FUNCTIONAL CHARACTERIZATION OF THE ATF2 TRANSCRIPTION FACTOR

CATHERINE MARY LIVINGSTONE

A thesis submitted for the degree of
Doctor of Philosophy
March, 1995

University College
Gower Street
London

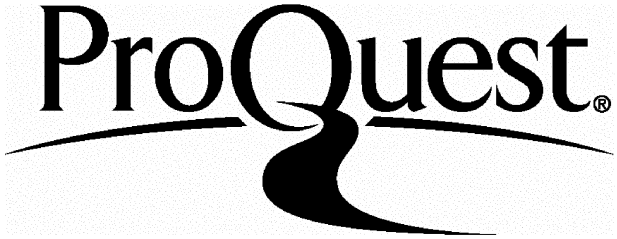
Imperial Cancer Research Fund
Lincoln's Inn Fields
London

ProQuest Number: 10055346

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.



ProQuest 10055346

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

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

ABSTRACT

ATF2 (CRE-BP1, CREB2) is a member of the bzip family of transcription factors which are characterized by their ability to bind to a CRE/ATF site. At least ten different mammalian cDNAs of this family have been cloned and the transcriptional and regulatory properties of some members determined by transient transfection assays. Such studies have implicated ATF2 in transcriptional activation by the adenovirus E1a protein.

Previous mutational analysis of ATF2 revealed the importance of an N-terminal region for activation by E1a. In this thesis I have carried out a more detailed investigation of this region and found that a segment consisting of amino acid residues 19-96 can act as a very potent transactivation domain when fused to a heterologous DNA binding moiety. This strong activation is seen in the absence of E1a and is not greatly enhanced by E1a. Since this domain does not activate transcription in the context of the full length protein, it appears to be "masked" in some, as yet, unknown way.

The activation domain contains a zinc finger and three potential MAPK sites at positions 69, 71 and 90. Mutating any of these motifs reduces the activity of the domain, with the threonine residues at positions 69 and 71 being the most critical for its function. The use of serum/UV stimulation experiments and MAPK site mutants has shown that phosphorylation of this domain is important for regulating its activity *in vivo* and therefore suggest that ATF2 plays a role in signal transduction.

The features required for activation by the N-terminus of ATF2 (zinc finger and phosphorylation sites) are also important for E1a activation of transcription via full length ATF2. The implications of these results with regard to a model for E1a transactivation are presented in this thesis.

TO MY PARENTS

ACKNOWLEDGEMENTS

I want to say a huge thank you to all who have helped and encouraged me during the course of this PhD. Biggest thanks must go to my supervisor, Nic Jones, for believing in my ability to carry out these studies and for helping me to do so. Thanks must also go to the past and present members of the Jones lab, especially my bay buddies Yan, Rick, Kate and Alex for the entertainment and friendship. I would also like to thank members of the Treisman lab for being an amazing source of knowledge and for always being willing to help. Finally within ICRF I would like to thank the photography department and the Oligonucleotide Synthesis Unit for their services.

I also want to thank all those friends who have helped me maintain some degree of sanity over the last four years; all the regulars at the George, in particular Nuz who could always be persuaded to have the odd drink or two, everyone from 113 St. Thomas's Rd, Grete-Marie and especially Christelle without whom I would never have coped.

TABLE OF CONTENTS

ABBREVIATIONS

APS	Ammonium persulphate
ATP	Adenosine triphosphate
bp	Base pair
BSA	Bovine serum albumin
cAMP	3', 5' cyclic adenosine monophosphate
CAT	Chloramphenicol acetyl transferase
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
HCl	Hydrochloric acid
HEPES	N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulphonic acid]
hr	Hours
IPTG	Isopropyl β-D-thiogalactopyranoside
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
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
Tween20	Polyoxyethylenesorbitan monolaurate
UTP	Uridine triphosphate
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
H	Histidine
I	Isoleucine
K	Lysine

L	Leucine
M	Methionine
N	Asparagine
P	Proline
Q	Glutamine
R	Arginine
S	Serine
T	Threonine
V	Valine
W	Tryptophan
Y	Tyrosine

CHAPTER 1:

INTRODUCTION

INTRODUCTION

Cell growth, development, differentiation and senescence are highly organized processes controlled by the pattern of gene expression within the cell. In eukaryotes, such regulation of gene expression occurs predominantly at the level of transcription as is evidenced by the estimate that up to 10% of human genes may encode transcription factors (Kingston and Green, 1994). The importance of correct gene expression to a cell becomes apparent in cases when such regulation breaks down or is altered e.g. in the disease cancer or during a viral infection. Thus the mechanism of gene transcription has been widely studied in order to try and understand the process and so develop strategies to protect against or correct faulty gene expression.

In this introduction I shall present the current model of eukaryotic transcription initiation and discuss the way in which this process is thought to be controlled. I shall also discuss the two proteins studied in this thesis, namely the transcription factor ATF2 and the viral protein E1a, and their roles in transcription. Finally, I shall briefly describe how transcription can be regulated by extracellular signals and also outline the reasons for embarking on this project and the information hoped to be gained from it.

1.1 Eukaryotic Transcription and Its Complexity

Cells synthesize three types of RNA; (i) ribosomal (ii) messenger and several small nuclear RNAs (iii) transfer, 5S and various other small cellular RNAs. These are synthesized from class I, II and III genes respectively (reviewed in Hernandez, 1993) by mechanisms which differ in prokaryotes and eukaryotes, the latter utilizing more complex processes. The first difference between the two types of cells is in the nature of the RNA polymerase used. In prokaryotes all three classes of genes are transcribed by a single enzyme, RNA polymerase (RNAP) holoenzyme. However, in eukaryotes three distinct RNA polymerases, RNAP I, II and III, are required for the synthesis of the three classes of RNA. The second level of complexity in eukaryotic transcription is seen in the process of initiation; bacterial RNAP holoenzyme is able to recognize specific promoter sequences and synthesize RNA from transcriptional start sites on its own whereas the eukaryotic RNA polymerases require a set of general transcription factors in order to achieve such sequence specific transcription.

During the past few years much progress has been made in understanding the complex nature of eukaryotic transcription and many of the general transcription factors involved have been identified, purified and the genes encoding the proteins cloned. This has enabled *in vitro* transcription systems to be reconstituted using either recombinant or purified factors and so allowed the role of the individual proteins to be studied. The results from such studies are presented below with most of the data discussed concerning RNAP II transcription since this is the process investigated in this thesis.

1.1.1 Structure of Class II Promoters

The promoters of class II genes contain sequences which position the start site of transcription (core promoter elements) as well as binding sites for regulatory proteins which control the level of transcription. Of the core promoter elements there are two types, the TATA box and the initiator (Inr) element, and one or both of these elements are present in all class II promoters (reviewed in Gill, 1994). The TATA box has the consensus sequence TATAA and is found approximately 30 base pairs upstream of the transcription start site in higher eukaryotes. It functions to bind the TATA-binding protein (TBP), the first step in the formation of the transcription initiation complex (see section 1.1.2). The initiator element is a sequence motif encompassing the transcription start site which directs transcription to a single start site in the absence of a TATA box. Its role is not well understood and although a consensus sequence has been derived (PyPyANT/APyPy where A is +1) (Javahery et al., 1994) it is not clear if there is a single initiator binding protein or whether several different proteins are able to bind depending on the exact sequence of the initiator as well as the promoter context e.g. TBP may bind via an intermediary protein(s) to the Inr element from the terminal deoxynucleotidyl transferase (TdT) gene (Zhou et al., 1992) whereas the transcription factor YY1 seems to bind to the Inr from the adeno-associated virus (AAV) P5 promoter (Usheva and Shenk, 1994). The interplay between the TATA box and the Inr in promoters containing both elements is not understood although it has been determined for the combination of TATA box and TdT Inr that the TATA box is dominant for determining the start site of transcription and the Inr functions to increase the level of transcription (O'Shea-Greenfield and Smale, 1992).

The regulatory regions of promoters are generally classified according to their position relative to the core elements (Dynan, 1989). Promoter proximal elements consist of one or more binding sites for regulatory (activator/repressor) proteins and are found at positions relatively close to the core promoter. Enhancers, on the other hand, are located many hundreds of base pairs up or downstream of the gene but again contain distinct sets of binding sites for transcription factors (transcriptional activators). Variations in the arrangement and numbers of these sites provides the potential for regulating the amount of activation. In addition, negative regulatory elements which contain binding sites for transcriptional repressors may also be present in the upstream gene sequences (Cowell, 1994). Thus the balance between the number of activators and repressors will determine the level of transcription from a given promoter.

1.1.2 Basal Transcription by RNA Polymerase II

Basal transcription occurs in the absence of regulatory elements, both positive and negative, and is supported by the core promoter (TATA box and/or Inr) which binds the general transcription factors (TBP, RNAP II, TFII-A, B, E, F, H, J) so forming an initiation complex. The assembly of the RNAP II initiation complex on TATA containing promoters is

now fairly well defined and is shown schematically in figure 1.1 (reviewed in Zawel and Reinberg, 1992; Zawel and Reinberg, 1993; Buratowski, 1994). The first step is binding of TFIID (TBP and TBP associated factors, TAFs, see section 1.1.3 (i)) to the TATA box to form an initial committed complex. Binding occurs via the TBP subunit which is the only general transcription factor with sequence specific DNA-binding activity. This complex is stabilized by binding of TFIIA; TFIIA's three subunits bind co-operatively to TBP and so enhance its binding to DNA. TFIIA is required for the stimulation of transcription via TFIID, at high concentrations TFIIA enhances basal transcription and at lower concentrations it enhances activated transcription. TFIIA is thought to achieve this stimulation by three mechanisms. Firstly, by stimulating TFIID binding. Secondly, by removing repressors from the TFIID complex and thirdly by actively stimulating transcription initiation via an unknown mechanism involving the TAFs (Ozer et al., 1994; Sun et al., 1994; Yokomori et al., 1994).

The next component to bind to the complex is TFIIB, binding to either TBP-DNA or TFIIA-TBP-DNA and recruiting RNAP II associated with TFIIF to give rise to the minimal initiation complex. The order in which the TBP-TFIIB-RNAP II interactions occurs is unclear; it may be that TFIIB is associated with RNAP II before it enters into the initiation complex or that TFIIB binds to TBP before recruiting RNAP II. However, either way a stable complex is assembled which under certain conditions is able to initiate transcription. Although this minimal initiation complex can initiate transcription, the factors TFIIE and TFIIH seem to be required for the synthesis of longer RNA transcripts (Goodrich and Tjian, 1994). Thus binding of TFIIE to the minimal complex and the subsequent binding of TFIIH, followed by TFIIJ binding, results in the formation of a complete initiation complex. Activation of this complex involves the hydrolysis of ATP or dATP distinct from that involved in the RNA polymerization process and is assumed to be carried out by TFIIH since this factor possesses a DNA-dependent ATPase activity. TFIIH also possesses a DNA helicase activity and so can contribute to the partial unwinding of the DNA template necessary to create the open elongation-competent complex. It also possesses a kinase activity which is capable of phosphorylating the repetitive C-terminal domain (CTD) of the largest subunit of RNAP II (for review on TFIIH see Drapkin and Reinberg, 1994). This seems to be important for modulating the association of the polymerase with the complex since the non-phosphorylated form of RNAP II is associated with TBP whereas phosphorylated RNAP II is found in the elongation stage of the reaction. Thus, it is proposed that phosphorylation of the CTD is a critical step in uncoupling RNAP II from the initiation complex so allowing the onset of transcription and elongation (Peterson and Tjian, 1992).

The above model applies to formation of the initiation complex at a TATA box and a similar model is thought to apply to TATA-less promoters. However, in this situation the Inr binding protein(s) are proposed to interact with one of the general transcription factors so enabling the remaining components of the complex to be brought to the promoter by a series of protein-protein interactions (Gill, 1994; Goodrich, 1994).

Figure 1.1 Schematic Diagram of the Assembly of the RNAPII Initiation Complex on a TATA Containing Promoter (Adapted From Buratowski, 1994)

Components of the initiation complex assemble onto the promoter as shown in the diagram. Basal transcription factors are represented by their letter designations;

D: TFIID

A: TFIIA

B: TFIIB

F: TFIIF

E: TFIIE

H: TFIIH

RNAPII: RNA polymerase

Double-headed arrows indicate protein-protein interactions.

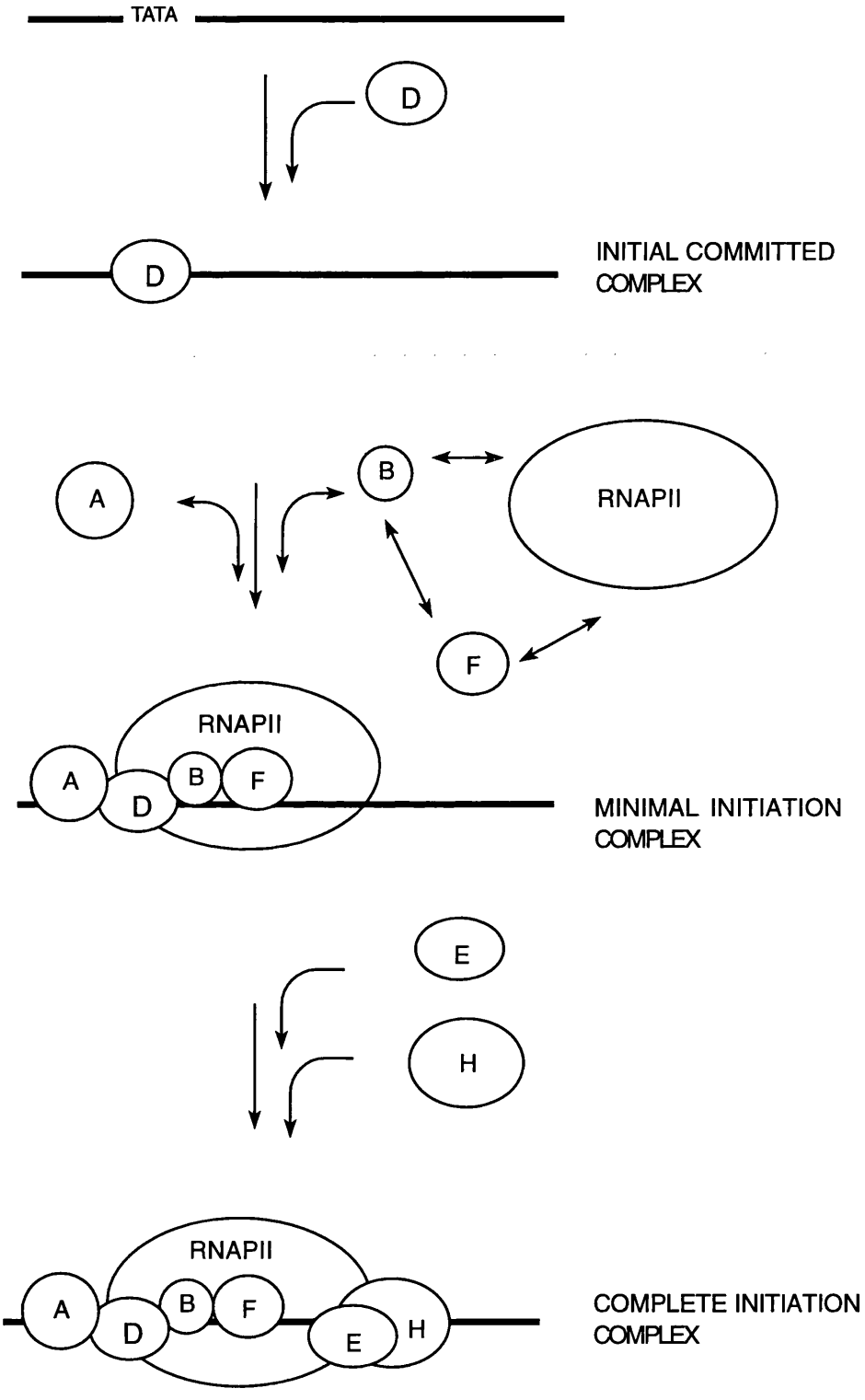


Figure 1.1 Schematic diagram of the assembly of the RNAPII initiation complex on a TATA containing promoter

1.1.3 The Many Roles of TBP

TBP is a 38kD protein which *in vivo* is complexed with a variety of other subunits or TBP-associated factors (TAFs). The association of TBP with different groups of TAFs produces four distinct units; TFIID, SL1, TFIIIB and SNAPc and each plays a unique role in eukaryotic transcription which TBP is unable to carry out alone (reviewed in Hernandez, 1993; Goodrich, 1994). TFIID is responsible for regulated, as opposed to basal, class II gene transcription; SL1 is essential for transcription of class I genes; TFIIIB is required for transcription of TATA-less class III genes; and SNAPc plays a role in transcription of both class II and III genes since it is involved in the synthesis of the small nuclear RNAs: U6 snRNA, a class III gene and U1 and U2 snRNA, class II genes. Since TBP is common to all these processes, it must be the TAF proteins which provide the specificity of function to these complexes presumably by interacting with proteins and DNA elements unique to the various different promoters. Specific TAFs must also interact differentially with the components of the RNAP I, II and III transcription machinery and modulate the activity of TBP e.g. TBP complexes which bind to class I and III TATA-less promoters (SL1 and TFIIIB respectively) must reduce the affinity of TBP in the complex for TATA boxes in order to prevent their binding to class II TATA-containing promoters. The recent cloning of many of the TAF proteins enables such roles for the TAFs to be determined and the various TBP-TAF complexes are discussed below:

(i) TFIID

Recombinant TBP is not able to support regulated i.e. activated or repressed RNAP II transcription in *in vitro* assays. However, TBP isolated from cells by immunoprecipitation with an anti-TBP antibody is complexed with at least 8 TAF proteins, the TFIID complex , and is capable of supporting such regulation (Dynlacht et al., 1991; Zhou et al., 1992). This shows the importance of the TAFs for TFIID function. So far, all of the 8 (dTAF$_{II}$-250, 150, 110, 80, 60, 40, 30α and 30β) Drosophila TFIID TAFs (Dynlacht et al., 1993; Goodrich et al., 1993; Hoey et al., 1993; Kokubo et al., 1993; Kokubo et al., 1993a; Kokubo et al., 1993b; Weinzierl et al., 1993a; Weinzierl et al., 1993b; Yokomori et al., 1993; Zhou et al., 1993; Kokubo et al., 1994; Verrijzer et al., 1994) and 3 (hTAF$_{II}$-250, 70 and 30) of the human homologs (Hisatake et al., 1993; Ruppert et al., 1993; Weinzierl et al., 1993b; Jacq et al., 1994) have been cloned and the human homologs (hTAF$_{II}$-130 and 100) have been partly characterized (Chen et al., 1994). Splice variants of some of these proteins have been identified and it is thought that cell type specific TAFs may also exist. These variations may increase the number of regulatory signals to which TFIID can respond.

TFIID is remarkably stable due to the numerous TAF-TAF and TAF-TBP interactions which occur in the complex. A model showing some of these interactions is shown in figure 1.2. Not all the TAFs interact with TBP directly but are held in the complex via interaction with other TAFs. As well as interacting with the general transcription machinery, TAFs have also

Figure 1.2 Schematic Diagram of the Drosophila TFIID Complex (Adapted from Chen et al., 1994)

Drosophila TFIID consists of dTBP and at least eight TAFs ($dTAF_{II}$-250, 150, 110, 80, 60, 40, 30α and 30β) associated with each other to form a multiprotein complex. The diagram indicates the TBP-TAF and TAF-TAF interactions identified so far and gives a picture of the structure of TFIID based on this data.

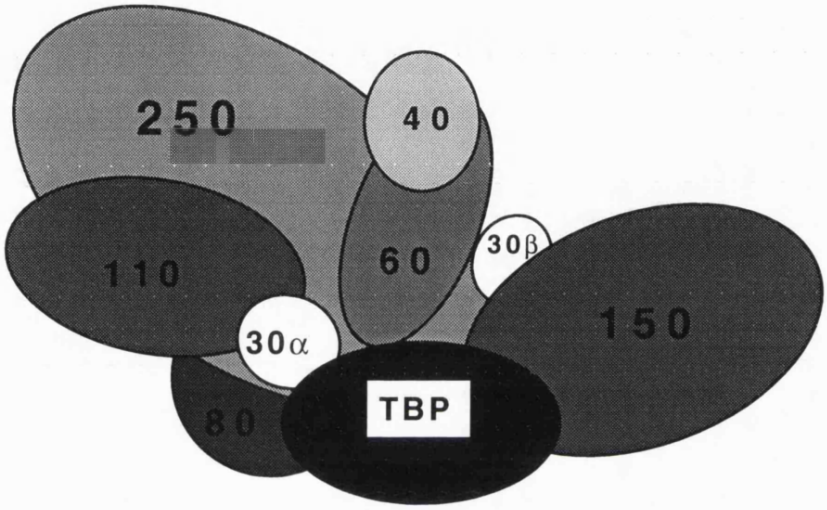


Figure 1.2 Schematic diagram of the Drosophila TFIID complex

been shown to interact with regulatory proteins. Interactions with transcriptional activators are the best documented and suggest that TAFs act as co-activators transmitting information from the upstream regulatory sites to the basal transcription machinery (Chen et al., 1994). This will be discussed in section 1.1.5.

(ii) SL1

SL1 contains TBP and three TAFs of 110, 63 and 48kD (Comai et al., 1992) which appear to be different to those in present in TFIID. These TAFs are specific for RNAP I transcription and may function to recruit TBP to the TATA-less class I promoters. An additional factor, upstream binding factor (UBF), is also required for RNAP I transcription and since this and the TAFs are cloned investigations into the mechanism of RNAP I transcription are feasible.

(iii) TFIIIB

TFIIIB is involved in transcription from TATA-less class III promoters containing 'A' and 'B' box elements downstream of the transcriptional start site e.g. transcription of tRNA promoters. It consists of TBP associated with 2-4 TAFs (Lobo et al., 1992; Taggart et al., 1992; White and Jackson, 1992) and it interacts with TFIIIC bound to the A/B boxes. Its role is presumed to be to recruit RNAP III to the promoter. Yeast homologs (BRF1 and a 90kD protein) of the mammalian TFIIIB TAFs have been identified (Kassavetis et al., 1992).

(iv) SNAPc

SNAPc is required for transcription of class III promoters containing a proximal sequence element (PSE) and a TATA box upstream of the transcriptional start site e.g. in the U6 promoter (Sadowski et al., 1993). SNAPc binds to the PSE not the TATA box where a second molecule of TBP is required to bind before transcription by RNAP III can occur. SNAPc is also involved in transcription from U1 and U2 class II promoters by RNAP II but in this case the second TBP molecule is not required.

1.1.4 Activators of RNA Polymerase II Transcription

Basal transcription can be greatly stimulated by promoter specific transcription factors (activators). These activators are generally modular in structure consisting of a DNA binding domain that recognizes a specific sequence of DNA and an activation domain that is responsible for stimulating transcription (Ptashne, 1988). Domain swap experiments have shown that the two functions are readily separable (Brent and Ptashne, 1985; Keegan et al., 1986) and the next two sections will describe the properties of such domains.

(i) DNA Binding domains

These domains are generally relatively small consisting of 40-100 amino acids and several distinct types have been determined (reviewed in Pabo and Sauer, 1992). The first

domain identified was the helix-turn-helix (HTH) motif in the prokaryotic proteins λ Cro (Anderson et al., 1981), λ repressor (Pabo and Lewis, 1982) and E. Coli CAP (McKay and Steitz, 1981). It consists of an α-helix, a turn often containing a glycine residue and a second α-helix which binds to the major groove of DNA and is referred to as the recognition helix (Steitz et al., 1982; Ohlendorf et al., 1983). Such a structure is not stable by itself but always occurs as part of a larger DNA-binding domain e.g. the DNA-binding domain of λ Cro contains β-sheet as well as the HTH.

The HTH motif also occurs in eukaryotic proteins forming part of the 60-residue homeodomain which was first identified in Drosophila developmental regulatory proteins. The structure of the Drosophila engrailed (Kissinger et al., 1990) and yeast Mat α2 (Wolberger et al., 1991) homeodomains have been determined and found to consist of an extended N-terminal arm followed by 3 α helices. The last 2 helices form the HTH unit and helix 3 is the recognition helix contacting the major groove of DNA. Even though the HTH structures of prokaryotes and eukaryotes are very similar, differences exist in the nature of contacts made by the recognition helices e.g. in prokaryotes the N-terminal portion of the recognition helix is closest to the bases whereas in eukaryotes the centre of the recognition helix is closest (Kissinger et al., 1990).

The POU domain defines a subgroup of homeodomain-containing transcription factors and was originally defined in the transcription factors Oct-1, Oct-2 and Pit-1. It is a 150-160 amino acid region consisting of an N-terminal 75-82 amino acid POU-specific domain (POU_S), a short variable linker region and a C-terminal 60 amino acid POU homeodomain (POU_{HD}) (Herr et al., 1988). The 2 domains fold independently, however, the POU_S domain is always found in association with the POU_{HD} domain and both domains are required for high affinity DNA recognition (Aurora and Herr, 1992). The POU_S domain, like the POU_{HD}, contains a HTH motif and its structure is very similar to that of the DNA binding domain of λ repressor (4 α-helices connected by short loops with helices 2 and 3 forming the HTH motif) (Assa-Munt et al., 1993; Dekker et al., 1993). This means that the complete POU structure consists of two HTH domains.

The second type of DNA binding motif is the zinc containing domain and three structural classes of this domain are now known; the zinc finger, the zinc twist and the zinc cluster and these are found in the transcription factor TFIIIA, the glucocorticoid receptor (GR) and the yeast GAL4 protein respectively (Vallee et al., 1991). The zinc finger contains a single Zn^{2+} molecule co-ordinated by 2 cysteines and 2 histidines separated by 12 amino acids which loop out to form the finger structure (Miller et al., 1985). The zinc twist contains 2 Zn^{2+} molecules each co-ordinated by 4 cysteines and it is the intervening region (the twist) which binds to DNA. Finally, the zinc cluster comprises 2 Zn^{2+} molecules co-ordinated by 6 cysteines such that the metals share 2 of the ligands (Vallee et al., 1991). In all three structures coordination of zinc exposes an α helix which binds to the major groove of DNA. Zinc binding occurs at either the amino or carboxy termini of these recognition helices.

A third type of DNA binding motif is the bzip motif which contains 2 subdomains; the zipper region which mediates dimerization and a basic region which contacts the DNA (Landschulz et al., 1988). Both regions are essential for DNA binding. Proteins containing such domains include the Jun, Fos and ATF/CREB families. The zipper monomer consists of an amphipathic α helix with a leucine at every seventh position. Dimerization occurs by interdigitation of the leucines so forming a coiled coil structure (O'Shea et al., 1991) which is stabilized by additional hydrophobic residues located between the leucines. The basic region is rich in arginines and lysines and when bound to DNA forms an α helical structure (Weiss et al., 1990). In the dimer, these helices splay away from each other and bind to the major groove on either side of the DNA (Ellenberger et al., 1992). Leucine zippers can form heterodimers and this is important for the regulation of activity since heterodimers may have different activities and DNA binding specificities compared to the homodimers. This will be discussed in section 1.2.5.

A fourth type of DNA binding domain is the helix-loop-helix (HLH) motif which like the leucine zipper motif has a basic region which contacts DNA and a neighbouring region which mediates dimer formation (Voronova and Baltimore, 1990). The dimerization region consists of a short amphipathic α helix containing hydrophobic residues at every third or fourth position, a loop region often containing one or more helix-breaking residues and a second amphipathic helix (Murre et al., 1989). HLH proteins include E12, MyoD and myogenin and as with the zipper proteins their activity is modulated by heterodimerization.

Further types of DNA binding domain exist e.g. the prokaryotic regulatory proteins MetJ and Arc repressors use an antiparallel β sheet for DNA binding (Breg et al., 1990). There are also DNA binding proteins which bear no obvious structural homology to the DNA binding domains described above e.g. SRF. Characterization of such DNA binding domains will lead to a further understanding of DNA sequence recognition by transcription factors.

(ii) Activation Domains

Several types of activators exist and these have generally been classified according to their amino acid composition. Thus acidic, glutamine rich, proline rich and more recently isoleucine rich activation domains have been identified. Examples of these different activators include the yeast activator GAL4 (Ma and Ptashne, 1987) and the herpes simplex virion protein VP16 (Triezenberg et al., 1988) whose activation domains are rich in aspartate and glutamate residues whereas Sp1 (Courey and Tjian, 1988), CTF (Mermod et al., 1989) and NTF-1 (Attardi and Tjian, 1993) are rich in glutamine, proline and isoleucine residues respectively. However, activation domains which do not fit into such a classification are also known e.g. the HOB activation domains of Fos, Jun and C/EBP were identified on the basis of sequence identity and predicted secondary structure rather than the nature of amino acids within the region (Sutherland et al., 1992).

The structural relationships and mechanisms of specificity of these different types of activation domains is unclear and recently it has been shown that not all activation domains of a given class interact with the same target (Hoey et al., 1993; Gill et al., 1994). This suggests that functional sub-classes within groups of activators may exist. Furthermore, mutagenesis studies suggest that the nature of the amino acids in the activation domain is not necessarily the key feature of the domain since mutation of up to 4 acidic residues at once in VP16 caused only modest reductions in activity (Cress and Triezenberg, 1991). Instead it seems that bulky hydrophobic residues interspersed between these predominant residues may, at least in some cases, be more important (Regier et al., 1993). On this basis, it has been proposed that the hydrophobicity provides the force for association between activator and target and that the nature of the residues within the domain provides the specificity. The structure of the activator is thought to be fairly flexible as shown by NMR and CD studies which do not detect any specific secondary structures in isolated activators. Upon interaction with its target, the activator is assumed to alter its conformation to fit the target so forming a specific, productive interaction (Hahn, 1993; Tjian and Maniatis, 1994).

1.1.5 Mechanism and Targets of Transcriptional Activators

The mechanism of transcriptional activation is unclear although activators are believed to function by interacting with and influencing one or more components of the transcription machinery (e.g. Choy and Green, 1993). Affinity chromatography together with co-immunoprecipitation studies have been used to identify the targets of activators and thus deduce the steps of transcription initiation they affect. Such studies have identified three distinct groups of activator targets; general transcription factors, TAFs and co-activators.

(i) General Transcription Factor-Activator Interactions

Many activators such as the acidic activation domain of VP16 (Stringer et al., 1990), CR3 of E1a (Lee et al., 1991), EBV Zta (Lieberman and Berk, 1991) and CMV IE2 (Hagemeier et al., 1992) protein have been shown to interact with TBP. This is expected given that TBP binding to DNA is the first step in the assembly of the initiation complex. The relevance of such interactions with regard to activation has been demonstrated for some activators; in the case of VP16, mutations in its activation domain which reduce its activation potential also decrease its ability to bind TBP (Ingles et al., 1991), EBV Zta has been shown to increase the binding of TBP to a non-consensus TATA box (Lieberman and Berk, 1991) although this is not true for all activators and finally, E1a has been shown to interact with holoTFIID, the form of TBP present in cells (Boyer and Berk, 1993).

TFIIB is also a target for the activator VP16 (Lin et al., 1991) and binding of TFIIB to the initiation complex has been shown to be enhanced by both acidic and non-acidic activators (Choy and Green, 1993). Mutations in TFIIB which prevent activated but not basal transcription also prevent an interaction with VP16 so qualifying the importance of the TFIIB

interaction for activation of transcription (Roberts et al., 1993). Finally, Zta has been shown to bind via its activation domain to TFIIA. Furthermore, Zta has been shown to enhance and stabilize the formation of a Zta-TFIID-TFIIA-promoter DNA complex, one of the first complexes to be formed during Zta-activated transcription initiation. This infers the relevance of the Zta-TBP/TFIIA interactions (Ozer et al., 1994).

