promoter, dependent upon the E2F sites, or as a GAL4 fusion. An interaction was seen with the adenovirus E4 protein, but whether this was direct or indirect was not clear from the nature of the experiment.

The availability of a cDNA for an E2F DNA binding activity enabled comparisons of its activity to be made with the products of the E4-interacting cDNA clones. These comparisons were carried out in both the yeast and mammalian systems that have already been used to characterise the E4-interacting clones.

4.5.1 RBP3 activity in yeast.

RBP3 was cloned by PCR into the SD06 vector to make a VP16 fusion (R. Fagan). It was tested for the following properties using lacz assays in yeast:

i) Interaction with the E2F site: VP16/RBP3 was transformed into the yeast strain Y4-1 and found to activate transcription. The activation was dependent upon the E2F sites because transformation into the yeast strain Y1-1, which contains an integrated reporter with mutated E2F sites, did not cause activation. (Fig4.12A).

ii) Interaction with E4: VP16/RBP3 was co-transformed with E4/SRF into the yeast strain S62L and found to activate transcription via the SRE. This was dependent upon the E4 part of the fusion since activation was not seen with SRF alone. The strength of the activation was the same level as that seen with clone D2. (Fig4.12B). In addition, VP16/RBP3 was tested for its ability to interact with the deletions of E4 fused to SRF. It

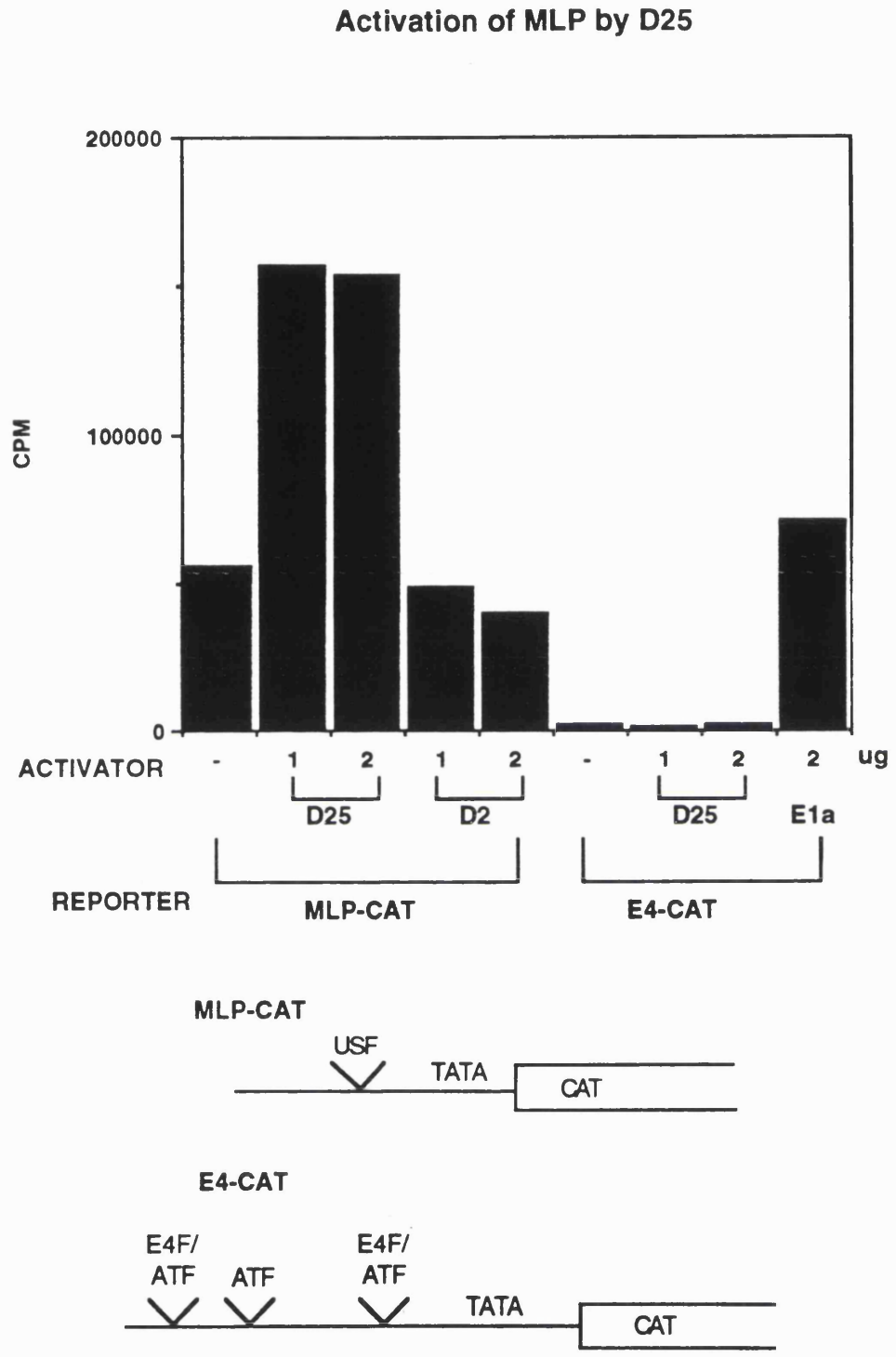


Fig 4.11
Constructs were transfected into NIH3T3 cells. 2μg of a CAT reporter plasmid consisting of either the Ad major late promoter (MLP) or the Ad E4 promoter was cotransfected with plasmids encoding activator proteins in the MLV expression vector, or E1a as indicated. Diagrams are shown of the CAT reporters.

Lacz activity of RBP3 in yeast assays.

	YEAST STRAIN	CONSRUCTS	β– GALACTOSIDASE ACTIVITY
A	Y4-1	VP16	3
		VP16/RBP3	18
	Y1-1	VP16	<1
		VP16/RBP3	<1
B	S62L	E4/SRF + VP16	3
		+ VP16/RBP3	21
		+ VP16/D2	23
C	Y699L	LexA/Rb + VP16	0.8
		+ VP16/RBP3	1.8
		+ VP16/E1a	20

D **VP16/RBP3**

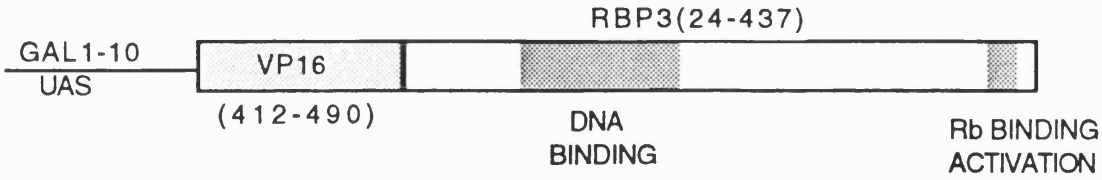


Fig4.12

Quantitative lacz assays were performed on yeast transformants as below.

A) VP16/RBP3 and VP16 alone were transformed into Y4-1 with wild type E2F sites and Y1-1 with mutated E2F sites.

B) The E4/SRF fusion was cotransformed with VP16 alone, or with VP16 fused to either RBP3 or the D2 clone.

C) The LexA/Rb fusion was cotransformed with VP16 alone, or with VP16 fused to either RBP3 or E1a.

D) Diagram of the VP16/RBP3 construct in a yeast expression vector. The positions of the regions of RBP3 involved in DNA binding and in Rb binding and transactivation are shown.

behaved identically to the four cDNA clones: it could still interact with the deletion of the C-terminal 10 amino acids, but failed to interact with the deletions of one or both putative helices.

iii) Interaction with Rb: VP16/RBP3 was cotransformed with LexA/Rb into the yeast strain Y35. A very low, but reproducible level of activation of transcription was seen, although the activation by VP16/E1a was much higher (fig4.12C).

The activity of RBP3 in yeast was therefore as expected from published results and the predicted properties of an E2F cDNA. It is able to interact with the E2F binding site, and with Rb (although weakly). The interaction of E4 with RBP3 can be compared to its interaction with the products of the cDNA clones obtained from the screen. RBP3 interacts with a similar region of E4 as the clones and with a similar affinity.

4.5.2 RBP3 activity in mammalian cells.

The VP16/RBP3 fusion in pSD06 was transferred to the MLV expression vector, and transient transfections performed in the same way as with the VP16/cDNA clone fusions.

It was found that VP16/RBP3 could activate the E2A promoter (E2A(-97)CAT). Activation seemed to be specific for the E2F sites since when a larger deletion of the promoter was used (E2A-59) which lacks the upstream E2F site, activation via RBP3 was lost. (Fig4.13). This would be expected from the published data, and served as a control for the negative results obtained with the cDNA clones.

The MLV-VP16/RBP3 construct was co-transfected with the GAL4/E4 fusion and the reporter with GAL4 binding sites (G5E4CAT). Despite the high basal level of activation given by the GAL4/E4 fusion alone, VP16/RBP3 could activate this fusion strongly (Fig4.14). The activation was in an E4-dependent fashion because the GAL4 DNA binding domain alone did not mediate such an effect. This result therefore indicated that RBP3 could interact with E4 and also confirmed the integrity of the GAL4/E4 fusion. Taken in conjunction with the ability of E4 and RBP3 to interact in yeast, it is a likely conclusion that the interaction is a direct one.

In the yeast assays, RBP3 had behaved in an identical way to clone D2 in its ability to interact with E4. Since RBP3 could also interact with E4 in mammalian cells, an explanation is required for why the D2 cDNA product was unable to behave similarly in mammalian cells.

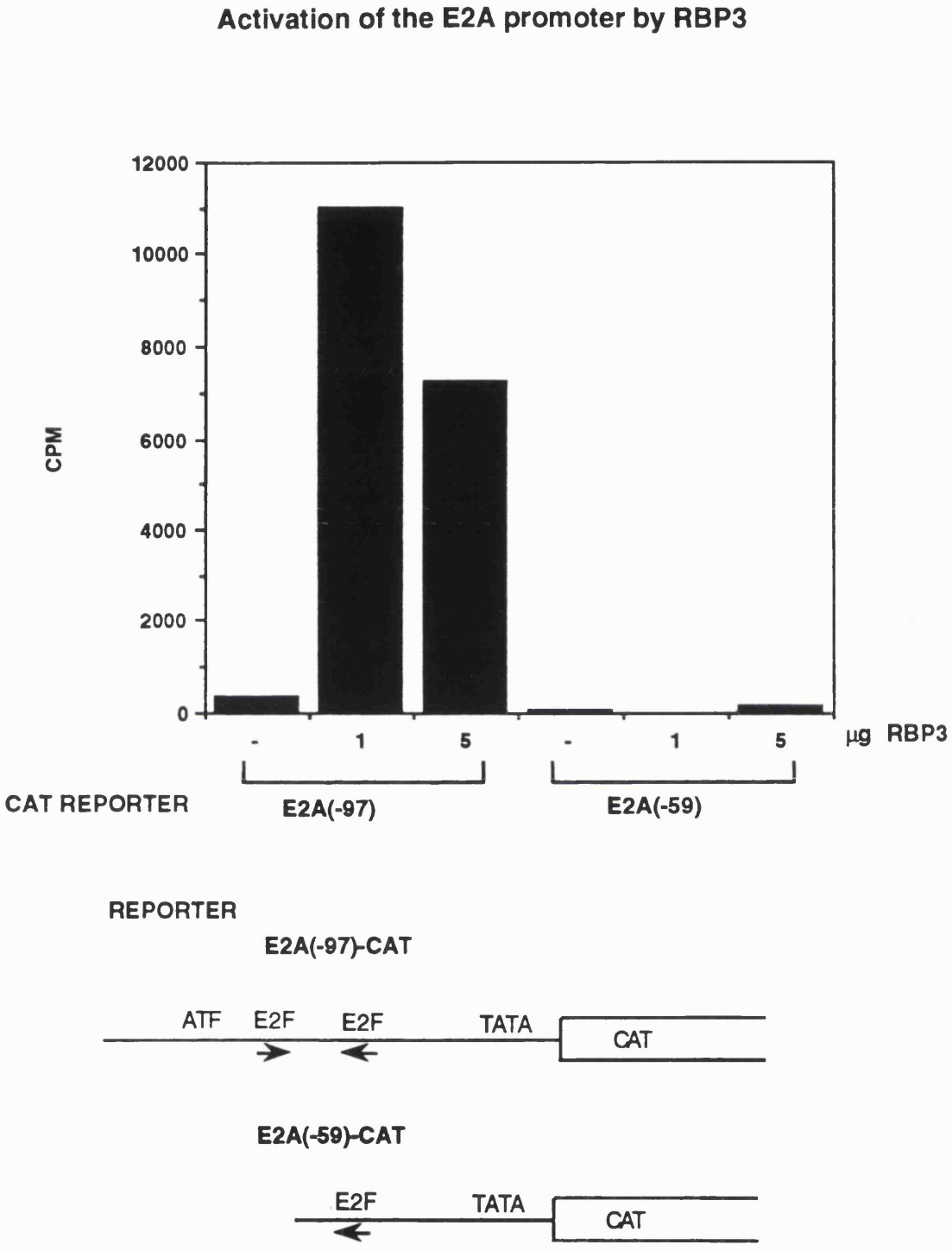


Fig4.13

NIH3T3 cells were transfected with 2µg of the reporter plasmids E2A-CAT as indicated, and 0, 1, or 5µg of the MLV vector expressing RBP3. Diagrams are shown of the CAT reporters.

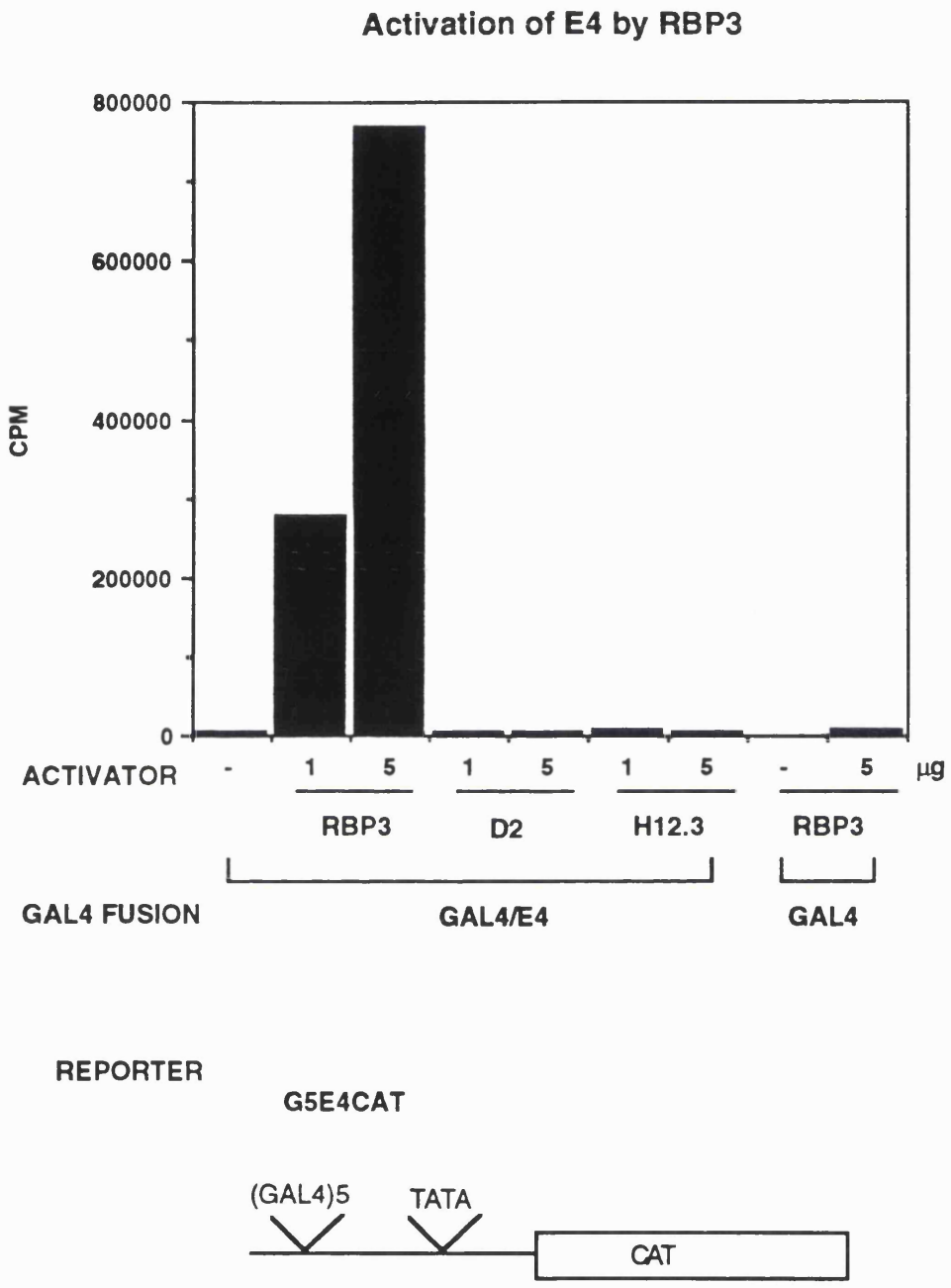


Fig4.14

NIH3T3 cells were transfected with 2µg of the G5E4CAT reporter plasmid, 5µg of the GAL4 fusions, and the indicated amounts of the VP16 activators in the MLV expression vector. A diagram is shown of the CAT reporter.

CHAPTER 5: DISCUSSION

5.1 E1a activation in yeast

E1a is a protein that can stimulate the transcription of a wide variety of viral and cellular genes. The mechanism whereby this stimulation occurs is thought to involve the interaction of E1a with the cellular complexes on the promoter. Considering the conservation in the transcription process between yeast and mammalian cells, a yeast system was used to investigate E1a transactivation further. E1a was expressed in the yeast *S. cerevisiae*, and its effect measured on reporters with different upstream binding sites. The upstream binding sites chosen were those that could mediate E1a transactivation in mammalian cells and that in yeast had endogenous activity due to factors binding to them. The yeast factors bind DNA with a similar specificity to the corresponding mammalian factors, and therefore it was possible that the regulatory potential of these sites could also be conserved. However, the expression of E1a in yeast did not alter the transcriptional activity of reporters containing either ATF, AP1 or E2F sites. This was in agreement with previous results in S. Pombe where it had been shown that the expression of E1a had no effect on transcription driven by the E2A promoter (Jones, 1990b). If E1a is assumed to be acting physically at the promoter during activation, then reasons for its inability to activate in yeast might include a failure to interact with and stimulate the yeast basal transcription machinery, or a failure to be targeted to a promoter via an interaction with yeast upstream transcription factors.

5.1.1 Interaction of E1a with basal transcription factors

When these experiments were begun, it was thought that the activation domain of E1a located in CR3 might be able to stimulate transcription in yeast. It was assumed to behave in a manner analogous to viral activation domains rich in acidic residues, such as those of VP16 or E2, both of which are potent activators in yeast. This assumption was based on the finding that a deletion of the N-terminal 10 amino acids of CR3 destroyed its ability to activate a promoter. This region contained four acidic amino acids and activation was restored when it was replaced with the VP16 activation domain (Lillie and Green, 1989). However, I found that in yeast, when the E1a activation domain was fused to a heterologous DNA binding domain, unlike VP16, it did not function as a strong activation domain. This indicated that the E1a activation domain may not be of the acidic type, since such activation domains can efficiently stimulate transcription in yeast.

Experiments from other laboratories have since confirmed that the CR3 activation domain of E1a does not behave as typical acidic one would. From the deletion studies of Lillie and Green (1989), it was found that CR3 is made up of two discrete subdomains, which have different targets. The N-terminal subdomain was thought to function as an activating region that might interact with the general transcription machinery. This subdomain contains the acidic residues mentioned above and a zinc finger. The C-terminal subdomain of CR3 was thought to function in directing E1a to the

promoter, perhaps by interacting with a DNA-bound transcription factor such as CREB2. To investigate the important features of CR3, mutational analysis was carried out in more detail than the initial deletions studies. This showed that CR3 is highly structured and most of the conserved amino acids contribute to this structure, whatever their charge. The four negatively charged residues in the N-terminal subdomain of CR3 were mutated, and although one of these acidic residues was critical, net negative charge was not important. In contrast, the importance of the zinc binding region was shown, since mutation of any of the cysteine residues of the zinc-binding finger to alanines, eliminated activity. When the C-terminal subdomain is mutated or deleted beyond residue 180, the product can no longer target E1a to a promoter containing ATF sites. It is also capable of repressing activation by the wild type protein. The N-terminal region must be binding a limiting cellular factor, presumably via the zinc finger. This factor could be an adaptor or a component of the basal transcription machinery. See fig 5.1. (Martin et al., 1990; Webster and Ricciardi, 1991).

The interaction of E1a with specific components of the basal transcription machinery has been investigated in a number of laboratories. There is recent evidence that E1a can directly interact with the essential factor TFIID (specifically, TBP). Horikoshi et al (1991) showed that *in vitro* translated hTBP was efficiently retained on an E1a affinity column, and yTBP only weakly. This could offer one explanation for the weak strength of the E1a activation domain seen in yeast. Their result is in contrast to the behaviour of VP16, which has been shown to bind to both human and yeast TBP *in vitro* (Stringer et al., 1990).

Lee et al (1991) used immunoprecipitations to show that the 13S E1a protein could interact with hTBP to a greater extent than the 12S version. They further showed that the interaction was mediated by CR3 of E1a by using GAL4 fusions. As shown by Lillie and Green (1989), the CR3 activation domain of E1a fused to GAL4 is a potent activator from a promoter with GAL4 binding sites. This fusion could bind to TBP *in vitro*, and two point mutations decreasing transactivation of E1a also diminished TBP binding. However, another point mutation of E1a that decreases transactivation still bound TBP. There must be criteria in addition to the ability to bind TBP for E1a activation. For example, E1a may have multiple targets and may interact with other basal factors besides TBP, such as TFIIB. Mutations of the cysteine residues of the zinc binding domain that severely impair E1a activation were not tested in these assays. The site on TBP that is involved in these interactions was mapped to a region in the conserved C-terminal domain that includes a repeat of basic residues, between the homologous direct repeats. Although yTBP cannot interact so strongly, the C-terminal domain of hTBP shares 80% homology with yTBP. The acidic activation domain of VP16 has also been shown to interact with TBP *in vitro*. Since E1a is not thought to have an acidic activation domain, this indicates that there may be additional targets involved that define the specificity of an activation domain. Further evidence for such additional targets is provided by *in vitro* systems where VP16 has been shown to interact with TFIIB (Lin et

Model for interactions of E1a CR3

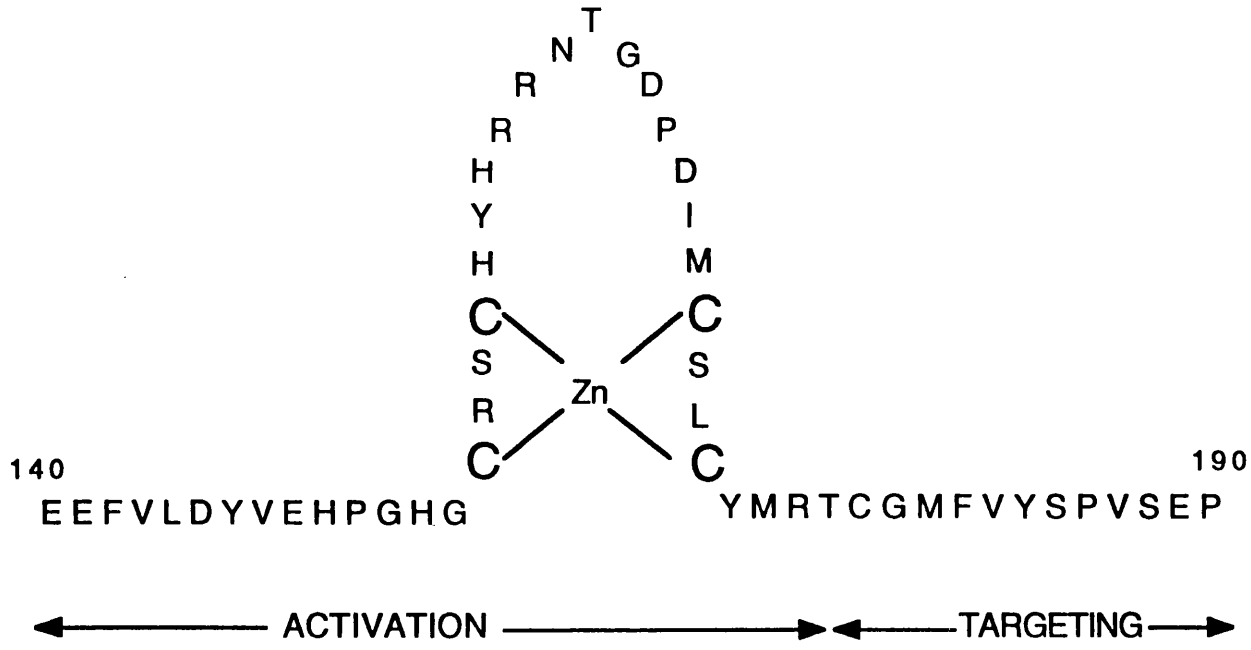


Fig5.1

CR3 of E1a is shown and the zinc-binding region indicated. The N-terminal activation region may interact with a component of the basic transcripional machinery, such as TFIID. The C-terminal targeting region may interact with a DNA-bound transcription factor. (Webster and Ricciardi, 1991).

al, 1991). Different activation domains may also interact with different motifs in TBP. At present all activator-TBP interactions have mapped to the conserved C-terminal domain of TBP. Additional targets for the activation domains may also be provided by the TAFs or adaptor proteins that are not be present in the *in vitro* systems used. They may be important in determining the specificity of an activation domain. Another drawback of these experimental approaches is that they rely on *in vitro* assays that involve overproducing the two proteins. The significance of these interactions to the *in vivo* situation remains to be shown.

Martin et al (1990) also addressed the identity of the target of the N-terminal region of CR3. In contrast to the *in vitro* binding results described above, they indicated that the activation domain of E1a may see a different target from the acidic activation domain of VP16. On a reporter with GAL4 binding sites, they found that expression of E1a inhibited transcription driven by GAL4/E1a, but not GAL4/VP16. It is interesting that VP16 does inhibit activation by GAL4/E1a: VP16 could be interacting with a factor essential for transcription, and E1a with an auxillary factor not required for activation by VP16 (or one upstream of the VP16-interacting factor).

The two sets of experiments described above are somewhat contradictory in their conclusions. However, they do offer some explanation for the weak activation ability of E1a CR3 in yeast. The *in vitro* binding experiments show that E1a can directly interact with hTFIID, and to a greater extent than with yTFIID. The transcriptional interference assays show that E1a has a different type of activation domain from VP16. This agrees with the evidence from the mutational analysis of E1a CR3, where (unlike VP16) the acidic residues are not the essential feature for its ability to activate. An acidic acivation domain is often of importance for the stimulation of transcription in yeast. The transcriptional activity of E1a may require specific TAFs or adaptor proteins that are not conserved in yeast.

5.1.2 Interaction of E1a with upstream factors

Once it was realised that the activation domain of E1a could not function efficiently in yeast, E1a was fused to the VP16 activation domain. This enabled me to study the question of whether the C-terminal subdomain of E1a CR3 interacts with upstream factors, the VP16 domain allowing any such interaction to activate transcription. Experiments involved seeing if E1a could interact with endogenous yeast factors and if an interaction between E1a and the mammalian factor CREB2 could be shown in yeast.

As discussed previously, several yeast transcription factors have DNA binding specificities that appear to be identical to corresponding mammalian factors. In addition, some aspects of regulation via these binding sites may be conserved. For example, protein kinase A can stimulate activation via the ATF/CREB site in mammalian cells and can also stimulate the binding of factors to this site in yeast. This lead to the possibility that E1a might modify the activity of yeast factors using the same mechanism as in

mammalian cells. However, this was not found to be the case. VP16/E1a was not seen to activate the reporters containing ATF, AP1 or E2F sites. There are several possible explanations for this negative result:

i) The level of activity of these reporters resulting from the binding of endogenous yeast factors may have been masking any activation by E1a. Results in Chapter 3 would indicate that this is unlikely, since three cDNA clones were isolated which could activate the ATF site in the wild type yeast strain.

ii) E1a may be incorrectly modified in yeast cells. Phosphorylation of E1a does occur in yeast, although E1a expressed in yeast has a different mobility on a denaturing gel from that of E1a from mammalian cells, indicating that the pattern of modification may be different. However, even if post translational modifications are different in yeast, such modifications have not been shown to be important for the transactivation ability of E1a.

iii) Most of the available evidence indicates that E1a acts at the promoter by physically interacting with factors bound there. However, it is possible that E1a may be activating mammalian reporters through an indirect mechanism that increases the activity of promoter-bound factors, for example by activating a kinase.

iv) Yeast factors may not have the conserved domains to allow a direct interaction with E1a. This is probably the most likely explanation for the negative result obtained.

Many aspects of the regulation of transcription have been conserved between yeast and humans. Despite this, it is becoming clear that although transcription factors may be conserved at the level of their DNA binding specificity, and therefore share homologous DNA binding domains, this does not necessarily mean that the mechanisms of regulation of these factors will also be conserved. It seems that in some cases, the interactions between transcription factors and signal transduction pathways have diverged in evolution between different species, although the fundamental aspects of transcription may be conserved. Therefore it may not be surprising that yeast and mammalian factors do not share the ability to interact with E1a.

The next set of experiments addressed whether E1a could directly interact with mammalian upstream factors. Precedents for such interactions have been shown in studies with the transcription factors CBF and Oct4. CBF binds to the CCAAT element of the human hsp70 gene promoter. While no one mutation of this promoter abolishes its activation by E1a, the CCAAT element is important, and a GAL4/CBF fusion can mediate activation by E1a. In vitro binding results show that bacterially produced CBF can bind to 13S E1a expressed in reticulocyte lysates more efficiently than 12S E1a (Lum et al., 1992). E1a and Oct4 can stimulate activation synergistically from an octamer element; both CR3 of E1a and the acidic activation domain of Oct4 are required for this. A labile complex between E1a and Oct4 can be detected *in vitro* although it is possible that an additional cellular factor may be required to stabilise it (Scholer et al., 1991).

Other viral proteins have also been shown to directly interact with mammalian transcription factors in order to activate transcription. The HSV VP16 complex with Oct1 has already been mentioned. An additional example is the X protein of the hepatitis B virus (HBV), which stimulates transcription from the HBV enhancer. This contains several overlapping binding sites for transcription factors, including a CRE-like element. CREB2 alone has a weak affinity for this element, but the X protein seems to stabilise its binding to DNA to generate a complex containing both proteins, as seen by a gel retardation assay. Labeled CREB2 will interact with X protein on a protein blot (Maguire et al., 1991).

The possible interaction between E1a and CREB2 has been defined in mammalian cells (Liu and Green, 1990). That an interaction occurs would be inferred from the transient transfection experiment where a fusion of GAL4 to E1a can be activated by CREB2 fused to VP16. The GAL4/E1a fusion was transcriptionally silent due to a mutation in the N-terminal subdomain of CR3 (see Introduction). If an interaction did not occur, then CREB2/VP16 would not be brought to the promoter to activate transcription. However, this experiment does not distinguish whether such an interaction is direct or mediated by another cellular protein. No interaction has been demonstrated clearly *in vitro*. E1a does not alter the gel retardation complex of CREB2 with DNA, and other assays such as the binding of one purified protein to the other immobilised on columns or blots have not yet yielded a clear answer of whether an interaction is occuring.

If the interaction is direct, then in the yeast system, by tethering one protein to the promoter via a binding domain and fusing the other to an activation domain, activation of the reporter should ensue. This was not the case in the experiments described in Chapter 2, so it is likely from this information that an intermediary factor, or one that stabilises the interaction is required. Alternatively, modifications of the two factors that do not occur in yeast may be important, so the results obtained from other systems must also be incorporated into any model. The yeast system would give an opportunity to directly clone an intermediary or stabilising factor if it was concluded that one is required. A mammalian cDNA library in a yeast vector could be introduced into the strain harbouring the CREB2 and E1a hybrid proteins, and colonies screened for elevated lacZ levels.

Such an intermediary would presumably be of importance in the transcriptional activity of CREB2 or other factors in an uninfected cell. It is likely to have the characteristics of the transcriptional adaptors that have so far only been defined biochemically in the mammalian cell. Therefore, identifying the direct target of E1a by means of a yeast genetic screen could lead to the isolation of a clone belonging to this new class of mammalian factors.

5.2 Screen for a yATF cDNA

There is a DNA binding activity in yeast distinct from that of AP1 and GCN4, which appears to have identical specificity to the mammalian ATF/CREB transcription factor family. Evidence for this involves gel retardation data and lacz assays comparing the ability of wild type and mutant binding sites to act as UAS elements (Jones and Jones, 1989). The aim of the screen in Chapter 3 was to isolate a *S. cerevisiae* cDNA for such an activity. Like mammalian ATF, the yeast ATF binding activity may have an important role in gene regulation. A comparison between the yeast and mammalian factors may also elucidate any conserved features of their regulation. Previous attempts to isolate yATF had included a binding site screen with a λgt11 library, and purification approaches. These were unsuccessful. The screen I adopted made use of a yeast cDNA library fused to the VP16 activation domain, which was transformed into a yeast strain with a reporter driven by ATF sites.