(ii) TAF-Activator Interactions

Activators also interact with TAFs which is not surprising given that TAFs are essential for activated transcription in *in vitro* transcription assays (Pugh and Tjian, 1990). Thus, Sp1 has been shown to bind to dTAF$_{II}$110 (Hoey et al., 1993; Gill et al., 1994), VP16 to dTAF$_{II}$40 (Goodrich et al., 1993) CTF-1 to hTAF$_{II}$55 (Ge and Roeder, 1994) and NTF-1 to dTAF$_{II}$150 and dTAF$_{II}$60 (Chen et al., 1994). Furthermore functional TBP-TAF complexes have been assembled and purified *in vitro* and used to demonstrate that the TAFs involved in certain of these specific TAF-activator interactions are also required for transcriptional activation e.g. GAL4-NTF-1 activates transcription via the partial TFIID complexes TBP-dTAF$_{II}$250-dTAF$_{II}$150 and TBP-dTAF$_{II}$250-dTAF$_{II}$60 whereas Sp1 acts via a TBP-dTAF$_{II}$250-dTAF$_{II}$150-dTAF$_{II}$110 complex (Chen et al., 1994). The same studies also found that different classes of activator target different TAFs although it should be noted that not all activators of a given class will necessarily interact with this target e.g. a number of glutamine-rich transcription factors failed to bind dTAF$_{II}$110 (Hoey et al 1993). Thus the basis for the specificity of activator-TAF interactions remains unclear.

(iii) Coactivator-Activator Interactions

Coactivators are essential to obtain significant levels of activated transcription *in vitro* using purified TFIID and transcriptional activators. These factors are distinct from the TAF proteins and include the positive cofactors PC1 (Meisterernst et al., 1991), PC2 (Kretzschmar et al., 1994), PC3 (Dr2, human topoisomerase I) (Kretzschmar et al., 1993; Merino et al., 1993), PC4 (Ge and Roeder, 1994; Kretzschmar et al., 1994) and CBP (Chrivia et al., 1993). PC2 acts as a cofactor for the synthetic acidic activator AH and seems to be specific for certain types of activator domains; PC4 is a general cofactor functioning with several types of activation domain including AH, VP16, CTF, Sp1, E1a and IE; and CBP is the cofactor for CREB and cJun (Arias et al., 1994; Kwok et al., 1994). Direct interactions between activator-coactivator have been demonstrated: PC4 with the activation domains of AH, CTF, IE and VP16 (Ge and Roeder, 1994); and CBP with phosphorylated CREB (Chrivia et al., 1993) and phosphorylated cJun (Arias et al., 1994). Coactivators also interact with the basal machinery; PC4 with TFIIA (Ge and Roeder, 1994) and CBP with TFIIB (Kwok et al., 1994). Thus, as is the case for at least some of the TAFs, coactivators may act as a link between upstream activators and the basal transcription machinery.

(iv) Mechanism of Activation

From the above sections it can be seen that activators are able to interact with many components of the transcription machinery although not all such interactions are necessarily significant *in vivo.* These numerous possible interactions are proposed to enable activators to influence many aspects of transcription initiation and in doing so increase the efficiency of assembly of transcription competent initiation complexes (reviewed in Hahn, 1993; Kingston and Green, 1994). This is evidenced by *in vitro* studies which have shown that activators stimulate transcription by increasing the number of functional initiation complexes present at promoters (e.g. Lin and Green, 1991). The ways in which an activator could influence initiation complex formation and function include recruitment of basal factors to the promoter (Choy and Green, 1993), stabilization of factor binding at the promoter (Ozer et al., 1994) and formation of productive interactions subsequent to binding possibly as a result of conformational changes induced in the target protein by the activator. This latter proposal is suggested by the detection of a two step process in TBP DNA binding, one stage is thought to involve conformational changes (Hoopes et al., 1992). In such schemes, activators may influence the basal factors directly or via the TAFs and coactivators. It is also possible that the TAFs and coactivators may act upon the activators so converting them into a form capable of influencing the basal machinery. However, no such case has so far been demonstrated. An additional method of activation has been described: structural activation (Tjian and Maniatis, 1994). This involves the formation of a highly specific three dimensional protein complex capable of activating transcription. Assembly of this complex is nucleated by architectural proteins such as LEF-1 and HMG I(Y) which bind to enhancers and facilitate subsequent transcription factor binding. These structural proteins also facilitate protein-protein interactions between the transcription factors once they are bound to the enhancer by bending the DNA and bringing the factors into close contact. This results in the formation of a stereospecific complex which is proposed to interact efficiently with the transcription machinery and so activate transcription.

1.1.6 Repression of Transcription

Basal transcription as well as being activated can also be repressed. This is brought about in two main ways by: (i) the effect of chromatin structure (ii) the action of repressor molecules. Both methods prevent the formation of functional initiation complexes at the promoter and must be overcome in order for basal transcription to proceed, a process known as antirepression. It is achieved by the action of transcription factors which are then able to stimulate transcription (true activation) by methods described above (section 1.1.5). A further form of repression exists. However, this does not interfere directly with the formation of initiation complexes but instead targets the activators. Such passive repression can be brought about by a variety of mechanisms such as titration of the activator away from its binding site at the promoter, competition with the activator for its binding site or formation of a

complex with the activator so masking its activation domain whilst leaving its DNA binding domain unaffected (for review see Cowell, 1994). This latter form of repression will not be discussed further.

In vivo DNA is packaged with histones to form chromatin. The basic unit of chromatin is the nucleosome which consists of a tetramer of two copies each of histones H3 and H4, two dimers of H2A and H2B, and 146 bp of DNA wrapped around this octamer core. Histone H1 binds to DNA in the linker sequences between the nucleosomes (Wolffe, 1994a). Such packaging at a promoter renders the DNA inaccessible to the basal transcription machinery so resulting in repression of gene expression (Perlmann and Wrange, 1991). Nucleosomes may also block upstream factor binding sites although in this case the regulatory factors are able directly or indirectly to reposition the nucleosomes and so bind to DNA. Certain factors e.g. the glucocorticoid receptor (GR) are able to bind to their respective sites even in the presence of nucleosomes since the nucleosomes are positioned within the promoter such that the GR binding sites are exposed towards solution (Pina et al., 1990; Archer et al., 1991). The binding affinity for such sites is decreased compared to naked DNA (Li and Wrange, 1993). Once upstream factors are bound, they can then recruit the SWI/SNF general activator complex (Yoshinaga et al., 1992). This complex is thought to act by destabilizing the chromatin structure via mechanisms including the release of histone H1 from linker DNA (Bresnick et al., 1992) or the weakening of the histone-DNA interactions within the nucleosome by either competing protein-protein or protein-DNA interactions (reviewed in Wolffe, 1994b). The SWI2/SNF2 subunit of the complex possesses a DNA-dependent ATPase activity (Laurent et al., 1993) which may enable the complex to track along the DNA to destabilize several nucleosomes including those at the core promoter. This then facilitates access of the basal machinery to this region and permits initiation of transcription to take place.

Once the promoter is accessible to the basal machinery active repression by negative regulators may occur. Several such regulators have been discovered; NC1 (one component of which is thought to be the high mobility group protein I, HMG1) (Ge and Roeder, 1994), NC2 (Meisterernst et al., 1991), Dr1 (Inostroza et al., 1992), Dr2 (PC3) (Merino et al., 1993) and ADI (Auble and Hahn, 1993). The first four regulators function by binding to TBP and preventing it from interacting with the other basal factors necessary to form an initiation complex. Such repression is proposed to be overcome by activators enhancing the action of TFIIA, presumably by directly interacting with this transcription factor. TFIIA binds to TBP and in doing so is thought to displace any inhibitors associated with it, therefore allowing initiation complex formation to proceed (Yokomori et al., 1994). ADI functions in a different way. Instead of blocking complex formation, it is thought to displace TBP from DNA in an ATP-dependent process (Auble and Hahn, 1993). However, as with the other repressors its action can be overcome by TFIIA since TFIIA is able to stabilize the TBP-DNA complex (Ozer et al., 1994). Thus in both types of repression TFIIA acts as an anti-repressor. However, TFIIA can

also participate in true activation and directly stimulate transcription. The mechanism of such activation is not understood but is known to involve the TAFs (Sun et al., 1994). It should be noted that PC3/Dr2 can act both as a positive cofactor enhancing activated transcription and as a negative regulator repressing basal transcription (Merino et al., 1993). This demonstrates the complex nature of the regulation of RNAPII transcription.

1.2 The ATF/CREB Family

Many of the factors that bind to the regulatory elements in promoters have been cloned and this led to the discovery that each site is capable of binding several related proteins which gave rise to the concept of transcription factor families. These groups of proteins not only bind to the same sequence but also share a common DNA binding domain. One such family is the activating transcription factor (ATF) family which binds to the consensus sequence TGACGTCA. This sequence is also found in the promoters of cAMP inducible genes and therefore the family is alternatively known as the cAMP response element binding (CREB) family (Ziff, 1990).

The ATF/CREB family is large and all members are bzip containing proteins (Hai et al., 1989). The sequence of this DNA binding domain is highly conserved throughout the family and outside of this region there is very little overall homology. However, recently certain family members have been identified which possess other limited regions of homology. These include CREB (Hoeffler et al., 1988; Gonzalez et al., 1989) which is related to both ATF-1 (Hai et al., 1989) and CREM (Foulkes et al., 1991), and ATF2 (also known as CRE-BP1 or CREB2) (Maekawa et al., 1989) which is related to both the ATFa (Chatton et al., 1993) and CRE-BPa (Zu et al., 1993) families. The number of ATF/CREB family members is larger than predicted from the number of genes isolated since many of these genes encode more than one protein by alternative splicing e.g. CREB encodes 8 products (Waeber et al., 1991; Ruppert et al., 1992) and CRE-BPa encodes 4 (Zu et al., 1993). The CREM gene is further complicated since it has two promoters each of which gives rise to several products by differential splicing; the P1 promoter encodes 9 products (Laoide et al., 1993) and the P2 promoter 4 products (Molina et al., 1993). Such an extensive transcription factor family enables transcription dependent upon the ATF/CRE site to be regulated by a variety of signals and to respond to the signals in a variety of ways. The roles of some of the family members in this regulation is beginning to be understood and it is found that the regulation occurs at many different levels including phosphorylation, viral stimulation, cooperation with other cellular factors and interaction with other bzip proteins. Such regulation is described in the following sections.

1.2.1 Regulation By Binding Site Specificity

The consensus sequence for ATF/CREB binding is TGACGTCA. However, in the promoters of genes ATF/CRE sites often diverge from this consensus. These altered sites

have varying affinities for the different family members as shown by gel shift analysis and so determine the type of ATF/CREB complex bound to the promoter. Sequences flanking the core binding site are also important for determining specificity which means that promoter context also influences binding specificity (Benbrook and Jones, 1994). The flanking sequences are also important for cooperation effects discussed in section 1.2.4.

1.2.2 Regulation By Phosphorylation

The best example of ATF/CREB regulation by phosphorylation is the stimulation of CREB activity by cAMP-dependent protein kinase A (PKA). The signalling pathway that leads to this stimulation is relatively well understood (reviewed in Lalli and Sassone-Corsi, 1994); peptide hormones bind to guanine nucleotide binding protein-coupled receptors and so stimulate the production of the second messenger cAMP. The resultant increase in cAMP concentration causes the catalytic subunit of PKA to dissociate from its regulatory subunit and so migrate into the nucleus (Hagiwara et al., 1993) where it phosphorylates CREB at serine 133 (Gonzalez and Montminy, 1989) in the kinase inducible domain (KID). This has two effects. Firstly, it is thought to cause a conformational change exposing the glutamine rich activation domain Q2 (Gonzalez et al., 1991) which then interacts with transcription machinery via a TAF (Q2 has been shown to bind dTAF$_{11}$110) (Ferreri et al., 1994). Secondly, it enables CREB to bind to the coactivator CBP (Chrivia et al., 1993) thus bringing CBP to the promoter where it can interact with the transcription machinery by contacting TFIIB (Kwok et al., 1994). These two effects increase the transactivation function of CREB and both seem to be required for the process. The activation of CREB is only transient and is reversed by phosphatases. The identity of the phosphatase responsible for dephosphorylating CREB is controversial. Both protein phosphatase I (Alberts et al., 1994) and IIA (Wheat et al., 1994) have been reported to be the primary phosphatase responsible for de-activating CREB.

Phosphorylation can also regulate ATF/CREB activity by altering the DNA binding properties of some of the family members. In the case of CREB, phosphorylation by PKA has been shown to increase its binding to low affinity CREs (Nichols et al., 1992) although this effect is minor compared to the effect it has on activation. ATF2 binding to a CRE is also stimulated by phosphorylation and the enzyme implicated is a mitogen-activated protein kinase (p42 or p54 MAPK) not PKA. Conformational changes induced by phosphorylation are suggested to account for the increases in DNA binding. In the case of ATF2, the unphosphorylated state is believed to be in a configuration which prevents dimerization and hence DNA binding. Phosphorylation is proposed to relieve the inhibition of dimerization (Abdel-Hafiz et al., 1992).

1.2.3 Regulation By Viral Proteins

The viral transactivator proteins, adenovirus E1a, hepatitis B virus (HBV) pX and human T-cell leukaemia virus type 1 (HTLV-1) Tax regulate transcription of a variety of viral and cellular genes via mechanisms not fully defined. However, one mechanism seems to involve the stimulation of transcription via ATF/CRE sites. Evidence for this comes from mutagenesis studies on certain E1a/pX/Tax sensitive promoters and the detection of *in vitro* associations between the viral activators and ATF/CREB proteins. The interaction with pX alters the DNA binding specificity of CREB and ATF2 so enabling them to bind to the CRE-like site (TGACGCAA) in the HBV promoter and thus recruit pX to the promoter where it can activate transcription (Maguire et al., 1991). However, no *in vivo* data to support this model has been produced. The interaction with Tax stimulates the formation of dimeric ATF/CREB proteins (DNA binding species) and so stimulates their DNA binding capacity (Wagner and Green, 1993). This will also stimulate Tax recruitment to the promoter. Tax has also been shown to alter the DNA specificity of CREB (Adya et al., 1994). The interaction with E1a does not alter the binding specificity of the ATF/CREB proteins but does again bring the viral activator to the promoter (Liu and Green, 1990; Liu and Green, 1994). Thus the viral activators appear to function in similar ways. The mechanism of E1a activation via ATF/CREB proteins is described further in section 1.3.2.

1.2.4 Regulation by Cooperation With Cellular Factors

ATF/CREB activity can also be influenced by cooperation with non bzip proteins and such cooperation has been shown to be important for gene regulation. One example of this is interleukin-1 (IL-1) induction of the E-selectin gene. By mapping regulatory elements in the promoter of the E-selectin gene the transcription factor NF-κB was found to be required but not sufficient for IL-1 induction (Whelan et al., 1991). An additional complex, NF-ELAM1, which contains ATF/CREB proteins e.g. ATFa, was also required. Since ATFa was shown to bind NF-κB *in vitro*, it suggests that cooperation results from protein-protein interactions formed when NF-κB is stimulated to bind to the promoter by IL-1 (NF-ELAM1 is constitutively bound) (Kaszubska et al., 1993).

A second example of ATF/CREB regulatory cooperation is virus induction of the human β-interferon (hu β-IFN) gene via the positive regulatory domain IV (PRDIV). This domain is a composite regulatory element containing binding sites for both ATF2 and the high mobility group protein HMG I(Y) and both sites are required for virus induction. HMG I(Y) does not function as a transcriptional activator but instead stimulates binding of ATF2 to PRDIV. It can achieve this by either stabilizing protein-DNA interactions, by inducing conformational changes in DNA or by increasing the affinity of ATF2 for its site as a result of protein-protein interactions between itself and ATF2. HMG I(Y) may also directly influence the transcriptional activity of ATF2. There is a second level of regulation at this promoter which again involves ATF2. HMG I(Y) is proposed to facilitate DNA looping so bringing the promoter regulatory

elements PRDIV and PRDII close to each other. This enables ATF2 bound to PRDIV to interact with NF-κB bound to PRDII and so cooperate functionally with it (Du et al., 1993).

Further examples of ATF/CREB cooperation include occupancy of the CCAAT site in the fibronectin (FN) promoter. In liver cells, binding to this site is stimulated by binding of ATF/CREB proteins to the adjacent CRE. Such cooperative binding is correlated with the stimulation of FN transcription in *in vitro* studies and may partly explain tissue specific regulation of this housekeeping gene since such cooperative binding is not seen in other cell types e.g. brain cells (Srebrow et al., 1993). ATF/CREB proteins also associate with nuclear matrix proteins which may regulate their activity (van Wijnen et al., 1993) and a specific family member, ATF2, has been shown to be activated by the retinoblastoma (Rb) protein. However, little is known about the mechanism of Rb activation and it may occur in an indirect way because although an *in vitro* association between ATF2 and Rb has been demonstrated the *in vivo* relevance of such an interaction has not been demonstrated (Kim et al., 1992).

1.2.5 Regulation By Heterodimerization

ATF/CREB proteins bind to DNA only as dimers with both homodimers and certain heterodimers being formed between the various family members e.g. CREB heterodimerizes with ATF1 (Hoeffler et al., 1991) and CREM (Foulkes et al., 1991), and ATF2 heterodimerizes with ATF3 (Hai et al., 1989). The formation of heterodimers depends on the interaction of specific residues in the leucine zipper region which either stabilize or destabilize the interaction of particular dimer partners. ATF/CREB proteins can also heterodimerize with members of other bzip families e.g. ATF2 can heterodimerize with cJun (Benbrook and Jones, 1990). Thus a diverse number of ATF/CREB complexes can be formed. This heterodimerization provides for complex regulation of the ATF/CREB family since interaction with a different partner can alter the properties of the individual members. Such influences are described in this section.

Firstly, heterodimerization could influence all the methods of regulation discussed above. Examples where this has been shown to be the case are:

(i) Binding Site Specificity

Heterodimerization of ATF2 with cJun still permits ATF2 binding to the consensus CRE (TGACGTCA) but now also enables ATF2 to bind to the non-consensus AP1 site (TTACCTCA) present in the cJun promoter (van Dam et al., 1993).

(ii) Phosphorylation

By engineering complementary CREB proteins that could only heterodimerize and combining these with phosphorylation mutants (serine 133 mutation), it was shown that a single phosphorylated CREB molecule in a dimer was sufficient to obtain a response to PKA. The response was decreased compared to fully phosphorylated dimers but shows that the introduction of CREB into a dimer has the potential to confer PKA responsiveness via phosphorylation upon the dimer (Loriaux et al., 1993).

(iii) Viral Proteins and Cellular Factors

As with phosphorylation, it can be imagined that one partner of the heterodimer can confer upon the other partner the ability to be influenced by certain viral or cellular proteins. This has not been demonstrated as elegantly as the phosphorylation study of CREB but examples do exist e.g. E1a represses the collagenase gene via Jun-Jun homodimers and Jun-Fos heterodimers whereas it stimulates the cJun gene via ATF2-cJun heterodimers (van Dam et al., 1993). This does not merely reflect the altered DNA binding specificity of cJun but is also thought to involve a change in its transactivation potential brought about by E1a induced phosphorylation (Hagmeyer et al., 1993).

Heterodimerization can have a more direct effect on regulation by producing a complex with altered transcriptional properties. This involves repressing or stimulating the activity of the individual components involved. This is best exemplified by heterodimerization of CREB and CREM. CREB and CREM are highly homologous proteins both composed of modular domains encoded by separate exons. The CREB gene consists of at least 11 exons of which 10 are coding (Ruppert et al., 1992). These include exons coding for the 2 glutamine rich activation domains, the kinase inducible activation domain (KID) and the bzip DNA binding domain (Brindle et al., 1993; Quinn, 1993). The CREM exon structure is virtually identical to that of CREB (Laoide et al., 1993), the main difference being the presence of two alternative DNA binding domains in CREM whereas only one is present in the CREB gene. There are also other short exons unique to either CREB or CREM although the function of these domains is not clear. Differential splicing of the exons produces a large number of gene products (8 for CREB and 9 for CREM) although not all isoforms have been detected in cells at the protein level. Additional CREM products arise by use of a second promoter, P2, located within the 3' portion of the gene. These isoforms are known as inducible cAMP early repressors (ICER) (Molina et al., 1993).

The role of the various CREB/CREM isoforms in transcription has been investigated by transient transfection analysis and it is found that the CREB proteins are all transcriptional activators whose activity can be stimulated by PKA. CREM, on the other hand, encodes both activators and repressors of transcription in the presence of PKA (Foulkes et al., 1992). The CREM repressors, CREMα, β and γ contain the KID and bzip region but lack the glutamine rich activation domains (Foulkes et al., 1991).The CREM repressor, S-CREM, lacks the KID (Delmas et al., 1992). These proteins can therefore bind to CREs but are not able to activate transcription. In fact they can repress transcription by competing with CREB or CREMτ activators for binding to CREs or by complexing with the activators to produce non-functional heterodimers. The latter mechanism is thought to be the predominant mechanism which means that the repressor is the dominant partner of the heterodimer (Laoide et al., 1993). This contrasts to the engineered heterodimers of CREB described above in which the activator was the dominant effector. The CREM isoforms, ICER, are also repressors. These proteins contain a DNA binding domain but lack both the KID and the glutamine rich

activation domains and so are powerful repressors functioning in the same way as the other CREM repressors.

Thus it is clear that both partners in a heterodimer play a role in determining the functional properties of the complex. This means that the level of transcription will be controlled by the relative concentration of the different partners within a cell. This is important for tissue specific gene expression since levels of activator/repressor proteins can vary from tissue to tissue e.g. CREM is expressed to differing degrees in different cell types whereas CREB expression is fairly constant (Foulkes et al., 1992). It is also important for inducible gene expression since the levels of the activators/repressors within a cell may be altered by signalling pathways e.g. ICER gene expression is stimulated by corticotroph-releasing factor (Molina et al., 1993).

Thus regulation of the ATF/CREB family is extremely complicated which makes deciphering the roles of the individual family members very difficult especially since all mammalian cells contain endogenous ATF/CREB proteins. The use of heterologous DNA binding domains has enabled certain roles to be established e.g. CREB was shown to be PKA responsive by this method and ATF2 was similarly shown to be E1a responsive (Flint and Jones, 1991). However, the interplay between factors and the signalling pathways which lead to gene regulation via CREs is very poorly understood with the exception of the PKA pathway. Thus this thesis concentrates on one family member ATF2 in order to gain further insight into gene regulation by the ATF/CREB family.

1.3 Adenovirus E1a

The goal of the DNA tumour viruses e.g. adenovirus, simian virus 40 (SV40) and human papilloma virus is to replicate and produce new virus via infection. To achieve this, these viruses reprogram the host cell by altering gene expression and so creating an environment that is favourable for DNA synthesis. If a productive infection does not ensue, oncogenic transformation results. Since these viruses affect key regulatory events in gene expression and cell cycle control, studying such viral regulation should lead to an understanding of the cellular processes involved and this has indeed been found to be true. This thesis looks at one aspect of such viral regulation namely transactivation by the adenovirus early region 1a (E1a) protein. In this introduction I shall present a summary of E1a and describe what is currently known about some of its regulatory properties.

E1a is the first gene to be expressed upon adenovirus infection and is capable of regulating gene expression and immortalizing primary rodent cells by itself (Nevins, 1987). Thus it is a key protein in adenovirus regulation. It encodes 5 mRNAs by differential splicing (sizes 13s, 12s, 11s, 10s and 9s) (Stephens and Harlow, 1987) with the major products being the 13s and 12s forms. These encode proteins of 289 and 243 amino acids respectively and both are phosphoproteins although phosphorylation does not appear to be required for E1a function (Tremblay et al., 1989). There are three regions in the 13s protein that are highly

conserved between the various adenovirus serotypes; amino acids 40-80, 121-139 and 140-188 (Kimelman et al., 1985). These are referred to as conserved regions 1, 2 and 3 (CR1, 2 and 3) respectively and they represent distinct functional domains in E1a; CR1 and 2 are required for transformation, transcriptional activation and repression whereas CR3 is solely involved in transcriptional activation (reviewed in Moran and Mathews, 1987).

1.3.1 Transactivation by E1a

E1a can activate several viral genes including E1b, E2a, E3 and E4 (Berk, 1986). It can also activate two cellular genes, the human heat shock gene and β-tubulin, as well as a variety of cellular promoters introduced into cells by transfection. Therefore, it is often regarded as a promiscuous transactivator. It does not bind to DNA in a sequence specific manner (Chatterjee et al., 1988) which means that it cannot activate transcription by binding to upstream elements in the promoters of E1a sensitive genes and behaving as a transcription factor. Instead it is thought to act by altering the activity of cellular transcription factors. Evidence for this has come from mutational analysis of E1a sensitive promoters which showed that certain transcription factor binding sites were required for E1a activation. These include the ATF/CRE site (Lee and Green, 1987), the E2F site (Manohar et al., 1990) and the TATA box (Wu et al., 1987). Furthermore, ATF/CRE sites and also E2F sites were found to confer E1a inducibility upon a heterologous promoter thus proving the involvement of these sites in E1a activation. With the cloning of the E2F family of transcription factors, however, it has become apparent that E1a activation via the two types of binding sites occurs by very different mechanisms.

1.3.2 Activation via the ATF/CREB Family

Once ATF/CRE sites were found to be important for E1a activation, evidence for a direct involvement of the ATF/CREB family was provided by showing that certain family members were capable of supporting an E1a response. This was achieved by transient transfection assays with GAL-ATF/CREB fusions. The use of GAL fusions enabled specific members of the family to be analyzed individually since the altered DNA binding specificity circumvents the problem of endogenous ATF/CREB proteins present in mammalian cells. Transcription stimulated by the fusions was measured from a GAL reporter and E1a activation was investigated by performing the assay either in the presence or absence of E1a. E1a was found to stimulate transcription through a specific family member namely ATF2 (CRE-BP1, CREB2). The other members examined e.g. CREB and ATF1 were unresponsive (Flint and Jones, 1991). More recently the ATF2 related proteins, ATFa1, ATFa2 and ATFa3, have also been shown to be E1a responsive (Chatton et al., 1993). This ability to stimulate transcription via ATF2 and ATFa1 is a property associated with 13s but not 12s E1a. These two proteins differ only by 46 amino acids comprising CR3 which is present in 13s E1a but

almost entirely missing in the 12s form. Thus, E1a activation of ATF2 and ATFa is mediated by CR3.

Many studies have been carried out on this unique region of 13s E1a to determine its functional properties. It has been shown to be necessary and, in some cases, sufficient to activate transcription e.g. a synthetic peptide comprised of CR3 residues can activate several early promoters *in vitro* and *in vivo* (Lillie et al., 1987) following its microinjection into cells. Furthermore, CR3 in the form of a GAL fusion is able to activate a construct containing only GAL binding sites upstream of a TATA box and reporter gene (Lillie and Green, 1989). Since this activation is dependent on both CR3 and the GAL sites in the reporter, it suggests that CR3 functions by binding to and acting at the promoter.

This is further suggested by deletion and mutagenesis studies which identified a promoter targeting region at the C-terminus of CR3, amino acids 183-188 (Martin et al., 1990; Webster and Ricciardi, 1991). Since E1a does not bind DNA in a sequence specific manner, this region is proposed to target promoters by interacting with specific promoter bound transcription factors rather than specific promoter sequences. The data discussed above suggests that ATF2 may be one such transcription factor and this has been shown to be the case. Mutating the glycine residue at position 180 to aspartic acid in 13s E1a not only destroys promoter targeting by CR3 but also destroys E1a activation via GAL-ATF2 (Liu and Green, 1990). Thus CR3 seems to function by interacting with promoter bound ATF2. This has been confirmed recently by the detection *in vitro* of a direct physical interaction between E1a and ATF2. The site of interaction on E1a was mapped to the promoter targeting region of CR3 which agrees with the *in vivo* functional data (Liu and Green, 1994). E1a has also been shown to interact with ATFa1, another E1a responsive member of the ATF/CREB family (Chatton et al., 1993). However, E1a has also been found to bind to unresponsive members of this family e.g. ATF1 (Liu and Green, 1994). This discrepancy will be discussed in section 6.4.

There is an additional reason for proposing the above mechanism of targeting based on studies of the herpes simplex virus (HSV) VP16 protein. Like E1a, VP16 is a strong transcriptional activator which cannot bind directly to specific sequences of DNA with high affinity. Instead VP16 forms a complex with the transcription factor Oct1 and the accessory family of polypeptides collectively known as host cell factor (HCF) (Stern and Herr, 1991). This complex is then able to bind to "tat-garat" sequences present in the enhancers of the HSV immediate early genes so bringing VP16 to the promoter where it can activate transcription. Thus, E1a and VP16 may function by analogous mechanisms.

Once E1a has been brought to the promoter, it then stimulates transcription and the region responsible for this has been delineated to the N-terminal portion of CR3, amino acids 147-177. This region contains a zinc finger and mutation of any one of the four cysteine residues forming this structure abolishes transactivation. Therefore, the zinc finger is essential for E1a activation. As well as being essential for activation, this region has also

been found to be essential for the trans dominant negative phenotype of E1a mutants which lack the promoter targeting region of CR3. This suggests that this region binds to and sequesters a limiting cellular transcription factor required for E1a activation and by doing so prevents wild type E1a from stimulating transcription (Webster and Ricciardi, 1991).

Several lines of evidence indicated that the limiting factor was the basal transcription factor TBP. Firstly, the transcriptional activity of a partially purified TFIID fraction from HeLa cells was increased by adenovirus infection (Leong et al., 1988). Secondly, a subset of E1a responsive promoters appear to only require the TATA box for transactivation (Wu et al., 1987; Simon et al., 1988). Thirdly, it was demonstrated that E1a CR3 binds *in vitro* specifically and stably to TBP and that the region of CR3 involved in the binding was the N-terminal zinc finger region (Horikoshi et al., 1991; Lee et al., 1991).

The E1a-TBP interaction was demonstrated to be functionally relevant in several ways. Firstly, the E1a-TBP interaction was shown to occur even when TBP was present as part of the TFIID complex (Boyer and Berk, 1993). This means that the surface of TBP which interacts with E1a must be accessible for such an interaction *in vivo.* Secondly, by making conservative substitutions at each of the 49 amino acids in CR3, it was shown that mutants which were defective in TBP binding were also defective in transactivation. These TBP binding mutations were all found to occur in the finger region of CR3. There were no examples of mutants that reduced TBP binding but did not reduce transactivation (Geisberg et al., 1994). The third method utilized an *in vitro* transcription system which was sensitive to E1a CR3. In this system, increasing the amount of CR3 from 0.06-0.6pmole resulted in an increase in the level of transcription. However, increasing the concentration of CR3 beyond 0.6pmole resulted in the inhibition of transcription, a phenomenon known as squelching and is thought to result from titration by an excess of CR3 of its own cellular target. Both activation and inhibition by CR3 were found to be sensitive to the same point mutations within CR3 suggesting that these processes utilize an identical region in CR3 for interacting with a common target. Since inhibition was only overcome by the addition of purified TFIID and not any other general transcription factor fraction, it appears that CR3 stimulates transcription through a direct physical association with TFIID (Boyer and Berk, 1993).