From this screen, I isolated three different cDNA clones as fusions to VP16, the products of which all shared the following property. They were highly specific in their activation of the ATF site and inability to activate either mutant ATF sites or any other site tested, including an AP1 site. However, it is unlikely that I succeeded in cloning yATF itself. This is based upon the finding that none of the proteins encoded by the cDNAs could be shown to be physically binding to the ATF site in yeast, or in any other system tested. One possible explanation for this could be that they are part of the complex that binds to the ATF site, but do not bind DNA directly themselves. For example they could be an ATF-specific adaptor protein. A weak interaction with the promoter might not be seen in the biochemical assays used. Alternatively they could be acting upstream of the ATF transcription factor in order to activate it in an enzymatic way, and therefore not be part of the actual ATF binding complex. Another explanation could be that the VP16/cDNA fusion product activates a gene which then influences yATF. It would therefore be informative to see if the VP16 portion of the cDNA fusion was required for activation of the ATF site; if it was then this might be evidence that the cDNA product was acting at the promoter rather than enzymatically. If VP16 was not required, then either the cDNA is not acting at the promoter, or else it could be acting at the promoter but contains its own activation domain.

The cDNAs obtained from this screen may define regulatory genes of importance in activation via the ATF site. The best method to determine if they have an important role in transcription via this site, is to delete the genomic copy of each of them and observe whether the deletion is lethal, and if not, its affect on the activity of the ATF site. Overexpression of the cDNA products as non-VP16 fusions on multicopy plasmids may also be instructive. Such experiments are being carried out by a post-doctoral fellow in the laboratory.

Recent evidence suggests that multiple proteins can bind to the ATF site in yeast, both activators and repressors. Two laboratories recently isolated a gene, ACR1 (or SKO1), that binds directly to the ATF site in yeast and can repress transcription (Nehlin et al., 1992; Vincent and Struhl, 1992). The gene has a basic, leucine zipper DNA binding domain that has 52% sequence identity to the mammalian factors CREB and CREM within the basic region. (CREM is also a transcriptional repressor: see Foulkes et al., 1991). Extracts from yeast strains in which the ACR1 gene has been deleted still contain ATF-like DNA binding activities. Also, increased activation through the ATF sites in these deleted strains occurs due to the absence of the repressor. These results indicate that other unidentified ATF binding proteins that can function as activators must be present in yeast.

5.3 E2F binding site screen

I showed that the E2F sites from the E2A promoter were active in *S. cerevisiae*, and that the levels of activity of the various wild type and mutant E2F sites seemed to mimic that seen in mammalian cells. However, the nature of the yeast activity binding to the E2A promoter remains under speculation. At present it is not known whether this is a specific yeast equivalent to mammalian E2F. Two DNA elements with homologies to the E2F site have been described in yeast that are bound by different transcription factor complexes (see fig5.2). The SCB element (SWI4/6 cell cycle box) is found upstream of the HO gene involved in mating type switching, and also upstream of the CLN1 and two genes which encode G1 cyclins. The SCB can be bound by a complex containing the SWI6 and SWI4 proteins. The MCB element (Mlu1 cell cycle box) is found upstream of DNA synthesis genes and is recognised by a complex that includes SWI6 and an unidentified 120kd protein. (Dirick et al., 1992). It is possible that one of these complexes is responsible for the E2F-specific binding activity. Alternatively, there could be a separate binding activity that is more specific for the E2F site, or one that is unrelated to either the E2F or the SWI6 complexes.

The activity of the E2F sites in yeast are not activated by adenovirus proteins as they are in mammalian cells. The VP16/E1a construct did not alter the activity of the lacz reporter driven by E2F sites. This reporter was also not activated by the adenovirus E4 protein fused to VP16, although when translated *in vitro*, this VP16/E4 construct can interact with the mammalian E2F activity as assayed by gel retardation.

The E2F sites from the E2A promoter placed upstream of the lacz gene were used as a reporter in the screen that was designed to isolate a mammalian cDNA binding to these sites. No cDNA clones whose products activated via the E2F sites were isolated from this screen. Reasons can be postulated why this was the case.

i) The yeast activity binding to the E2F sites could have resulted in a basal activity of the promoter that was too high, so that weaker positive clones would not be

Two binding sites in yeast have homologies to the E2F site

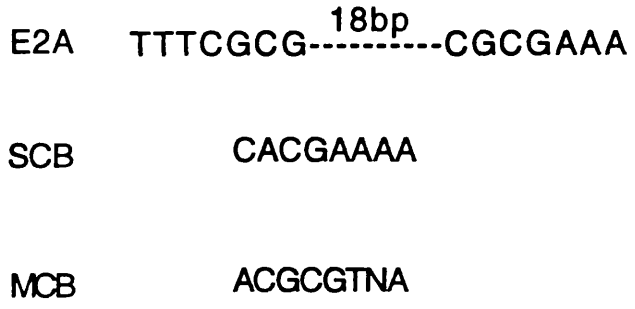


Fig5.2

The E2F sites from the adenovirus type 5 E2A promoter compared to two binding sites from S. cerevisiae, the SCB and the MCB.

seen. As well as masking the transcriptional activity of any mammalian clone that may bind, the yeast activity could sterically hinder such a mammalian factor from binding.

ii) Although E2F has been purified so that it has the appearance of a single polypeptide on a denaturing gel, the binding activity may comprise of a heterodimer where each momoner cannot bind DNA alone. There are examples of transcription factors where the formation of a homodimer or its binding to DNA is of a low affinity or non-existent. For example, the formation of cFos:cJun heterodimers is required for high affinity DNA binding. cJun and cFos homodimers are much less stable than the heterodimer, nor can either bind DNA as a monomer (O'Shea et al., 1989). This screen would rely upon E2F being able to bind to DNA as either a monomer or a homodimer. Most transcription factors so far described bind DNA as a dimer, although most can bind as a homodimer. The ATF family would exemplify this, where most factors can bind as homodimers and heterodimer formation is selective. There are also a few examples of transcription factors that can bind DNA as a monomer, such as Sp1. However, it is still possible that E2F binding could require the interaction of more than one polypeptide in order to form a DNA binding complex.

iii) E2F may require post-translational modification in order to bind DNA and this modification may not occur when the protein is expressed in yeast. The only known modification that may occur to E2F is phosphorylation. E2F binding activity can be seen in extracts from adenovirus-infected HeLa cells. Incubation of these extracts with phosphatase results in decreased binding to the E2F sites. Binding can be restored by incubation of the phosphatase-inactivated extract with cAMP-dependent protein kinase. (Bagchi et al., 1989). However, there is no evidence that this is an *in vivo* method of regulation. If such phosphorylation is necessary, then this should occur in yeast since it contains such a conserved cAMP-dependent kinase. Therefore, lack of phosphorylation is an unlikely reason for this screen to have failed.

iv) E2F binding activity is of very low abundance. There are an estimated 50 to 100 fold fewer active molecules of E2F in a virus infected cell compared to factors such as Sp1 or Ap1 (Yee et al., 1989). The abundance of the E2F mRNA may be similarly low, and therefore it may not be represented in the library.

A potential problem with the library for all the screens, is that it requires a fusion between VP16 and the cDNA in the correct reading frame and with the cDNA in the correct orientation (a 1 in 6 probability). If the required domain of a fused cDNA product is near the N-terminus, then there may be a lower probability of obtaining it. This is because the fusion may include the 5' untranslated region of the gene in which there are likely to be termination codons. It may be advantageous for all screens to use simultaneously a library containing a C-terminally fused activation domain. Although a C-terminal tag may not be expressed if the cDNA includes termination codons, the use of both libraries may increase the probability of obtaining the desired cDNA. Such a library was not available. As explained later, there is now some evidence that the failure of this screen was due to the complexity of the library.

5.4 Screen with the E4 protein

The second screening method for an E2F cDNA used the adenovirus E4 protein as a target for the VP16-tagged cDNA library, in the hope that requirements for E2F binding to E4 would be different from its requirements to bind DNA. In order for the strategy of the E2F binding site screen to succeed, it was necessary that the cDNA bound to DNA as a homodimer or monomer. This problem may still exist, however it is possible that a single molecule of E2F may be able to interact with E4. The modification requirements of either of the interacting proteins may also still be a problem with the E4 screen, but possibly the interaction of a viral protein with E2F is not so dependent upon the particular modified state of E2F. This second strategy may also depend upon isolating a different domain or polypeptide of E2F to the one with DNA binding ability, which may be represented better in the library than the DNA binding domain, since the library does not necessarily contain full length cDNAs. The problem still remained of the possible low copy number of an E2F cDNA in the library. However, as well as the potential of isolating the E2F binding activity with this screen, any cellular protein that interacts with the adenovirus E4 gene product could potentially be isolated, and may be of interest even if not related to E2F. cDNA clones were isolated whose products interacted specifically with the E4 protein in yeast. Other clones were also isolated whose products were not dependent on the E4 protein for their positive phenotype.

5.4.1 E4-specific clones

The main aims of subsequent experiments with the clones that were dependent upon E4 were to see if any of them bore any features that could be ascribed to an E2F-like activity, and also to investigate if their ability to interact with E4 was specific and could be seen in other systems. An initial analysis of these clones consisted of testing whether they interacted with the same region of E4 that is known to be required for E2F binding.

The interaction of E4 with E2F has been described in some detail (Neill and Nevins, 1991; O'Connor and Hearing, 1991). These results are based upon gel retardation experiments with HeLa cell extracts that contain E2F, and E4 expressed *in vitro*, and substantiated by transient transfection assays testing the activation of the E2A promoter by E4. Deletion and mutation analysis of E4 indicated that the region important for interaction with E2F and forming a stable complex on the E2A promoter involved the putative HLH domain and regions of the protein surrounding it. A critical region seemed to be the very C-terminal 10 amino acids of E4: a deletion of this region or an insertion to change the reading frame generated a protein that was not detected to interact with E2F. Transient transfection experiments showed that the same regions of E4 were required for its ability to transactivate the E2A promoter.

In the E4/SRF construct used for the screen, fusing SRF to the C-terminus of E4 did not prevent its interaction with E2F, as analysed by a gel retardation assay. However, the specificity of interaction of the four cDNA products with E4 was not as might be expected for an E2F-like clone. It was expected that the deletion of the C-terminal 10 amino acids of E4 would eliminate its ability to bind to the clones. However, this was not the case, and this deletion made no difference to the lacz signal obtained with any of the clones. The importance of the putative HLH domain of E4 in interacting with the cDNA products was shown by deletions of either the C-terminal helix, or both helices, which did abolish the interaction when assayed in yeast. These results show a discrepancy in the importance of the C-terminal 10 amino acids between the cDNA clones assayed in yeast and the published results with E2F. It seems that the cDNA products do not interact with exactly the same regions of E4 as E2F does, implying that they are not related to E2F. Alternatively, it could be speculated that the integrity of the HLH domain of E4 is the critical structural requirement for the interaction with E2F, and although the C-terminus is involved in this interaction, it may not be essential. The discrepancy in behaviour of the C-terminal deletion of 10 amino acids could result from the differences in the assay systems used. Deletion of the C-terminal 10 amino acids may interfere with the HLH domain but still allow an interaction to take place which can be seen in the yeast assay. Analysis of the regions of E4 required for interaction with E2F relied on gel retardation and transfection assays. The yeast assay could be less sensitive to small decreases in the strength of interaction of proteins with E4 than the other assays are. Another reason for the difference could be that the E4 deletions assayed in yeast are fused N-terminally to the SRF DNA binding domain. The presence of the extra SRF sequences at the C-terminus of E4 could eliminate the requirements for the C-terminal E4 residues. This is unlikely, since the E4D/SRF fusion translated *in vitro* does not seem to interact with the E2F DNA binding activity. However, recent studies suggest that for interaction with E2F in yeast, the C-terminal 10 amino acids of E4 are dispensible (see later).

Other E2F-like properties that the cDNA clones might be predicted to have were also investigated using the yeast system. They did not appear to interact with the E2F binding site in yeast, nor with the Rb protein. Both of these interactions might be an expected property of an E2F cDNA. However, as previously discussed, the domain interacting with E4 may not be the same one exhibiting these other properties of E2F, and the cDNAs are not complete clones. Also there may be more than one E2F polypeptide, each with different properties. The yeast assay for DNA binding has the same problems as stated above for the screen for an E2F clone by this method of detection. The assay for Rb binding could also be affected by incorrect modifications in yeast. Although Rb is able to bind to E1a in yeast, its interaction with E2F appears to be cell cycle dependent, and so the exact nature of its phosphorylation may be critical.

In vitro systems were used in an attempt to find another assay to analyse the interaction between the cDNA clones and E4. Neither co-immunoprecipitations or gel retardation experiments yielded any confirmation that the cDNA clones were interacting with E4. Such assays are generally more stringent than the *in vivo* yeast assay, so the interaction may not be strong enough to occur in these conditions.

Transient transfections in mammalian cells also did not yield any results to confirm that the cDNA clones were either part of the E2F complex, or interacting with E4. However, there is no evidence that the cDNA products were able to reach the nucleus of the transfected cells, so no firm conclusions can be drawn from the lack of activation.

5.4.2 Comparison of E4-specific clones with RBP3

RBP3 (or RBAP1 or Ap12) was recently isolated independently by three different laboratories and has properties that indicate it resembles E2F (Helin et al., 1992; Kaelin et al., 1992; Shan et al., 1992). Studies showed that it is a transcriptional activator, either on the E2A promoter via the E2F sites, or as a GAL4 fusion. It also activated transciption in yeast; its activation domain has 19 acidic residues which may form an amphipathic a-helix. RBP3 interacted directly with Rb, the interaction being competed by E1a peptides, and also interacted with E4 *in vitro*. To show the interaction with E4, cellular extracts were incubated with either GST-Rb or GST-E4. In both cases, a specific protein of 60kd was detected by probing a filter of electrophoresed, bound proteins with an anti-RBP3 antibody. This did not prove that the interaction between E4 and RBP3 was a direct one: other proteins present in the extract could have been required to mediate the interaction. The size of RBP3 protein approximates that predicted for E2F. RBP3 RNA was detected in a wide variety of cell and tissue types, indicating that expression patterns are similar to that of E2F. Antibodies raised against RBP3 were able to precipitate E2F binding activity. However, other experiments showed that RBP3 is perhaps not the only polypeptide to interact with the E2F site. Helin et al (1992) could not use their anti-RBP3 antibody to either supershift an E2F-DNA complex or to clear cell extracts of E2F binding activity. This could be the fault of the antibody, or indicate that polypeptides unrecognised by the antibody are present in the E2F complexes. Kaelin et al (1992) were able to supershift two out of four complexes containing E2F binding activity with their anti-RBAP1 antibody. One of these was thought to be uncomplexed E2F and the other, E2F complexed with Rb. Of the complexes that could not be supershifted, one contained p107, and the other was probably also uncomplexed E2F. This might indicate that not all of these E2F complexes contain RBP3. However, from the data of all three groups, it can be concluded that RBP3 is a member of the E2F family, if not the only one. Its behaviour was therefore characterised in the same assays I had used with the E4-interacting clones, to attempt to draw some conclusions about the screens I had carried out.

When RBP3 was expressed in yeast as a VP16 fusion, it could interact with the E2A promoter and with E4. These abilities had been predicted from the known properties

of E2F, and were the basis of the screens that I had carried out. Therefore the surmises made in setting up both yeast screens in order to isolate an E2F binding activity were correct and either strategy could have worked. It would seem that RBP3 is capable of binding to DNA or to E4 as a monomer or a homodimer and unaided by any other cellular factor. This presumes that a yeast factor is not involved in mediating either interaction, the possibility of which has not yet been eliminated by any experiment, but seems remote. The question remains why the RBP3 cDNA was not isolated in either screen. To check if the VP16 tagged library used had contained the RBP3 cDNA, it was probed by PCR with several different primers for the RBP3 cDNA. No specific product was isolated, although these primers did generate the correct product with another library. Therefore, the problem lay with the complexity of the library, rather than the strategy of the screen. A cDNA for RBP3 is not represented, or only at very low levels, in the library used.