Thus from the above evidence TBP appears to be a physiological target of the zinc finger region of CR3. However, it is now apparent that it is not the only target of this region and also that it is not the limiting cellular factor suggested by the phenotype of the transdominant negative mutants of CR3. This arises from several observations. Firstly, when the series of conservative mutations throughout CR3 were tested for their ability to bind TBP and transactivate an E1a responsive promoter *in vivo* (the adenovirus E3 promoter), five mutants were found in the zinc finger region which were defective in transactivation but which were still capable of binding TBP. Therefore, it seems that TBP binding, although necessary, is not sufficient for activation. These five zinc finger mutants did not exhibit a transdominant negative phenotype even in the context of a mutant promoter targeting region which means

that TBP binding is not sufficient for negative transdominance as was originally suggested by the promoter targeting mutants i.e. TBP does not seem to be the limiting cellular factor sequestered by the negative transdominant mutants (Geisberg et al., 1994). Finally, evidence for an additional factor is also provided by studies using the E1a sensitive *in vitro* transcription assay. In this system, squelching caused by high concentrations of CR3 was not able to be overcome by increasing the concentration of TBP in the assay which again demonstrates that TBP by itself is not sufficient for activation by CR3 (Boyer and Berk, 1993).

All the above data suggests that the N-terminal activation domain of CR3 binds two factors; the general transcription factor TBP and an additional unknown factor referred to as limiting factor 1 (LF1). Both factors seem to be required for CR3 activation of transcription. At present the nature of LF1 is unclear although one possibility is that it may be a TAF. This is suggested by the observations that squelching by CR3 is overcome by TFIID rather than TBP *in vitro* (Boyer and Berk, 1993) and similarly that only TFIID, not TBP, is able to support E1a stimulated transcription *in vitro* (Zhou et al., 1992). Thus it is clear that TAFs are essential for CR3 transactivation but how they contribute to this process is not known and whether LF1 corresponds to a known TAF remains to be determined. Alternatively, LF1 may be a coactivator. This is suggested by the requirement of PC4 for E1a activation *in vitro* (see section 1.1.5 (iii)).

A general model for E1a CR3 activation via ATF/CRE sites has been suggested based on all the above data and is shown schematically in figure 1.3. E1a is brought to the promoter via an interaction of its CR3 carboxyl targeting region with promoter bound ATF2. This then enables the N-terminal activation region of CR3 to interact with the basal machinery by contacting TBP and to stimulate transcription by the mechanisms described in section 1.1.5 (iv). Interaction of CR3 with the additional factor LF1 is presumably required to either transmit further stimulatory signals from E1a to the transcription machinery or to allow E1a to function at a later stage during transcription initiation than TBP binding. In this model, ATF2 is viewed as a passive player in the process being merely required to tether E1a. This aspect will be discussed further in section 6.4.

This model offers the scope for regulation of a variety of promoters by E1a providing that the promoters contain sites for transcription factors which are capable of interacting with E1a especially if these transcription factors are not members of the ATF/CREB family. This has been found to be the case e.g. a direct interaction between E1a and CBF has been demonstrated (Lum et al., 1992; Agoff, 1994). This partly explains the promiscuous nature of E1a activation.

1.3.3 Activation via the E2F Family

The E2F binding site was originally identified in the adenovirus E2a promoter in two copies by mutational analysis. Such analysis determined the importance of these sites for both basal and E1a induced transcription of the E2 gene (Manohar et al., 1990). The amount

Figure 1.3 Model of E1a activation via ATF/CRE sites

E1a is brought to the promoter via its interaction with ATF2. Its activation region then contacts TBP as well as a limiting cellular factor (LF1) in order to stimulate transcription. ATF2 plays a structural role in this model merely acting as an anchor for E1a.

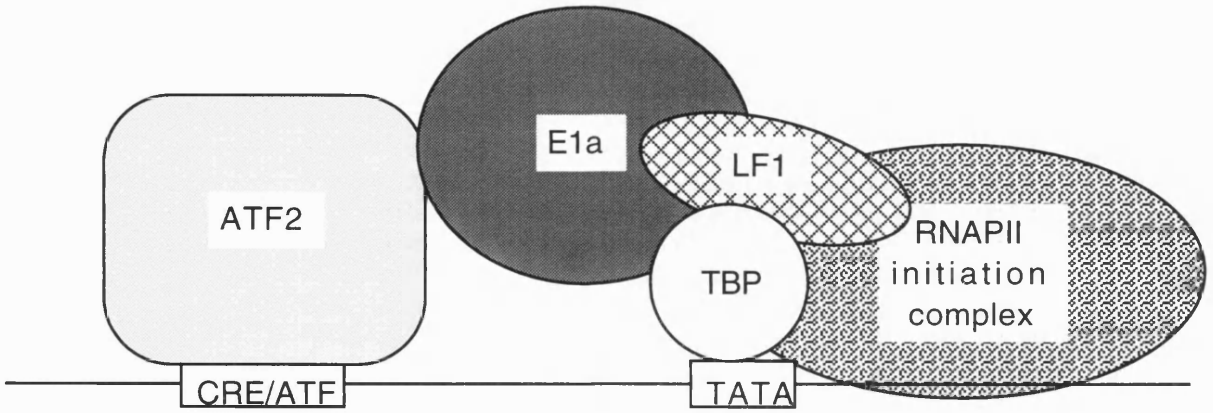


Figure 1.3 Model of E1a activation via ATF/CRE sites

and type of protein bound at these sites was found to alter greatly during an adenovirus infection and much work has since been carried out in order to understand the role of these sites in transcriptional regulation and E1a activation.

Two approaches contributed towards the understanding of E2F regulation by E1a. The first was the identification of E1a cellular targets by coimmunoprecipitating E1a associated proteins from radioactively labelled, infected cells using an E1a specific antibody (Yee and Branton, 1985; Harlow et al., 1986). These protein complexes were run on SDS-PAGE and the individual polypeptides identified as a series of bands in the gel and classified according to their molecular weights. The major polypeptides isolated in this way were p300, p130, p107, p105, p60 and p33. These proteins have since been identified as being key cellular regulatory proteins.

The first protein identified, p105, was found to correspond to the product of the retinoblastoma tumour supressor gene (Rb) (Whyte et al., 1988). Its E1a binding region is referred to as the pocket domain (Kaelin et al., 1991) and consists of two subdomains A and B separated by a spacer. It is the subdomains which are necessary and sufficient for E1a binding. The E1a associated proteins, p130 and p107, are highly related to Rb since both contain the pocket domain and possess Rb homologous A and B regions (Ewen et al., 1991; Hannon et al., 1993; Li et al., 1993). The spacers of the pocket domains, however, are unrelated to pRb but are related to each other. This means that p107 and p130 are able to interact with cyclins A and E and the associated kinase, cyclin-dependent kinase 2 (cdk2) (Dyson et al., 1993; Li et al., 1993), via this region but that Rb cannot. It also means that cyclin A (p60), cyclin E and cdk2 (p33) are E1a associated proteins. The region of E1a required for binding these Rb related and associated proteins has been mapped to CR1 and 2 with primary binding requiring the site in CR2 (Barbeau et al., 1994). The other major E1a associated protein, p300, is discussed in section 1.3.4.

The second approach taken to try and understand E1a regulation via E2F was to identify the proteins bound to the E2F sites in the E2a promoter. Several such factors have now been cloned (the E2F family) and they include the transcription factors E2F-1, -2, -3 (Helin et al., 1992; Ivey-Hoyle et al., 1993; Lees et al., 1993) and DP-1, -2, -3 (Girling et al., 1993; La Thangue, 1994). These HLH zipper proteins bind to DNA as heterodimers composed of one E2F-like and one DP-like partner (Bandara et al., 1993). As well as a DNA binding domain, these proteins also contain a potent transactivation domain at their C-terminus and they therefore function to stimulate transcription from promoters containing E2F sites (Kaelin et al., 1992; Bandara et al., 1993). Another property of these E2F heterodimers is their ability to interact, via the E2F partner of the dimer, with the pocket domain of the proteins pRb, p130 and p107 (Shirodkar et al., 1992; Cobrinik et al., 1993; Lees et al., 1993). Indeed it was on the basis of Rb binding that the first family member (E2F-1) was cloned. Due to the complexity of the E2F family, it is thought that each of the Rb related proteins interacts with different E2F-DP heterodimers. This has been demonstrated for pRb and p107; the pRb

complex contains E2F-1 whereas the p107 complex lacks this family member (Dyson et al., 1993).

The binding site for the Rb related proteins in E2F maps to the C-terminus of the protein and is found to overlap with its activation domain. This suggests that interaction of E2F with one of the Rb related proteins would mask the activation domain and so prevent stimulation of transcription by E2F. It is possible to measure E2F dependent transcription by transient transfection in Rb^- cell lines of a reporter gene under the control of the E2a promoter. Using this system Rb has been shown to repress E2F activated transcription and furthermore this repression was found to depend upon the ability of Rb to interact with E2F (Helin et al., 1993). Similarly, p107 has been shown to repress E2F dependent transcription (Schwarz et al., 1993) and p130 is presumed to function in a similar manner.

Thus the two approaches taken to study E1a transactivation via E2F identified the pocket proteins, pRb, p130 and p107, as common targets for E1a and E2F. The final link in understanding how this related to the mechanism of activation was provided by the discovery that E1a was able to release E2F from E2F-pRb/p130/p107 complexes (Chellappan et al., 1992). This release is mediated by CR1 and CR2 of E1a which are the same regions used by E1a to bind the Rb related proteins. This suggests that E1a dissociates E2F by direct competition for binding to the pocket proteins. Further mapping studies found that CR1 of E1a and the pRb binding domain of E2F bind to the same or an overlapping region on pRb. CR2, on the other hand, binds to a distinct region of pRb but with a 10-fold higher affinity. This leads to the model of disruption whereby E1a binds to pRb-E2F via CR2 to give a pRb-E2F-E1a intermediate from which E2F is competitively released by CR1 binding (Fattaey et al., 1993). The result of such dissociation is free E2F which is transcriptionally active and able to stimulate the transcription of genes whose promoters contain E2F sites.

Therefore, E1a transactivation via E2F sites occurs by a CR3 independent process which releases E2F from transcriptionally repressed complexes so converting E2F into a transcriptional activator. An additional level of control is imposed on this mechanism to ensure that the released E2F preferentially activates viral rather than cellular genes. This is achieved by the action of a second adenoviral protein, E4 orf6/7. E4 orf6/7 binds to two molecules of transcriptionally active E2F and so produces a DNA binding complex which recognizes simultaneously two copies of an E2F site in a certain orientation and spacing (Hardy and Shenk, 1989; Huang and Hearing, 1989). Such an arrangement is found in the promoters of certain adenoviral genes e.g. in the E2a promoter but is not found in the promoters of cellular genes. This means that the released E2F will preferentially bind to and activate the viral promoters so ensuring the expression of the viral genes within infected cells.

1.3.4 Repression via p300

E1a is able to repress a number of transcriptional enhancers and promoters. These include the viral enhancers of SV40 and polyoma virus and the enhancers and promoters of

tissue-specific cellular genes associated with differentiation (Bayley and Mymyrk, 1994). The region of E1a required for repression maps to the same region required to bind p300 (Wang et al., 1993), a phosphoprotein first identified by its association with E1a. p300 has recently been cloned and shown to be able to overcome E1a repression of the SV40 enhancer (Eckner et al., 1994). It is thought to function as a transcriptional coactivator transmitting signals from enhancer bound transcription factors to the transcription machinery. Evidence for this is provided by the fact that p300 can associate with TBP and also that it possesses features (a bromodomain and cys/his rich motifs) which are found in other coactivators. The mechanism by which E1a represses the activity of p300 is unclear at present.

1.3.5 Repression and Activation via AP-1

E1a differentially regulates AP-1 responsive genes; it represses collagenase and stromelysin but activates cJun. Both functions are mediated by E1a CR1 acting via TPA-responsive elements (TREs) in the promoters of these genes (Hagmeyer et al., 1993; van Dam et al., 1993). In the case of the collagenase gene, the TRE binds cJun/cJun and cJun/cFos dimers and E1a acts by blocking binding of these factors thereby repressing transcription. The cJun TRE, on the other hand, binds cJun-ATF2 dimers and E1a does not prevent binding of this dimer. Instead, E1a increases phosphorylation of cJun's activation domain and hence stimulates transcriptional activation via the jun2TRE. The way in which E1a distinguishes between the various cJun dimers is unknown.

1.3.6 Transformation

E1a transforms cells as a consequence of its interactions, via CR1 and CR2, with the cellular proteins described above (sections 1.3.3 and 1.3.4). These proteins have been found to be key regulators of the cell cycle. pRb and p107 are negative regulators of cell cycle progression and p130 and p300 are thought to exert similar effects. pRb, p130 and p107 are proposed to function at different times during the cell cycle by binding and inactivating E2F (Cobrinik et al., 1993; Schwarz et al., 1993). This transcription factor is involved in the progression of quiescent cells into S phase as evidenced by the fact that genes activated in late G1 and those encoding S phase functions contain E2F sites in their promoters (Nevins, 1992). Thus, by releasing active E2F from pRb, p130 and p107, E1a stimulates progression of cells into S phase. E1a must also alter p300 activity in order to transform cells. p300 is proposed to be a coactivator of genes required for G0/G1 arrest (Eckner et al., 1994) and since E1a represses p300 function, it can overcome this growth block. These actions of E1a are necessary to create an environment favourable for viral replication. In situations where a lytic infection does not ensue, then growth stimulation by E1a results in host cell transformation.

1.4 Signal Transduction

Gene expression is regulated at the level of transcription in response to a wide variety of extracellular stimuli. During such a response, external signals have to be transmitted from the cell surface to the nucleus, a process known as signal transduction. Until recently only one signalling pathway was understood in detail from cell surface receptor to transcription factor i.e. the stimulation of CREB activity as described in section 1.2.2. However, biochemical studies in mammalian cells have delineated a second pathway, the mitogen activated protein kinase (MAPK) pathway.

1.4.1 The MAPK Pathway

The MAPK pathway is stimulated by growth factor binding to the extracellular domain of receptor tyrosine kinases (RTKs). This is followed by receptor dimerization, stimulation of protein tyrosine kinase activity and autophosphorylation (Ullrich and Schlessinger, 1990). The phosphotyrosine and neighbouring amino acids function as a binding site for the Src homology domain, SH2, of adaptor molecules such as Grb2 which in turn recruit the guanine-nucleotide exchange factor, Sos, to the cytoplasmic surface of the plasma membrane (McCormick, 1993). There Sos catalyzes the dissociation of GDP from Ras so facilitating GTP binding (Aronheim et al., 1994). This converts Ras into its active form and allows it to interact with the next component of the pathway, Raf, thereby recruiting it to the plasma membrane. Membrane association of Raf, a MAPK kinase kinase (MAPKKK/MEKK), activates its serine/threonine kinase activity (Leevers et al., 1994) and results in phosphorylation and activation of the next component in the pathway, MAPK kinase (MAPKK/MEK) (Kyriakis et al., 1992). MAPKK is dual specificity kinase which phosphorylates MAPKs on both tyrosine and threonine residues (Nakielny et al., 1992). This dual phosphorylation is required to activate the MAPKs which then phosphorylate substrates at the minimal consensus sequence serine/threonine-proline (Davis, 1993).

Thus, the MAPK pathway can be seen to consist of a kinase cascade triggered by Ras. The pathway can be stimulated in ways other than the RTK scheme described above e.g. certain G protein coupled receptors are able to stimulate Ras and Raf activity through the action of their effectors (reviewed in Blumer and Johnson, 1994). This extra input of signals allows stimulation of the MAPKs in response to a wider variety of external stimuli.

1.4.2 MAP Kinase Homologs

Cloning of the MAPKs has led to the identification of two subfamilies. These are the extracellular regulated kinases e.g. Erk1, 2 and 3 (Boulton et al., 1991) and the stress activated protein kinases, SAPKα, β and γ (SAPKγ is also referred to as JNK1) (Derijard et al., 1994; Kyriakis et al., 1994). The two families are 40-45% identical in sequence and both require dual phosphorylation on key regulatory tyrosine and threonine residues in order to be activated. Furthermore, both families are activated by Ras. There are, however, distinct

differences which suggest that despite the similarities the families belong to separate signalling pathways.

Firstly, known MAPK stimuli differentially activate the two families. Mitogens such as epidermal growth factor (EGF), fibroblast growth factor (FGF) or phorbol ester strongly activate the Erks but have only a weak stimulatory effect on the SAPKs. Conversely, agents which induce cellular stress such as hydrogen peroxide, UV light and heat shock strongly activate the SAPKs but only weakly or moderately activate the Erks. Secondly, the MAPKK responsible for activating these proteins is different in the two cases; MEK1 and MEK2 phosphorylate and activate the Erks (Zheng and Guan, 1993) whereas SEK1 activates the SAPKs (Sanchez et al., 1994). Thirdly, the motif phosphorylated by the relevant MAPKK differs in sequence in the two groups; in the Erks it is Thr-Glu-Tyr but in the SAPKs it is Thr-Pro-Tyr. Finally, the two families differ in their substrate specificity e.g. the SAPKs phosphorylate GST-cJun fusion proteins containing the N-terminal activation domain of cJun much more efficiently than they phosphorylate myelin basic protein (MBP) whereas the opposite is true for the Erks.

A further MAPK homolog has been identified, p38/RK/p40 (Freshney et al., 1994; Han et al., 1994; Rouse et al., 1994), which appears to be distinct from the Erks since it is not activated by EGF or MEK1 and it does not phosphorylate MBP. Instead it is activated by heat shock which stimulates the SAPKs. p38, however, is distinguished from the SAPKs because it is unable to phosphorylate GST-cJun and its sequence of regulatory phosphorylation by MAPKK is Thr-Gly-Tyr (in the SAPKs it is Thr-Pro-Tyr). Therefore, p38/RK/p40 appears to be a novel MAPK and may thus represent the first identified member of a third MAPK subfamily.

1.4.3 Several MAPK Pathways In Mammalian Cells

The identification of three potential MAPK subfamilies, two of which (Erks by MEK1 and 2; SAPKs by SEK1) are activated by different upstream MAPKKs, suggests that different but related MAPK signalling pathways exist in mammalian cells. This is further suggested by the identification of three distinct mammalian MAPKKKs: Raf-1 (Kyriakis et al., 1992), MEKK1 (Lange-Carter et al., 1993) and Mos (Nebreda et al., 1993). Two of these have recently been shown to stimulate independent pathways *in vivo*; Raf-1 activates MEK1 thus leading to stimulation of Erk activity whereas MEKK1 activates SEK1 thereby stimulating SAPK activity (Yan et al., 1994). The pathway leading to p38 stimulation has not yet been established. However, since the upstream activator of p38 is not MEK1, it suggests that p38 activation also occurs via a distinct signalling pathway.

Further support for the existence of independent mammalian MAPK cascades is given by the fact that three distinct MAPK-like cascades have been identified in yeast. One of these is involved in coordinating the physiological changes involved in mating, another is involved in cell wall construction and integrity, and the third is involved in adaptation to osmotic stress (reviewed in Ammerer, 1994). Erk1 and 2 are mammalian homologs of the yeast proteins

FUS3/KSS1 which lie on the mating pathway whereas p38 is a homolog of HOG1 which lies on the osmoregulation pathway (Rouse et al., 1994). This suggests that the different mammalian MAPK regulatory cascades have different physiological functions which is in agreement with the observation that the different MAPK subfamilies are generally activated by different stimuli and phosphorylate different substrates. However, there is also a certain degree of cross talk between the different pathways e.g. Ras is required for both Erk and SAPK activation, although Ras independent activation of the SAPKs is known. Furthermore, in certain cases the same stimuli can activate more than one pathway although to differing extents e.g. EGF strongly stimulates the Erks and also weakly activates the SAPKs. Only when all the components of the different signalling pathways have been identified can the true extent of overlap and interplay between the various cascades be determined.

1.5 Aims and Summary

The transcription factor ATF2 has been studied in this thesis with the aim of understanding the role it plays in the activation of transcription and also its role in E1a transactivation. To this end, deletion and point mutants of ATF2 were constructed and analyzed in *in vivo* transcriptional assays in the absence and presence of E1a. Such studies established the presence of a transcriptional activation domain within ATF2 which was required for the response to E1a. Further work demonstrated that this domain was responsive to certain extracellular stimuli and implies that ATF2 is an important component of a signalling pathway which regulates transcription via ATF/CRE sites. These studies are described in detail in this thesis.

CHAPTER 2:

MATERIALS AND METHODS

MATERIALS AND METHODS

MATERIALS

2.1 Reagents

Reagents were obtained from the following companies unless otherwise stated; BDH, BRL, Stratagene, Fisons, Pharmacia and Sigma.

2.2 Enzymes

DNA and RNA modifying enzymes were obtained from Boehringer Mannheim or New England Biolabs.

2.3 Radiochemicals

All radiochemicals were purchased from Amersham.

2.4 Bacterial Strains

DH5 : *E. Coli* K12 strain. F^-, recA, endA1, gyrA96, thi-1, hsdR17(r-k, m+k), supE44

2.5 Bacterial Media

Bacteria were cultured in Luria Broth (LB); 10g bacto-tryptone (Difco), 5g bacto-yeast extract (Difco) and 10g NaCl per litre, pH7.5. 15g of agar (Difco) was added per litre of LB for agar plates. Ampicillin was added to 100μg/ml for selection purposes where necessary and bacteria were grown at 37^oC unless otherwise stated.

2.6 Mammalian Cell Lines

CHO cells	: D422hprt$^-$
293 cells	: Graham et al., 1977
NIH 3T3 cells	: Jainchall et al., 1969

2.7 Mammalian Cell Media

293 cells and NIH 3T3 cells were cultured in E4 (Dulbecco's modified Eagle's medium, Gibco) containing 10% foetal calf serum (FCS). CHO cells were cultured in Minimum Essential Medium Alpha (MEMα) without ribonucleosides and deoxyribonucleosides (Gibco) containing 10% FCS.

2.8 Antibodies

ATF2 antibody (NJ2) :	Antibody raised against a peptide comprising amino acids 85-96 of human CREB2 protein.
LEXA antibody :	Neosystem, S. A,, France

2.9 Oligonucleotides

All oligonucleotides were synthesized by Iain Goldsmith and the Oligonucleotide Synthesis Unit at the Clare Hall Laboratories, ICRF

Mutagenesis and cloning primers:

1. **L8** CCGGGTCGACATGAGTGATGACAAACCCTTTCTATG
2. **P2** GGCCTCTAGAGGCATTTTTTTAAT
3. **P3** CCGGGTCGACTTTCTATGTACTGCGCCTGGA
4. **P4** GGCCTCTAGAGAATTCATTCTCAAATGGACT
5. **P5** GGCCTCTAGAAGATCTATGTTTATGGACAGCCAAATG
6. **P6** CCGGGTCGACAGATCTGCTGATCAGACCCCAACACCA
7. **P7** GGCCTCTAGAAGATCTCACAATGACACTGTCATTACG
8. **P8** CCGGGTCGACAGATCTGAAGAAGTGGGTTTGTTTAAT
9. **P9** GGCCTCTAGAAGATCTACAGTTTTTCAAGAATCTTGT
10. **P10** CCGGGTCGACAGATCTAAGAAAGCTTCAGAAGATGAC
11. **104T** AATTCAAGAAAGCTTCAGAAGATGACATTTAATGAGGATCCT
12. **104B** CTAGAGGATCCTCATTAAATGTCATCTTCTGAAGCTTTCTTG
13. **MIXT** AATTCGTATACCGAGAAATAGCAGGACTATAATGAGGATCCT
14. **MIXB** CTAGAGGATCCTCATTATAGTCCTGCTATTTCTCGGTATACG
15. **98T** AATTCAAGAAATAATGAGGATCCT
16. **98B** CTAGAGGATCCTCATTATTTCTTG
17. **96T** AATTCTAATGAGGATCCT
18. **96B** CTAGAGGATCCTCATTAG
19. **T>A IS** GATCAGGCACCAGCACCAACAAGA
20. **T>A IA** TCTTGTTGGTGCTGGTGCCTGATC
21. **S>A OA** GGCCTCTAGAGAATTCATTCTCAAATGGAGCCGCCAA
22. **T1 IS** GCTGATCAGGCACCAACACCA
23. **T1 IA** TGGTGTTGGTGCCTGATCAGC
24. **T2 IS** CAGACCCCAGCACCAACAAGA
25. **T2 IA** TCTTGTTGGTGCTGGGGTCTG
26. **L4** CCGGCCATGGTGACTGGTCGAACGAGTGGGACTGC

27. **CR2ΔREP1** TTTGAGAATGAATTCCCTCTTGCAACACCTATC

28. **CR2ΔREP2** AGGTGTTGCAAGAGGGAATTCATTCTCAAATGG

29. **GST GAL 5'** CCGGGGATCCATGAAGCTACTGTCTTCT

30. **GST GAL 3'** CCGGGAATTCCGGCGATACAGTCAACTG

31. **GST STOP C2T** AATTCCTCGAGATCGATTAGAGATCTC

32. **GST STOP C2B** TCGAGAGATCTCTAATCGATCTCGAGG

Sequencing Primers:

33. **CREB2SEQ** TAATGACAGTGTCATTGTGGCTGAT

34. **CR2SEQXBA** TTTAATGAGTTGGCGAGTCCATTTGAG

35. **CR2SEQNCO** CAGGCACCATCCTCTAACAGG

36. **GAL-PRIM** GCGACATCATCATCGGAAGAGAGTAG

37. **LEX-PRIM** CAGCAGAGCTTCACCATTGAAGGGC

38. **GST-PRIM** GCATGGCCTTTGCAGGG

EMSA Oligonucleotides (Ruden, 1990)**:**

39. **GAL4BS A** CTGCAGTCGGAGGACAGTACTCCGACCCGGG

40. **GAL4BS B** AATTCCCGGGTCGGAGTACTGTCCTCCGACT

2.10 Constructs

All DNA manipulations were carried out by standard techniques and the plasmid structures verified by sequencing.

2.10.1 Mammalian Expression Vectors

a) GAL4 N-terminal ATF2 Constructs:

All GAL4-ATF2 amino terminal fragments were constructed in the vector pAG147. This is based on pSG424 (Sadowski and Ptashne, 1989) in which the SV40 promoter has been replaced with the β-actin promoter.

Δ9: The Bgl II-NcoI fragment of Δ9pSG424 (Flint and Jones, 1991) was blunted into HindIII-XbaI pAG147.

C1, C34, C56, C78, C2: The following PCR fragments were generated using the primers indicated with Δ9 as the template; C1 (3+2), C2 (1+4), C3 (1+5), C4 (6+2), C5 (1+7), C6 (8+2), C7 (1+9), C8 (10+2). These fragments were then cloned into Δ9 either as Sal I-Xba I

fragments (C1, C2) or as Sal I-Bgl II+Bgl II-Xba I fragments (C3 +C4, C5 +C6, C7+C8) to give the final constructs.

C2-104, MIX, 98, 96: annealed oligonucleotides as indicated which encoded different 3' ends were cloned into EcoRI-XbaI of C2; C2-104 (11+12), C2-MIX (13+14), C2-98 (15+16), C2-96 (17+18)

C2pmZn: primers 1+4 were used for PCR on pm27/32 (Flint and Jones, 1991). The resultant fragment was inserted into Sal I-Xba I C2.

C2-ALL, -T1T2, -S, -T1, -T2, -T1S, -T2S: PCR was performed on C2 using oligonucleotides that introduced the desired mutation as indicated; C2-ALL (1, 19, 20 +21), C2-T1T2 (1, 19, 20 +4), C2-S (1+21), C2-T1 (1, 22, 23 +4), C2-T2 (1, 24, 25 +4), C2-T1S (1, 22, 23 +21), C2-T2S (1, 24, 25 +21). The resultant SalI-XbaI fragments were inserted into C2.

b) GAL4 Full Length ATF2 Constructs:

Full length ATF2 constructs were constructed in MLV GAL4-CREB2 (Flint and Jones, 1991), called GAL4-ATF2 in this thesis.

ΔRep: PCR using primers 1, 26, 27+28 on GAL-ATF2 produced a fragment which was subsequently cloned into Sal I-Nco I of GAL-ATF2.

ATF2-ALL, -T1T2, -S, -T1, -T2, -T1S, -T2S: the XhoI-EcoRI fragment of the appropriate C2 mutant was cloned into MLV GAL4-CREB2 to produce the full length derivative.

ATF2pmZn: pm27/32 (Flint and Jones, 1991)

c) Other GAL constructs:

GAL-E1a: Amino acids 121-223 of E1a fused to GAL(1-147) (Lillie and Green, 1989)

GAL-VP16: Amino acids 413-490 of VP16 fused to GAL(1-147) (Sadowski et al., 1988)

GAL-CTF: Amino acids 1-499 of CTF1 fused to GAL(1-147) (Martinez et al., 1991)

GAL-Sp1: Based on Gal4-Sp1 (Southgate and Green, 1991) and containing Sp1 region B instead of region A

GAL-Tat: Residues 1-48 of Tat fused to GAL(1-147) (J. Blau unpubl.) based on GAL4-Tat (Southgate and Green, 1991)

GAL-Elk: amino acids 307-428 of Elk in MLV.GAL147plink (Marais et al., 1993)

d) LEX Constructs:

These are based on MLV LEXplink which contains amino acids 3-202 of LEX plus a nuclear localization signal in the MLV plink expression vector (Marais et al., 1993; Marais & Treisman unpubl.).

LEX-C2/LEX-C2ALL: the SalI-EcoRI fragment of C2/C2-ALL was cloned into MLV LEX202plink.

LEX-Jun: amino acids 1-193 of cJun cloned into MLV LEX 202plink (R. Treisman unpubl.)

e) E1a Constructs:

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

JF12: 12s E1a cDNA expression plasmid (Schneider et al., 1987)

JN20: 13s E1a cDNA expression plasmid (Schneider et al., 1987)

150PG: E1a activation mutant (Webster and Ricciardi, 1991)

147VL: E1a activation mutant (Webster and Ricciardi, 1991)

f) Reporter Constructs:

G5E4CAT: 5 GAL binding sites upstream of the E4 TATA box driving the CAT gene (Flint and Jones, 1991)

LEXOP2: A high affinity double LexA operator upstream of a minimal TK promoter driving the CAT gene (Marais et al., 1993)

pJATLAC: β-Galactosidase gene under the control of the rat β-actin promoter (Ellis, M. unpubl.)

2.10.2 *In Vitro* Transcription/Translation Plasmids:

Sp6 E1a: 13s E1a expression plasmid (Slavicek et al., 1988)

Myc: c-Myc inserted into pKSM (Amati et al., 1993)

GAL: T7plink GAL(1-147) (R.Marais unpubl.)

GAL-C2: Xho I-Xba I(filled) fragment of C2 was cloned into Xho I-Cla I(filled) T7 plink GAL(1-147)

2.10.3 GST Constructs:

GST-TBP: All but the first three codons of TBP cloned into pGEX-3X (Caswell et al., 1993)

GST-TFIIB: TFIIB inserted into pGEX-2T (Caswell et al., 1993)

The following were all based on pGEX-KG (provided by S. Goodbourn):

GST-GAL: The GAL portion was synthesized from pAG147 by PCR using primers 29+30. It was then cloned into BamH I-EcoR I pGEX-KG.

GST-GAL C2: the fragment produced by PCR on C2 using primers 29+4 was cloned into BamH I-Xba I pGEX-KG.

GST-GAL C2ALL: as GST-GAL C2 but using primers 29+21 with C2ALL as the template.

GST-C2, -T1, -T2, -S, -ALL: GST-C2/C2-ALL were constructed by cloning the Sal I(filled)-XbaI fragment from C2/C2-ALL into Xma I(filled)-Xba I of pGEX-KG and then inserting an oligonucleotide (31+32 annealed) into EcoR I-Xba I at the 3' end of the fusion to make the 5' Sal I site unique. GST-C2T1T2, -C2S, -C2T1 and -C2T2 were then constructed by cloning in Sal I-EcoR I fragments from ATF2-T1T2, -S, -T1 and -T2 into GST-C2.

GST-ATF2: constructed by cloning the Sal I-Xho I fragment from CR2AM2 (A.Eccleston unpubl.) into Xba I-Xho I of pGEX-KG.

GST-Δ9: constructed by cloning the Sal I-Nco I fragment from Δ9pSG424 into Xba I-Nco I of pGEX-KG.

GST-ΔN: constructed by cloning the Xba I-Xho I fragment from CR2AM2 into pGEX-KG.

GST-X: constructed by cloning the Nco I-Sac I fragment of pSVX (Spandau and Lee, 1988) into pGEX-KG.

GST-E1a: constructed by cloning the Ase I(filled)-Xho I fragment of E1a (V32, G. Micklem unpubl.) into Xma I(filled)-Xho I of pGEX-KG.

GST-Elk: amino acids 307-428 Elk in pGEX30X (Marais et al., 1993)

GST-^{32}P ATF2: The BamH I-Sac I(filled) fragment of GST-ATF2 was cloned into BamH I-EcoR I (filled) pGEX-TK (Kaelin et al., 1992).

METHODS

2.11 Bacterial Transformation

a) Preparation of Competent Cells

An overnight culture of E. Coli was diluted 1:100 into fresh LB and grown at 37°C with good aeration until a density OD_{590} 0.3-0.5 was reached. The culture was then rapidly cooled on ice and the cells harvested by centrifugation (3Krpm, 10min, 4°C) in pre-cooled tubes. The bacterial cell pellet was resuspended in 0.6 of the original cell culture volume in ice cold CTG (50mM Ca Cl_2, 10%glycerol, 0.05mg/ml thymidine) and incubated for 20min on ice. After this time, the cells were harvested as before and resuspended in 0.04 volumes of ice cold CTG. The cell suspension was then aliquoted into 200µl amounts, frozen on dry ice and stored at -70°C.

b) Transformation of E. Coli

Frozen competent cells were thawed on ice and used immediately. 1-10µl of DNA were added to 100µl of cells and incubated for 15 min on ice. This was followed by a heat shock of 42°C for 2 min and a further 5min incubation on ice. 100µl of LB was then added and the cell/DNA suspension transferred to 37°C for 20 min before being plated onto LB plates containing the appropriate selection medium.

2.12 DNA Techniques

2.12.1 Plasmid Preparation

a) Small Scale (Boiling Method)

Cells from a 1.5ml overnight culture were harvested by centrifugation in a microfuge (13Krpm, 30 seconds, RT). The media was aspirated off and the cell pellet resuspended in 300µl of STET (0.1M NaCl, 10mM Tris pH 8.0, 1mM EDTA pH 8.0, 5% Triton X-100). 25µl of lysozyme (10mg/ml) was added, the mixture vortexed briefly and then placed in a boiling water bath for 45 seconds. The resultant bacterial lysate was centrifuged for 10 min at 13Krpm in a microfuge at room temperature and the pelleted debris removed from the tube with a sterile toothpick. 35µl of 3M NaOAc pH 5.2 and 420µl of isopropanol were added to the supernatent which was then mixed and left at room temperature for 5 mins. The precipitated DNA was recovered by centrifugation (13Krpm, 5 mins, RT), washed with 80% ethanol and resuspended in 50µl TE (10mM Tris pH 8.0, 1mM EDTA pH 8.0).

b) Large Scale (Alkaline Lysis Method)

Cells were harvested from a 400ml overnight culture by centrifugation at 5Krpm for 10 mins. The cell pellet was resuspended in 10mls of solution 1 (50mM glucose, 25mM Tris pH 8.0, 10mM EDTA pH 8.0) and the cells lysed by the addition of 20mls of solution 2 (0.2M NaOH, 1% SDS). This was left to stand for 5 mins at room temperature. Cell debris and

chromosomal DNA were then precipitated from the lysate by the addition of 15mls of solution 3 (5M KOAc pH4.8) and pelleted by centrifugation at 7Krpm for 10 mins. The supernatent was filtered through gauze to remove any remaining debris and plasmid DNA precipitated by the addition of 27ml of isopropanol to the filtered solution. The DNA was recovered by centrifugation at 8Krpm for 10 mins and then resuspended in 4ml of TE. To this was added 4.8g of CsCl and 30µl of 10mg/ml EtBr. The solution was then loaded into a 5ml Beckman polyallomer tube, heat sealed and spun in a Beckman VTi65 rotor at 20°C for either 4hr at 64Krpm or 16hr at 55Krpm. The resulting plasmid band was removed from the tube with a hypodermic needle and syringe and the EtBr removed by repeated extraction with water saturated isobutanol. The volume of the plasmid solution was increased 3 fold with TE in order to reduce the CsCl concentration, 2.5 volumes of ethanol were added and the precipitated DNA recovered by centrifugation at 10Krpm for 10 min in a Sorval SS-3B rotor. The DNA was resuspended in 400µl of TE plus 5µl of 10mg/ml RNase A and incubated at 37°C for 10mins. This was followed by phenol/chloroform/isoamyl alcohol (25:24:1) extraction, chloroform/isoamyl alcohol (24:1) extraction and ethanol precipitation. The purified DNA was resupended in 100-400µl of TE and the concentration determined by spectrophotometric absorbance at 260nm ($1OD_{260}$=50µg/ml of double stranded DNA).

2.12.2 Restriction Digests

5µl of miniprep DNA or 1-3µg of maxiprep DNA was digested in a 20-50µl volume with approximately 10 units of enzyme for 90mins. 1/10 volume of 10X loading buffer (25% Ficoll, 0.02% bromophenol blue, 0.02% xylene cyanol) was added and the sample run on an agarose gel. All digests were carried out in the reaction buffers recommended and supplied by the enzyme manufacturers.

2.12.3 Agarose Gel Electrophoresis

Agarose gels were prepared and electrophoresed in 1X TBE (100mM Tris.Borate, 1mM EDTA pH8.3). The gels also contained EtBr (0.5µg/ml) to permit DNA detection under ultraviolet light. The concentration of agarose used varied from 0.8-2.0% according to the size of the DNA fragment being examined.

2.12.4 Recovery of DNA fragments From Agarose Gels

a) Frament Size >500bp

DNA fragments of size 500bp and greater were recovered from agarose gels by using the "GENECLEAN II" Kit (BIO 101 Inc.). The fragment of interest was excised from the gel, placed in an eppendorf tube and 4.5X the volume of the gel slice of NaI stock solution plus 0.5X the volume of TBE modifier were added. The tube was heated to 45-55°C for 5 mins to dissolve the agarose, 5µl of "GLASSMILK" was added for 5µg or less of DNA and the DNA allowed to bind to the "GLASSMILK" for 5mins at room temperature with occassional mixing.

The matrix with the bound DNA was pelleted by spinning in a microfuge for 5 secs, washed with 500µl of ice cold "NEW WASH" and the process repeated twice. The pellet was reuspended in water equal in volume to the amount of "GLASSMILK" used, incubated at 45-55°C for 2 mins to elute the DNA. and the "GLASSMILK" removed by centrifugation in a microfuge for 1min. The elution procedure was repeated and the eluates pooled.

b) Fragment Size<500bp

Fragments of less than 500bp were recovered by using the QIAEX kit (QIAGEN Inc.). 3X the volume of buffer QX1 and 0.1X the volume of 1M mannitol were added to the excised gel slice. This was heated to 50°C for 10 mins to solubilize the agarose and 10µl of QIAEX suspension was added per 5µg of DNA. The mixture was tumbled at room temperature for 10 mins, centrifuged for 30 secs in a microfuge and the pellet washed with 500µl of buffer QX2. This washing step was repeated and followed by two 500µl washes with buffer QX3. Care was taken to remove all the supernatent after the final spin. The pellet was resuspended in 20µl of water, incubated for 5 mins at room temperature, spun in a microfuge for 30 secs and the supernatent containing the eluted DNA transferred to a clean tube. Elution was repeated and the two DNA eluates combined.

2.12.5 Blunt End Formation

a) Filling In Protruding 5' Ends

1µl of a 2mM solution of all four dNTPs and approximately 2 units of Klenow were added to a completed restriction digest and incubated at room temperature for a further 15-30 mins.

b) Removing Protruding 3' Ends

The same procedure was used as in a) except on this case the enzyme used was T4 DNA polymerase.

2.12.6 Ligation

Ligations were generally performed with 100ng of purified vector and 500ng of purified insert in a volume of 20µl containing 2µl of 10X ligation buffer (200mM Tris pH7.6, 50mM $MgCl_2$, 50 mM DTT) plus 1µl of 5mM ATP. The reaction was allowed to proceed for 30 mins-4 hours after which time 5µl was transformed into E. Coli.

2.12.7 Polymerase Chain Reaction (PCR)

PCR reactions were carried out in a 100µl volume containing 1µg of template DNA, 1µg of each primer (generally about 20 bp in length), 4µl of a 5mM dNTP mix, 10µl of 10X reaction buffer (100mM Tris pH8.3, 500 mM KCl, 15mM $MgCl_2$, 0.1% gelatin) and 2.5 units of AmpliTaq polymerase. This was overlayed with 100µl of mineral oil to prevent evaporation during heating. The following amplification protocol was routinely used; template denaturation at 94°C for 1 min, primer annealing at 50°C for 1 min and primer extension at 72°C for 1 min.

This was repeated for a total of 30 cycles and followed by an additional extension step at 72°C for 5 mins.To ensure the ends were completely filled in, 1µl 5mM dNTPs, 1µl 0.1M DTT and 1µl T4 DNA polymerase were added to the amplified sample which was then incubated at 37°C for 1 hour. This ensured efficient cleavage of the fragment ends.

2.12.8 Mutagenesis/Deletion By PCR

PCR was used to introduce mutations or deletions into DNA by the method depicted in figure 2.1. PCR steps 1 and 2 were performed as described above and the products were then annealed by heating to 65°C for 5 mins and cooling at room temperature for 10 mins. The DNA overhangs were filled in by addition of 4µl 5mM dNTPs, 2.5 units of AmpliTaq polymerase and incubation at 37°C for 30 mins. PCR step 3 was carried out on the product by the standard procedure.

2.12.9 Sequencing

Sequencing was carried out using the Sequenase Version 2.0 kit (US Biochem. Corp.) which provides the following buffers;

5X Reaction buffer: 200mM Tris pH7.5, 100mM $MgCl_2$, 250 mM NaCl

Labelling Mix: 7.5µM each of dGTP, dCTP, dTTP

Termination Mix: 50mM NaCl, 80µM of three of the four dNTPs and 8µM of the ddNTP equivalent of the fourth dNTP e.g. ddA termination mix contains 80µM dCTP, dGTP, dTTP and 8µM ddATP

Stop Solution: 95% formamide, 20mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol

a) Template Denaturation

3-5µg of plasmid DNA was denatured by adding NaOH to 0.2M and EDTA to 0.2mM in a total volume of 20µl and incubating at room temperature for 5 mins. The solution was neutralized by adding 2µl of 2M NH_4Ac pH4.6 and the DNA precipitated by adding 60µl ethanol, mixing and placing the sample on dry ice for 5 mins. The sample was centrifuged in a microfuge for 6 mins, the DNA pellet was washed once with 250µl 80% ethanol, dried and resuspended in 7µl of water. All the denatured DNA was used for sequencing.

b) Annealing

5-10ng of primer and 2µl 5X reaction buffer were added to the denatured DNA to give a final volume of 10µl. This was heated to 65°C for 5 mins and then cooled by placing at room temperature for 10 mins.

c) Labelling

To the annealed DNA mixture was added 1µl 0.1M DTT, 2µl labelling mix diluted 1:5, 0.5µl ^{35}S dATP and 2µl sequenase diluted 1:8. This was mixed and incubated at room temperature for 2-5 mins.

Figure 2.1 Mutagenesis/deletion by PCR

PCR using the indicated primers was performed as described in the text.

OS: outside sense primer

IS: inside sense primer

IA: inside antisense primer

OA: outside antisense primer

X: desired mutation or deletion contained in the primer

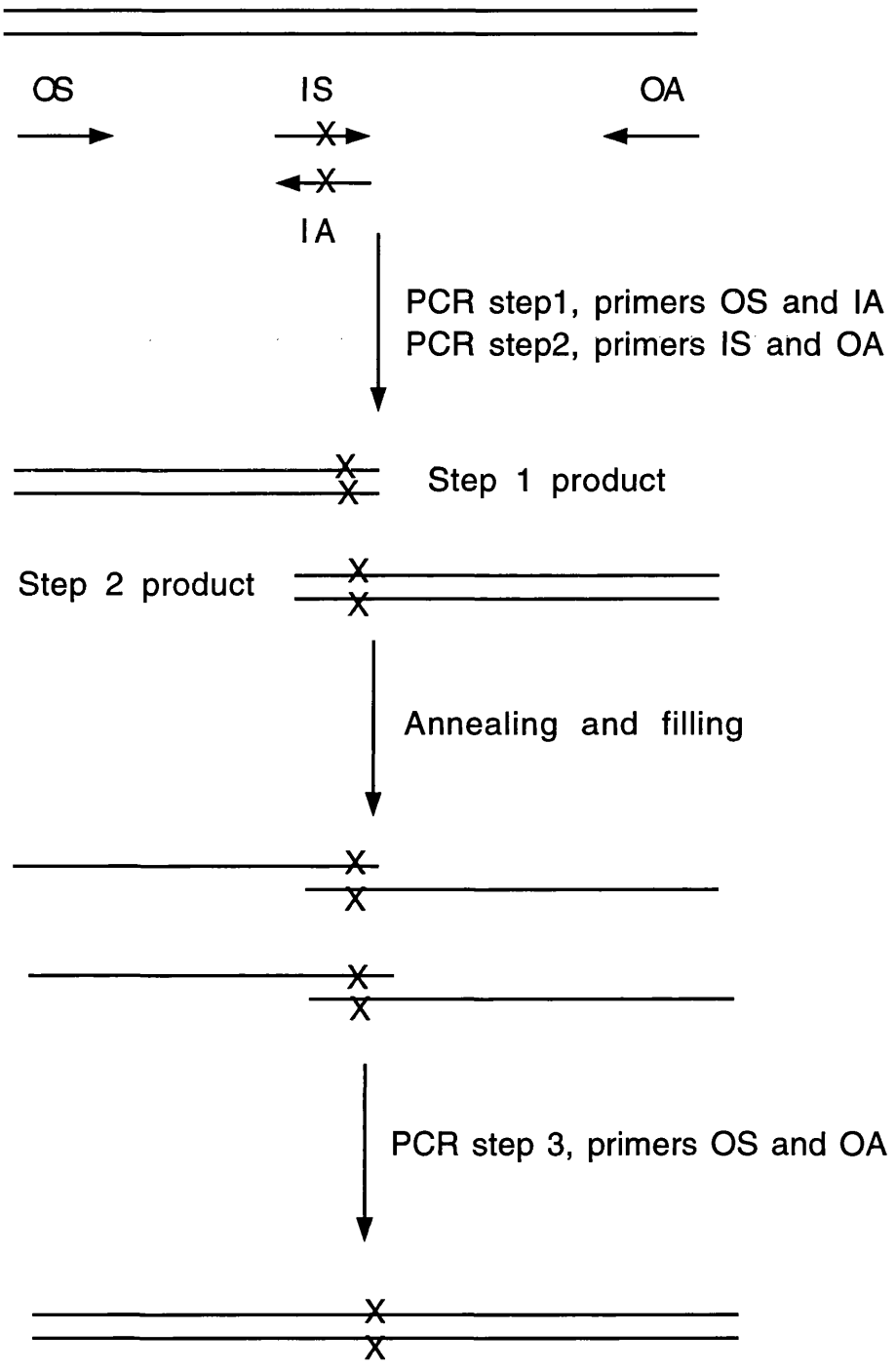


Figure 2.1 Mutagenesis/deletion by PCR

d) Termination

3.5µl of the labelling reaction was transferred to each of four tubes containing 2.5µl of a different termination mix, the samples were mixed and incubated at 37°C for 2-5 mins. The reaction was stopped by the addition of 4µl of stop solution to each tube, heated in a boiling water bath for 2 mins and immediately loaded onto a denaturing sequencing (polyacrylamide) gel.

2.12.10 Denaturing Polyacrylamide Gel Electrophoresis

(i) Straight Gel

The following stock solutions were purchased from National Diagnostics and used according to the manufacturers instructions;

SequaGel Concentrate: 237.5g acrylamide, 12.5g methylene bisacrylamide, 500g urea per litre (19:1 acrylamide:bisacrylamide)

SequaGel Diluent: 500g urea per litre

SequaGel Buffer: 500g urea per litre of 10X TBE

4-8% acrylamide gels were poured depending on the length of the DNA fragments to be separated and gel polymerization was initiated with 400µl 10% ammonium persulphate (APS) and 20µl TEMED per 50ml acrylamide mix. Electrophoresis was carried out at 40W for 2-5 hours in 1X TBE buffer. Gels were fixed in 10% glacial acetic acid, 10% methanol for 15 mins before being dried on 3MM paper (Whatman) and exposed to X-ray film (XAR Kodak).

(ii) Gradient Gel

The following solutions were used to prepare a 6% buffer gradient polyacrylamide gel;

5X Mix: 15ml 40% acrylamide (19:1 acrylamide:bisacrylamide), 50ml 10X TBE, 46g urea, 5mg bromophenol blue in a total volume of 100ml

0.5X Mix: 37.5ml 40% acrylamide (19:1 acrylamide:bisacrylamide), 12.5 ml 10X TBE, 115g urea in a total volume of 250ml

To 8ml of 5X Mix was added 40µl 10% APS and 16µl TEMED and to 42ml 0.5 Mix was added 210µl 10% APS and 84µl TEMED. 6ml of 0.5X Mix was drawn up into a 25ml pipette, followed by 6ml of 5X Mix and a rough gradient between the two solutions formed by drawing 2-3 air bubbles through the two solutions. The mix was poured between two glass plates separated by 0.35mm spacers and the remainder of the 0.5X Mix was added to complete the gel. After polymerization the gradient gel was treated exactly the same as for straight gels.

2.12.11 Electrophoretic Mobility Shift Assay (EMSA)

(i) Annealing Oligonucleotides

Single stranded complimentary oligonucleotides were annealed by mixing 50µl of each strand (1mg/ml solution) in a total volume of 150µl containing 30µl 5X annealing buffer (335mM Tris pH 7.5, 65mM $MgCl_2$, 33.5 mM DTT, 6.5mM spermidine, 65mM EDTA). This

mix was heated for 2 mins at 90°C, 10 mins at 65°C, 10 mins at 37°C and 5 mins at room temp. After annealing, the oligonucleotides were ethanol precipitated and resuspended at a concentration of 1mg/ml.

(ii) End Labelling Annealed Oligonucleotides

Annealed oligonucleotides were end labelled by filling in the 5' overhangs using reverse transcriptase. 100ng annealed oligonucleotide was added to 2µl 5X reverse transcription buffer (50mM Tris pH7.5, 75mM KCl, 10mM DTT, 3mM $MgCl_2$), 1µl 10mM dGTP, 1µl 10mM dTTP, 1µl (10µCi) α-^{32}P dATP, 1µl (10µCi) α-^{32}P dCTP and 1µl MMLV-reverse transcriptase in a total volume of 10µl. This mix was incubated at 37°C for 30 mins, 1µl dATP and 1µl dCTP were added and the incubation continued for a further 10 mins. 30µl TE buffer was added to the reaction and unincorporated nucleotides were removed from the probe on a G-50 sephadex column.

(iii) G-50 Spin Column

Sephadex G-50 (Pharmacia) was hydrated by the addition of TE buffer and hydration was allowed to proceed at room temp overnight after which time the slurry was autoclaved. A 1ml syringe was plugged with polymer wool (Interpet) and filled with the G-50 slurry. The column was placed in a 15ml Falcon tube and centrifuged at 1.5Krpm for 4 mins at room temp. The 50µl labelling reaction was loaded onto the column and the column spun as before. The labelled probe was collected in an Eppendorf tube which had been placed in the bottom of the Falcon tube prior to centrifugation.

(iv) Binding Reactions

20µl reactions containing 10µl 2X Chasman buffer (50mM HEPES pH7.6, 100mM KCl, 10mM $MgCl_2$, 0.2mM EDTA, 0.2% NP40, 20% glycerol), 0.5mM PMSF, 5µg BSA, 1µg poly[dI-dC] (Pharmacia), 0.2ng labelled probe and 1-5µl nuclear extract were set up on ice. The amount of nuclear extract varied to ensure that each reaction contained an equal amount of protein. Probe was added last and binding of protein to probe was allowed to proceed for 20 mins at room temperature. The reaction was then loaded onto a 4% non-denaturing polyacrylamide gel. Binding reactions were also performed in the presence of antibody in order to verify the nature of any complexes formed. In this case, 1µl of the required antibody was added to the binding reaction 5 mins before probe addition.

(v) Non-Denaturing Gel Electrophoresis

A 50ml mix containing 4.5ml acrylamide solution (44:0.8 acrylamide: bisacrylamide) (Severn Biotech Ltd.) and 2.5ml 10X TBE buffer was used to pour 4% polyacrylamide gels. Gel polymerization was initiated by the addition of 300µl 10% APS and 50µl TEMED and the gels were electrophoresed at 200V for 90 mins in 0.5X TBE buffer. Gels were dried onto 3MM paper and exposed to Kodak XAR film at -70°C with an intensifying screen.

2.13 Protein Techniques

2.13.1 *In vitro* Transcription and Translation

In vitro transcription was carried out using the Stratagene mCAP kit according to the manufacturers instructions. 4μg plasmid was linearized at the 3' end of the cDNA gene to be transcribed by digestion with the appropriate restriction enzyme, phenol/chloroform extracted and ethanol precipitated. This was then used in a 100μl transcription reaction containing 20μl 5X transcription buffer (200mM Tris pH7.5, 250mM NaCl, 40mM $MgCl_2$, 10mM spermidine), 4μl 10mM rUTP, 4μl 10mM rATP, 4μl 10mM rCTP, 4μl 1mM rGTP, 4μl 0.75M DTT, 0.5mM CAP analogue (5'7meGppp5'G Pharmacia), 1μl RNAguard (Pharmacia) and 10U T7/Sp6 RNA polymerase. The reaction was incubated at 37°C for 60 mins after which time 10U DNase I were added to remove the DNA template and the incubation continued for a further 30 mins at 37°C. The RNA was phenol/chloroform extracted, ethanol precipitated and resuspended in 30μl water. 1.5μl was run on an agarose gel to determine the level of RNA synthesis.

The RNA was translated in rabbit reticulocyte lysate from Promega. Typically 4μl RNA was added to a mix of 35μl reticulocyte lysate, 1μl RNAguard, 1μl 1mM amino acid mixture minus methionine and 5μl ^{35}S-methionine (10μCi/μl). Translation was performed at 30°C for 30-60 mins and the products analyzed by SDS-polyacrylamide gel electrophoresis.

2.13.2 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)

10-12% separating gels were prepared using the stock solutions; 30% acrylamide (29:1 acrylamide: bisacrylamide) and 3X buffer (68.1g Tris pH8.8, 5ml 10% SDS in 500ml). For a minigel, 9ml of gel was prepared and polymerized using 90μl 10% APS and 3.6μl TEMED. For large gels, 36ml of gel was prepared and polymerized using 360μl 10% APS and 14.4μl TEMED. The gel mix was poured between two glass plates and the surface overlaid with isobutanol to remove any bubbles and to ensure a straight interface with the stacking gel. After polymerization, the isubutanol was washed off and the 4% stacking gel (13ml 30% acrylamide (29:1 acrylamide: bisacrylamide), 12.5ml 1M Tris pH6.8, 1ml 10% SDS in 100ml) was poured on top. 2ml/15ml stack was used for mini/large gels respectively and polymerization initiated by the addition of 10μl 10% APS and 1μl TEMED per ml stack solution.

4X Laemlli buffer (0.5ml 1M Tris pH6.8, 0.8ml glycerol, 1.6ml 10% SDS, 0.4ml β-mercaptoethanol, 0.05% bromophenol blue and 4ml water) was added to the sample to give a final concentration of 1X, mixed and boiled for 2 mins to denature the proteins before loading onto the gel. Electrophoresis was carried out at 200V in Tris-Glycine buffer (14.4g glycine, 3g Tris, 10ml 10% SDS per litre). Proteins were visualized by Coomassie staining. The gel was fixed in destain solution (250ml MeOH, 70ml glacial acetic acid per litre) for 15 mins, placed in staining solution (2.5g Coomassie Brilliant Blue, 500ml MeOH, 100ml glacial

acetic acid per litre) for a minimum of 30 mins and then shaken in destain solution until the protein bands were clearly visible. To visualize ^{32}P-labelled proteins, the gel was fixed for 15 mins, dried onto 3MM paper and exposed to Kodak-XAR film to visualize the bands. ^{35}S-methionine labelled proteins were enhanced by soaking in Amplify (Amersham) for 15-30 mins after fixing and then treated as for ^{32}P-labelled proteins. Gels were not fixed before Western blotting.

2.13.3 Western Blotting

Proteins were transferred from SDS-PAGE to nitrocellulose by placing the gel onto nitrocellulose and sandwiching them between 3MM paper. This was placed into a blotting apparatus (minigel, Biorad; large gel, Hoeffer Scientific) according to the manufacturers instructions. Transfer was carried out in 25mM Tris, 192mM glycine, 20% MeOH at 100V for 60 mins (minigel) or 120 mins (large gel). The nitrocellulose filter was blocked in PBS-T:5% (PBS plus 0.1% Tween20 (PBS-T) and 5% Marvel) for 1hr at room temp or overnight at 4°C. After blocking, the filter was washed 3x10 mins in PBS-T and incubated for 1hr at room temp in PBS-T:1% (PBS-T containing 1% Marvel) with a 1:1000 dilution of the appropriate antibody. The filter was washed 3x10 mins in PBS-T and incubated with the second antibody, biotinylated anti-rabbit IgG (Amersham) at a dilution of 1:1000 in PBS-T:1% for 1hr at room temp. The filter was washed 3x10 mins in PBS-T then incubated in PBS-T:1% containing streptavidin-biotinylated peroxidase conjugate (Amersham) at a 1:1000 dilution for 1hr at room temp. The filter was washed 3x10 mins in PBS-T followed by a final wash in PBS for 10 min. The peroxidase conjugate bound to the filter was detected using the ECL kit (Amersham) by following the manufacturers instructions.

2.13.4 Affinity Purification of GST Fusion Proteins

This was performed essentially as described in (Kaelin et al., 1991) and is a modification of the original Smith and Johnson protocol (Smith and Johnson, 1988). An overnight culture of E. Coli transformed with the GST fusion plasmid of interest was diluted 1:10 into fresh LB plus amp and grown for 1 hr at 37°C with shaking. IPTG (Sigma) was added to a final concentration of 0.1mM to induce expression of the fusion protein and the culture incubated for a further 90 mins-4 hours. The bacteria were harvested by centrifugation at 5Krpm for 5 mins at 4°C and all subsequent steps were performed at 4°C using pre-chilled solutions. The bacterial pellet was resuspended in 1/10 the original culture volume of NETN (100mM NaCl, 1mM EDTA, 20mM Tris pH8.0, 0.5% NP-40) and sonicated with 3X5-10 sec bursts. The sonicate was centrifuged at 10Krpm for 5 mins to remove the cell debris and the clarified supernatent was incubated for a minimum of 30 mins with 25µl glutathione sepharose suspension (1:1 glutathione sepharose 4B (Pharmacia) : NETN plus 0.5% Marvel) per ml of sonicate. The beads were washed 3X in NETN to dislodge any contaminating bacterial proteins and were then stored either at 4°C in an equal volume of NETN (short term

storage) or at -20°C in an equal volume of NETN containing 50% glycerol (long term storage). In either case, the beads were washed in fresh NETN or the required buffer prior to use. Levels of protein purified were estimated by boiling the bound beads in Laemli buffer, running the proteins on SDS-PAGE and visualizing the proteins by Coomassie staining. Once the level of expression of a particular GST protein had been determined, it was possible to adjust the ratio of glutathione-sepharose suspension: sonicate in order to obtain a known concentration of protein bound to the beads.

2.13.5 Elution and Protease Cleavage of Purified GST Fusion Proteins

To elute the bound fusion protein, the beads were incubated in 1-10X the bead volume of a freshly prepared solution of 20mM glutathione (Sigma) in 100mM Tris pH8.0, 120 mM NaCl for 30 mins at 4°C with agitation. The beads were pelleted out of solution (14Krpm, 15 secs) and the supernatent containing the eluted protein carefully removed to a fresh tube.

To cleave the GST moiety from the fusion, the eluted protein solution was adjusted to contain 150mM NaCl and 2.5mM $CaCl_2$ and 100ng human thrombin was added per 50µg fusion protein. Cleavage was performed at 25°C and allowed to proceed for 1 hr.

2.13.6 *In Vitro* Phosphorylation of Purified GST Fusion Proteins Using PKA

The following method describes phosphorylation of GST beads prepared from a 10ml induced culture (i.e. 1ml sonicate) and was scaled up as required. Approximately 12µl beads loaded with fusion protein were washed 1X with HMK buffer without DTT (10X HMK buffer: 200mM Tris pH7.5, 1M NaCl, 120mM $MgCl_2$) and then resuspended in 30µl reaction mix containing 3µl 10X HMK buffer plus 10mM DTT, 2µl (20µCi) ^{32}P-γ ATP and 1µl HMK (Sigma 250U/vial, 25µl 40mM DTT was added to the vial to resuspend the enzyme) per ml of sonicate. Labelling was carried out at 4°C for 30 mins and the tube was flicked occasionally to resuspend the beads in the reaction mix. After this time 1ml HMK Stop buffer (10mM Na phosphate pH 8.0, 10mM Na pyrophosphate, 10mM EDTA, 1mg/ml BSA) was added, the solutions mixed and the beads recovered from the suspension by centrifugation (14Krpm, 30 secs). The supernatent was carefully removed and discarded and the labelled protein eluted from the beads as described in section 2.13.5. If necessary the GST portion was cleaved from the fusion (section 2.13.5). Labelling efficiency was determined by counting a 25µl aliquot for ^{32}P and generally 5-25µl of labelled protein solution gave 250,000cpm. Labelled proteins were also visualized by autoradiography following SDS-PAGE.

2.13.7 *In Vitro* Phosphorylation of Purified GST Fusion Proteins Using Erk1/2

Glutathione-sepharose beads bound with equal amounts of GST proteins were used as the substrate for *in vitro* MAPK phosphorylations. 1µl of these beads containing 0.1-5µg of protein were added to the reaction mix of 5µl 2X phosphorylation buffer A (40mM β-glycerophosphate, 3mM EDTA, 0.8mM Na_3VO_4, 10mM NaF), 1µl 100mM $MgCl_2$, 1µl 10X protease/phosphatase inhibitor mix (10mM DTT, 0.5mg/ml PMSF, 50µg/ml aprotinin, 50µg/ml

leupeptin, 50μg/ml pepstatin A, 100mM benzamidine, 1μg/ml okadaic acid), 1μl ATP (1:10 dilution ^{32}P-γATP [6000Ci/mM]:1mM ATP) and 1μl Erk1/2 preparation (kindly provided by M.A. Price). The reaction was incubated at 30°C for 20 mins before being boiled in Laemlli buffer, run on SDS-PAGE and the phosphorylated proteins visualized by autoradiography.