It is interesting that the interaction of RBP3 with E4 in yeast was of identical strength as clone D2. The behaviour of the E4 deletions would also indicate that the interaction with RBP3 involved the same region of E4 as all of the E4-specific clones I isolated. Therefore the hypotheses about the importance of the C-terminal 10 amino acids of E4 in its interaction with E2F cannot be extended to the *in vivo* yeast assays. This presumes that RBP3 has the same characteristics as the E2F binding activity used in the *in vitro* binding experiments and transfections with E4. Therefore it seems that the four clones I isolated do have similar characteristics as RBP3 in binding to E4. The region required is the same in terms of the three deletions tested, and D2 in particular has a similar affinity of binding. A prediction from this is that there could be a site in common between RBP3 and D2 that serves as a binding domain for E4. From the sequence data that is known for D2, no homologies can be seen to RBP3. Mutagenesis studies would be required to determine the precise regions of RBP3 and D2 involved in their interaction with E4 in order find any common motifs.

VP16/RBP3 interacted with LexA/Rb only very weakly in the yeast assay. However, this Rb fusion did interact strongly with the E1a and (HPV) E7 proteins in yeast. This contrasts with the situation *in vitro*, where bacterially-expressed Rb is able to bind efficently to both RBP3 and E1a. It is possible that the difference in the strength of these interactions is due to the phosphorylation state of Rb in yeast. E1a is capable of binding to Rb whatever its phosphorylation state, therefore during any stage of the cell cycle. On the other hand, E2F is only thought to interact with Rb when Rb is in the underphosphorylated state. This is the state that it is predominately found in during the G1 phase of the cell cycle. In S phase, when Rb becomes hyperphosphorylated, the interaction with E2F is lost. Perhaps in yeast, only a small proportion of Rb is in the correct phosphorylation state to be able to bind RBP3. However, a report suggests that E7 (like the SV40 T antigen) is also only able to bind to the underphosphorylated form of Rb (Imai et al., 1991). The fact that E7 is seen to interact with Rb in yeast more strongly than with RBP3 indicates that either the report is incorrect, or that the

phosphorylation state of Rb is not the reason for its weak ability to interact with RBP3. If this is the case, possibly the phosphorylation state of RBP3 itself is the reason for the weak interaction with Rb.

The activity of RBP3 in mammalian cell transfections confirmed what had been seen in yeast. As a VP16 fusion, RBP3 could activate the E2A promoter. This would be expected from published results. VP16/RBP3 could also activate the E4 gene product when E4 was expressed as a GAL4 fusion, using a reporter with GAL4 binding sites. This was despite the high basal activity of the E4 protein as a GAL4 fusion.

An interaction could not be detected in mammalian cells between the D2 or 29 cDNA products and the E2A promoter nor with E4. The negative result on the E2A promoter was less surprising because the DNA binding assay in yeast was also negative for these clones. However, since the cDNA clones and RBP3 behaved identically in yeast in terms of their interaction with E4, it was surprising that they were unable to interact with E4 in mammalian cells. At the moment the most likely explanation for this negative result is that the cDNA products are unable to reach the nuclei of the transfected cells. Until this question of nuclear localisation is resolved it is difficult to draw conclusions about whether the cDNA clones can behave exactly as RBP3 in their interaction with E4 in systems other than yeast. It is also difficult to conclude if they are part of the complex on the E2A promoter without having isolated full length cDNA clones. This would be worthwhile if they could be shown to interact with E4 in mammalian cells as RBP3 does.

5.4.3 Sequence homologies of the E4-specific clones

Four different cDNA clones were isolated that encoded polypeptides that specifically interacted with E4. It is notable that one of them (D2) was obtained six times. Since the sequence at the junction between the cDNA and VP16 was identical for all the clones, this may reflect a low complexity of the library. Two of the four clones were not homologous to any sequence in the databank. However, D2 had 90% amino acid identity over a 96 amino acid region corresponding to a 289bp mouse expressed sequence tag mRNA (Tsg67x. C. Hoog, unpublished). See fig5.3. This high degree of conservation between a human and mouse gene might indicate the functional importance of the product; however its function is unknown.

The only E4-specific clone to be identical to a sequence in the data bank was clone D4, which was the previously described H12.3 gene (Guillemot et al., 1989). The H12.3 amino acid sequence is completely conserved between an avian and mammalian species, which again could indicate functional importance. The highest similarity of amino acid sequence is to the bovine transducin β subunit and the human G protein β subunit (GPβ2), although it is unknown whether H12.3 has a similar function to these proteins. These G proteins belong to a family of proteins which have a repeat structure of about 40 to 45 amino acids, with specific residues conserved among the repeats (see fig5.4). The yeast STE4 protein is in this family and also serves as a b subunit in the

Comparison of the amino acid sequence of D2 with that of the mouse expressed sequence tag mRNA Tsg67x

```
D2:      37 LALLAVPDMSSLACGYLRNLTWTLSNLCRNKNPAPPIDAVEQILPTLVRLLHHDDPEVLA 96
            LALLAVPD+S+LACGYLRNLTWTLSNLCRNKN+APP+DAVEQILPTLVRLLHH+DPEVLA
Tsg67:    1 LALLAVPDLSTLACGYLRNLTWTLSNLCRNKNSAPPLDAVEQILPTLVRLLHHNDPEVLA 60

D2:      97 DTCWAISYLTDGPNERIGMVVKTGVVPQLVKLLGAS 132
            D+CWAISYLTDGPNERI MVVK GVVPQLVKLLGA+
Tsg67:   61 DSCWAISYLTDGPNERIEMVVKKGVVPQLVKLLGAT  96
```

Fig5.3 The part of the deduced D2 amino acid sequence that is homologous to Tsg67x is shown. The consensus sequence is shown between the two aligned sequences. Identical amino acids are indicated. + represents biochemically similar amino acids and a gap indicates non-conserved residues.

Repeated domains of the H12.3 protein

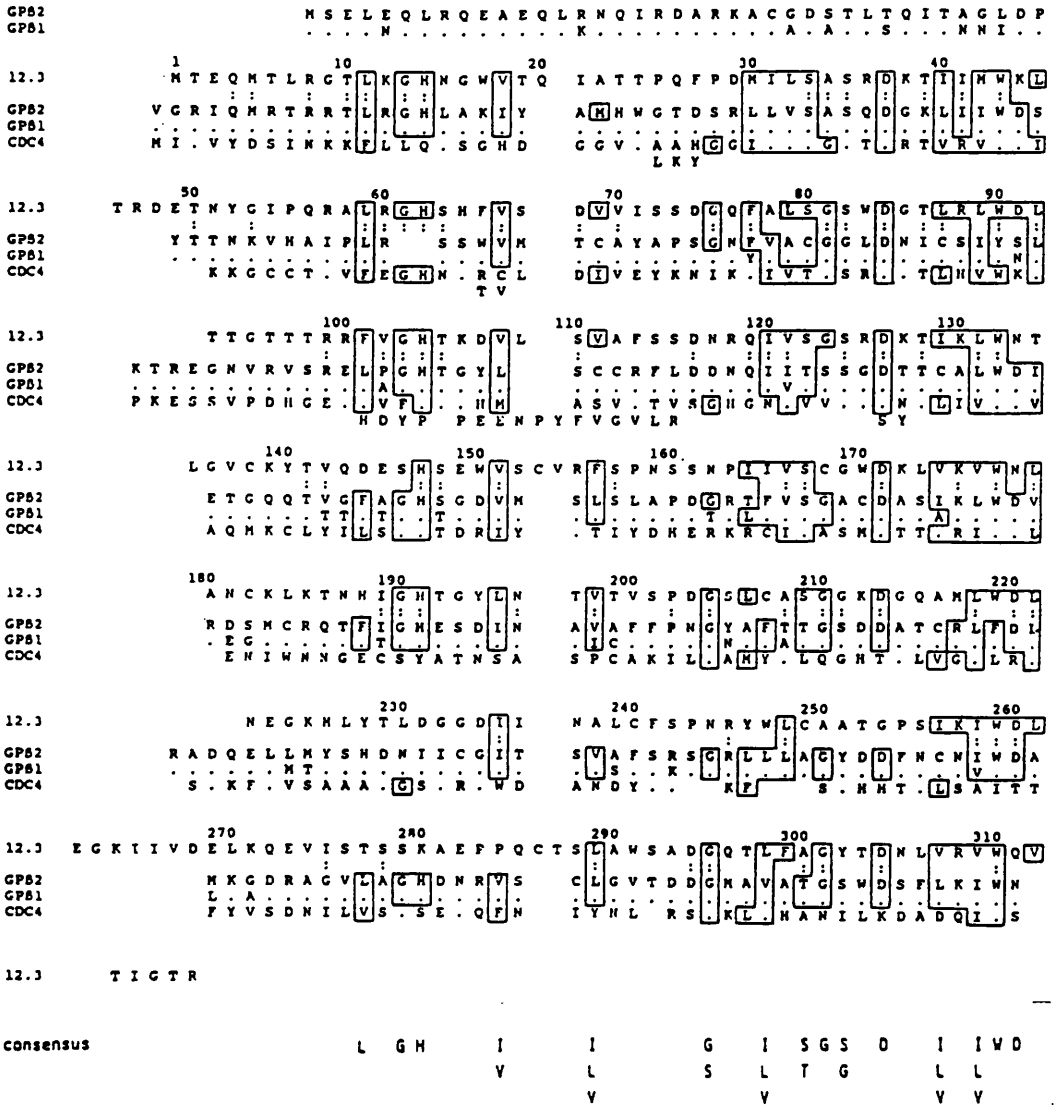


Fig5.4
Comparison of the amino acid sequence of H12.3 with that of bovine transducin β (GPβ1), human G protein β2 subunit (GPβ2), and CDC4. Residues of GPβ1 and CDC4 identical to those of GPβ2 are replaced by dots. Residues shared by H12.3 and at least one other sequence are indicated with a colon. The 7 repeated homologous segments in each sequence are aligned, with gaps introduced. Identical and highly conserved residues that align at least 5 times in one of the 4 sequences are boxed. The consensus sequence is shown below, and notes positions where a similar amino acid or group of biochemically similar amino acids predominates. (Guillemot et al, 1989).

heteromeric G-protein complex (Whiteway et al., 1989). STE4 and GPβ2 are approximately the same size as H12.3 and all three consist almost entirely of the repeat domains. Another yeast protein, PRP4, has five repeats of the β subunit structure, but is not part of the G protein complex: it is involved in RNA splicing. It is slightly larger than H12.3 with 250 residues bearing no homology to the repeats. (Whiteway et al., 1989). Other members of this family that share the same repeat structure are larger and have six copies of the repeat structure at their C-terminus. There is no significant homology outside the repeated domains. Examples in this category include the yeast CDC4 and TUP1 gene products and the Drosophila Enhancer of split gene product (E(spl)). None of these are thought to form part of a G-protein complex, and their biochemical roles are still largely undefined. (Williams and Trumbly, 1990). The CDC4 gene product is a nuclear protein involved in progression beyond the G1 phase of the cell cycle. TUP1 was isolated as a mutation which reduced glucose repression and has been found to be involved in the repression of many genes by interacting together with another protein, SSN6, and forming a complex on the promoter (Keleher et al., 1992). E(spl) is implicated in neurogenesis (Hartley et al., 1988). The amino acid sequence similarities between all the family members with this repeat are quite low. For example, comparing H12.3 to the other sequences, it has 22% overall amino acid identity to GPβ2, 12% identity to CDC4 and 18% identity to STE4.

The function of the repeat sequence is unknown. The larger proteins mentioned that share this repeat structure are not likely to be G proteins. This is clearly the case for TUP1, and it might be misleading to attribute this motif solely to G proteins. Therefore the sequence analysis of H12.3 does not clarify its function. It is not known if H12.3 is part of the G protein complex or where it is localised in the cell. The β subunit itself is found in the G protein complex on the membrane. Together with the γ subunit, it serves as a regulator of the GTP-binding (α) subunit. In some cases it may also serve as a downstream mediator of the signal. If the function of H12.3 was found to be as a β subunit of a G protein, then it would be difficult to speculate why E4 would be interacting with it in the nucleus. However, all the functions of the β subunit are consistent with it being involved in protein-protein interactions. For example, the α subunit has the catalytic functions of the complex and the γ subunit is responsible for membrane localisation. Therefore the conserved repeat motifs of the β subunit could have a general role in protein-protein recognition, which could be a common feature of all the proteins in this family. If so, this repeat could be similar to the ankyrin repeats which are also thought to mediate such protein-protein interactions, and are similarly found in a whole family of proteins with diverse functions.

Further experiments that might define the role of H12.3 would be aided by raising antibodies to the protein. This would enable studies of the cellular localisation of H12.3, and coimmunoprecipitations from cell extracts and *in vitro* would be possible to see if an interaction occurs between H12.3 and an α or γ subunit of a G protein complex.

A recent report showed that H12.3 had been isolated from a screen for cDNA products that could interact with Rb (Y-H Lee, unpublished). Therefore it seems to be a second example of a protein that may be able to interact with both Rb and E4, since RBP3 also seems to display these properties. However, in the case of RBP3 the functional significance *in vivo* of its interaction with these two proteins has been defined. This is not the case with H12.3. At present, its interaction with E4 has only been shown in yeast, and its interaction with Rb only *in vitro* by bacterial overexpression of the proteins. It remains to be seen if these interactions can be seen *in vivo*. Neither clone D4 nor H12.3 were seen to interact with Rb in yeast. However, the interaction of RBP3 with Rb in yeast could only be seen very weakly, so as previouly discussed, Rb could be in the wrong phosphorylation state in yeast to mediate these interactions. Further experiments to investigate the putative interaction between H12.3 and Rb could involve analysis of its interaction *in vitro* with mutants of Rb unable to bind E2F or E1a. Antibodies to H12.3 could be used to see if H12.3 in cellular extracts binds to GST-Rb. Informative *in vivo* experiments would be to test whether H12.3 could affect the block of the cell cycle seen when Rb is added to Rb^- cells. H12.3 may also affect the ability of Rb to repress promoters, by analogy with the TUP1 repressor.

5.4.4 E4-Independent clones.

Two copies of SAP1 were isolated from this screen. This was not surprising, since in the original screen for SRF-interacting proteins (Dalton and Treisman, 1992), out of the 500,000 colonies screened, 2 copies of the SAP1 cDNA were obtained.

The third E4-independent clone characterised (D25) encoded a protein almost identical to FIP (Blanar and Rutter, 1992). FIP was isolated in a screen for proteins that interact with the leucine zipper of cFos, and is a member of the of the basic, HLH, zipper family of proteins (see fig5.5). The *in vivo* activity of FIP is not clear because a full length cDNA has not been obtained, nor have its dimerisation interactions been defined fully *in vivo*. The D25 cDNA terminates immediately before the final leucine residue of the leucine zipper of FIP. It has a divergent sequence from FIP at its very N-terminus. FIP is almost identical to USF in the basic, HLH domain, although there is no similarity N-terminal to the basic domain. Published results show that FIP can bind to the USF site in a gel retardation assay. I therefore tested the activity of D25 as a VP16 fusion on a USF site-driven reporter in mammalian cells. An approximate 3-fold activation was seen. A similar 3-fold level of activation was seen by cotransfecting FIP and cFos and measuring activation on an AP1 site indicating that FIP and cFos may interact *in vivo* and bind DNA (Blanar and Rutter, 1992). Since the D25 fusion had a VP16 tag, it might have been expected to activate the USF site more highly. Perhaps the lack of the final leucine of the zipper reduced the DNA binding affinity of the D25 product.

The activation of the USF site by D25 may explain why it was isolated in this screen. Its ability to directly bind to DNA would enable it to interact with the promoter

Comparison of the bHLH-Zip region of FIP to other bHLH-Zip proteins

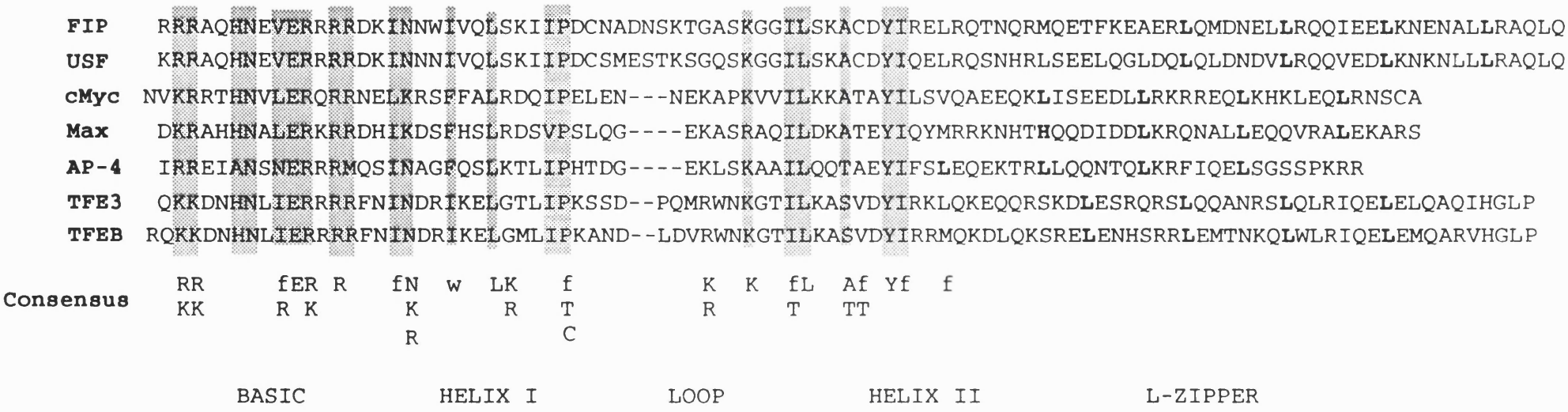


Fig 5.5

The bHLH-IZIP region of FIP is aligned with those of USF, Myc, Max, AP4, TFE3 and TFEB. Shaded regions identify conserved residues. Residues that can form a zipper region are in bold. A consensus sequence for the bHLH family is shown. f=L,I,V, and w=F,L,I,Y. (B. Amati; Blanar and Rutter, 1992)

region of the reporter, in which, as mentioned previously, there is an element that fits the consensus sequence for this family of transcription factors. It would be informative to isolate a full length cDNA for clone D25 and FIP, to investigate whether alternative splicing of the mRNA occurs, and what properties the unique sequences confer upon the products. Although few examples of heterodimerisation are known between members of the basic, HLH, zipper family, it would be interesting to see if FIP could heterodimerise with any of them. The high degree of homology between FIP and USF in their dimerisation domains may indicate that USF could be a potential partner for FIP. If so, FIP may be involved in a diverse set of interactions.