2.13.8 GST Binding Assays

(i) Column Chromatography: ^{32}P-Labelled ATF2 as Probe

A blue tip (Gilson, France) was plugged with polymer wool. 100μl suspension of glutathione-sepharose beads loaded with fusion protein (1:1 beads:5% BLOTTO (50mM Tris pH7.5, 50mM NaCl, 1mM EDTA, 1mM DTT, 5% Marvel)) was packed into the tip to generate a 50μl column. An equivalent amount of protein was used per set of columns (typically 5μg/column) and this was achieved by either mixing non-bound gluathione-sepharose beads with gluathione-sepharose beads containing protein to give bead suspensions of equal protein concentrations or by altering the ratio of sonicate:beads used during protein purification (section 2.13.4). The columns were washed twice with 1% BLOTTO (the percentage indicates the concentration of Marvel in the buffer).

^{32}P-labelled ATF2 was prepared as in section 2.13.6 and 10μl (~250Kcpm) in 40μl of 1% BLOTTO per column was heated to 55°C for 7 mins and immediately chilled on ice. It was loaded onto each of the columns, recycled 5X and allowed to bind for 1 hour. The columns were then washed 3X with 0.5ml 0% BLOTTO to remove unbound ATF2. Any remaining bound ATF2 was eluted by washing the columns firstly with 6X50μl 0% BLOTTO containing 0.5M NaCl and then with 6X50μl 0% BLOTTO containing 1M NaCl. The collected fractions (flow through, washes and elutions) were run on SDS-PAGE and any ATF2 present in the fractions was visualized by autoradiography.

(ii) Column Chromatography: ^{35}S-Labelled E1a as Probe

The same basic procedure, as described for column chromatography using ^{32}P-labelled ATF2, was used for ^{35}S-labelled E1a but with the following modifications:

-- BLOTTO was replaced with MJ buffer (0.1M KCl, 10mM Hepes-KOH pH7.8, 5mM $MgCl_2$, 0.1mM $ZnCl_2$, 0.1mM EDTA, 2mM DTT)

-- 10μl of *in vitro* translated ^{35}S-E1a (prepared as in section 2.13.1) in 100μl of MJ buffer plus 10μl 100mg/ml BSA was loaded onto each 50μl column

-- After binding the columns were washed once with 1ml MJ buffer and once with 50μl MJ buffer (last wash)

-- Bound ^{35}S-E1a was eluted from the columns by recovering the beads from the columns after washing, boiling them in Laemlli buffer and running the eluted proteins on SDS-PAGE

(iii) Far Western Blotting

GST fusion proteins were purified, eluted and cleaved as described in sections 2.13.5 and 2.13.6. Equal amounts of protein were run on SDS-PAGE (generally 2-5μg) and then transferred to nitrocellulose (section 2.13.3). The filter was blotted in 5ml 5% BLOTTO with 100μl (~1250Kcpm) ^{32}P-labelled ATF2 (section 2.13.6) as the probe i.e. 250Kcpm labelled protein/ml hybridization buffer. The probe was heated to 55°C for 7 mins and immediately chilled on ice before being used. The blot was probed overnight and then washed in 3X100ml TNE-50 (10mM Tris pH7.5, 50mM NaCl, 1mM EDTA, 1mM DTT) for approximately 1 hour per wash. To detect any bound ^{32}P-ATF2 present after washing, the blot was covered in Saran wrap and exposed to X-ray film at -70°C with an intensifying screen. The following variations were used:

-- Binding was performed at room temperature or 4°C

-- BLOTTO was replaced with Hyb buffer (20mM HEPES-KOH pH7.7, 75mM KCl, 0.1mM EDTA, 2.5mM $MgCl_2$, 1% Marvel with or without 0.05% NP-40)

-- The proteins on the filter were renatured prior to probing using the following protocol (Kaelin et al., 1992). After SDS-PAGE, the proteins were transferred to nitrocellulose in a buffer lacking methanol (25mM Tris, 192mM glycine, 0.01% SDS). All subsequent steps were performed at 4°C using a sufficient volume of buffer to amply cover the filter (100-250ml). Firstly, the filter was placed in blocking buffer (1X HBB, 5% Marvel, 1mM DTT, 0.05% NP-40 where 10X HBB is 250mM Hepes-KOH pH7.7, 250mM NaCl, 50mM $MgCl_2$) and rocked overnight at 4°C. The blocking solution was then discarded, denaturation buffer (1X HBB, 6M GdHCl, 1mM DTT) added and the filter rocked in this buffer for 10 mins. This buffer was discarded, replaced with fresh denaturation buffer and the filter washed for a further 10 mins. The buffer was decanted from the filter into a cylinder and diluted with an equal volume of 1X HBB containing 1mM DTT thereby reducing the GdHCl concentration in the buffer twofold. Half of this solution was used to wash the filter for 10 mins and it was then discarded. The remaining solution was diluted 1:1 with 1X HBB plus 1mM DTT to reduce the GdHCl concentration to 1.5M and half of this solution was again used to wash filter. This procedure was repeated twice. Thus the filter was exposed to buffer containing 6M, 6M, 3M, 1.5M, 0.75M, 0.187M GdHCl in all so ensuring gradual renaturation of the proteins bound to it. The filter was then washed 2X10 mins in 1X HBB plus 1mM DTT, 1X60 mins in 1X HBB, 5% Marvel, 1mM DTT, 0.05% NP-40 and 1X 30 min in 1X HBB, 1% Marvel, 1mM DTT, 0.05% NP-40. The filter was then placed in the required hybridization buffer and probe added as before.

-- Nuclear extract (section 2.14.9) prepared from a confluent 100mm dish was included in the hybridization step.

(iv) Pull Down Assay

This was a modification of the assay described in (Hagemeier et al., 1992). 25µl of 50% glutathione-sepharose containing approximately 5µg of the required GST protein was washed once in EBC-M (140mM NaCl, 50mM Tris pH8.0, 0.5% NP-40). The beads were then resuspended in 200µl EBC-M plus 2µl BSA (100mg/ml) and tumbled on a wheel for 5 mins at room temperature. 2-5µl *in vitro* translated ^{35}S-labelled protein, depending on the efficiency of translation, was added to each batch of beads to give an approximately equal amount of protein in each case. This mix was tumbled for 1 hour at romm temperature. The beads were washed 3X in NETN, resuspended in Laemmli buffer and boiled to release and denature any bound proteins. The samples were run on SDS-PAGE and any labelled proteins bound to the beads detected by autoradiography.

2.14 Mammalian Tissue Culture Techniques

All cell lines were cultured as monolayers in 10 ml media (section 2.7)/100mm plate (Nunclon) and were incubated at 37°C with 10% CO_2 unless otherwise stated. Cells were split every 3-4 days1:10-15 into fresh media.

2.14.1 Transient Transfection of 293 Cells

Near confluent 293 cells were split 1:10 24 hours prior to transfection in order to obtain a cell density of 30-40% confluency at the start of transfection. Calcium phosphate-DNA precipitates were set up by preparing a mix of plasmid DNA (5-10µg), 360µl 0.1X TE (1mM Tris pH8.0, 0.1mM EDTA) and 40µl 2.5M $CaCl_2$. This mix was added dropwise to 400µl HBS (70mM HEPES pH7.1, 280mM NaCl, 1.5mM sodium phosphate) whilst bubbling air through the HBS to obtain a fine precipitate. The precipitate was allowed to stand at room temperature for 15-20 minutes before being added dropwise to the cell culture. The cells were then incubated for 8-18 hours after which time the media was aspirated off and the cells washed 2-3 times 8ml PBS to fully remove the DNA preciptate. 10ml E4 containing 10% FCS was added to the cells which were cultured for 40-48 hours before harvesting.

2.14.2 Transient Transfection of CHO Cells

1.2-1.5X10^6 CHO cells were seeded per 100mm dish and cultured for 16-24 hours. The cells were then washed once with PBS, 4ml warm DEAE-dextran mix (4ml MEMα plus 10% FCS, 20µl 10mg/ml DEAE-dextran, 10-15µg plasmid DNA) added and the cells incubated for 6 hours. The DEAE-dextran mix was aspirated off and the cells shocked by the addition of 2ml MEMα containing 10% DMSO for 2 mins. After this time the DMSO mix was aspirated off and the cells washed twice in PBS. 4ml MEMα plus 10% FCS and 0.1mM chloroquine (diphosphate salt) was added and the cells incubated in this mix for 1 hour. The mix was then removed, the cells washed once in PBS and 8ml MEMα plus 10% FCS added. The cells were incubated for 48 hours before harvesting.

2.14.3 Transient Transfection of NIH 3T3 Cells

The CHO transfection protocol was followed with the following modifications. $0.8X10^6$ NIH 3T3 cells were seeded per 100mm dish 16 hours before transfection. The DEAE-dextran mix consisted of 4ml PBS, 10µl 10mg/ml DEAE-dextran and 10-15µg plasmid DNA and was left on the cells for only 1 hour. The chloroquine incubation (4ml E4 plus 10% FCS and 0.1mM chlororquine) was carried out for 4 hours and the final incubation before harvesting performed in 8ml E4 plus 10% FCS.

2.14.4 Serum and UV Induction of Transfected Cells

Following transfection NIH 3T3 cells were incubated in media containing 0.5% serum for 40 hours. Cells were then incubated for a further 8 hours without treatment (control) or were supplemented with FCS to a final concentration of 20% (serum induction) or were irradiated with UV-C ($40J/m^2$) (UV induction). Induced cells were harvested 8 hours after stimulation.

2.14.5 Harvesting and CAT Assay Analysis of Transfected Cells

Transfected cells were washed twice with PBS at room temperature. 300µl lysis buffer (10mM Tris pH8.0, 1mM EDTA, 150mM NaCl, 0.65% NP-40) per dish was then added and left for 2 mins to ensure all cells were lysed. The lysate was transferred to an Eppendorf tube, cooled and centrifuged (14Krpm, 2 mins, 4°C) to pellet out any cell debris. The supernatent was removed to a fresh tube and stored at -20°C.

The relative protein concentration of each lysate was measured as described in section 2.14.7. An equivalent amount of protein was used per CAT assay and lysis buffer was added, if necessary, to ensure that the lysate volume was constant at 60µl. A blank consisting of 60µl lysis buffer only was also set up to measure the background level of acetyl CoA conversion. This volume of extract/buffer was heated to 68°C for 10 mins to inactivate the deacetylases and then cooled on ice. To it was added 20µl 8mM chloramphenicol, 20µl 0.5mM acetyl coenzyme A (1:10 ^{14}C-acetyl CoA [54mCi/mmol]:0.5mM cold acetyl CoA) and 4µl 2M Tris pH8.0. This was mixed and incubated at 37°C for at least 5 hours. The reaction was then chilled on ice and the acteylated chloramphenicol extracted from the unreacted acetyl CoA by adding 130µl cold ethyl acetate and vortexing vigorously. The mix was spun in a microfuge for 1 min to separate the two phases and 100µl of the upper, organic phase was removed to a scintillation vial containing 5ml Aquasol scintillant and counted.

2.14.6 β-Gaiactosidase Assay

2µg of the plasmid pJATLAC which contains the β-galactosidase gene under the control of the rat β-actin promoter was included in every transfection. This enables the level of β-galactosidase and hence the efficiency of each transfection to be measured. Equivalent protein amounts of transfected cell extracts were added to 400µl lacZ buffer (60mM $Na_2HPO_4.7H_2O$, 40mM $NaH_2PO_4.H_2O$, 10mM KCl, 10mM$gSO_4.H_2O$), 100µl 4mg/ml ONPG

and the final volume made up to 600µl using lysis buffer where required. The reaction was mixed and incubated at 37°C until it turned pale yellow in colour. The reaction was then stopped by the addition of 250µl 1M Na_2CO_3 and spun at 14Krpm for 5 mins. The supernatent was transferred to a 1ml cuvette and the OD_{420} measured. CAT activities were not corrected for β-galactosidase activity. However, within a comparable set of transfections the β-galactosidase levels were checked to ensure there was no more than a twofold variation.

2.14.7 Protein Concentration Assay

Protein concentrations were determined using the Bio-Rad Protein Assay as described in the manufacturers instructions for the microassay procedure (1-20µg protein). 800µl water and 200µl Dye Reagent Concentrate were placed in a 1ml cuvette (Kartell microcuvettes). To this was added an appropriate volume of cell lysate, typically 2µl of CHO or 5µl NIH 3T3 lysate, or an equivalent volume of lysis buffer (the blank). The solutions were mixed by inverting the cuvette, sealed with Parafilm 'M' (American National Can.), several times. The resultant solution was incubated at room temperature for 5-10 mins and the OD_{595} was measured against the reagent blank. A comparison of the lysate's ODs gives their relative protein concentrations. Furthermore, by determining and plotting the ODs of standard protein solutions against their known concentrations, the absolute concentrations of the lysates could be determined from the standard curve.

2.14.8 Preparation of Whole Cell Extracts

Cells were washed twice in ice cold PBS and then lysed in 500µl immunoprecipitation (IP) lysis buffer (400µl 2.5X RIPA (50mM Tris pH8.0, 250mM NaCl, 1.25% Na deoxycholate, 1.25% NP-40, 1.25% SDS), 100µl glycerol, 20µl 0.5M EGTA, 100µl protease/phosphatase inhibitor mix, 10µl Na_3VO_4, 1µl 0.1mg/ml okadaic acid, 10µl NaF in 1ml final volume). The lysate was sonicated to shear the chromosomal DNA and then clarified by centrifugation (14Krpm, 5 mins, 4°C). The supernatent was removed and stored at -20°C.

2.14.9 Preparation of Nuclear Extracts

Nuclear extracts were prepared at 4°C using ice cold buffers according to the micropreparation tecnique (Andrews and Faller, 1991). All centrifugations of less than 30 secs were carried out in a microfuge at room temperature. Cell monolayers were washed twice in PBS, scraped into 1ml PBS and the cells pelleted by centrifugation for 10 secs. The cells were resuspended in 400µl buffer A (10mM HEPES-KOH pH7.9, 1.5mM $MgCl_2$, 10mM KCl, 0.5mM DTT, 0.2mM PMSF) and allowed to swell on ice for 10 mins. They were then vortexed for 10 secs followed by centrifugation for 10 secs. The supernatent (cytoplasmic extract) was discarded and the pellet resuspended in 20-100µl (2-3 volumes) of buffer C (20mM HEPES-KOH pH7.9, 25% glycerol, 420mM NaCl, 1.5mM $MgCl_2$, 0.2mM EDTA, 0.5mM DTT, 0.2mM PMSF). This was incubated on ice for 20 mins for high-salt extraction

and then centrifuged for 2 mins at 4°C to remove cellular debris. The supernatent (nuclear extract) was transferred to a fresh Eppendorf tube and stored at -70°C. Extract protein concentration was measured as in section 2.14.7 and equal proteins amounts were used either in Western Blotting analysis or EMSA to detect transfected LEX or GAL fusions respectively.

2.14.10 *In vivo* Labelling

CHO cells were transfected by the calcium phosphate method (section 2.14.1) and the precipitate exposed to the cells overnight. 24 hours after washing off the precipitate, the cells were transferred into 10 ml phosphate free MEMα for 1 hour. This was then replaced by 1.5ml phosphate free MEMα containing 10% FCS (dialysed extensively against 10mM HEPES pH7.0, 137mM NaCl, 2.7mM KCl) and 2.0 mCi of ^{32}P-orthophosphate was added. Labelling was allowed to proceed for 3h on a rocker table to ensure all cells were covered with the labelling media. The following steps were all performed on ice. The media was aspirated off and the cells washed twice in 5ml cold PBS. 500µl IP lysis buffer (section 2.14.8) was added and the cell lysate scraped to one end of the dish before being transferred to an Eppendorf tube. The lysate was then passed 10X through a yellow tip (Gilson, France) to shear the chromosomal DNA and centrifuged for 10 mins in a microfuge to pellet out the cell debris. The supernatent was transferred to a fresh Eppendorf tube and immunoprecipitated (section 2.14.11) using the LexA antibody (Neosystem, S. A., France).

2.14.11 Immunoprecipitation

1µl of an appropriate antibody was added to the protein preparation of interest (whole cell extract, section 2.14.8 and 2.14.10; nuclear extract, section 2.14.9; eluted GST protein, section 2.13.5) and the mix tumbled for 1.5 hours at 4°C. 25µl of protein A sepharose suspension was added and the mix tumbled for a further 1.5 hours at 4°C. The beads were then washed 3X 1ml 1X RIPA, boiled in Laemmli buffer, run on SDS-PAGE and detected by autoradiography (labelled proteins) or Western blotting (cold proteins).

CHAPTER 3 (RESULTS):

IDENTIFICATION OF AN ACTIVATION DOMAIN WITHIN ATF2

IDENTIFICATION OF AN ACTIVATION DOMAIN WITHIN ATF2

3.1 INTRODUCTION

The ATF/CREB family of transcription factors regulate transcription in response to a wide variety of different stimuli e.g. cAMP, Ca^{2+} (Liu et al., 1993), adenovirus E1a (Flint and Jones, 1991). The various family members have been extensively studied both *in vivo* and *in vitro* in order to ascertain which members are involved in which regulatory responses. The *in vivo* studies are complicated by the fact that all mammalian cells contain endogenous ATF/CREB proteins. Therefore, fusion proteins with altered DNA binding specificities have been employed to study the regulation of particular family members. In most of the studies, the DNA binding domain of the yeast transcription factor GAL4 was fused to the protein of interest. This GAL4 fusion was then transiently transfected into cells along with a reporter construct containing GAL4 binding sites in its promoter. Since there is no mammalian equivalent of the GAL4 protein, any transcription of the reporter results from the activity of the transfected fusion protein. Using such a system, the transcription factor ATF2 (also known as CRE-BP1 or CREB2) was found to be one member of the ATF/CREB family that is E1a responsive (Flint and Jones, 1991).

The regions of ATF2 which are required for activation by E1a were determined by analysing deletion mutants of ATF2, again as GAL fusions. It was found that the N-terminus of ATF2 was required for an E1a response. Furthermore it was shown that this region, amino acids 19 -112, was sufficient for E1a activation (Flint and Jones, 1991). Since this N-terminal region of ATF2 is so important for the stimulation of transcription by E1a, I decided to analyze the region in greater detail in order to try and understand the role it plays in the activation process and to determine the features such as post-translational modifications involved in the activation.

3.2 RESULTS

3.2.1 Identification of an N-terminal activation domain

The N-terminal region of ATF2 shown to be required for E1a activation comprises of amino acids 19 to 112. Therefore, the first approach I took to studying this region was to construct a series of small in-frame deletions of between 5 and 20 amino acids within it (figure 3.1B). These GAL fusions were transfected into CHO cells either in the presence or absence of an E1a expression vector. The resultant level of activity was determined by co-transfection of a CAT reporter construct containing five GAL4 binding sites upstream of the E4 TATA box and assaying the extracts of transfected cells for the level of CAT expression. Constructs C1, C34, C56 and C78 did not stimulate transcription when transfected alone (figure 3.1A), giving a basal level of transcription similar to that seen with full length ATF2. However, C2 was found to be a very potent transcriptional activator, increasing the basal level of transcription

Figure 3.1 ATF2 contains an N-terminal activation domain

CHO cells were transfected with 10μg of GAL fusion, 2μg of the reporter G5E4CAT and where indicated 5μg of pCE, a 13s E1a expression plasmid.

A. Graph of transcription levels measured from the reporter.

B. Diagram of the GAL fusions used.

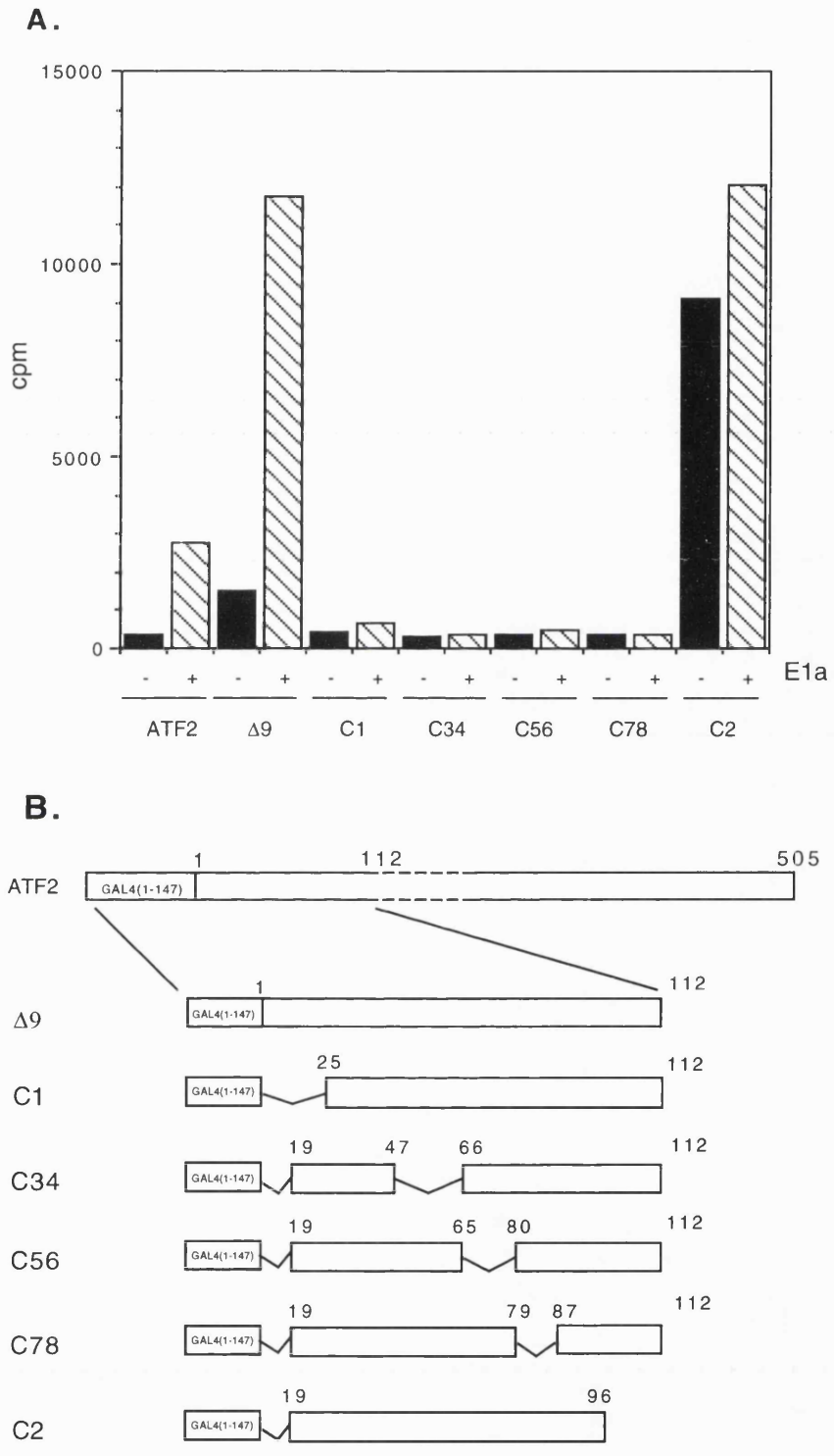


Figure 3.1 ATF2 contains an N-terminal activation domain

about 15 fold above that seen with ATF2. The basal level of transcription of C2 was found to be variable but was consistently higher than that of the full length protein. Therefore, the results suggest that the N-terminal region can function as an activation domain. However, since this high basal transactivation is not observed with ATF2, it suggests that the N-terminal activation domain is masked in some way in the context of the full length protein.

The previously identified E1a responsive domain of ATF2, construct Δ9 (amino acids 1-112), has a basal level of transcription 5 fold higher than that of ATF2 (figure 3.1A) and is therefore intermediate in basal transactivation potential between ATF2 and C2. This can be explained in two ways. Firstly, the activation domain in Δ9 may be only partially masked. Thus a modest increase in basal transcription is seen with this construct but it is not as great as in the fully exposed domain of C2. Alternatively, Δ9 may contain inhibitory sequences which are not present in C2 i.e. amino acids 96-112. In this case the activation domain is proposed to be fully unmasked but is not fully active because of the action of an adjacent repressor region. A combination of the two models is also possible.

It was important to determine that the differences seen in the basal levels of transcription were not due to differences in the expression of the various GAL fusions. Therefore, the protein levels of the GAL fusions were determined by transiently transfecting 293 cells with the appropriate constructs, preparing nuclear extracts from the transfected cells and using the extracts in electrophoretic mobility shift assays (EMSAs).The probe used in the assays contained GAL4 binding sites and therefore GAL fusion proteins present in the nuclear extracts would bind to the probe causing a shift in its mobility through a polyacrylamide gel upon electrophoresis. From figures 3.2 A and B it can be seen that all the fusion proteins are expressed; complex II corresponds to GAL(1-147) bound to the probe whereas complex I which has a decreased mobility represents the larger complex formed by the fusion proteins. Slight variations in complex I mobility are observed with the different cell extracts and this reflects the variation in GAL fusion size. A further complex was often observed in the assay (the band labelled with the asterisk), however, this is a non specific complex as it is present in the reaction using extract from cells transfected with Bluescript.

Although all the fusion proteins are expressed, it can be seen that the level of expression varies greatly; Δ9 is the most highly expressed whereas C2 and C78 are the poorest and are often difficult to detect e.g. C78 cannot be observed in fig. 3.2A nor can C2 in fig. 3.2B. However, the differences in the protein levels do not correlate with the differences in the level of transcription obtained with the constructs e.g. C56 is transcriptionally inactive yet is expressed to a reasonably high level in contrast to C2 which is transcriptionally very active but is very poorly expressed; its protein level is at least 5 times lower than C56 (fig. 3.2A). Therefore, the very high level of transcriptional activation seen with C2 is not merely a reflection of protein over-expression but must be due to the exposure of an activation domain in this protein.

Figure 3.2 Detection of GAL Fusion Proteins by EMSA

Nuclear extracts prepared from transfected 293 cells were incubated with a radiolabelled GAL probe (oligos 39 and 40 annealed). The reaction mixtures were electrophoresed on a 4% acrylamide gel and DNA-protein complexes detected by autoradiography.

A. Labels above the lanes indicate the constructs used in the transfections; GAL = GAL(1-147); Δ9, C1, C2, C34, C56 and C78 correspond to the GAL fusions shown in figure 3.1B; and BS = the Bluescript KS^- plasmid. Equal protein concentrations were used in each reaction.

B. As for A.

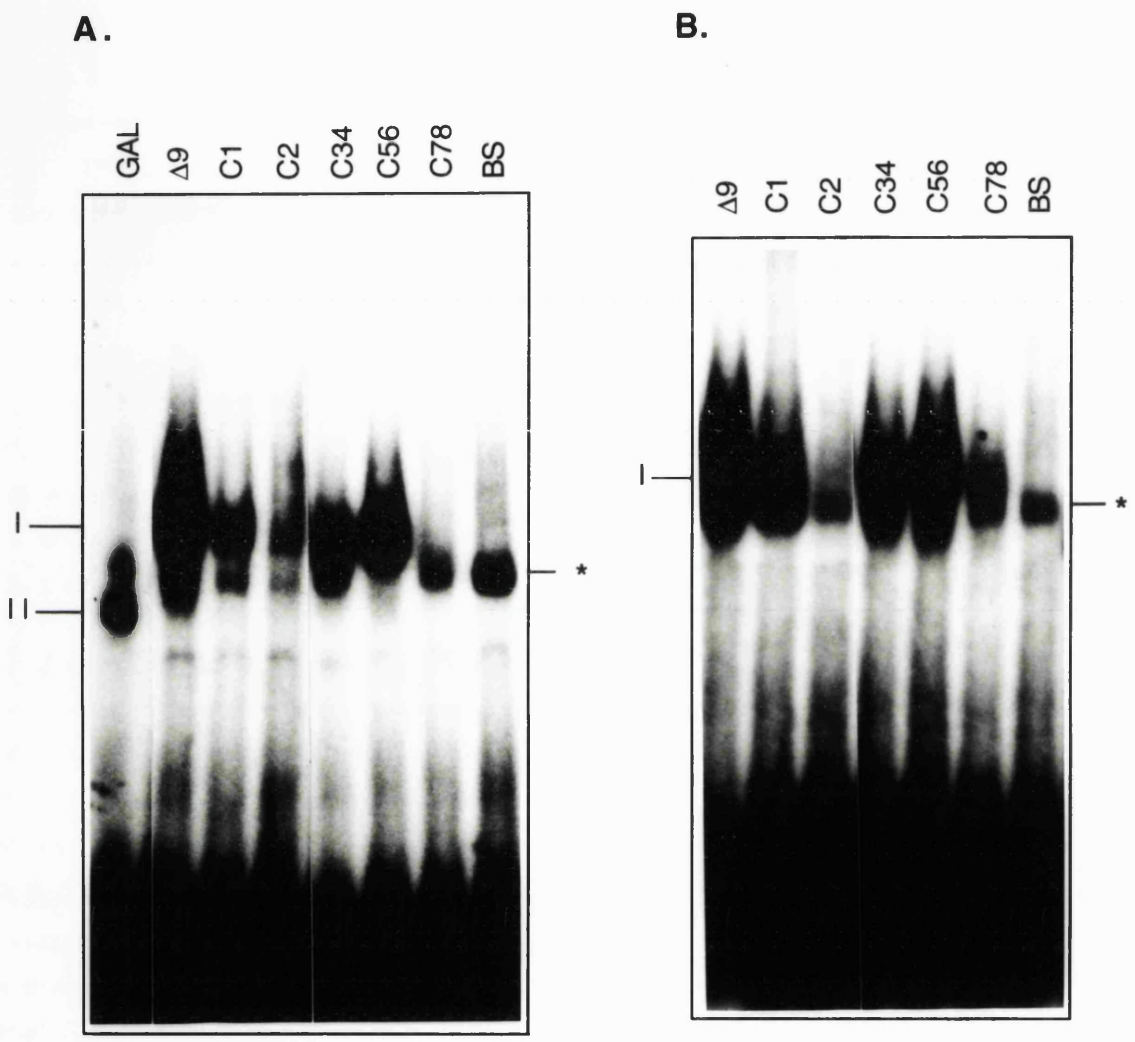


Figure 3.2 Detection of GAL fusion proteins by EMSA

To ensure that complex I did correspond to the GAL fusion proteins bound to probe, the EMSA reactions were repeated but in the presence of NJ2, an antibody specific for the N-terminus of ATF2. This antibody should bind to the ATF2 portion of the fusion and result in a higher molecular weight species bound to the probe. This antibody-fusion-probe complex will therefore have a decreased mobility and produce a "shifted" band on the gel. This is seen in figure 3.3. No such shift is seen with GAL or C78 because both these proteins lack the epitope recognized by NJ2.

The transactivation potential of the N-terminal deletions in the presence of E1a was also determined. C1, C34, C56 and C78 were all E1a unresponsive (figure 3.1A) whereas C2 was slightly enhanced by E1a. This stimulation of C2's activity was variable depending upon the basal level of C2; a high basal level meant only a very small increase in activation by E1a, conversely a lower basal level meant a correspondingly greater increase. However, in general, the fold increase in activation of C2 by E1a was about 5 fold and the actual level of E1a activated transcription was generally found to be of the same magnitude for Δ9 and C2. Since all the other deletions (C1, C34, C56, C78) were unaffected by E1a, it appears that E1a is only able to stimulate transcription through the N-terminal regions of ATF2 that are themselves transcriptionally active.