It has become increasingly clear over the last few years that the regulation of transcription initiation involves numerous specific protein-protein interactions. Methods have been developed to investigate such interactions in both *in vitro* and *in vivo* conditions. The two-hybrid system first described by S. Fields for studying protein-protein interactions in yeast is particularly attractive and versatile. In this thesis I have attempted to use this approach in three different studies. Its versatility has been exploited both in investigating the direct interaction between two previously defined proteins, and also in providing a system whereby new interactions are identified and cDNAs isolated that encode these interacting proteins. These studies demonstrate how this yeast system can address very specific questions but also highlight pitfalls that can occur. Despite these, however, it is likely to prove to be an important tool in studying such a complex process as transcription initiation.

CHAPTER 6: MATERIALS AND METHODS

6.1 *E. coli* STRAINS

TG1 K12, Δ(lac-pro), supE, thi, hsdD5/F' traD36, proA+B+, lacq, lacZΔM15

SCS1 F^-, recA1, endA1, gyrA96, thi^-1, hsdR17, supE44, relA1

6.2 YEAST STRAINS

Y700 MATα his3-11,15 trp1-1 ade2-1 leu2-3,112 ura3-52 can1-100 (W303-1B: (Thomas and Rothstein, 1989))

Y700b Y700 with 3xATF-lacz reporter integrated at ura3 locus (S. Kuge)

Y4-1/Y1-1 Y700 with E2F(4-1)-lacz or E2F(1-1)-lacz reporter integrated at ura3 locus

Y657 MATa his3-11,15 trp1-1 ade2-1 leu2-3,112 ura3-52 can1-100 his4::HIS3 (Newman and Norman, 1991)

Y35 Y657 with LexAop-lacz reporter integrated at ura3 locus (G. Micklem)

S62L S50 (MATα his3-11,15 trp1-1 ade2-1 leu2-3,112 can1-100) with SRE-lacz reporter integrated into URA3 locus (Dalton and Triesman, 1992)

YT6::171 YT6 (ΔGAL4, Δgal80, ura3, his3, leu2) withUAS_G upstream of GAL1-lacz reporter integrated at ura3 locus (Gill and Ptashne, 1987)

6.3 MATERIALS

Unless otherwise stated, all reagents were obtained from BDH, BRL, Stratagene, Fisons or Sigma, except enzymes which were from New England Biolabs, Boehringer Mannheim or Anglian Biotechnology. All radiochemicals were from Amersham. All media was sterilised by autoclaving unless mentioned otherwise.

6.4 BACTERIAL MEDIA

Bacteria were grown at 37°C in LB (Luria-Bertani) medium (per litre: 10g Bacto-tryptone, 5g Bacto-yeast extract (Difco), 10g NaCl, pH7.5). For plates 15g of Bactoagar was also added. For selection purposes ampicillin was added to 100μg/ml.

6.5 DNA PREPARATION FROM E.COLI

Both large scale and small scale preparation of DNA employed the alkaline lysis method of DNA preparation essentially as described in Maniatis et al., 1982.

Small Scale

1.5mls of an overnight culture was pelleted and the bacteria resuspended in 100µl of solution 1 (50mM glucose, 25mM Tris pH7.4, 10mM EDTA). Cells were lysed by addition of 200µl solution 2 (200mM NaOH, 1% SDS) and placed on ice after which 150µl solution 3 (60ml 5M KOAc, 11.5mls glacial acetic acid, 28.5mls H_2O) was added. This was kept on ice for 5 mins and then centrifuged at 4°C for 5 minutes. The supernatant was decanted, phenol extracted and ethanol precipitated. The pellet was washed in 70% ethanol, dried and resuspended in 50µl TE buffer (10mM Tris pH8.0, 1mM EDTA).

Large Scale

400ml of a bacterial culture grown overnight was pelleted by centrifuging at 6Krpm for 10min. Pellets were resuspended in 10ml of solution 1. Cells were lysed in 20ml of solution 2 and chromosomal DNA was precipitated by the addition of 15ml of solution 3. This was centrifuged at 8Krpm for 10min. To the supernatant was added 24ml of isopropanol and the DNA/RNA was pelleted by centrifugation at 10Krpm for 10min. The pellet was resuspended in 4.5ml TE, 5.0g of caesium chloride and 30µl (10mg/ml) of ethidium bromide. This was loaded into a 5ml Beckman polyallomer tube which was sealed and spun in a Beckman Vti65 rotor for 4hr at 64Krpm. The resulting band gained from the caesium chloride gradient was removed with a needle and syringe. 3x volume of TE was added and the ethidium bromide was extracted with repeated washes of isobutanol. DNA was ethanol precipitated, washed in 70% ethanol, dried and resuspended in TE. DNA concentration was determined by OD260nm spectrophotometry.

6.6 DNA MANIPULATIONS

Restriction Enzyme Digestions

1-5µl of miniprep DNA or 1-3 µg of maxiprep DNA was digested with about 10units of restriction enzyme at 37°C in a 20µl volume for 1 to 2 hours before adding 1µl loading dye (25% Ficoll, 0.02% bromophenol blue, 0.02% xylene cyanol) and loading onto an agarose gel. All digests were carried out in the reaction buffers recommended and supplied by the enzyme manufacturers.

Filling in of recessed ends of DNA:

Recessed 3' termini were filled out by adding all four dNTPs (0.1mM final concentration) to a complete restriction digestion, adding 2 units of Klenow fragment DNA polymerase, and incubating at room temperature for 10 to 40min.
Overhanging 3' termini were treated similarly except that 1µl of T4 polymerase was used instead of Klenow.

Dephosphorylation of DNA

The 5' terminal phosphate of vector fragments was removed using calf intestinal alkaline phosphatase (CIP) to prevent vector recircularisation upon ligation. Restricted DNA was phenol extracted, precipitated and resuspended in 40µl of phosphatase buffer (0.1mM EDTA, 50mM TrisCl pH8.0), and incubated with 20 units of CIP at 37°C for 45 minutes. The appropriate fragment was then gel purified.

Phenol and Phenol /chloroform extractions.

Phenol (BRL) was equilibrated with 0.1M Tris pH8.0. Calibrated phenol was added to chloroform and isoamyl alcohol in a ratio of 25:24:1. A standard extraction involved adding to the sample an equal volume of phenol: chloroform: isoamyl alcohol, vortexing and cetrifuging in a microfuge for 2 mins. The upper aqueous phase was then removed and the process repeated if necessary.

Fragment Purification

DNA fragments were purified on 150ml 1-2% agarose, 1xTBE gels with 0.5µg/ml ethidium bromide (1xTBE: 100mM Tris borate pH8.3, 1mM EDTA). Gels were run at 80 mA after which the gel was removed and the DNA visualised on a long wave UV transilluminator. The appropriate fragment was excised, the gel slice placed into an eppendorf and prepared for the ligation step by use of the "Gene Clean" kit (Bio 101 Inc). A volume of 6M NaI solution was added that was 4.5x the volume of the gel slice, plus 0.5x volume of TBE modifier, and the mixture heated to 50°C to dissolve the agarose. 6µl of "Glassmilk" was added to bind the DNA and left for 5min. The Glassmilk containing the DNA was spun down for 10secs and the pellet washed 3 times with cold "NEW wash". The DNA was eluted by mixing the pellet with 10-20µl TE, incubating at 50°C for 3min and spinning to remove the Glassmilk.

Ligation

Ligations were generally carried out in a 20µl volume containing 2µl 10x ligase buffer (0.66M Tris pH 7.5, 50mM MgCl2, 50mM DTT and 10mM ATP), 100ng of purified vector, 400ng of purified fragment and 10units of T4 DNA ligase. Ligations were left for 1-16 hrs after which they were transformed into E.coli as below.

PCR

1µg of the DNA to be amplified was mixed with 5µl of a 20µM solution of each primer (normally 20bp long), 10µl 2mM solution of 4 dNTPs, 10µl 10xPCR buffer (100mM Tris pH8.3, 500mM KCl, 15mM MgCl2, 0.1% gelatin), 2.5units AmpliTaq polymerase in a 100µl total volume. 100µl mineral oil was laid on top of this. Amplification was usually for 12 cycles with denaturation at 94°C, annealing at 50-60°C, and extension at 72°C, each for 1min. At the end of the last cycle, extension was continued for an additional 5min. To additionally fill in any single stranded ends for

cloning purposes, the DNA was incubated at 37°C for 1hr with 5units of DNA polymerase, 1µl 5mM dNTPs and 1µl 0.1M DTT. The PCR product was run on an agarose gel and purified with GeneClean before cutting with restriction enzymes.

6.7 TRANSFORMATION OF *E. COLI.*

Competent Cells

An overnight culture grown in LB was diluted 100 fold into fresh medium and grown shaking at 37°C until the OD 600nm was 0.5 to 1.0. All the following steps were performed at 4°C. Cells were harvested for 10min at 3Krpm, and gently resuspended in 0.6 of the original culture volume of ice-cold CTG (50mM CaCl2, 0.05mg/ml Thymine, 10% glycerol). The cells were allowed to stand on ice for 20min, then spun as before and resuspended in 1/15 of the original culture volume of CTG. They were either used immediately or aliquoted into microfuge tubes, frozen on dry ice, and stored at -70°C.

Transformation of Competent Cells

Frozen competent cells were thawed on ice and used immediately. 100µl of cells were added to 1-10µl of DNA in a tube pre-chilled on ice, mixed and left on ice for 10 to 50 minutes, then transfered to a 42°C waterbath for 2 minutes. 400µl of LB was added and incubation was continued for a further 30min. After a brief spin and removal of excess supernatant, the remaining contents of the tube were spread on a selective plate.

6.8 DNA SEQUENCING

Template Preparation For Single Stranded Sequencing

Plasmids containing the M13 origin of replication were rescued as phage particles containing single strand DNA by infection of F^+ bacteria (eg TG1) with the helper phage M13 K07, as follows: 1.5 ml cultures were innoculated from fresh patches or colonies, using medium containing helper phage (0.5-1.0 x 10^9 PFU per ml M13 K07 final concentration, from G.Micklem), and were grown shaking at 37°C for 3 to 6 hours. This culture was transfered to a microfuge tube and centrifuged for 5min. The phage in 800µl of supernatant were precipitated with 200µl of 20% PEG 6000, 2.5M NaCl for 15 min at room temperature, and collected by centrifugation for 4min. The PEG/ NaCl was aspirated off, and again after the tube was respun for 30seconds to collect residual supernatant. The phage pellet was resuspended in 100µl TE, phenol extracted by vortexing for one minute with 80µl phenol (equilibrated with TE buffer) and the single stranded DNA precipitated from the aqueous phase with 6µl 3M Na Acetate pH5.5 and 200µl 95% Ethanol. After spinning for 3-4 min, the pellet was washed with 95% ethanol, allowed to dry and resuspended in 100µl of TE. 2-7µl of this was used for sequencing.

Template Preparation For Double Stranded Sequencing

5µg DNA was denatured with 20µl of 0.25M NaOH for 5 minutes. Neutralisation with 10µl of 1.5M ammonium acetate pH4.5 was followed by precipitation with 90µl of ethanol and centrifugation for 3-4 min. The pellet was washed with 70% ethanol and resuspened in 7µl of TE. All of this was used for sequencing.

Sequencing

Sequenase Version 2.0 kit (US Biochem. Corp.) was used.

Annealing: per clone, 10ng (in1µl) of primer, 7µl template and 2µl of Reaction Buffer (100mM MgCl2, 200mM TrisCl pH7.5, 250mMNaCl) was mixed. This gave approximately stoichiometric quantities of template and primer. The mix was heated in a beaker of water at 65°C for 2min, then left to cool slowly to <35°C.

Labelling: The following components were added to the annealed template and primer: 0.5µl [^{35}S] dATP, 1µl 0.1mM DTT, 2µl labelling mix (5.5µM of each of dGTP, dCTP, dTTP, diluted 1in 9), and 1µl Sequenase enzyme (diluted 1 in 8 in TE). This reaction was incubated at room temperature for 2-5min.

Termination: 4 tubes were prewarmed to 37°C, each containing 2.5µl of a different termination mix (50mM NaCl and 80µM of 3 dNTPs. A different dNTP was missing from each of the 4 and replaced by ddNTP, eg ddA: 80µm dGTP, dCTP and dTTP, 8µM ddATP). 3.5µl of labelled reaction was added to each stop mix and incubated at 37°C for 5min. 4µl stop mix (95% formamide, 0.5% Xylene cyanol, 0.05% bromophenol blue, 20mM EDTA) was then added.

The reactions were denatured by incubation at 80°C for 2min, immediately before being run on a 50cm 6% acrylamide buffer gradient gel at 40W for 2-5 hours.

Buffer Gradient Gels

Buffer gradient gels were poured using 5x gel mix (30ml of 40% stock acrylamide (38:2 acrylamide: NN-methylene bisacrylamide), 100ml 10x TBE, 92g urea, 10mg bromophenol blue, in 200ml), and 0.5x gel mix (as above but no bromophenol blue, and 0.5x TBE final concentration). To a 7ml aliquot of 5x mix, 12.5µl of 25% APS and 12.5µl of TEMED were added to initiate polymerisation, and to a 40ml aliquot of 0.5x mix 75µl of APS and 75µl of TEMED were added. 6ml of 0.5x mix was taken up in a pipette followed by all of the 5x mix. 2-3 bubbles were taken up to form a rough gradient, and the mix was poured into two taped 20cm by 50cm glass plates separated with 0.35mm spacers. This was followed by the rest of the 0.5x mix, and then a 24 slot comb clamped in position. After being allowed to set for at least 1 hour the gel was run in a vertical apparatus with 1xTBE. After running, the gel was fixed in 10% glacial acetic acid,10% methanol for 15min while still attached to one of the gel plates, before being transferred to 3MM paper (Whatman), dried at 80°C on a flatbed vacuum drier, and exposed to X-ray film (XAR Kodak).

6.9 YEAST MEDIA

Yeast were grown at 30°C. Recipes are given for liquid medium, to which 0.1 volume of 20% glucose or galactose and any other supplements are added before use. For plates, 20g Bacto Agar (Difco) was added and this was cooled to 50-60°C after autoclaving before adding supplements and a carbon source.

Where the CUP1 promoter was used, expression was induced with 1mM $CuSO_4$ added to the medium.

Rich Medium (YEPD): 11g Yeast extract (Difco), 22g Bacto Peptone (Difco), 55mg Adenine Sulphate per litre.

Minimal Medium: 0.67% Difco Nitrogen Base (without amino acids).

Selective Media: As minimal medium, and all contain 55mg/litre each of adenine sulphate, uracil, and tyrosine, unless stated otherwise.

-URA: no uracil, plus 0.5% casamino acids (Difco), supplemented with 10ml/litre 0.5% tryptophan and 20ml/litre 0.5% leucine.

-TRP: plus 0.5% casamino acids, supplemented with 20ml/litre 0.5% leucine.

-HIS: supplemented with 10ml/ litre drop-out solution.

Drop out solutions: (per 100ml) 0.2g arginine, 0.1g histidine, 0.6g isoleucine, 0.6g leucine, 0.4g lysine, 0.1g methionine, 0.6g phenylalanine, 0.5g threonine, 0.4g tryptophan, with the amino acid for which selection is required being omitted.

Sorbitol top agar: Sorbitol to 1M was added to the relevant medium.

6.10 YEAST TRANSFORMATION

Spheroplasting method (for integrations)

Exponentially growing cells (OD600nm 0.5-1) were harvested by centrifugation, washed once in water, once in 1M Sorbitol then resuspended in 20ml of SCEM (1M Sorbitol, 0.1M Na Citrate pH5.8, 10mM Na2EDTA, 30mM β-Mercaptoethanol). 1000units of lyticase was added and after incubation at 30°C for 10min, spheroplast formation was checked under phase contrast after mixing an aliquot of cells with 10% SDS. When about 80% of cells are spheroplasts, they were spun at 4°C and resuspended gently in 20ml STC (1M Sorbitol, 10mM Tris pH7.5, 10mM $CaCl_2$). After pelleting again, cells were finally resuspended in 1ml of STC. 0.1ml aliquots were distributed to 5ml tubes at room temperature, already containing DNA. For plasmid transformations, 5µg of DNA was used. For transplacements 5µg of excised fragment, and 0.5µg of co-transforming plasmid (Trp marker) was used. The cells and DNA were left for 10min after gentle mixing, and then 1ml of PEG (20% polyethylene glycol-8000, 10mM CaCl2, 10mM Tris pH7.5) was added, mixed, left for 10min, spun at 1Krpm for 3min, and the cells resuspended in 0.15ml of SOS (1M sorbitol, 6.5mM CaCl2, 0.25% yeast extract, 0.5% bactopeptone). Appropriate selective molten top agar held at 55°C was mixed with cells and plated onto a selective plate. After allowing the top agar to set, plates were incubated for 2-3 days.

Lithium Acetate method

Exponentially growing cells were harvested by centrifugation, washed once in water, once in TEL (10mM Tris pH8.0, 1mM EDTA, 0.1M LiOAc), resuspended in TEL at approximately 10^9cells/ml, and incubated for 30min with gentle agitation. 100µl cells were mixed with about 5µg DNA in 50µl TE and incubated at 30°C for a further 30min. 700µl 40%PEG4000 in TEL was added and incubation continued for 60min. After a 42°C heat shock for 5min, cells were centrifuged, washed once in water, resuspended in water and plated onto selective medium.

6.11 YEAST BETAGALACTOSIDASE ASSAY

Qualitative

Fresh colonies of yeast on a nylon filter (Amersham) were immersed in liquid nitrogen for about 5 seconds, then the filter allowed to thaw and dry on tissues. It was then transferred to the lid of a petri dish containing a Whatman No. 1 filter soaked in 2ml of Z buffer (1x Z buffer: 16.1g $Na_2HPO_4.7H_2O$, 5.5g $NaH_2PO_4.H_2O$, 0.75g KCl, 0.246g $MgSO_4.7H_2O$, pH to 7.0, per litre.) containing 1mg/ml BCIG (stock solution 100mg/ml in dimethylformamide). This colourless synthetic lacZ substrate is cleaved to give an insoluble blue product. Care was taken to avoid bubbles between the two filters. Strongly lacZ positive colonies became obviously blue after about 10 minutes. (Breeden and Nasmyth, 1985)

Quantitative

2 to 5 ml cultures were grown to OD 600nm 0.5 - 1.0 (exponential phase), in suitable medium. Cells were pelleted for 2 minutes at 2.5Krpm, washed once in water and resuspended in 200µl buffer (0.1M Tris pH7.5, 0.05% Triton X-100), then permeabilised by freeze thawing at -70°C. LacZ enzymatic activity was measured by the increase in OD 420nm caused by the cleavage of the colourless LacZ substrate ONPG as follows: 0.4ml aliquots of Z buffer and 0.1ml 4mg/ml ONPG were put in 1.5ml disposable cuvettes and prewarmed in a water bath at 30°C. The time was recorded as 100µl lysate was added to a cuvette, and again when the reaction was stopped by addition of 250µl of 1M Na_2CO_3. The reactions were stopped when a slight yellow colour was observed. (Breeden and Nasmyth, 1987).

Units of LacZ activity were calculated as 1000 x OD420nm units / OD600nm x volume of lysate in assay(ml) x time(min). The assay is not particularly reproducible, but serves to indicate within an order of magnitude approximately how much LacZ activity is present. Data shown are from a typical assay that has been repeated at least twice.

6.12 LIBRARY SCREENING

The lithium acetate method was used as above, normally using 2.5µg library DNA and 5µg carrier DNA (eg sonicated salmon sperm DNA) per $20cm^2$ plate. Transformed cells were plated on nylon filters (Amersham) laid on selective medium. After growth for 24 to 36hrs, when colonies were visible but still small, the filters were transferred to galactose medium and grown for a further 12 to 16 hrs.

The colony colour assay was performed as above by immersing filters in liquid nitrogen. The $20cm^2$ filters were developed by laying them on Whatmann 3MM paper moistened with 25ml Z buffer/BCIG solution. Positive (blue) colonies were picked and patched onto selective glucose plates. For colony purification, cells from each patch were innoculated into medium and a dilution spread on filters on 9cm plates. The lacz assay was repeated by transferring the filters to galactose plates and retesting for colony colour. Single blue colonies were now picked.

Yeast plasmid DNA preparation

15ml cultures were grown to stationary phase, harvested and washed once in 1M sorbitol. The pellet was resuspended in 0.2ml 1M sorbitol, 50mMTris pH7.5, 10mM β-mercaptoethanol and 4000units/ml lyticase and incubated for 30min at 30°C with gentle shaking. To acheive complete lysis, 0.4ml 0.2M NaOH/1% SDS was added, mixed and left on ice 10min. 0.3ml 3M NaOAc, pH5.0 was then added, mixed and left on ice 10min, then this was spun for 10min at 4°C. 0.6ml isopropanol was added to the supernatant in a fresh tube, and left for 20min at room temp to precipitate the DNA. After spinning for 10mins and washing the pellet with 70% ethanol, the pellet was resuspended in 0.2ml TE.

Electroporation of yeast preparation DNA into *E. coli.*

It is necessary to purify the crude yeast DNA preparation via *E.coli* to remove the endogenous yeast 2µ plasmid. $CaCl_2$ transformation will not work with this crude preparation. Cells for electroporation were prepared by growing a culture to 0.4 OD600nm. All the following steps were at 4°C. 100ml cells were harvested and washed twice in 100ml 1mM Hepes, pH 7.9, then resuspended in 200µl. 1µl DNA was electroporated with 40µl of this suspension using a Biorad electroporator (2.5kV, 25µF, 200ohms). 1ml SOC medium was immediately added, and the cells transferred to a 5ml tube to shake at 37°C for 1hr. (SOC: LB medium plus 2.5mMKCl, 10mM MgCl2, 10mM MgSO4, 20mM glucose.) Cells were plated and several single colonies picked, mini-prep DNA prepared and transformed into yeast. This DNA could also be used for sequence analysis once the phenotypes obtained previously were confirmed.

6.13 YEAST MATING

Fresh patches of haploid yeast strains of opposite mating type were used. One strain was suspended thickly in 200µl YPD and spread on a selective plate. The other

strain was patched very lightly on top of this. Plates were incubated for 2 to 3 days. To select for the diploids, his3$^-$ haploids were mated to his4$^-$ haploids; only the resulting diploids were HIS$^+$ and therefore able to grow in the absence of exogenous histidine.

6.14 YEAST EXTRACT PREPARATION (GEL RETARDATION)

All steps were performed at 4°C with ice-cold solutions. 30-50ml of exponentially growing cells were harvested by centrifugation, washed once in water, once in breakage buffer (20mM Tris pH7.9. 10mM MgCl2, 1mM EDTA, 5% glycerol, 1mM DTT, 1mM PMSF, 1x protease inhibitor (PI) mix), and resuspended in 1-200µl of this buffer. Acid-washed glass beads (0.5mm) were added to the meniscus of the cell suspension and the tube vortexed for 4-5min, cooling intermittently on ice. The beads were washed twice by adding 0.1ml breakage buffer, vortexing and centrifuging. The pooled supernatants were centrifuged again to remove remaining cell debris, and 1-2µl of this extract was used per gel retardation reaction. (100x PI mix: 50mM TPCK, 5mM TLCK, 2mg/ml pepstatin A)

6.15 YEAST EXTRACT PREPARATION (WESTERN ANALYSIS)

All steps were performed at 4°C with ice-cold solutions. Exponentially growing cells (30-50ml, grown overnight in galactose medium) were harvested by centrifugation, washed once in water, once in lysis buffer (50mM Hepes pH7.9, 5mM EDTA, 5mM EGTA, 1mM DTT, 1mM PMSF), and resuspended in 350µl of this buffer in 5ml tubes. Acid-washed glass beads (0.5mm) were added to the meniscus of the cell suspension and the tube was vortexed for 4-5min (cooling intermittently on ice). 0.2ml of 2xRIPA (50mM Tris pH7.4, 150mM NaCl, 1% Triton X-100, 1% Na deoxycholate, 0.1% SDS for 1x) was added, vortexed, and the supernatant removed after spinning. A further 200µl RIPA was added, vortexed and left on ice for 20min before spinning and pooling this supernatant with the first. The pooled supernatants were centrifuged again (5min, microfuge) to remove remaining cell debris and loading buffer added.

6.16 YEAST GENOMIC DNA PREPARATION (SOUTHERN ANALYSIS)

5ml of an overnight culture of yeast was harvested, washed with 1M sorbitol in an eppendorf and resuspended in 0.25ml SCEM with 5000units lyticase. This was incubated for 15-45min at 30°C, spun for 5sec, then resuspended in 0.15ml GuHCl solution (4.5MGuCl, 0.1MEDTA, 0.15M NaCl, 0.05% sarkosyl pH8.0). After incubating for 10min at 65°C, 150µl cold ethanol was added and the tube spun for 5min. The pellet was resuspended in 0.3ml TE and 2µl 10mg/ml RNase A, and incubated at 37°C for 1hr. 6µl 10mg/ml proteinase K was added and incubation continued at 65°C for 30min. After cooling, the DNA was extracted twice with phenol: chloroform: isoamylalcohol then precipitated with ethanol. The pellet was washed in 70% ethanol and resuspended in 100µl TE. 5µl of this was used per restriction digest.

6.17 SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS

Protein gel dimensions were 18cm x 20 cm x 0.8 mm. A 30ml resolving gel was poured (0.375M Tris pH8.8, 0.1% SDS, 8-10% acrylamide from 30:0.8 acrylamide:bis stock. 30µl TEMED (Gibco), and 250% of 10% ammonium persulphate (APS) was added for polymerisation). Once this had set, a stacking gel (100mM Tris.HCl pH6.8, 0.1% SDS, 5% 30:0.8 acryl:bis, plus TEMED and APS) was poured on top. An equal volume of sample and 2x loading buffer (10% glycerol, 2% SDS, 60mM Tris.HCl pH6.8, 100mM DTT, 0.001% pyronin Y) were mixed and boiled for 2min before loading. Electrophoresis was performed at 50-200V in running buffer (6g Tris, 28.8g Glycine and 1g SDS in 1 litre). For enhancing of 35S methionine-labelled bands, gels were fixed (20min in 40% methanol, 10% acetic acid) then soaked in Amplify (Amersham) for 20min. For all labelled samples, gels were dried on a flat-bed vacuum dryer at 80°C and exposed to Kodak-XAR film at -80°C (or RT for 35S-labelled, unamplified samples).

6.18 *IN VITRO* TRANSCRIPTION AND TRANSLATION

Reagents and protocol for transcription were as described in the Stratagene kit. 4µg plasmid containing a cDNA insert in the T7 vector was linearised (normally with Spe1), then treated with Proteinase K and phenol/chloroform extracted. Capped mRNA was synthesised at 37°C for 30min in a 100µl reaction with 1mM rATP, rUTP, rCTP, 0.1mM rGTP, 0.5mM of the cap analogue m7GpppG, 20µl 5x transcription buffer (200mM Tris pH7.5, 250mM NaCl, 40mM MgCl2, 10mM Spermidine), 4µl 0.75M DTT, 1µl RNasin and 10units T7 RNA polymerase. The DNA template was removed with DNase 1 (10units, 30min at 37°C) and the RNA phenol/chloroform extracted, ethanol precipitated and resuspended in 100µl TE. 5µl was checked on an agarose gel.

This RNA was translated in rabbit reticulocyte lysates (Promega) as described by the manufacturers. 5µl template was heated to 67°C for 5min and immediately put on ice to destroy any secondary structure. To this was added 35µl lysate, 1µl RNasin, 1µl 1mM amino acid mixture (minus methionine if using labelled methionine) and 4µl 35S-methionine (10µCi/µl). The reaction was incubated at 30°C for 30-60min and the translation products analysed on an SDS polyacrylamide gel. For immunocoprecipitations, two mRNA templates were translated in the same reaction.

6.19 IMMUNOCOPRECIPITATION

Reactions (27µl) contained 20µl bufferE (20mM HEPES pH7.9, 100mM KCl, 0.2mMEDTA, 10% glycerol, 1mM DTT, 1% NP40, 0.5mg/ml BSA), 0.1µl 9E10 ascites fluid, 1µl rabbit antimouse IgG and 5µl reticulocyte lysate. This was gently mixed with 10µl Protein A Sepharose, spun out from a previously swollen 50% (v/v) stock equilibrated with buffer E. The solution was tumbled for 3hr in a cold room and immune complexes washed by addition of 1ml cold buffer E without BSA, mixing rapidly by vortexing and inverting the tube, and recovering the sepharose by a 15sec spin. This

wash was repeated twice before boiling the sepharose in protein gel loading buffer and analysing the immunoprecipitated products by SDS-PAGE. Gels were fixed and amplified before autoradiography at -80°C.

6.20 SOUTHERN HYBRIDISATION

Blotting

Gels were run as in DNA manipulations, and then soaked in 0.25M HCl for 20min to partially hydrolyse the DNA and assist its transfer. This was followed by denaturing of the DNA with 0.2M NaOH, 1.5M NaCl for 30min twice, and then neutralising with 0.5M TrisCl pH8.0, 1.5M NaCl twice for 30min, all gently rocking. The blot was set up as follows with no air bubbles between layers. The gel was placed on 2 pieces of 3MM (Whatman) paper soaked in, and acting as wicks from a reservoir of 500ml 20xSSC (175.3g NaCl, 88.2g Trisodium citrate, pH 7.0-7.2, per litre). A nylon membrane (Hybond-N, Amersham) pre-soaked in 6xSSC, was placed on top of the gel, followed by saran wrap. After cutting a window in the saran wrap slightly smaller than the size of the nylon filter, two pieces of 3MM dampened in 2xSSC were placed on top, followed by 10 further dry ones and about 5cm thickness of paper towels with a weight on top. After leaving the blot overnight, the nylon membrane was unpacked, rinsed several times in 6xSSC, dried on 3MM paper, and baked at 80°C under vacuum for 2 hours.

Hybridisation

The filter was sealed in a heat-sealable plastic bag with prehybridisation mix (4xSSC, 50% formamide, 10x Denhardt's, 0.5% SDS, 100µg/ml denatured salmon sperm DNA) and no air bubbles, and incubated at 65°C for 2 to 3 hours. This was removed, the hybridisation mix added, and the bag sealed and incubated overnight at 65°C. Hybridisation mix was the same as for prehybridisation except that probe, boiled for 5min with the salmon sperm DNA was included. After hybridisation the filter was washed twice in 1xSSC at room temperature, and then twice for 20min at 65°C. It was wrapped in saran-wrap to prevent drying out, and exposed at -70°C with a screen.

Denhardt's: For 100ml of 100x: 2g polyvinyl-pyrrolidone, 2g bovine serum albumen, 2g Ficoll 400 (Pharmacia), dissolved in 1mM Na2EDTA.

Probe: "Random primed DNA labeling kit" (Boehringer Mannheim) was used. DNA excised in a low melt agarose gel fragment was heated to 95°C for 10 min and then cooled to 37°C for 5 minutes. Approximately 50-100ng of this was mixed with 5µl (50µCi) of [α-32P] dCTP, 3µl of unlabelled dGTP, dATP and dTTP mixture (each one at 16mM), 2µl reaction mixture, 2units Klenow, and made up to 20µl with H_2O. This was incubated for 1 hour at 37°C, heated at 65°C for 5 min to stop the reaction, and passed over a G-50 column (see below) to separate unincorporated nucleotides.