3.2.2 Strength of ATF2's Activation Domain

In order to assess the strength of the activation domain of ATF2, C2's ability to stimulate transcription was compared to that of other known transcriptional activators. CHO cells were transfected with GAL-activator fusions and transcriptional activation was measured by co-transfecting a GAL-CAT reporter and assaying the transfected cell extracts for CAT activity. The results are presented in figure 3.4. The weak activators Sp1, CTF and Tat (figure 3.4A) do not stimulate transcription greatly above the level seen with GAL(1-147) which itself contains a weak activation domain (only GAL1-94 is transcriptionally inert). Compared to these activators, the activation domain of ATF2 is much stronger; Δ9 activates transcription approximately 4 times more efficiently than GAL(1-147) and C2 approximately 20 times. However, in comparison to the viral activators, VP16 and E1a (figure 3.4B), ATF2's activation domain seems rather weak; VP16 activates transcription 180 fold above GAL(1-147) and E1a 60 fold. This result is not that surprising since these viral activators are known to be extremely potent transcriptional activators. Therefore, it appears that C2 is of an intermediate ability to stimulate transcription in comparison with other activators.

3.2.3 Proposed repressor region in ATF2

Δ9 has a basal level of transcription about 3 fold lower than C2. As explained earlier this may be due to an inhibitory domain (amino acids 96 - 112) in Δ9 which is absent in C2. However, it may also be due to conformational differences between the two proteins and may therefore not represent a specific inhibitory effect. In order to address these possibilities, this

Figure 3.3 Detection of GAL Fusion Proteins by EMSA in the Presence of Antibody

Nuclear extracts prepared from transfected 293 cells were incubated with a radiolabelled GAL probe (oligos 39 and 40 annealed) in the presence of NJ2, an antibody specific for amino acids 85-96 of ATF2. The reaction mixtures were electrophoresed on a 4% acrylamide gel and DNA-protein complexes detected by autoradiography. Labels above the lanes indicate the constructs used in the transfections; GAL = GAL(1-147); Δ9, C1, C2, C34, C56 and C78 correspond to the GAL fusions shown in figure 3.1B; and BS = the Bluescript KS^- plasmid. Equal protein concentrations were used in each reaction.

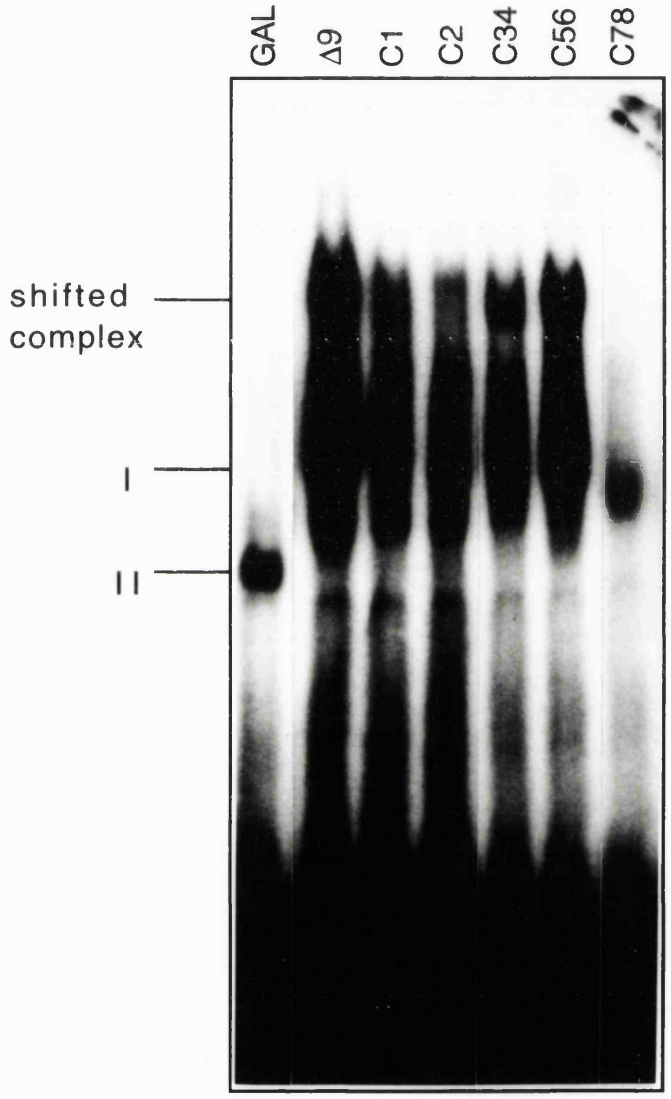


Figure 3.3 Detection of GAL fusion proteins by EMSA in the presence of antibody

Figure 3.4 Strength of the Activation Domain of ATF2

CHO cells were transfected with 5μg of a GAL activator and 2μg of the reporter G5E4CAT.

A and B. Graphs showing the level of transcription stimulated by the GAL fusions indicated.

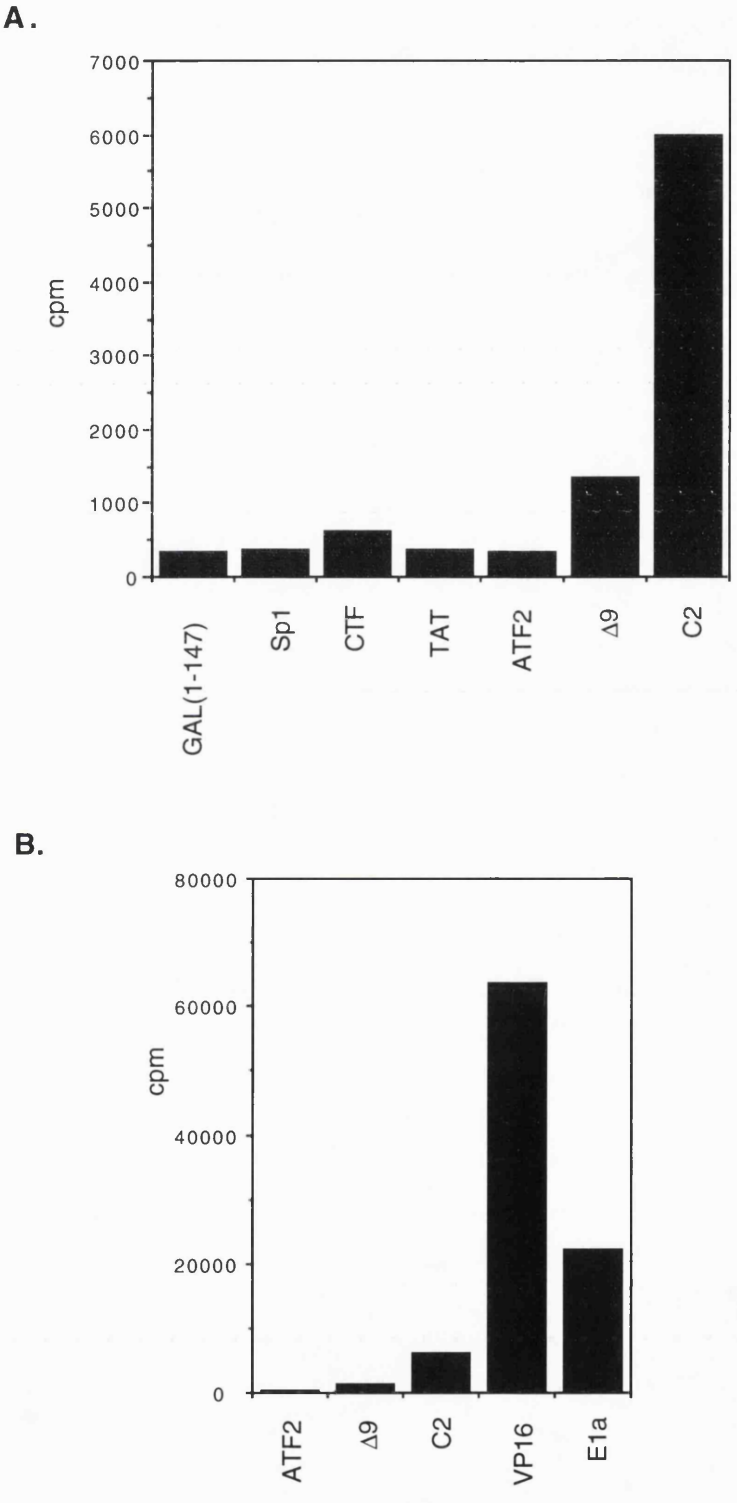


Figure 3.4 Strength of the activation domain of ATF2

putative repressor region was removed in the context of the full length protein so generating the construct ΔREP (fig. 3.5B). When transfected into CHO cells, ΔREP is found to have an increased basal level of transcription, about 2 fold greater than ATF2 (fig. 3.5A). This increase is very small but is a reproducible effect. However, in this experiment the protein ΔREP contains a leucine zipper and is therefore able to dimerize with endogenous ATF/CREB proteins in cells. This means that the level of transcription determined from the reporter construct may not just be a measure of ΔREP homodimers but also a measure of ΔREP/wild type ATF2 heterodimers. Since the wild type ATF2 protein still contains the repressor region, it may actively inhibit the activity of ΔREP if it is able to work in trans. Alternatively, the presence of the wild type ATF2 in the dimer may just limit the effect of ΔREP seen. In either case the full effect of removing the proposed repressor region will not be detected and this may account for the fact that only a very small increase in the basal activity is seen. This experiment could be improved by designing full length mutants of ATF2 which were unable to heterodimerize with wild type ATF2 but were still able to homodimerize with each other. Similar approaches have been developed to study Myc and Max *in vivo* (Amati et al., 1993) and also to study the influence of CREB phosphorylation upon its dimeric partner (Loriaux et al., 1993). The removal of the repressor region could then be made in the context of these full length mutant ATF2 proteins so avoiding the problem of heterodimerization with wild type proteins and subsequent limitation of the assay.

One other possible way to test the repressor hypothesis would be to transfect Δ9 into cells along with C2. Both proteins are GAL fusions and can therefore heterodimerize through the GAL domain. If the repressor region can function in trans it would be expected to reduce the level of transcription determined in the co-transfection experiment as compared to that obtained when C2 is transfected into cells alone. However, C2 would have to be transfected in excess over Δ9 in order to avoid studying the effect of competition of Δ9 homodimers with C2 homodimers for binding to the GAL sites in the reporter construct. It should be noted that this experiment would only reveal whether the repressor region was able to function in trans and in the context of a heterodimer. An alternative method for confirming the existence of a repressor region within ATF2 is possible if this region acts by binding an inhibitory protein (see section 3.3 for other possible mechanisms of repression). In this situation, over-expression of a plasmid which contains the inhibitory region but lacks the activation domain should titrate the inhibitory protein away from a cotransfected wild type construct such as Δ9 or ATF2 and thus relieve repression.

Δ REP was also transfected into cells along with E1a. In this situation the same result was obtained as for basal transcription i.e. a 2 fold increase in the level of transcription upon removal of the inhibitory sequences (fig 3.5A). In the case of Δ9 which is proposed to contain inhibitory sequences the E1a activated levels of transcription were the same as those of C2 which lacks these sequences. Thus E1a seems to be able to overcome the inhibition imposed by the repressor region in the context of these truncated ATF2 proteins. If E1a was also

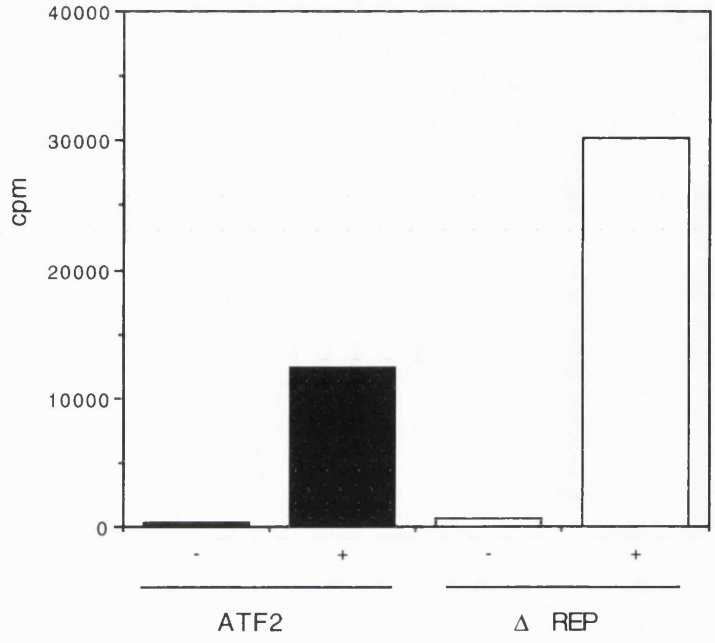
Figure 3.5 Proposed Repressor region in ATF2

CHO cells were transfected with 5μg of GAL fusion, 2μg of the reporter G5E4CAT and where indicated 5μg of pCE.

A. Graph showing transcription levels.

B. Schematic diagram of the GAL fusions used.

A.



B.

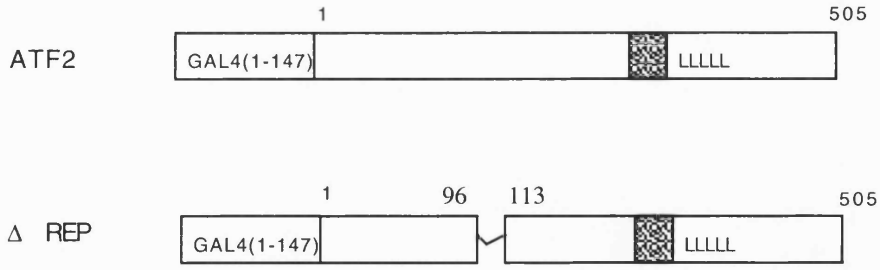


Figure 3.5 Proposed repressor region in ATF2

capable of overcoming repression in full length ATF2 then no difference in the transcription levels of wild type ATF2 and ΔREP in the presence of E1a should be observed. Yet, E1a activation of ΔREP was found to be greater than that of ATF2 which suggests that E1a is unable to overcome the full effect of repression in wild type ATF2. Thus it is probable that in this experiment a more complicated situation than merely the effect of removing a repressor region is being investigated e.g. conformational changes may also be involved.

Since full length ATF2 is transcriptionally inactive it is thought to be folded in such a way that the N-terminal activation region is masked. Upon deletion of amino acids 96-112, the proposed repressor region, the conformation of ATF2 may be altered so partially exposing the activation domain and consequently converting ATF2 into a more active configuration. This conformational change alone may account for the increase in basal transcription seen with ΔREP and may not therefore be the result of deleting a repressor region although it is possible that both factors contribute. Stimulation of this activated ATF2 with E1a may then result in further unmasking of the activation domain and formation of a "super activated" molecule which is only possible because of the altered conformation. It is also feasible that an inhibitory region has been deleted in ΔREP which in combination with the conformational change again gives rise to "super activation" by E1a.

Therefore, it is not clear from this deletion analysis whether a repressor region truly functions in full length ATF2 and the situation needs to be clarified by performing the experiments described above. However, in the case of the truncated versions of ATF2 the presence of inhibitory sequences appears to be more obvious although again whether these sequences function by actively repressing the function of the activation domain or merely play a conformational role is not clear. This will be discussed later (see section 3.3).

3.2.4 Delineation of the activation domain

C2 was constructed to consist of GAL(1-147) fused in-frame to amino acids 19-96 of ATF2. Immediately C-terminal to this, however, there are an extra eight amino acids of non-ATF2 sequence before a termination codon is reached (figure 3.6C). I wanted to confirm that these extra sequences were not contributing to the transactivation potential of C2. Therefore, a further series of plasmids based on C2 were constructed to address this question (figure 3.6B and C). Firstly, the extra non-ATF2 sequence was replaced with the corresponding adjacent eight amino acids of ATF2 i.e. residues 96-104 to give the construct C2-104. When this was transfected into cells it was found to behave in a similar manner to C2; it had a high basal level of transcription (about 15 fold above ATF2) which was increased about 5 fold by E1a (figure 3.6A). The additional sequences were also replaced with eight random amino acids to give construct C2-MIX. This again behaved identically to C2 (figure 3.6A). Therefore, the unrelated eight amino acids at the C-terminus of C2 do not appear to be contributing to its activity in a sequence specific manner. Since this was the case, all further studies on this N-terminal activation domain of ATF2 were carried out in the context of C2. It should also be

Figure 3.6 Transcriptional Activity of C2 Variants

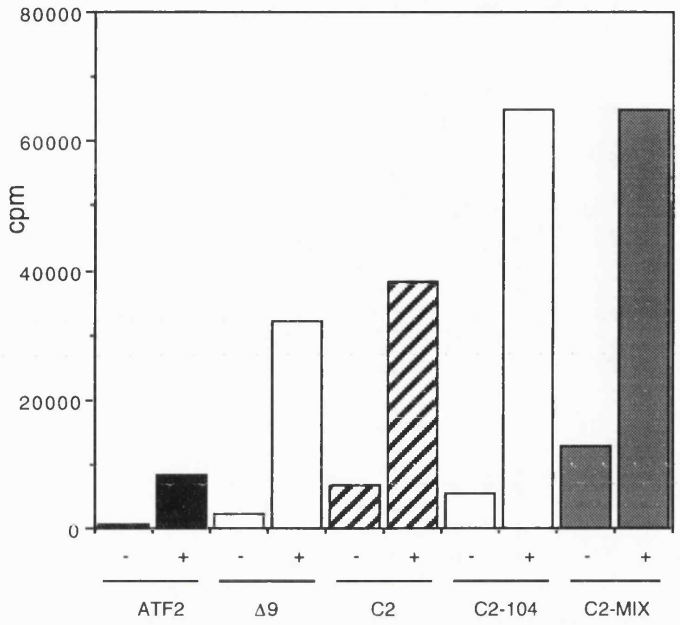
CHO cells were transfected with 5μg of GAL fusion, 2μg of the reporter G5E4CAT and where indicated 5μg of pCE.

A. Graph showing transcription activation measured from the reporter.

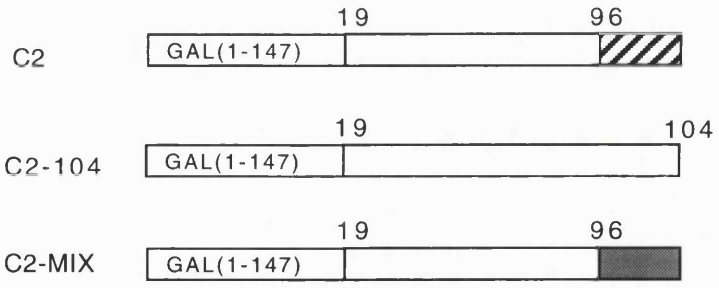
B. Schematic diagram of the GAL fusions used.

C. Amino acid sequence at the 3' end of the fusions. Residues underlined differ from the correct ATF2 sequence.

A.



B.



C.

C2	Glu Phe Ser Arg Phe Ile Leu Met Thr Met STOP
C2-104	Glu Phe Lys Lys Ala Ser Glu Asp Asp Ile STOP
C2-MIX	Glu Phe Val Tyr Arg Glu Ile Ala Gly Leu STOP

Figure 3.6 Transcriptional activity of C2 variants

noted that since these extra sequences did not alter the activity of C2, they must not be part of the proposed repressor region. This, therefore, limits the boundary of the putative repressor region to between amino acids 104 and 112 of ATF2. In order to further clarify this, this smaller region should be deleted in the context of the full length protein to see if there is a modest relief of repression as was shown to be the case for the larger deletion. However, again the affect of an internal deletion upon the conformation of ATF2 will have to be taken into account when interpreting the result obtained.

Variants of C2 were constructed which stopped precisely at either amino acid 98 or 96 of ATF2 with no extra sequences C-terminal to this, constructs C2-98 and -96 respectively (figure 3.7B and C). These were found to behave differently to C2 and were much more like Δ9 (figure 3.7A). They had a basal level of transcription only about 5 fold greater than ATF2 (compare with 15 fold for C2) and were stimulated 15-20 fold by E1a (compare with 5 fold for C2). They are therefore weaker transcriptional activators but can be stimulated up to maximal levels by E1a. Thus, deletion up to amino acid 98 of ATF2 destroys the strong activation domain found in the N-terminus although a reasonable degree of transactivation is still maintained in this smaller region.

This result is surprising given that in these constructs the proposed repressor region (amino acids 104-112) has been deleted and therefore fully active constructs would have been expected. However, removal of the inhibitory sequences may be balanced by the fact that the deletion impinges in some way on the activation domain. It may be that residue 98 is right at the edge of the activation domain and that constructs lacking any extra sequences do not fold into the correct conformation necessary for a fully active domain or that the stability of the domain is altered. However, since I have already been shown that protein expression and transactivation are not correlated, I believe that the extra sequences influence protein folding rather than stability.

3.2.5 Interaction of the activation domain with the transcription machinery

Since C2 is such a strong activation domain it was possible that it functioned by interacting directly with the basic transcription machinery. This has been shown to be the case for other transcriptional activators e.g. E1a, VP16 and Fos which contact the TATA-binding protein (TBP) (Stringer et al., 1990; Lee et al., 1991; Metz et al., 1994) and Myc and VP16 which bind to TFIIB (Lin et al., 1991; Kouzarides, T. pers. comm.). It is thought that such interactions aid in overcoming a rate limiting step in transcriptional initiation by either stabilizing the transcription complex, aiding in its formation or enabling certain components of the complex to assume an active conformation (see introduction section 1.1.5 (iv)).

Therefore, I tested to see whether C2 could interact with either TBP or TFIIB. This was achieved by performing a GST binding assay; GST-TBP or -TFIIB prepared from bacteria were used in a binding reaction with *in vitro* translated GAL-C2. Slight binding of GAL-C2 to GST-TBP was detected (figure 3.8). However, the same degree of interaction was observed

Figure 3.7 Transcriptional Activity of Shortened Forms of C2

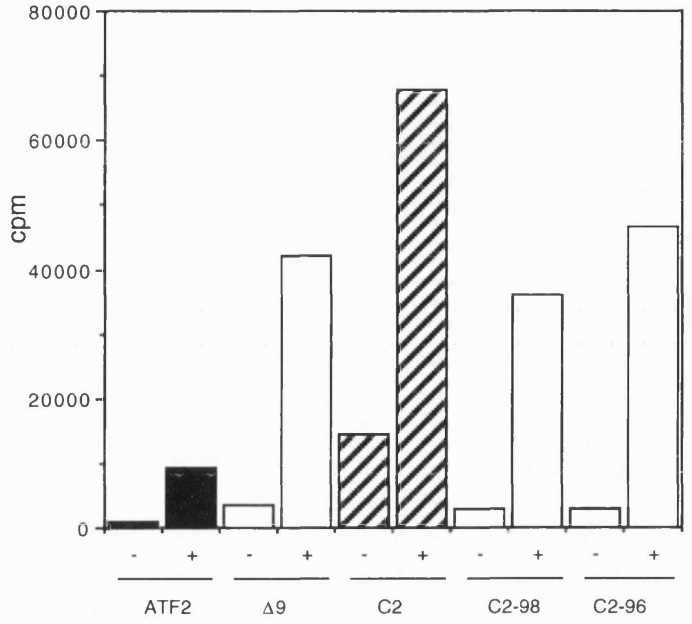
CHO cells were transfected with 5μg of GAL fusion, 2μg of the reporter G5E4CAT and where indicated 5μg of pCE.

A. Graph showing transcription activation measured from the reporter.

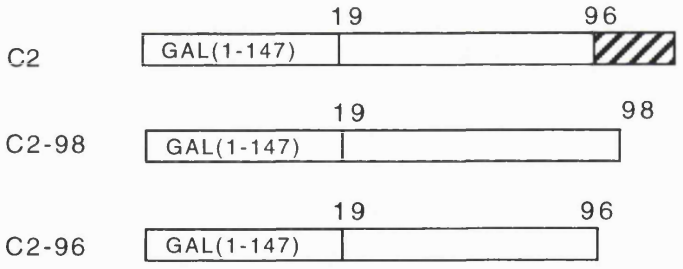
B. Schematic diagram of the GAL fusions used.

C. Amino acid sequence at the 3' end of the fusions. Residues underlined differ from the correct ATF2 sequence.

A.



B.



C.

C2	Glu Phe Ser Arg Phe Ile Leu Met Thr Met STOP
C2-98	Glu Phe Lys Lys STOP
C2-96	Glu Phe STOP

Figure 3.7 Transcriptional activity of shortened forms of C2

Figure 3.8 GST Pull Down Assay to Detect Protein-Protein Interactions Between C2 and TBP/TFIIB

In vitro translated ^{35}S-labelled proteins were incubated with GST proteins bound to glutathione sepharose beads at room temperature. After incubation any unbound protein was recovered, the beads washed three times and boiled in Laemmli buffer to elute any bound protein. 1/4 of the unbound sample and all the bound sample were run on a 12% SDS-PAGE and any proteins present were detected by autoradiography.

M= Molecular weight markers and their sizes are shown in kD at the left of the gel.

Labels directly above the gel indicate the translated protein used in the binding assay.

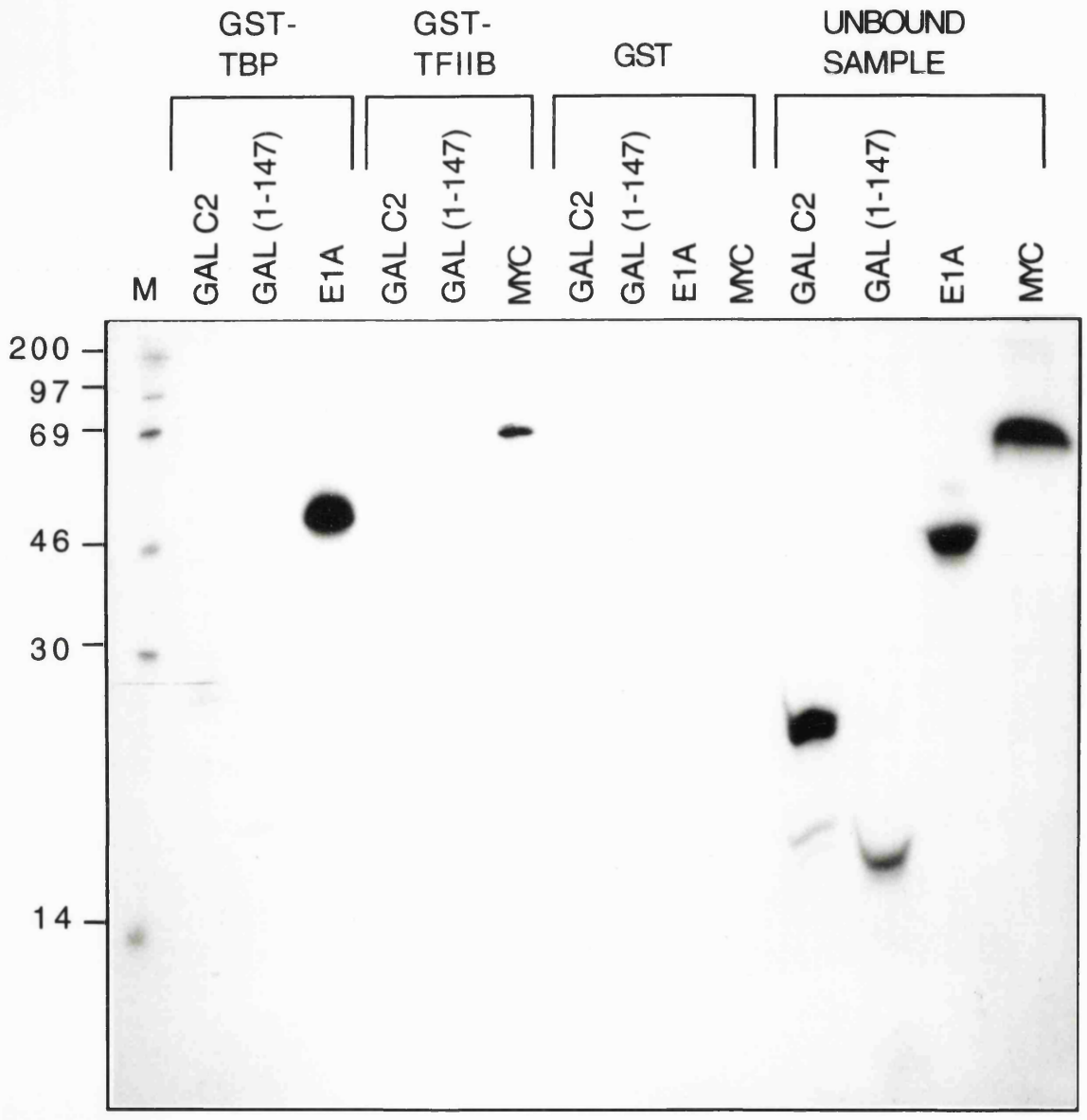


Figure 3.8 GST pull down assay to detect protein-protein interactions between C2 and TBP/TFIIB

with GAL(1-147) thus implying that the interaction seen with GAL-C2 arises via the GAL portion of this fusion. No interaction was observed between either GAL-C2 or GAL(1-147) and GST-TFIIB. Thus, C2 does not appear to interact directly with either TBP or TFIIB. Positive controls for binding were E1a for TBP and Myc for TFIIB and as can be seen (figure 3.8) these bound well with E1a binding the most efficiently. None of the proteins bound to GST alone showing that any interactions detected were mediated through the TBP/TFIIB portion of the GST fusion protein.

Therefore, I have shown by use of an *in vitro* assay that the activation domain of ATF2 does not bind to either TBP or TFIIB. However, this does not necessarily mean that this interaction does not occur *in vivo*. It may be that the activation domain has to be modified in order to be able to interact e.g. it may have to be phosphorylated (see chapter 4 for details on phosphorylation of C2) and such a modification may not occur during the *in vitro* synthesis of GAL-C2. It may be possible to find conditions during translation which allow the modifications to take place e.g. addition of okadaic acid may allow complete phosphorylation of C2. However, such studies have not yet been carried out. It is also feasible that C2 does not contact the transcriptional machinery through either TBP or TFIIB but via another factor (see discussion for details).

3.3 DISCUSSION

By analysing deletion mutants of ATF2 a very strong transcriptional activation domain (C2) has been identified in its amino terminus which is to some extent E1a responsive. The role of C2 with regard to E1a activation of ATF2 will be discussed in a later chapter (see chapter 5). This domain is comprised of amino acids 19-104, with residues 96-104 proposed to be necessary for the protein to fold into the correct conformation and not for the activation function *per se*. All other N-terminal ATF2 deletion mutants were found to be transcriptionally inactive (with the exception of Δ9 which is moderately active) suggesting that the entirety of C2 is required for its function since the deletions cover regions present in C2. The lack of activity of these mutants is not attributable to their lack of expression since all the mutants were detected by EMSA using extracts from transfected cells. However, the protein level of C78 was found to be considerably lower than that of the other mutants and this may therefore account, in part, for its transcriptional inactivity. However, this is balanced by the fact that the level of C2 was also found to extremely low and yet this construct is highly active. Thus transcriptional activity does not necessarily reflect the level of protein expression.