6.21 WESTERN BLOT ANALYSIS

To transfer protein from an SDS polyacrylamide gel to nitocellulose, the gel was laid onto a sponge and 3MM paper cut to size all soaked in transfer buffer (25mM Tris base, 190mM aminoglycine, 20% methanol). The nitrocellulose was laid on top of the gel, followed by 3MM paper and a sponge. The sandwich was put in an electronic transfer apparatus (Hoefer Scientific Instruments) with buffer and run at 100V for 90min. The nitrocellulose was blocked for 30min at room temperature in TBS (50mM Tris pH 7.5, 200mM NaCl) with 7% BSA, rinsed in TBS then 500µl of a mouse monoclonal E1A antibody (M73) preparation was added to 20mls of RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% Deoxycholate, 0.1% SDS, 50mM Tris pH 8.0) and 1% BSA and incubated for 60 min. After washing the filter three times for 10 min each in fresh 50ml aliquots of RIPA buffer, 60µl of biotinylated antimouse rabbit IgG (Amersham) was added in 25mls of TBS and incubated for 60 min at room temperature. The filter was then washed 3 times as above after which 60µl of streptavidin peroxidase complex (Amersham) was added in 25mls of TBS and incubated with the filter for 30 min at room temperature. The filter was subsequently washed again as above and the conjugate visualised with a substrate consisting of a mix of 10µl 30% hydrogen peroxide and 12mg 4-chloro-1-naphthol in 4ml methanol.

6.22 GEL RETARDATION

The binding of proteins to DNA can be detected by forming a complex between radiolabelled DNA and protein and separating this complex from the free probe by native gel electrophoresis. The protein appears as a band of retarded migration, and binding to DNA is sequence-specific if unlabelled, non-specific competitor DNA is added to the binding reaction.

Reactions

Typical binding reactions contained 0.2ng labelled probe, 1-2µg poly [dI-dC] (Pharmacia), 20mM Tris.HCl pH7.5, 100mM NaCl, 2mM EDTA, 5mM MgCl2, 10% glycerol, 1mM DTT with lysate or extract (1-5µl) in 25µl final volume. For E2F complexes, 2µg salmon sperm DNA was used per binding reaction instead of poly dI-dC, and the following buffer: 20mM HEPES pH7.5, 100mM KCL, 0.08mM EDTA, 10% glycerol, 0.1% NP40, 5mM MgCl2, 0.5mM DTT. Any unlabelled competitor DNA (50-100 fold excess) was added before the probe; probe was always added last. Reactions were mixed and incubated for 30min, then loaded onto a non-denaturing gel.

Probe

Single-stranded oligonucleotides were annealed by mixing equal amounts of each complementary oligonucleotide and heating to 95°C for 2min, 65°C for 10min, 37°C for 10min and RT for 5min.

To fill in a 5' overhang, 100ng of oligonucleotide was mixed with 2µl 5x reverse transcription buffer (50mM Tris pH7.5, 75mM KCl, 10mM DTT, 3mM MgCl2), 1µl

10mM dGTP and dTTP, 1µl (10µCi) α-32P dCTP and dATP, and 1µl MMLV-reverse transcriptase in a 10µl reaction. This was incubated at37°C for 30min, then the volume made up to 100µl and the unincorporated nucleotides separated from the labelled probe on a G-50 sephadex column.

G-50 sephadex column: Sephadex G-50 (Pharmacia) was added to TE and left to stand overnight to let the beads hydrate. The hydrated beads were then autoclaved. A 1ml syringe (Sabre) was plugged with polymer wool (Interpet) and filled with hydrated G-50. The column was placed in a 15ml tube and centrifuged at 1.5Krpm for 4 min. Fluid in the tube was discarded and an eppendorf tube placed at the bottom of the tube. The syringe was replaced and the labeling reaction in 100µl volume placed into the syringe. This was then spun under the same conditions as before. The labelled DNA (now in the eppendorf) was stored at -20°C until required.

Non-denaturing gel

A 40ml mixture was prepared consisting of 7.2 mls acrylamide (44% acrylamide:0.8% bis), 2ml 10x TBE, 30.6ml H_2O, 40µl TEMED and 250µl 10% APS and poured into a 18cm x 20cm x 0.8mm gel apparatus. This was allowed to set for 1 hr after which samples were loaded and electrophoresed at 200V for 90min. For the visualisation of E2F-E4 complexes, gels were run at 4°C.

6.23 *IN VIVO* TRANSIENT EXPRESSION ASSAYS

Monolayer cells were grown to 30–40% confluency (3-4x10^6 cells/100mm plate) in 10ml E4 medium containing 10% foetal calf serum (FCS) at 37°C. (E4: Eagle's/Dulbecco's modified medium; Gibco). Calcium phosphate-DNA coprecipitates were prepared. To obtain fine precipitates, the DNA in 360µl TE mixed with 40µl CaCl2 was added dropwise while gently bubbling air into 400µl HBS (280 mM NaCl, 70mM HEPES pH7.1, 1.5mM Na2HPO4). Precipitates were allowed to stand at room temperature without agitation for 20min before being added dropwise to the tissue culture cells, and the cells were incubated overnight. HeLa cells were then shocked for 2min in 20% DMSO in E4 medium. After washing cells twice in PBS (150mM NaCl, 12mM phosphate buffer pH7.8), 10ml E4 (containing 10% FCS) was added. For NIH3T3 cells, DMSO shocking was not necessary; the cells were washed in PBS 2 or 3 times until the DNA precipitate could no longer be seen, then E4 (with 10% FCS) added as for HeLa cells. Cells were harvested approximately 48hr later. Prior to harvesting, cells were washed twice with PBS buffer (150mM NaCl, 12mM phosphate buffer, pH 7.8) and 0.25ml lysis buffer (0.64% NP40, 10mM Tris pH8.0, 1mM EDTA, 150mM NaCl) added. After allowing 2min for the cell walls to lyse, the lysate was transfered to a microfuge tube, spun for 1 min, and 60µl of the supernatant assayed as below.

CAT Assay

To measure the activity of chloramphenicol acetyl transferase (CAT). 60µl of extract was heated to 68°C for 5min to prevent degredation of the substrate. To this was added the assay mixture containing 4µl 2M Tris pH7.8, 20µl 8mM chloramphenicol, 20µl acetyl coenzyme A (made by diluting 14C acetyl CoA [54 mCi/mmol] with unlabelled 0.5mM stock at 1 in 10) and the reaction mixture incubated at 37°C for 3-6 hrs then chilled on ice. The acetylated chloramphenicol was extracted from the unreacted acetyl CoA by adding 130µl cold ethyl acetate, vortexing vigorously, spinning for 1min, then removing 100µl of the upper, organic phase to a scintillation vial containing 5ml Aquasol scintillant. This was counted in a scintillation counter. A blank reaction containing lysis buffer was also taken through this procedure to give the CAT background. To normalise the counts obtained, a β-galactosidase assay was used. All transfections contained 2µg of a plasmid (pJATLAC; from M.Ellis) containing the β-galactosidase coding sequence under the control of the rat β-actin promoter. 100µl extract or blank lysis buffer was added to 400µl Z buffer and 100µl ONPG (4mg/ml) and incubated until pale yellow. The reaction was stopped with 250µl Na2CO3, the tube spun for 5min and the OD420nm of the supernatant measured. Results are expressed as CAT value subtracted of CAT background/Lacz OD. Figures given are a typical example of an assay that has been repeated at least twice.

6.24 OLIGONUCLEOTIDES

Synthesised by the Oligonucleotide Synthesis Unit at Clare Hall Laboratories, ICRF.

ATF(FN)

```
5'  GATCCGGCTCCCGGGTGACGTCACGGGG      3'
3'      GCCGAGGGCCCACTGCAGTGCCCCTAGG  5'
```

3X ATF

```
5' TCGACTTAACCGTTACGTCATTTTTATTAACCGTTACGTCATTTTTAT
3'     GAATTGGCAATGCAGTAAAAATAATTGGCAATGCAGTAAAAATA

TAACCGTTACGTCATTTGCATG 3'
ATTGGCAATGCAGTAAAC     5'
```

E2F(4-1)

```
5'TCGACGTAGTTTTCGCGCTTAAATTTGAGAAAGGGCGCGAAACTAGCATG 3'
3'    GCATCAAAAGCGCGAATTTAAACTCTTTCCCGCGCTTTGATC      5'
```

6.25 PLASMID CONSTRUCTS

Yeast LacZ Reporters and Indicator strains

pB3E4, pB1E4, pB3AP1, pB1AP1, p4-1: The parent plasmid was pLG559D-312 (Guarente and Mason, 1983), a 2μ (high copy number) CYC1/LacZ reporter, with URA3 marker. The Sma1-Xho1 fragment containing the wild type CYC1 UAS was replaced with an oligonucleotide linker, and into this, oligonucleotides containing the ATF, AP1 or E2F binding sides were subsequently inserted (R. Jones).

Y4-1 and Y1-1 yeast strains: The UAS/lacz reporter was excised from the above 2μ plasmids as a Sma1-Nco1 fragment and inserted into the Stu1 site of pURA3 (the URA3 gene in Bluescript, G. Micklem). Each URA3 disruption plasmid (5μg) containing the reporter was digested with HindIII to release URA3 from the plasmid backbone, and co-transformed with a TRP1 marker plasmid (pRS314, 1μg) into Y700 by the spheroplast transfomation method. Trp+ transformants were assayed for lacz activity by on HybondN filters and blue colonies picked and purified (4-1 gave a darker blue phenotype than p1-1). Southern blot analysis was carried out on 3 colonies from each, to ensure that in the disruption of the Ura locus only one copy of the reporter had integrated. The Trp1 cotransforming plasmid was lost by growing in medium containing tryptophan.

Yeast Expression Plasmids

GAL4: pMA424 (Ma and Ptashne, 1987). This contains amino acids (aa) 1-147 of GAL4 which includes the DNA binding and dimerisation domain but has no sequences involved in activation. GAL4 is under control of the constitiutive ADH1 promoter, and the plasmid has a 2μ origin and the His3 marker.

GAL4/E1a: E1a(aa86-221) was excised from a GAL4(1-147) fusion in the mammalian vector pSG424 (Sadowski and Ptashne, 1989) as a Xho1-Bcl1 fragment, and cloned into the yeast vector pMA424 cut with Xho1 and BamH1.

GAL4/P150L: As above, but with a mutant E1a cDNA where the Pro at aa 150 has been changed to Leu.

GAL4/CBF: Lum et al, 1992 (Gal-CSSP). N-terminal 192 residues of CBF in pSG424.

All the following yeast expression plamids are based on the pRS series of vectors (Sikorski and Hieter, 1989). They have CEN sequences and pRS314 has the Trp1 marker, pRS316 the URA3 marker.

VP16/Rb: The CUP1 promoter from -246 to -1 (Butt et al., 1984) was cloned by PCR as a Kpn1-BamH1 fragment into pRS314 which had a new polylinker and CYC1 termination sequences inserted between the Kpn and Sac1 sites. Rb (aa300-928) was inserted 3' of VP16 (aa412-490) as an EcoR1-HindIII fragment and this fusion was inserted into the CUP1 plasmid.

VP16/E1a: E1a (aa21-289) was inserted as a PvuII-HindIII fragment 3' of VP16 (aa412-490). This fusion was then inserted into pRS314 containing a new polylinker between the Kpn1 and Sac1 sites and the GAL1-10 promoter and CYC1 termination sequences. A version was also made with the URA3 marker by cloning into pRS316.

pSD06 (VP16): Dalton and Triesman, 1992. The VP16 activation domain (aa412-490) in the Trp1 plasmid (pRS314) with the GAL1-10/CYC1 promoter.

LexA: LexA(aa1-201) was inserted into pSD04a (Dalton and Triesman, 1992), which is pRS314 with the GAL1-10/CYC1 promoter, a new polylinker, and CYC1 termination sequences (G.Micklem)

LexA/Rb: Rb(aa1-928) was cloned as a (HinP)f - HindIII fragment into the LexA plasmid so that Rb was 3' to the LexA coding region.(G. Micklem).

pSD09: Dalton and Triesman, 1992. SRF(aa1-412) in Trp1 plasmid containing the GAL1-10/CYC1 promoter and CYC1 3' end sequences.

SRF/CREB1 and 2: Full length CREB1(aa1-327) or CREB2(1-505) coding sequences were cloned into pSD09 by PCR using BamH1 and Xho1 sites so that CREB was 5' of SRF. (L. Clark.)

VP16/SAP1: Dalton and Triesman, 1992. SAP1 residues 1-171 (as in clone 40 from screen) in pSD06.

VP16/LexA/CREB2 (VLCR2): VP16 (412-490) as a Sal1 fragment was inserted upstream of LexA (1-201) as a Hpa1 fragment, which was inserted upstream of CREB2(1-505) in pSD04b (Dalton and Triesman, 1992; based on pRS316) so that the ORF was driven by the GAL1-10/CYC1 promoter and had CYC1 3' end sequences. (L. Clark).

VP16/CREB2 (VCR2): The LexA coding sequences from plasmid VLCR2 were removed as a Hpa1 fragment, and upon recircularisation, CREB2 sequences were in the same ORF as VP16 sequences. (L.Clark)

LexA/CREB2 (LCR2): The VP16 coding sequences from plasmid VLCR2 were removed as a Sal1 fragment, and upon recircularisation, CREB2 sequences were in the same ORF as LexA sequences. (L.Clark)

LexA/E1a: E1a CR3(aa139-189) was cloned by PCR into the LexA plasmid via the BamH1 and Sal1 sites, so that E1a was 3' to the LexA coding region.

E4/SRF: E4 (aa1-150) was cloned by PCR as a BamH1-Xho1 fragment, and inserted into pSD09 so that E4 was 5' to SRF.

E4D/, E4DH/, E4D2H/SRF: As E4/SRF, but with different PCR primers at the 3' end of E4 so that deletions were made of 10, 35 and 56 residues respectively.

VP16/RBP3: RBP3(aa24-437) was cloned into pSD06 as an EcoR1- Xho1 fragment so that VP16 was upstream of RBP3 sequences (R.Fagan).

VP16-tagged *S. cerevisiae* cDNA library: G. Micklem, unpublished. The cDNA was inserted into pSD06 with polylinker so that the cDNA was 3' to VP16.

In Vitro Translation Plasmids

pT7 was as Dalton and Triesman, 1992 (pT7βplink). pT7T was the same vector, with a myc tag consisting of a 14 residue insertion encoding the 9E10 epitope of myc inserted 5' of the polylinker. E4 was inserted into T7 and T7T as an EcoR1-Xho1 fragment. cDNA clones D2 and 29 and the yA clones were inserted into the T7 vectors as BamH1-Spe1 fragments which included the VP16 domain.

Mammalian Expression Plasmids

MLV plasmids as Dalton and Triesman, 1992. The same polylinker as T7 vector is present downstream of the enhancer element of the MLV LTR (MLV-E) and the promoter and 5' non-coding sequences of the human β-globin gene (β-UT). Constructs from the cDNA library or T7 vectors could be transferred as VP16 fusions with BamH1 and Spe1. D25 and RBP3 have internal BamH1 sites, so were cloned without the VP16 portion as EcoR1-Spe1 fragments into an MLV plasmid already containing VP16.

E1a(pCE): (Schneider et al., 1987).

GAL4: pSG424 (Sadowski and Ptashne, 1989)

GAL4/E4: E4(aa1-150) was cloned into pSG424 via the EcoR1 and Xba1 sites so that E4 was 3' to the GAL4 DNA binding domain.

Mammalian CAT reporter plasmids

E4-CAT: Promoter sequences from the Ad E4 gene upstream of the CAT gene. (Lillie and Green 1989).

E2A-CAT(-97), (-59): pEC series (Imperiale et al., 1985). 5' E2A promoter deletion derivatives cloned into pCAT3M.

MLP-CAT: Ad major late promoter upstream of the CAT gene (C. Goding).

G5E4-CAT: Flint and Jones,1991. 5 copies of the GAL4 binding site upstream of the TATA element and flanking sequences from the E4 promoter.

REFERENCES

Abe, A., Wada, T., Handa, H., Nogi, Y. and Fukasawa, T. (1988). Agric Biol Chem **52:** 2035-2041.

Allison, L. A., Wong, J. K., Fitzpatrick, V. D., Moyle, M. and Ingles, C. (1988). Mol Cell Biol **8:** 321-329.

Bagchi, S., Raychaudhuri, P. and Nevins, J. R. (1989). Proc Natl Acad Sci U S A **86:** 4352-6.

Bagchi, S., Raychaudhuri, P. and Nevins, J. R. (1990). Cell **62:** 659-69.

Benbrook, D. M. and Jones, N. C. (1990). Oncogene **5:** 295-302.

Berger, S. L., Cress, W.D., Cress, A., Triezenberg, S. J. and Guarente, L. (1990). Cell **61:** 1199-1208

Berger, S. L., Pina, B., Silverman, N., Marcus, G. A., Agapite, J., Regier, J. L.,

Triezenberg, S. J. and Guarente, L. (1992). Cell **70:** 251-65.