The N-terminal activation domain is not active in the full length protein since ATF2 is found to be transcriptionally inert and it is only exposed in our system by removal of the C-terminus of ATF2. This idea of activation domains being hidden in the context of full length transcription factors is starting to become a more general theme e.g. the transcription factor Myo D has a very low basal activity which is greatly increased upon deletion of its C-terminal sequences (Weintraub, H. unpubl.). This masking is presumably necessary in order to ensure

that transcription factors are not constitutively active and allows another level of regulation to be applied to transcriptional control i.e. the unmasking of the activation domain under suitable conditions. This control mechanism would function in addition to other already established mechanisms such as regulation of transcription factor binding to DNA e.g. cJun's binding to DNA is inhibited by phosphorylation (Lin et al., 1992) whereas ATF2's binding is stimulated by phosphorylation (Abdel-Hafiz et al., 1992), or control by regulation of dimerization e.g. heterodimerization of CREB with CREM converts CREB homodimers from being transciptionally active into inactive complexes (Foulkes et al., 1991).

The lack of basal transcription with ATF2 may not only be a question of a masked activation domain but may also be due to active repression of the domain. Sequences between amino acids 104 and 112 of ATF2 have been identified which partially inhibit the activity of the N-terminus and so may constitute a repressor region in truncated forms of ATF2. However, deletion of such sequences in the context of the full length protein led to inconclusive results as to whether this region did function as a true repressor. Further work is necessary to clarify the situation. However, the presence of inhibitory domains within transcription factors has been observed before e.g. the activation domains in Fos and Jun are regulated by negative acting sequences (Baichwal et al., 1992; Kouzarides, T. pers. comm.)

These inhibitory domains may function by binding a repressor molecule and indeed a cell-specific repressor has been suggested for controlling the activity of Jun (Baichwal and Tjian, 1990). Alternatively, the repressor region may act directly on the activation domain itself in an as yet unknown manner, although one could imagine that the charge or structure of the repressor region could easily influence the nature/activity of neighbouring sequences. The influence of the repressor region could be controlled by post-translational modification. In this regard, it is interesting to note that right next to the proposed repressor region of ATF2 there is a potential PKA phosphorylation site (ser 100). This site was found not to be required for E1a activation of transcription via full length ATF2. However, this does not mean that it is a redundant feature of ATF2. One way to assess its potential importance in negative regulation would be to mutate the serine residue at position 100 in the context of Δ9. Δ9 is moderately transcriptionally active suggesting that the activation domain is unmasked, at least partially. However, it is proposed not to be fully active due to the action of the repressor sequences. Therefore, mutating the PKA site may relieve some of the repression observed with Δ9 if it is involved in controlling the inhibition. However, construct C2-104 also contains the PKA site and yet this construct is fully active. This can be explained because in this case the PKA site is at the very C-terminus of the protein and so may no longer be able to be phosphorylated. Furthermore, even if it was able to be phosphorylated, it may function in conjunction with the neighbouring sequences (amino acids 104-112) which are lacking in this construct.

Finally, it has been shown that although C2 is a potent transcriptional activator, it does not *in vitro* directly contact either TBP or TFIIB, two components of the basal transcriptional machinery. As discussed in the results, this may be due to incomplete post-translational

modifications of C2 in the *in vitro* system used. Alternatively, it may be that C2 interacts with the transcription machinery through proteins in the initiation complex other than TBP or TFIIB.

In this regard, it is important to note that TBP in the cell is found complexed with a large number of accessory proteins, TBP-associated proteins or TAFs (Dynlacht et al., 1991). Together TBP and the TAFs constitute holo-TFIID, the form of TBP which is capable of supporting activated transcription *in vitro* (TBP by itself is only capable of supporting basal transcription *in vitro*) (Zhou et al., 1992). It is therefore possible that C2 interacts with a TAF protein rather than the basic machinery itself. This has been shown to be the situation for other activators e.g. the transcription factors Sp1 and CREB have been shown to interact with Drosophila $TAF_{II}110$ ($dTAF_{II}110$) (Hoey et al., 1993; Ferreri et al., 1994) and the viral activator VP16 to interact with $dTAF_{II}40$ (Goodrich et al., 1993). In the case of Sp1, mutants of Sp1 which were no longer able to activate transcription did not bind to $dTAF_{II}110$ (Gill et al., 1994). Thus these interactions appear to be relevant and important for transcriptional activation and not merely an *in vitro* artefact. Since most of the TAF proteins have now been cloned, it should be possible to perform *in vitro* binding assays to determine which, if any, of the TAFs the transcriptional activation domain of ATF2 interacts with. Alternatively, binding assays with holo-TFIID and the activation domain of ATF2 could be performed. This is possible since a stable cell line expressing an epitope tagged TBP has been established. Immunopreciptation using a monoclonal antibody specific for the epitope allows purification of tagged TBP from the cells complexed with its associated factors (TAFs). Such a complex has been shown to be capable of supporting activated transcription and also of binding E1a (Boyer and Berk, 1993). Thus it appears to be functional and suitable for binding studies with C2.

Another way in which C2 could stimulate transcription, beside acting through a TAF, would be to function through a co-activator i.e. a protein which facilitates transcription by acting as an intermediate between an upstream sequence specific DNA-bound activator and the general transcription machinery (section 1.1.5 (iii)). Such a situation has been reported for CREB which stimulates transcription in response to an increase in intracellular cAMP levels. This response is mediated via PKA which phosphorylates CREB at ser 133 so converting it into an activator. This active phosphorylated form binds, via its kinase inducible (KID) domain, a nuclear protein called CREB-binding protein or CBP (Chrivia et al., 1993). CBP when brought to the promoter is then capable of stimulating transcription in a PKA dependent manner (Kwok et al., 1994). Other co-activators have been identified e.g. the yeast ADA2 protein which is an adaptor for certain acidic activators such as VP16 and GCN4 (Berger et al., 1992). Screens are currently being undertaken in the laboratory in order to try and isolate proteins which interact with the activation domain of ATF2 in the hope of identifying a co-activator molecule(s). These screens may also isolate any TAFs that ATF2 can associate with.

It is possible that ATF2 activates transcription by interacting with both a TAF and a co-activator. This would be in keeping with the situation for CREB which is able to interact with dTAF$_{II}$110 and CBP. These interactions are mediated via different regions of CREB; the glutamine rich region Q2 (amino acids 160-283) contacts dTAF$_{II}$110 whereas it is the KID (amino acids 101-160) which interacts with CBP. A CREB mutant lacking the Q2 region is PKA unresponsive even though it can still bind CBP (Chrivia et al., 1993) which means that in order to obtain a PKA response CREB must interact with dTAF$_{II}$110 as well as CBP. Whether such interactions can occur simultaneously has not yet been addressed but such a model provides an interesting framework in which to study activation by ATF2 especially since E1a can be regarded as a viral co-activator of ATF2.

CHAPTER 4 (RESULTS):

REGULATION OF THE ACTIVATION DOMAIN OF ATF2

REGULATION OF THE ACTIVATION DOMAIN OF ATF2

4.1 INTRODUCTION

Having shown that ATF2 contains an activation domain (C2) in its amino terminus, it was next important to determine what structural features in this region were important for its activity. This would provide valuable information on the unregulated (basal) and E1a regulated activity of the domain as well as possibly providing an explanation as to why the basal activity was variable in different transfection experiments. Therefore, the amino acid sequence of the activation domain was inspected to identify any known sequence motifs, the key residues within such motifs were mutated and transient transfection assays using GAL fusions were performed, as before, in order to determine the effect of such mutations on transactivation. Any essential regions within the domain were further studied to establish/confirm their role in activation.

4.2 RESULTS

4.2.1 Importance of the zinc finger

The N-terminus of ATF2 contains a series of cysteine and histidine residues between amino acids 27 and 49 that can potentially form a zinc finger. This amino terminal region has been shown to bind zinc and mutation of the cysteine at position 32 to a serine severely compromises this ability (Abdel-Hafiz et al., 1993). Furthermore, the cysteine residues at positions 27 and 32 have been shown to be essential for E1a activation of ATF2 (Flint and Jones, 1991). Therefore, it was important to see if this structure was also involved in basal activation by C2. To achieve this, the cysteine residues at positions 27 and 32 were replaced with alanines in the context of C2 to give the construct C2pmZn (figure 4.1B) which was transfected into CHO cells along with a GAL-CAT reporter. The resultant level of basal transcription was found to be reduced down to the level of Δ9 but was not completely abolished (figure 4.1A). Thus, an intact zinc finger is required for maximal activation by C2 but does not seem to be essential to obtain modest levels of transcription with the N-terminal activation domain. This suggests that the zinc finger is not a critical structure for the abiliity of C2 to stimulate the basic transcription machinery.

When C2pmZn was assayed for transcriptional activation in the presence of E1a, it was found to be fairly unresponsive (figure 4.1A). There is only an average 5 fold increase in transcription as compared to a 15-20 fold increase with Δ9 which has a comparable level of basal activity. Thus, the zinc finger appears to be critical for E1a activation whilst it is less important for basal transcription.

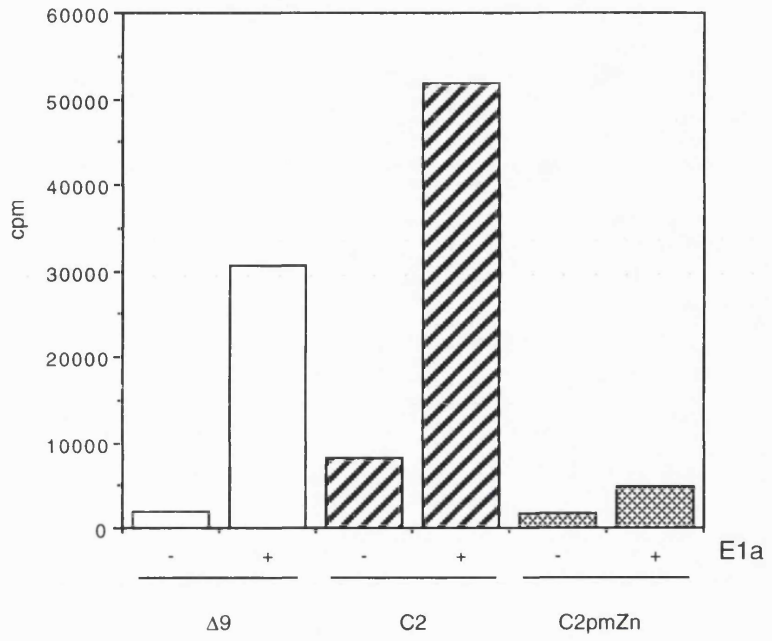
Figure 4.1 An Intact Zinc Finger is Required for C2 Transactivation

CHO cells were transfected with 5μg of GAL fusion, 2μg of the reporter G5E4CAT and where indicated 5μg of pCE.

A. Graph of transcription levels measured from the reporter.

B. Sequence showing the residues mutated in the zinc finger.

A.



B.

Figure 4.1 An intact zinc finger is required for C2 transcativation

4.2.2 Potential MAP kinase sites and activation

Another discernible feature in C2 are three potential mitogen-activated protein (MAP) kinase sites. These phosphorylation sites have the consensus sequence T/S P (Davis, 1993) and such sequences are found at residues 69, 71 and 90 of ATF2. The reason for investigating these sites was twofold. Firstly, as already mentioned, the basal level of transactivation seen with C2 was found to be variable. Although variability was detected in transcriptional activation with all the constructs used, its range was relatively small for ATF2 (over 11 experiments range = 229-356 cpm, average = 288, standard deviation = 36) and Δ9 (range over 11 experiments = 875-2,185 cpm, average = 1,412, standard deviation = 418) but was considerably larger for C2 (range over 11 experiments = 1,895-10,269 cpm, average = 4,775, standard deviation = 2,309). This was not due to differences in transfection efficiencies because the β-galactosidase assay used to measure this parameter gave relatively constant OD_{420} absorbances (generally only a two fold variation between samples). Therefore, I wondered if the variation in C2's activity somehow reflected the growth state of the cells during the transfection. This is reasonable if C2 is regulated by phosphorylation at the identified MAPK sites since MAPKs are themselves activated by growth factor signalling pathways. Secondly, full length ATF2 has already been shown to be phosphorylated *in vitro* by MAP kinases although the sites of phosphorylation have not been mapped (Abdel-Hafiz et al., 1992). Therefore, to ascertain if the N-terminal phosphorylation sites were involved, the serine or threonine residues at positions 69, 71 and 90 were mutated to alanines. These mutations were carried out either singly or in combination with each other (figure 4.2B). The resultant GAL constructs were assayed by transfection into CHO cells along with a GAL-CAT reporter.

When all three potential MAP kinase sites were mutated, C2-ALL, the basal level of transcription was drastically reduced so that it was only very slightly higher than the level of ATF2 and only 5-8% the level of C2 (figure 4.2A). This construct is therefore essentially silent with regard to transcription. Similarly, mutating the threonine residues either singly (C2-T1, C2-T2), both together (C2-T1T2) or in combination with the serine residue (C2-T1S or C2-T2S) leads to a drastic reduction in basal transcription with these constructs being only 5-14% active as compared to C2 and are therefore fairly inert domains. However, mutation of the serine residue on its own, C2-S, results in a fusion which is 30-45% as active as C2 and comparable to Δ9. Therefore, the serine residue within the potential SP phosphorylation motif does not appear to be critical for basal transcription itself. However, it does appear to be required to obtain maximal levels of transcription, as was the case for the zinc finger.

Thus from these results it appears that it is the threonine residues present in the TPTP motif which are essential for the high levels of transcription seen with C2. Since these residues are potential targets for phosphorylation by the MAPKs, it is presumably this post-translational modification that potentiates C2's activity. If this is so (the remainder of the

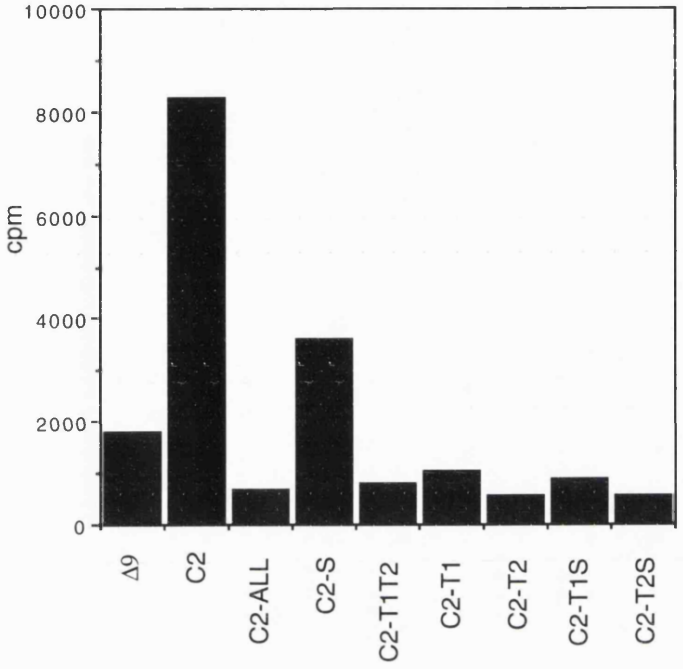
Figure 4.2 Basal Transactivation Levels of C2 Phosphorylation Mutants

CHO cells were transfected with 5μg GAL fusion and 2μg of the reporter G5E4CAT.

A. Graph of transcription levels measured from the reporter.

B. Diagram of the C2 phosphorylation mutants used in the transfections. Mutated residues are highlighted.

A.



B.

	69		71			90	
C2	---- T	P	T	P	--- S	P	----
C2-ALL	---- A	P	A	P	--- A	P	----
C2-S	---- T	P	T	P	--- A	P	----
C2-T1T2	---- A	P	A	P	--- S	P	----
C2-T1	---- A	P	T	P	--- S	P	----
C2-T2	---- T	P	A	P	--- S	P	----
C2-T1S	---- A	P	T	P	--- A	P	----
C2-T2S	---- T	P	A	P	--- A	P	----

Figure 4.2 Basal transactivation levels of C2 phosphorylation mutants

chapter describes experiments which address this issue), it should be determined whether it is solely the negative charge of the phosphate group that is important for the high activation potential of phosphorylated C2 or whether the phosphate group itself plays a specific role in the transcriptional activation process (this is discussed in section 4.3). This can be tested by replacing the threonine residues with negatively charged amino acids such as aspartic or glutamic acid and assessing whether these mutants are as active as the wild type version. However, such an experiment has not yet been performed.

The effect of mutating the potential phosphorylation sites on E1a activation was also determined by transfecting the mutants into CHO cells along with an E1a expressing plasmid and a reporter construct. The resultant levels of activation are shown in figure 4.3A. C2-ALL was found to be E1a unresponsive as was the double threonine mutant C2-T1T2. Constructs containing a single mutated threonine residue (C2-T1, T2, T1S and T2S) varied in their response but generally their induction was about 5 fold. This is a weak response when compared to Δ9 or ATF2 which are stimulated 15 fold or more. Thus, the proposed phosphorylation of the threonine residues appears to be important for both basal transcription and E1a activation. Mutating the serine residue, C2-S, did not prevent E1a activation since transcriptional levels with this mutant in the presence of E1a were almost as high as those obtained with Δ9 or C2. Thus, the SP phosphorylation motif does not appear to be critical for the function of the activation domain and the decrease in basal activation caused by mutating this site was able to be overcome by E1a. However, E1a is not able to overcome the effect of mutating the threonine residues. Since these mutated proteins are virtually transcriptionally inactive, this reinforces the idea that E1a activation only occurs through ATF2 derivatives that are transcriptionally active. This result has been seen previously with the series of deletions in the N-terminus of ATF2 (section 3.2.1). The only exception to this rule is the mutation of the zinc finger (section 4.2.1) which has a comparable basal level of transcription to Δ9 and C2-S but unlike these proteins has low E1a inducibility. This further highlights the importance of the zinc finger to E1a activation.

4.2.3 Regulation of the activation domain by serum and UV light

In the previous section the importance of three potential MAP kinase sites for transcriptional activation by C2 was established; the TPTP motif at residues 69-72 was identified as the major contributor to the activation potential. Since levels of phosphorylation are controlled by the action of kinases and phosphatases within a cell, it was interesting to see under conditions of enhanced kinase activity whether the activation potential of C2 could be increased. One way to activate a subfamily of the MAP kinases, the Erks, is to serum stimulate quiescent cells (Marais et al., 1993). This leads to activation of cell surface growth factor receptors which in turn activate the MAP kinase pathway via intracellular signalling molecules such as Ras and Raf. Another pathway which also stimulates proline-directed kinases is that activated in response to ultraviolet (UV) light irradiation of cells and again it

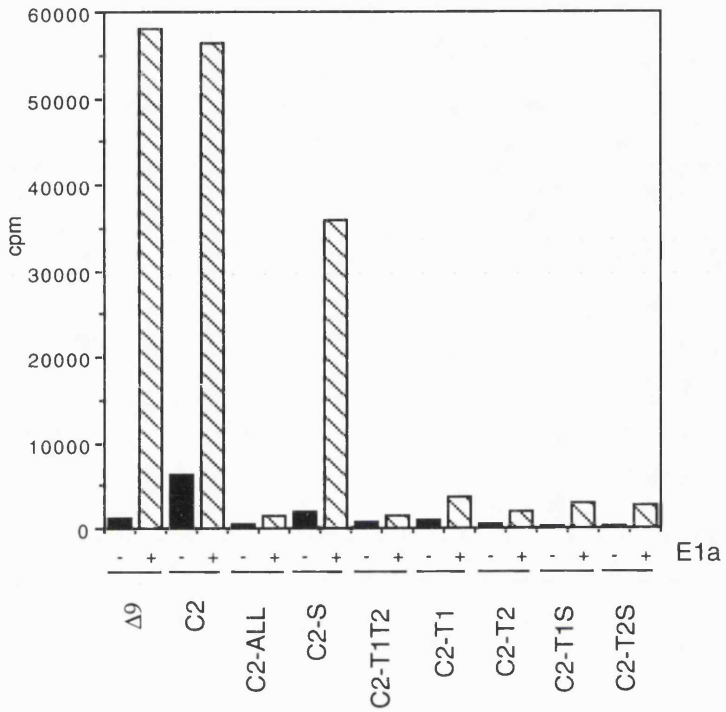
Figure 4.3 E1a Activation of C2 Phosphorylation Mutants

CHO cells were transfected with 5μg GAL fusion, 2μg of the reporter G5E4CAT and where indicated 5μg of pCE.

A. Graph showing transcription levels measured from the reporter.

B. Phosphorylation mutants used in the transfections (as figure 4.2B)

A.



B.

	69	71	90
C2	---- T P	T P ---	S P ----
C2-ALL	---- A P	A P ---	A P ----
C2-S	---- T P	T P ---	A P ----
C2-T1T2	---- A P	A P ---	S P ----
C2-T1	---- A P	T P ---	S P ----
C2-T2	---- T P	A P ---	S P ----
C2-T1S	---- A P	T P ---	A P ----
C2-T2S	---- T P	A P ---	A P ----

Figure 4.3 E1a activation of C2 phosphorylation mutants

signals via Ras. This signalling pathway induces kinases which are 40-45% identical to the Erks. However, these UV inducible kinases are stimulated by cellular stress rather than mitogens or phorbol esters and so have been named stress-activated protein kinases (SAPKs) by certain groups (Derijard et al., 1994; Kyriakis et al., 1994). SAPKs can also be stimulated by Ras independent pathways (Minden et al., 1994a).

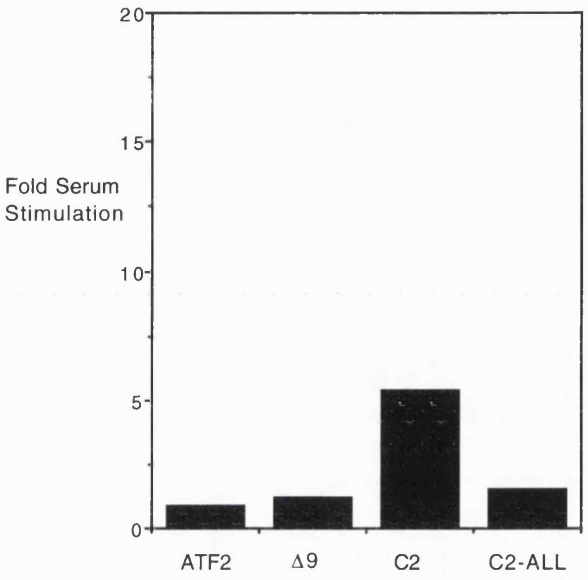
To determine the effect of serum on the activity of C2, NIH 3T3 cells were transfected with GAL fusions along with a GAL-CAT reporter construct. The cells were then maintained under serum starvation conditions before being stimulated with serum as described in materials and methods. Transfections in which the cells were continually maintained under starvation conditions were also performed. A positive control for the induction, GAL-Elk, was included in the experiment. This construct contains the C-terminus of Elk (amino acids 307-428) fused to the DNA binding domain of GAL (amino acids 1-147) and it has previously been shown to be serum responsive via phosphorylation of two key SP motifs at positions ser-383 and ser-389 (Marais et al., 1993). As can be seen in figure 4.4, the activity of this construct is increased 49 fold upon serum stimulation. In the same experiment, the activity of GAL-C2 was also increased by serum athough not by such a large extent, 5.4 fold. This increase was mediated through the three phosphorylation sites previously identified since if these sites are mutated, C2-ALL, the serum response is significantly reduced (1.5 fold).

Transcription by full length ATF2 was found not to be activated by serum (fold stimulation of 0.9). This could be for two reasons. Firstly, the phosphorylation sites in the activation domain may not be accessible to the kinases which would be consistent with the proposal that the activation domain in ATF2 was masked in some way. Alternatively, these sites may be phosphorylated but since the domain is in its masked conformation and hence not active, the effect of the phosphorylations would not be detected. If this latter proposal is true it means that phosphorylation within the N-terminal activation domain is not sufficient to cause unfolding and exposure of the domain. Δ9 was also found to be unresponsive to serum (fold stimulation 1.2). This construct is thought to be partially unmasked and yet its activity is not increased by stimulation of the MAP kinase pathway. It may be that in this construct the phosphorylation sites are still inaccessible to the kinases or alternatively the proposed repressor region may negate the effect of any phosphorylations which occur.

The effect of UV light on C2's activity was investigated in the same way as serum stimulation; NIH 3T3 cells were transfected with LEX-fusions, a LEX-CAT reporter (LEX-OP2) and maintained under serum starvation conditions before being irradiated with short wave UV light i.e. UVC. As with GAL, there is no mammalian equivalent of the bacterial LEX A repressor protein so the reporter only measures transcriptional activation arising from the LEX-fusion. The positive control for the experiment was LEX-cJun which consists of the DNA binding domain (amino acids 3-202) of LEX A fused to the amino terminus (amino acids 1-194) of cJun. cJun has been shown to be transcriptionally activated in response to UV. This is due to phosphorylation of two serine residues (ser-63 and ser-73) in its activation domain

Figure 4.4 C2 Transactivation is Serum Inducible

NIH 3T3 cells were transfected with 5µg GAL fusion and 2µg of the reporter G5E4CAT. 40 hours post-transfection the cells were either left untreated or serum stimulated. The results are presented as fold induction of transcription i.e. cpm serum stimulated/cpm untreated as determined from the CAT reporter from harvested cell extracts.



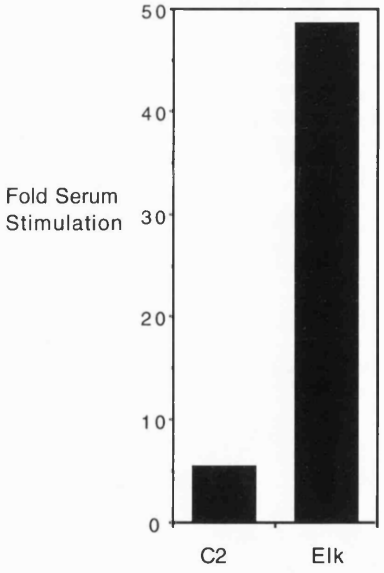


Figure 4.4 C2 transactivation is serum inducible

which potentiates its activity (Hibi et al., 1993). As shown in figure 4.5, UV irradiation was indeed found to stimulate the activity of LEX-cJun, the fold increase in activity being 17.1. The activity of LEX-C2 was also increased upon UV irradiation with the fold increase in this case being 5.9. As before, this stimulation of activity was dependent on the three proposed phosphorylation sites in C2. This is evidenced by the lack of UV induciblity (1.2 fold increase in activity) with the construct C2-ALL .

Thus the results demonstrate that both serum and UV light increase the activation potential of the amino terminus of ATF2 and that this increase is dependent on the presence of the three T/S P motifs in ATF2's activation domain. This suggests that the increase in activity is due to phosphorylation of these motifs by kinases activated by these extracellular stimuli, namely members of the MAPK family. The precise nature of the kinases involved in ATF2 phosphorylation is discussed in section 4.3.

4.2.4 Further definition of the sites involved in regulation of C2's activity

In the previous sections, the three potential MAP kinase sites in ATF2's activation domain have been shown to be important for basal transcription, E1a activation and also the regulation of its activity. The three sites differ in the extent to which they are involved in the first two processes, therefore, it was possible that there was also a difference in their involvement in regulation. To test this, serum stimulation experiments were performed as before but using the set of point mutants which disrupt the phosphorylation sites at positions 69, 71 and 90 either singly or in combination (figure 4.2B). When these were transfected into NIH 3T3 cells along with a GAL-CAT reporter the levels of transcription after serum stimulation were determined. The level of stimulation obtained with the mutants was variable (figure 4.6) but in general it was found that mutation of the threonine residues was detrimental; the fold stimulation was reduced from about 6 fold with wild type C2 to 3 fold or less with any of the threonine mutants (C2-T1T2, T1, T2, T1S, T2S). In contrast mutation of the serine residue had no effect upon the serum induction with transcriptional activation by C2-S being stimulated to levels similar to those seen with the wild type. This parallels what has been observed before namely that the SP phosphorylation is not critical for the function of the activation domain. The same results were obtained with UV induction i.e. only the threonine residues were important for stimluation of the activity of C2 by UV light (van Dam et al., 1995). Thus the two stimuli (serum and UV light) act on the same phosphorylation sites in C2 to increase its activity.

I also decided to investigate whether the zinc finger was required for stimulation of C2's activity by serum. The zinc finger cannot directly contribute to the phosphorylation level, and hence regulation, of the activation domain since it is not able to be phosphorylated itself. However, it can be envisioned to play an indirect role e.g. it may be necessary to maintain the correct conformation required for the interaction of kinases with the domain and so allow subsequent phosphorylation of the region. However, as shown in figure 4.6 mutating the zinc

Figure 4.5 C2 Transactivation is UV Inducible

NIH 3T3 cells were transfected with 5μg LEX fusion and 2μg of the reporter LEXOP2. 40 hours post-transfection the cells were either left untreated or UV-C irradiated. The results are presented as fold induction of transcription i.e. cpm UV treated/cpm untreated as determined from the CAT reporter from harvested cell extracts.

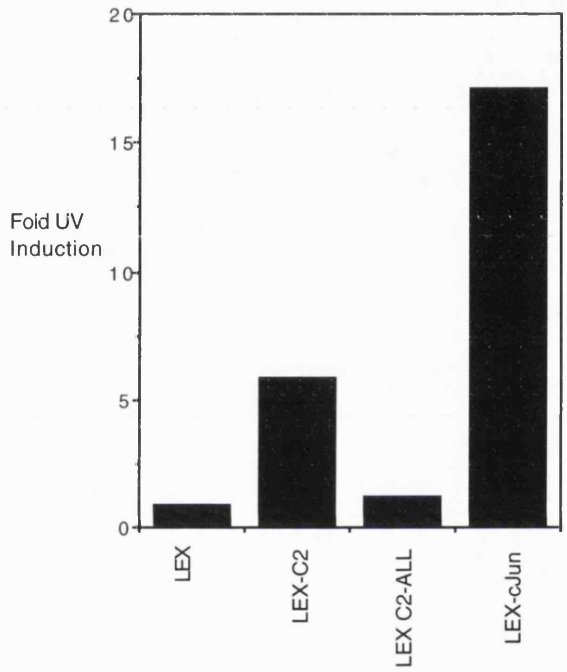


Figure 4.5 C2 transactivation is UV inducible

Figure 4.6 Effect of Mutating the Potentiai Phosphorylation Sites Upon Serum Induction of C2.

NIH 3T3 cells were transfected with 5μg of GAL fusion and 2μg of the reporter G5E4CAT. The mutants used as are those described in figure 4.2B. The graph shows transcription levels determined from the reporter in the presence and absence of serum stimulation as indicated.

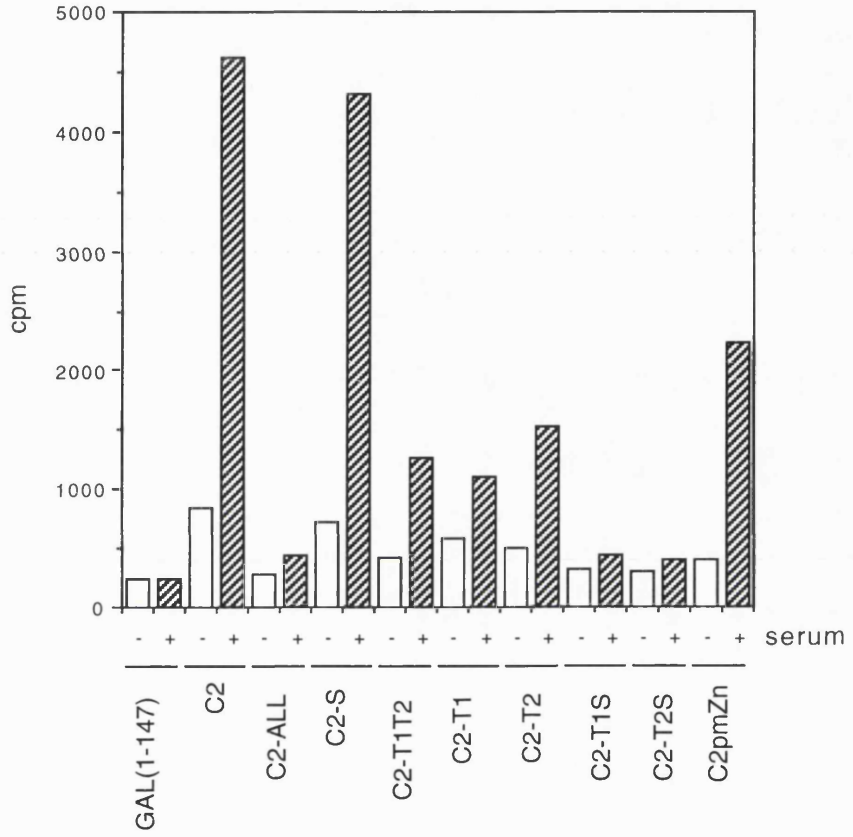


Figure 4.6 Effect of mutating the potential phosphorylation sites upon serum induction of C2

finger has no effect on stimulation of transcription by serum (fold serum stimulation approximately 6 fold i.e. the same as for wild type C2). Thus, it seems that an intact zinc finger is not required to obtain phosphorylation of key residues in the activation domain in response to serum. It has not yet been tested whether the zinc finger is required for the reponse to UV light. This will be discussed later, section 6.3.