Blanar, M. A. and Rutter, W. J. (1992). Science **256:** 1014-1018.

Brandl, C. J. and Struhl, K. (1989). Proc Natl Acad Sci USA **86:** 2652-2656.

Breeden, L. and Nasmyth, K. (1985). Cold Spring Harbour Symp Quant Biol **50:** 643-650.

Breeden, L. and Nasmyth, K. (1987). Cell **48:** 729-736.

Brent, R. and Ptashne, M. (1985). Cell **43:** 729-36.

Buratowski, S., Hahn, S., Sharp, P. A. and Guarente, L. (1988). Nature **334:** 37-42.

Butt, T. R., Strenberg, E., Gorman, J., Clark, P., Hamer, F., Rosenberg, M. and Crooke, S. (1984). Proc Natl Acad Sci USA **81:** 3332-3336.

Chellappan, S.P., Hiebert, S. Mudryj, M., Horowitz, J.M. and Nevins, J.R. (1991). Cell **65:** 1053-61

Chien, C. T., Bartel, P. L., Sternglanz, R. and Fields, S. (1991). Proc Natl Acad Sci U S A **88**: 9578-82.

Chittenden, T., Livingston, D. M. and Kaelin, W. J. (1991). Cell **65**: 1073-82.

Chodosh, L. A., Olesen, J., Hahn, S., Baldwin, A. S., Guarente, L. and Sharp, P. A. (1988). Cell **53**: 25-35.

Comai, L., Tanese, N. and Tjian, R. (1992). Cell **68**: 965-976.

Courey, A. J. and Tjian, R. (1988). Cell **55**: 887-898.

Cousens, D. J., Greaves, R., Goding, C. R. and O'Hare, P. (1989). EMBO J **8**: 2337-42.

Culp, J. S., Webster, L. C., Friedman, D., Smith, C., Huang, W.-J., Wu, F., Rosenberg, M. and Ricciardi, R. P. (1988). Proc Natl Acad Sci USA **85**: 6450-6454.

Dalton, S. and Treisman, R. (1992). Cell **68**: 597-612.

Devoto, S. H., Mudryj, M., Pines, J., Hunter, T. and Nevins, J. R. (1992). Cell **68**: 167-76.

Dirick, L., Moll, T., Auer, H. and Nasmyth, K. (1992). Nature **357**: 508-13.

Dynan, W.S. (1989). Cell **58**: 1-4

Dyson, N. and Harlow, E. (1992). Cancer Surveys **12**: 161-194.

Eisenmann, D. M., Dollard, C. and Winston, F. (1989). Cell **58**: 1183-91.

Emami, K.H. and Carey, M. (1992). EMBO J **11**: 5005-12

Evans, R. M. and Hollenberg, S. M. (1988). Cell **52**: 1-3.

Ewen, M., Xing, Y. J. L. and Livingstone, D. (1991). Cell **66**: 1155-1164.

Felsenfeld, G. (1992). Nature **355**: 219-224.

Fields, S. and Jang, S. K. (1990). Science **249**: 1046-1049.

Fields, S. and Song, O., (1989). Nature **340**: 245-6.

Fikes, J. D., Becker, D. M., Winston, F. and Guarente, L. (1990). Nature **346**: 291-4.

Flint, K. J. and Jones, N. C. (1991). Oncogene **6**: 2019-26.

Foulkes, N.S., Borrelli, E. and Sassone-Corsi, P., (1991). Cell **64**: 739-49

Gill, G. and Ptashne, M. (1987). Cell **51**: 121-6.

Gill, G. and Ptashne, M. (1988). Nature **334**: 721-4

Giniger, E. and Ptashne, M. (1987). Nature **330**: 670-2.

Giniger, E., Varnum, S. M. and Ptashne, M. (1985). Cell **40**: 767-74.

Gonzalez, G. A. and Montminy, M. R. (1989). Cell **59**: 675-680.

Green, M., Loewenstein, P. M., Pusztai, R. and Symington, J. S. (1988). Cell **53**: 921-926.

Guarente, L. and Birmingham-McDonogh, O. (1992). TIG **8**: 27-32.

Guarente, L. and Mason, T. (1983). Cell **32**: 1279-1286.

Guillemot, F., Billault, A. and Auffray, C. (1989). Proc Natl Acad Sci USA **86**: 4594-4598.

Hahn, S., Buratowski, S., Sharp, P. A. and Guarente, L. (1989a). Embo J **8**: 3379-82.

Hahn, S., Buratowski, S., Sharp, P. A. and Guarente, L. (1989b). Cell **58**: 1173-81.

Hai, T. and Curran, T. (1991). Proc Natl Acad Sci U S A **88**: 3720-4.

Hai, T. W., Liu, F., Coukos, W. J. and Green, M. R. (1989). Genes Dev **3**: 2083-90.

Handa, H., Toda, T., Tajima, M., Wada, T., Iida, H. and Fukasawa, T. (1987). Gene **58**: 127-36.

Harlow, E., Franza, B. J. and Scholey, C. (1985). J Virology **55**: 533-546.

Harlow, E., Whyte, P., Franza, B. R. and Schley, C. (1986). Mol Cell Biol **6**: 1579-1583.

Hartley, D. A., Preiss, A. and Artavanis, T. S. (1988). Cell **55**: 785-95.

Helin, K., Lees, J. A., Vidal, M., Dyson, N., Harlow, E. and Fattaey, A. (1992). Cell **70**: 337-350.

Herr, W., Sturm, R. A., Clerc, R. G., Corcoran, L. M., Baltimore, D., Sharp, P. and al, e. (1988). Genes and Dev **2**: 1513-1516.

Hiebert, S. W., Blake, M., Azizkhan, J. and Nevins, J. R. (1991). J Virol **65**: 3547-52.

Himmelfarb, H. J., Pearlberg, J., Last, D. H. and Ptashne, M. (1990). Cell **63**: 1299-309.

Hipskind, R. A., Rao, V. N., Mueller, C. G., Reddy, E. S. and Nordheim, A. (1991). Nature **354**: 531-4.

Hoeffler, W. K., Kovelman, R. and Roeder, R. G. (1988). Cell **53**: 907-20.

Hoffman, A., Sinn, E., Yamamoto, T., Wang, J., Roy, A., Horikoshi, M. and Roeder, R. G. (1990). Nature **346**: 387-90.

Horikoshi, M., Carey, M.F., Kakidani, H. and Roeder, R.G. (1988a). Cell **54**: 665-9

Horikoshi, M., Hai, T., Lin, Y.S., Green M.R. and Roeder, R.G. (1988b). Cell **54**: 1033-42.

Horikoshi, M., Maguire, K., Kralli, A., Maldonado, E., Reinberg, D. and Weinman, R. (1991). Proc Natl Acad Sci USA **88**: 5124-8

Huang, M. M. and Hearing, P. (1989). Genes Dev **3**: 1699-710.

Imai, Y., Matsushima, Y., Sugimura, T. and Terada, M. (1991). J Virology **65**: 4966-4972.

Imperiale, M. J., Hart, R. P. and Nevins, J. R. (1985). Proc Natl Acad Sci USA **82**: 381-385.

Ingles, C. J., Shales, M., Cress, W. D., Triezenberg, S. J. and Greenblatt, J. (1991). Nature **351**: 588-90.

Inostroza, J., Mermelstein, F. H., Ha, I., Lane, W. S. and Reinberg, D. (1992). Cell **70**: 477-89.

Jones, N. (1990a). Cell **61**: 9-11.

Jones, R. H. (1990b). PhD Thesis. University College London.

Jones, R. H. and Jones, N. C. (1989). Proc Natl Acad Sci U S A **86**: 2176-80.

Kaelin, W. G., Krek, W., Sellers, W. R., DeCaprio, J. A., Ajchenbaum, F., Fuchs, C. S.,

Chittenden, T., Li, Y., Farnham, P. J., Blanar, M., Livingstone, D. M. and Flemington, E. K. (1992). Cell **70**: 351-364.

Kakidani, H. and Ptashne, M. (1988). Cell **52**: 161-7.

Kao, C. C., Lieberman, P. M., Schmidt, M. C., Zhou, Q., Pei, R. and Berk, A. J. (1990). Science **248**: 1646-1650.

Keleher, C. A., Red, M. J., Schultz, J., Carlson, M. and Johnson, A. D. (1992). Cell **68**: 709-719.

Kelleher, R.J., Flanagan, P.M. and Kornberg, R.D. (1990). Cell **61**: 1209-15

Kraus, V. B., Moran, E. and Nevins, J. R. (1992). Mol Cell Biol **12**: 4391-4399.

Kristie, T. M. and Sharpe, P. A. (1990). Genes Dev **4**: 2383-2396.

Lambert, P. L., Dostatni, N., McBride, A. A., Yaniv, M., Howley, P. M. and Arcangoli, B. (1989). Genes Dev **3**: 38-48.

Landschulz, W. H., Johnson, P. F. and McKnight, S. L. (1988). Science **240**: 1759-1764.

Lech, K., Anderson, K. and Brent, R. (1988). Cell **52**: 179-184.

Lee, K. A., Hai, T. Y., SivaRaman, L., Thimmappaya, B., Hurst, H. C., Jones, N. C. and Green, M. R. (1987). Proc Natl Acad Sci USA **84**: 8355-8359.

Lee, W.S., Kao, C.S., Bryant, G.O., Liu, X. and Berk, A.J. (1991). Cell **67:** 365-76

Lees, E., Faha, B., Dulic, V., Reed, S. I. and Harlow, E. (1992). Genes Dev **6:** 1874-1885.

Leuther, K. K. and Johnston, S. A. (1992). Science **256:** 1333-5.

Lillie, J. W. and Green, M. R. (1989). Nature **338:** 39-44.

Lillie, J. W., Loewenstein, P. M., Green, M. R. and Green, M. (1987). Cell **50:** 1091-1100.

Lin, Y.-S., Ha, I., Maldonado, E., Reinberg, D. and Green, M. R. (1991). Nature **353:** 569-571.

Lin, Y. S. and Green, M. R. (1989). Proc Natl Acad Sci U S A **86:** 109-13.

Lin, Y. S. and Green, M. R. (1991). Cell **64:** 971-81

Liu, F. and Green, M. R. (1990). Cell **61:** 1217-24.

Loeken, M. R. and Brady, J. (1989). J Biol Chem **264:** 6572-9.

Long, R. M., Mylin, L. M. and Hopper, J. E. (1991). Mol Cell Biol **11:** 2311-4.

Lum, L. S., Hsu, S., Vaewhongs, M. and Wu, B. (1992). Mol Cell Biol **12:** 2599-605.

Ma, J. and Ptashne, M. (1987). Cell **51:** 113-119.

Ma, J. and Ptashne, M. (1988). Cell **55:** 443-6.

Maekewa, T., Matsuda, S., Fujisawa, J., Yoshida, M. and Ishii, S. (1991). Oncogene **6:** 627-32

Maguire, H. F., Hoeffler, J. P. and Siddiqui, A. (1991). Science **252:** 842-4.

Maniatis, T., Fritsh, E. F. and Sambrook, J. (1982). Cold Spring Harbour Laboratory Press .

Martin, K. J., Lillie, J. W. and Green, M. R. (1990). Nature **346:** 147-52.

Mermod, N., O'Neill, E. A., Kelly, T. J. and Tjian, R. (1989). Cell **58**: 741-753.

Metzger, D., White, J. H. and Chambon, P. (1988). Nature **334**: 31-6.

Miller, J., McLachlan, A. and Klug, A. (1985). EMBO J **4**: 1609-1614.

Moran, E. and Matthews, M. B. (1987). Cell **48**: 177-178.

Moye-Rowley, W., Harshman, K. D. and Parker, C. S. (1989). Genes Dev **3**: 283-92.

Mueller, C. G. and Nordheim, A. (1991). EMBO J **10**: 4219-29.

Murre, C., McCaw, P., Vaessin, H., Caudy, M., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauachka, S. D. and Lassar, A. B. (1989). Cell **58**: 537-544.

Nehlin, J. O., Carlberg, M. and Ronne, H. (1992). Nuc Ac Res **20**: 5271-5278.

Neill, S. D. and Nevins, J. R. (1991). J Virol **65**: 5364-73.

Nevins, J. R. (1989). Adv in Vir Res **37**: 35-83.

Newman, A. and Norman, C. (1991). Cell **65**: 115-123.

O'Connor, R. J. and Hearing, P. (1991). Nucleic Acids Res **19**: 6579-86.

O'Shea, E. K., Rutkowski, R. and Kim, P. S. (1992). Cell **68**: 699-708.

O'Shea, E. K., Rutkowski, R., Stafford, W. F. and Kim, P. S. (1989). Science **245**: 646-648.

Oehler, T. and Angel, P. (1992). Mol Cell Biol **12**: 5508-5515.

Oliviero, S., Robinson, G., Struhl, K. and Spiegelman, B. S. (1992). Genes Dev **6**: 1799-1809.

Oliviero, S. and Struhl, K. (1991). Proc Natl Acad Sci U S A **88**: 224-8.

Patel, G. and Jones, N. C. (1990). Nuc Ac Res **18**: 2909-2915.

Pei, R. and Berk, A. J. (1989). J Virol **63**: 3499-506.

Pinto, I., Ware, D. E. and Hampsey, M. (1992). Cell **68**: 977-88.

Ptashne, M. (1988). Nature **335**: 683-9.

Rasmussen, R., Benvegnu, D., O'Shea, E. K., Kim, P. S. and Alber, T. (1991). Proc Natl Acad Sci U S A **88**: 561-4.

Raychaudhuri, P., Bagchi, S., Devoto, S. H., Kraus, V. B., Moran, E. and Nevins, J. R. (1991). Genes Dev **5**: 1200-11.

Raychaudhuri, P., Bagchi, S. and Nevins, J. R. (1989). Genes Dev **3**: 620-7.

Richter, J.D., Slavicek, J.M., Schneider, J.F., Jones, N.C. (1988). J Virol 62: 1948-1955.

Rochette-Elgy, C., Fromental, C. and Chambon, P. (1990). Genes Dev **4**: 137-150.

Sadowski, I. and Ptashne, M. (1989). Nuc Ac Res **17**: 7539.

Schneider, J. F., Fisher, F., Goding, C. R. and Jones, N. C. (1987). EMBO J. **6**: 2053-2060.

Scholer, H. R., Ciesiolka, T. and Gruss, P. (1991). Cell **66**: 291-304.

Shan, B., Zhu, X., Chen, P., Durfee, T., Yang, Y., Sharp, D. and Lee, W.-H. (1992). Mol Cell Biol **12**: 5620-5631.

Shepherd, J. C., McGinnis, W., Carrasco, A. E., DeRobertis, E. M. and Gehring, W. J. (1984). Nature **310**: 70-71.

Sikorski, R. S. and Hieter, P. (1989). Genetics **122**: 19-27.

Stringer, K. J., Ingles, C. J. and Greenblatt, J. (1990). Nature **345**: 783-786.

Struhl, K. (1988). Nature **332**: 649-50.

Tanese, N. B., Pugh, B. F. and Tjian, R. (1991). Genes Dev. **5**: 2212-2224.

Tasset, D., Tora, L., Fromental, C., Scheer, E. and Chambon, P. (1990) Cell **62**: 1177-87

Taylor, I. C. and Kingston, R. E. (1990). Mol Cell Biol **10:** 176-83.

Thomas, B. J. and Rothstein, R. (1989). Cell **56:** 619-630.

Tremblay, M.L., Dumont, D.J. and Branton, P.E. (1989). Virology 169: 397-407

Vincent, A. C. and Struhl, K. (1992). Mol Cell Biol **12:** 5394-5405.

Webster, L. C. and Ricciardi, R. P. (1991). Mol Cell Biol **11:** 4287-96.

Whiteway, M., Hougan, L., Dignard, D., Thomas, D. Y., Bell, L., Saari, G., Grant, F., O'Hara, P. and MacKay, V. L. (1989). Cell **56:** 467-477.

Whyte, P., Buchkovich, K., Horowitz, J. M., Friend, S. H., Raybuck, M., Weinberg, R. A. and Harlow, E. (1988). Nature **334:** 124-129.

Williams, F. E. and Trumbly, R. J. (1990). Mol Cell Biol **10:** 6500-6511.

Workman, J. L., Taylor, I. C. and Kingston, R. E. (1991). Cell **64:** 533-44.

Yee, A. S., Raychaudhuri, P., Jakoi, L. and Nevins, J. R. (1989). Mol Cell Biol **9:** 578-85.

Zawel, L. and Reinberg, D. (1992). Curr Opin Cell Biol **4:** 488-95.