4.2.5 *In vitro* phosphorylation of C2

Having identified three important phosphorylation sites in the activation domain of ATF2, I decided to see if they could be phosphorylated *in vitro*. As evidenced above, the kinase responsible for phosphorylating these sites is probably one or more of the MAPK family. Therefore, the enzyme chosen for the *in vitro* phosphorylation of C2 was a preparation of Erk1 (p44) and Erk2 (p42) (Marais et al., 1993). The experiment was performed in the following way; C2 was synthesized in bacteria as a GST fusion, purified on glutathione sepharose beads and then phosphorylated in the presence of γ^{32}P-ATP using Erk1/2 partially purified from U937 cells treated with TPA (kindly provided by M. A. Price). This preparation is enriched for Erk1/2 although it is possible that other kinases e.g. the SAPKs are present but at considerably lower quantities. After phosphorylation the proteins were run on an SDS-PAGE and any labelled proteins visualized by autoradiography.

GST-GAL C2 was first used as the substrate for phosphorylation since it was the GAL version of C2 that had been used for the *in vivo* studies. This protein was able to be phosphorylated by Erk1/2 *in vitro* (figure 4.7) and this was not due to either phosphorylation of the GST moiety since this on its own was not labelled at all or to phosphorylation of the GAL portion since this was only labelled to a very low level. GST-GAL C2ALL was also used as a substrate and the level of labelling in this construct was found to be much lower than that obtained with GST-GAL C2 even though equivalent amounts of protein were used in the experiments. Thus the phosphorylation of C2 seen *in vitro* must occur at one or more of the three potential MAP kinase sites since these are the sites mutated in C2-ALL. However, there is still a certain level of phosphorylation with GST-GAL C2ALL which is greater than that of GST-GAL and this is discussed below.

GST Elk was also phosphorylated being the positive control for phosphorylation by Erk1/2 (Marais et al., 1993). It was found to be phosphorylated very efficiently and, even taking into account the fact that Elk contains more potential phosphorylation sites than C2 (9 sites in Elk versus 3 in C2), it seems to be labelled to a greater degree than C2 (approximately 20 fold greater incorporation of ^{32}P). This reflects the fact that *in vivo* the activity of GAL-Elk is stimulated to a much greater extent by serum (48.6 fold) than is the activity of GAL-C2 (5.4 fold), see figure 4.4. These results together with the fact that C2 is not phosphorylated very efficiently considering the amount of protein used in the phosphorylation experiments (ng-μg amounts have been used) suggest that Erk1/2 may not be the most suitable kinase for phosphorylating these sites. This is in agreement with a previous report

Figure 4.7 ***In Vitro*** **Phosphorylation of GST-GAL C2 By Erk1/2**

GST proteins were synthesized and phosphorylated by Erk1/2 as described in the materials and methods using equal amounts of protein in each reaction. After phosphorylation, the whole reaction was boiled in Laemmli buffer and run on a 10% SDS-PAGE. Phosphorylated proteins were detected by autoradiography. Substrates phosphorylated are indicated above the gel.

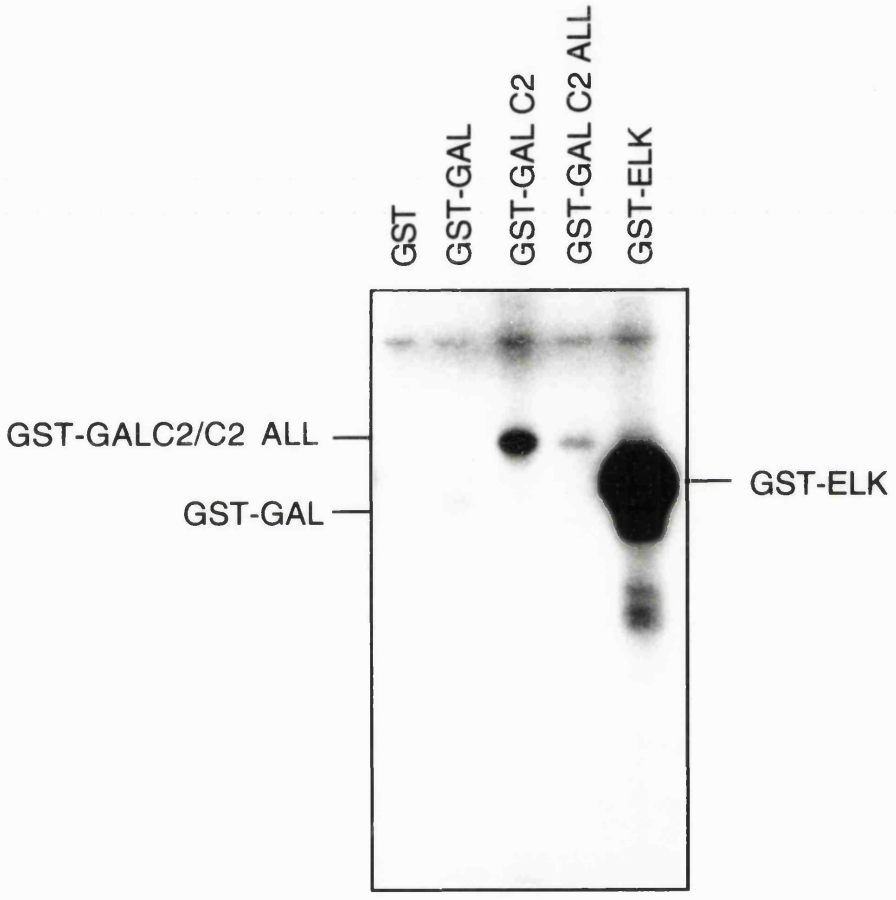


Figure 4.7 *In vitro* phosphorylation of GST-GAL C2 by Erk1

which has shown that *in vitro* SAPK (p54) rather than Erk1/2 is the most appropriate kinase for efficiently phosphorylating ATF2 (Abdel-Hafiz et al., 1992).

GST C2 fusions lacking the GAL moiety were also constructed and the same *in vitro* phosphorylation experiments performed on the purified proteins. GST C2 is again found to be phosphorylated in such an experiment (figure 4.8) although the level of phosphorylation is still very low considering the amount of protein used. The point mutations in the individual phosphorylation sites were introduced into this GST context and the effect on the level of phosphorylation determined to see if all three of these sites were being targeted. The amount of protein used in each phosphorylation reaction was the same, thus, any reduction seen in the degree of labelling was due to a lack of phosphorylation. In each of the mutants C2-T1, T2 and S the level of labelling was decreased compared to wild type C2. Therefore, each of the sites appears to be phosphorylated *in vitro*. The decrease in labelling was of a similar magnitude in each of the mutants which suggests that there is no preference for a particular site in this assay. When all three sites were mutated simultaneously (GST-C2ALL) the level of phosphorylation was decreased even further. However, there is still a small amount of phosphate incorporation into this protein as was found with GST-GAL C2ALL. This is feasible since in these proteins only the potential MAP kinase sites have been mutated which still leaves other phosphorylatable residues untouched (there are several other serine and threonine residues present in this N-terminal region although none are followed by a proline). Given that the MAP kinase preparation is not pure and so may contain other kinases it is possible that these other residues may therefore be phosphorylated. However, this phosphorylation only occurs at a very low level and is relatively insignificant in comparison to the MAP kinase phosphorylations.

Phosphorylation of GST-C2 by Erk1/2 *in vitro* resulted in a form of the fusion which had a decreased mobility as evidenced by the upperband in figure 4.8. This decrease in mobility is thought to reflect an alteration in the conformation of the fusion protein brought about by phosphorylation and has been detected in other proteins e.g. Elk (Marais et al., 1993). Since this decrease was not seen with GST-C2T2, it must arise from phosphorylation of the threonine at position 71. This is consistent with the observation that GST-C2 ALL which also contains a mutated thr-71 does not change in mobility upon phosphorylation. However, the decrease in mobility was not detected with any of the GST-GAL fusions described above even those containing an intact thr-71 (see figure 4.7). This may reflect the way in which these fusions are folded which prevents a conformational change from occurring. Alternatively, it is possible that the larger size of the GST-GAL fusions may have made the decrease in mobility more difficult to detect. Finally, the conditions used in the original phosphorylation experiments may not have been optimal since the level of labelling was much lower than that obtained with the later GST experiments and thus the mobility shift may not have been detected.

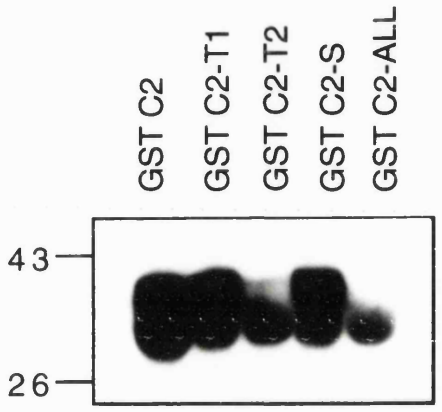
Figure 4.8 ***In Vitro*** **Phosphorylation of GST-C2 By Erk1/2**

The same procedure was followed as described in figure 4.7 but using GST fusions lacking a GAL domain.

A. Autordiograph of the phosphosphorylated proteins. Labels above the gel indicate the GST substrate used. Numbers to the left are the sizes in kD of the molecular weight markers.

B. Phosphorylation mutants used. Mutated residues are highlighted.

A.



B.

	69		71			90		
C2	---- T	P	T	P	---	S	P	----
C2-ALL	---- A	P	A	P	---	A	P	----
C2-S	---- T	P	T	P	---	A	P	----
C2-T1	---- A	P	T	P	---	S	P	----
C2-T2	---- T	P	A	P	---	S	P	----

Figure 4.8 *In vitro* phosphorylation of GST-C2 by Erk1

4.2.6 *In vivo* phosphorylation of C2

The activation domain of ATF2 has been shown in the previous section to be phosphorylated by Erk1/2 *in vitro*. However, as this is an artificial situation, it was important to show that this region is also phosphorylated *in vivo*. This was achieved by isolating proteins from cells labelled with ^{32}P orthophosphate. To do this LEX-C2 and LEX-C2ALL were transiently transfected into CHO cells. 40 hours after transfection, the medium was changed to phosphate free medium for an hour, the cells UV irradiated and then labelled by addition of ^{32}P orthophosphate to the medium. LEX proteins synthesized after transfection were immunoprecipitated from the cell extracts using a LEXA antibody, run on an SDS-PAGE and visualized by autoradiography. LEX fusions were used in this experiment for two reasons. Firstly, LEX A antisera which is efficient at immunoprecipititation is commercially available whereas there is no equivalent antisera available for GAL. Secondly, labelling experiments performed to investigate Elk phosphorylation *in vivo* have demonstrated that the LEX portion of LEX fusions is not itself phosphorylated (Marais et al., 1993). Therefore any phosphorylation detected in labelling experiments using such fusion proteins must be due to phosphorylation of the fused protein. The same is not true for GAL fusions since in the *in vitro* studies described above the GAL moiety was found to be phosphorylated albeit to a very low level. Thus LEX not GAL fusions were used in these labelling studies.

The proteins recovered by immunoprecipitation and visualized by autoradiography are shown in figure 4.9A. A non-specific phosphoprotein was immunoprecipitated by the LEX antibody in all samples, including the LEX control transfection, as indicated by the asterisk labelled band. However, a specific labelled band was also immunoprecipitated from the LEX-C2 transfection but not the LEX-C2ALL transfection. This presumably corresponds to phosphorylated LEX-C2. This means that LEX-C2 must be phosphorylated *in vivo* on one or more of the potential MAP kinase sites identified at positions 69, 71 and 90 since these are the sites mutated in C2-ALL. However, it cannot be determined exactly which of the sites are phosphorylated from this experiment. To address this problem, the labelling experiment needs to be repeated but using LEX-C2 constructs in which the potential phosphorylation sites have been mutated individually, not all together. Then the contribution of the three sites to *in vivo* phosphorylation can be assessed.

The protein levels of LEX-C2 and LEX-C2ALL were checked in order to ensure that the differences in phosphorylation detected between these two proteins did not merely reflect a difference in their expression. Therefore, CHO cells were transfected with LEX-C2 or LEX-C2ALL. Whole cell extracts were prepared 48 hours later, samples of which were run on SDS-PAGE, transferred to nitrocellulose and then probed with the LEX A antisera. The levels of the two proteins were found to be comparable (figure 4.9B) and therefore the lack of phosphorylation of LEX-C2ALL (figure 4.9A) could not be accounted for by its lack of

Figure 4.9 ***In Vivo*** **Phosphorylation of C2**

A. CHO cells were transfected with 10μg LEX fusion, labelled with ^{32}P orthophosphate and the LEX proteins recovered by immunoprecipitation. The immunoprecipitated proteins were run on 10% SDS-PAGE and any labelled proteins visualized by autoradiography.

B. Whole cell extracts (500μl) were prepared from CHO cells transfected with 10μg LEX fusion. After transfection the cells were either left untreated or were UV-C irradiated. 50μl of each sample was run on 12% SDS-PAGE, transferred to nitrocellulose and any LEX proteins present detected by Western blotting using the LEX antibody. Numbers to the left of the gel indicate the sizes in kD of the molecular weight markers.

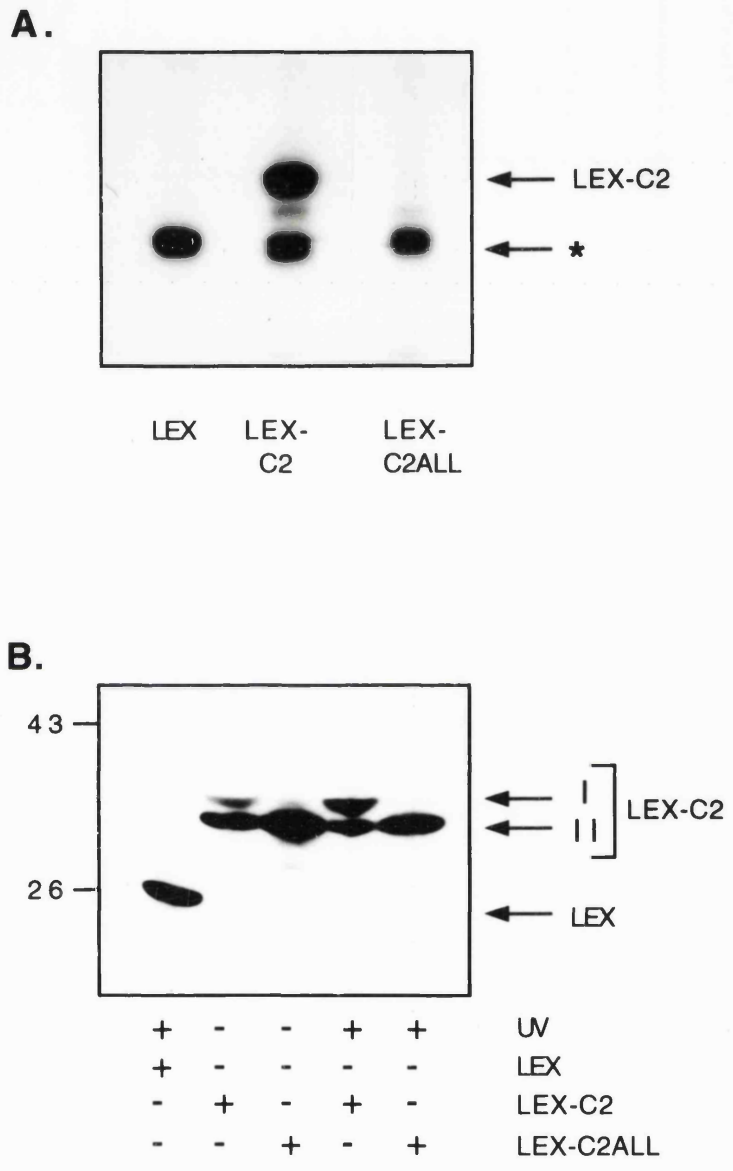


Figure 4.9 *In vivo* phosphorylation of C2

expression. This result also demonstrates that the differences in transactivation potential seen with C2 and C2-ALL does not arise because of differences in their protein levels.

LEX-C2 was found to exist in two forms; a slower and a faster migrating form (bands I and II respectively in figure 4.9B) and LEX-C2ALL possessed the same mobility as the faster (band II) migrating species. This parallels the situation found in the *in vitro* phosphorylation assays where phosphorylation of GST-C2 caused a shift in mobility not obtained with LEX-C2ALL. Thus the slower migrating species of LEX-C2 seen in the Western blot presumably corresponds to a phosphorylated form in which at least thr-71 is phosphorylated since this phosphorylation is known to induce the mobility shift of C2. Extracts were also prepared from UV irradiated cells and again two forms of LEX-C2 were detected. However, in this case the amount of the slower migrating, phosphorylated species (band I) was increased compared to the amount in untreated cells. This increase in phosphorylated LEX-C2 presumably reflects an increase in activity of kinases in response to UV light. This therefore supports the proposal that the activation domain of ATF2 is phosphorylated and further suggests that this phosphorylation can be carried out by the SAP kinases, in agreement with the data presented in section 4.2.3.

4.3 DISCUSSION

The activation domain of ATF2 has been analysed and found to contain two potential key features; the first is a zinc finger between residues 27-49 and the second consists of three phosphorylation sites at residues 69, 71 and 90. By mutating critical residues within these motifs their importance with regard to basal transcriptional activation by this region and also E1a activation was addressed.

Mutation of the two cysteine residues within the zinc finger decreased the level of basal transcription of C2 to an intermediate level, comparable to that obtained with Δ9. Therefore, the zinc finger mutant can be regarded as a transcriptionally active protein which is not as powerful an activator as the wild type form. This means that the zinc finger is dispensable for transcription although it is required for full efficiency. The same situation was found for mutation of one of the phosphorylation sites, namely the serine at position 90. When this was mutated, the level of basal transcription was again reduced to the level of Δ9 but as with the zinc finger mutant still remained a transcriptionally active protein. This is in contrast to the other phosphorylation sites (thr-69 and thr-71) since mutation of either of these residues renders the domain essentially transcriptionally inactive; the level of basal transcription is reduced to lower than Δ9 and is more like that of ATF2 which is transcriptionally silent. Thus, the threonine residues are essential for the N-terminal domain of ATF2 to function as a transcriptional activator. Since these threonine residues have been shown to be phosphorylated it is presumably this post-translational modification that is critical for activation by C2.

The mechanism by which phosphorylation of an activation domain increases its activation potential is not yet understood although several cases of such regulation have been described e.g. phosphorylation of c-Jun at ser-63 and ser-73 increases its activation potential (Smeal et al., 1991) and the same is true for c-Myc upon phosphorylation of thr-58 and ser-62 (Gupta et al., 1993). It is possible that the phosphate groups are required to make essential contacts with protein(s) of the transcription machinery or alternatively they may contribute to the overall structure of the activation domain so ensuring it is in the correct conformation for such an interaction to occur. This is thought to be the case for CREB which is only transcriptionally active when phosphorylated on ser-133. This phosphorylation is proposed to alter the conformation of the activation domain (Gonzalez et al., 1991) so allowing it to interact with the co-activator CBP (Chrivia et al., 1993). Phosphorylation of cJun also allows it to interact with CBP (Arias et al., 1994) and again this may involve a conformational change although as for CREB no evidence for such a model exists.

It is also possible that phosphorylation of the TPTP motif in the activation domain of C2 is not required for the interaction of C2 with a component of the transcription machinery but instead is necessary to obtain a "productive" interaction. In this scenario C2 can interact with the transcription machinery whether phosphorylated or not. However, phosphorylation is now required to activate a rate limiting step in transcription subsequent to binding e.g. it may be required to alter the conformation of another protein in the intiation complex so converting it from an inactive state into an active one. Such a structural change would not occur with the unphosphorylated C2. It is not possible to test this hypothesis until the site of interaction between C2 and the transcription machinery has been determined. Then it can be assessed whether C2 needs to phosphorylated in order to interact or not.

When the mutants in the activation domain were studied with respect to E1a activation the relative importance of the zinc finger and the phosphorylation sites was not found to mirror the situation observed for basal activation. The zinc finger and ser-90 phosphorylation are not essential for moderate basal activation whereas the zinc finger was found to be critical for E1a activation (the serine phosphorylation was not required at all). Since it has been found that E1a activation generally occurs via transcriptionally active forms of ATF2 (sections 3.2.1 and 4.2.2) the zinc finger mutation appears to be an anomaly. This stresses the importance of the zinc finger in E1a activation. This has been further confirmed in the laboratory by studying other members of the CREB family, ATFa1/2/3 and CRE-BPaα. These proteins are highly homologous to ATF2 but only the ATFa sub-family is E1a responsive (Chatton et al., 1993; Nomura et al., 1993). When the sequence of the zinc finger in these proteins was compared to that of ATF2 the amino acid composition of ATFa1/2/3 was found to be identical between cys-27 and his-45 whereas the sequence of CREB-Paα had 4 amino acid substitutions (figure 4.10). When these changes were introduced into an amino terminal portion of ATF2 they were found to decrease its ability to respond to E1a further stressing the

Figure 4.10 Differences in the Amino Terminal Sequences of ATF2, ATFa1/2/3 and CRE-BPaα

Amino acids 24-51 of ATF2, 6-33 of ATFa1/2/3 and 15-42 of CRE-BPaα are compared and the cysteine and histidine residues which form the zinc fingers in these regions are underlined. The ATFa proteins are identical in their N-termini but diverge from each other at residue 90.

Differences from the ATF2 sequence are depicted by hollow letters and those residues of CRE-BPaα which lie in the zinc finger motif are further highlighted by asterisks. These are the mutations introduced into ATF2 (see text).

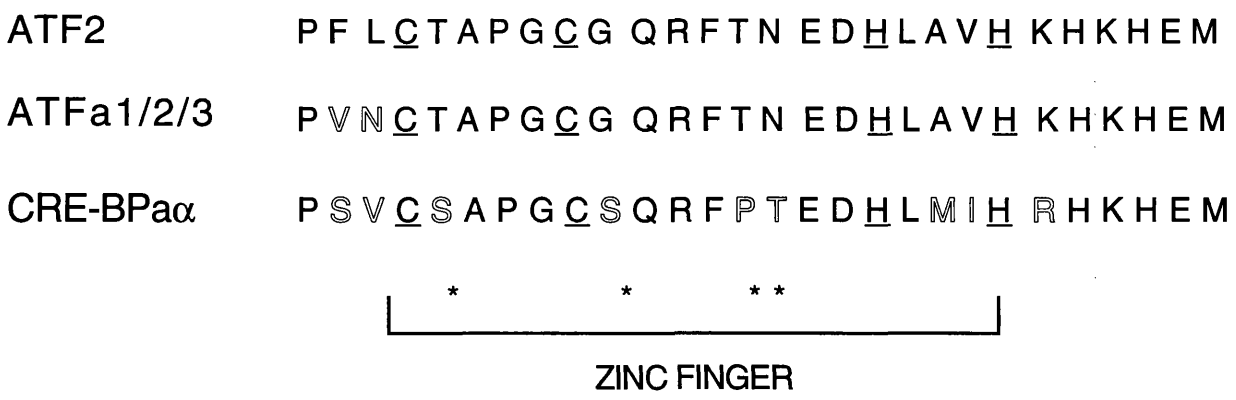


Figure 4.10 Differences in the amino terminal sequences of ATF2, ATFa1/2/3 and CRE-BPaα

importance of the zinc finger in E1a activation (A. Clarke unpubl. obs.). The possible role of this structure in such activation will be discussed later (section 6.3 and 6.6).

The three MAP kinase sites in the activation domain of ATF2 have been shown to be phosphorylated both *in vitro* and *in vivo*. From the experiments I have described, the precise nature of the kinase involved in such a phosphorylation is not at present clear. From the serum and UV irradiation experiments performed in NIH 3T3 cells it appears that both stimuli are capable of potentiating the activity of C2 by phosphorylation of these sites. Although serum and UV light stimulate predominantly different families of enzymes (serum stimulates the Erks whereas UV stimulates the SAPKs), the amount of crosstalk and overlap between the two pathways is not fully determined. That overlap does occur is evidenced by the fact that both pathways are stimulated by Ha-Ras. Thus it is reasonable to find that both serum and UV light can activate the activation domain of C2. However, it also means that it is not clear from these experiments which kinases are actually responsible for phosphorylating and regulating ATF2 *in vivo*. In order to better define this and also the pathway leading to the activation of such kinases, stimulation experiments with transfected cells should again be carried out but using agents whose activities are better defined with respect to the two sets of proline-directed kinases than serum. For example, treatment of NIH 3T3 cells with PMA or FGF activates Erk1 and 2 but does not stimulate the SAPKs (Kyriakis et al., 1994). Another approach would be to use mutants known to interfere with or activate specifically one of the MAPK pathways in transfection experiments and investigate their effect on C2 activation by e.g. UV light. One such mutant, a constitutively active MEK1 (S222D MEK1), is described in Pages et al., 1994. Since MEK1 only activates the Erks and not the SAPKs (section 1.4.2), this mutant should lead to differential activation of the two subfamilies of MAPKs when transfected into cells. Similarly, dominant negative forms of Erk 1 and 2 have been described (K71R Erk1, K52R Erk2) (Frost et al., 1994). These are kinase deficient mutants which interfere with endogenous Erk activity. The mechanism of interference is unclear; it may be that these mutants bind to Erk substrates so preventing their access to active, endogenous Erks or alternatively these mutants may out compete the endogenous Erks for activation by MEK1 and 2. In either case, the activity of the SAPKs should not be affected by these mutants unless the Erks and SAPKs have identical substrate specificity. This means that these Erk mutants can be used to distinguish between Erk/SAPK activation of ATF2.

Like experiments can be performed using mutants which influence the activity of the SAPKs rather than the Erks and with the identification and cloning of the kinases responsible for and leading to SAPK activation (see sections 1.4.2 and 1.4.3) several such mutants have recently been described. These include a dominant negative SEK1 (S220A T224L SEK1) which has been shown to block SAPK activation *in vivo* (Yan et al., 1994) and since SEK1 causes little or no activation of Erk1 *in vivo*, this mutant should specifically inhibit only the SAPK pathway. Similarly, a kinase inactive SEK1 mutant (K129R SEK1) has been described and shown to block SAPK but not Erk activity (Sanchez et al., 1994). A kinase inactive mutant

of SAPK (K55A SAPKα) has also been constucted (Yan et al., 1994). This presumably will act as a dominant negative inhibitor of SAPK activity *in vivo* by mechanisms analogous to those described above for the Erk mutants. However, this has not yet been tested. Finally, since Raf and MEKK have been shown to activate the Erk and SAPK pathways respectively (Yan et al., 1994) these MAPKKKs should also prove useful in identifying *in vivo* the ATF2 kinase (see section 6.3).

A different strategy for determining the correct kinase of ATF2 involves *in vitro* phosphorylation experiments although this will only give an indication of the preferred enzyme and any results obtained would have to be further substantiated by *in vivo* experiments. Erk1/2 has already been tested for its ability to phosphorylate GST C2 and is found to be able to phosphorylate the appropriate sites although not very efficiently. At the time of carrying out this experiment the SAP kinases had not been cloned. However, several members of this family have now been published e.g.SAPKα, β, γ (Kyriakis et al., 1994) and so it should be possible to use these enzymes for *in vitro* phosphorylation assays. The efficiency of phosphorylation of GST C2 by these various enzymes can be compared to each other as well as to their known substrates e.g. Erk1/2 are known to phosphorylate Elk (Marais et al., 1993) whereas SAPKγ (JNK1) phosphorylates cJun (Kyriakis et al., 1994). These experiments have since been performed by G. Patel and are discussed in chapter 6.

CHAPTER 5 (RESULTS):

ATF2 AND E1A ACTIVATION

ATF2 AND E1A ACTIVATION

5.1 INTRODUCTION

ATF2 has been widely studied because it is one member of the CREB family that is E1a responsive. By investigating the properties of this transcription factor it is hoped that it will lead to a better understanding of the mechanism of E1a activation. The converse is also true i.e. that by studying the mechanism of E1a activation an understanding of the normal cellular function of ATF2 will be gained.

The studies described so far have concentrated on the N-terminal activation domain of ATF2. This domain is inactive in the context of the full length protein and so the experiments focused on investigating the truncated and hence active versions of ATF2. Such studies have identified a zinc finger and phosphorylation sites as being important for the function of this domain both with regard to basal transcription and also E1a induction. However, it was essential to assess the importance of these features in the context of the full length protein. Since ATF2 is transcriptionally inert under standard transfection conditions (the activity of GAL-ATF2 is similar to the DNA binding domain of GAL) these N-terminal features in the full length protein can only be studied in the presence of E1a. This remains the case until the normal physiological stimulus or conditions for ATF2 activation are established.

At the beginning of these studies the generally accepted model for E1a activation of transcription via ATF2 was a simple one in which ATF2 merely functioned as a scaffold to bring E1a to the promoter (section 1.3.2). Transcription was then proposed to be activated by CR3 of E1a. The specificity of ATF2 for the E1a response was assumed to be due to the necessary physical interactions between the two proteins. However, at that time a direct physical interaction between ATF2 and E1a had not been demonstrated. Therefore, experiments were performed to address this aspect of the activation model.

5.2 RESULTS

5.2.1 A functional activation domain of ATF2 is required for E1a activation

It has already been shown that a GAL-ATF2 construct in which the amino terminus of ATF2 is missing (residues 8-94 construct Δ4) is E1a unresponsive (Flint and Jones, 1991, see also figure 5.3A). This essentially deletes the activation domain of ATF2 and shows the requirement of this domain for E1a activation. It was next important to determine the way in which this region of ATF2 was involved i.e. did it need to be functionally active in order to allow an E1a response or did it merely perform a physical role. The first possibility was addressed by introducing mutations, which from the studies on C2 were known to reduce or destroy the activity of the activation domain, into the context of full length ATF2. The effects of such mutations on E1a activation were then determined by co-transfection of the mutant

GAL-ATF2 fusions into CHO cells along with an E1a expressing plasmid and a GAL-CAT reporter construct.

Phosphorylation of the amino terminus of ATF2 was investigated first since this has been shown to be critical for the function of the activation domain. Therefore, the point mutations in the three potential MAP kinase sites were introduced into full length ATF2 as shown in figure 5.1B and their effect upon E1a induction is shown in figure 5.1A. When all three sites were mutated together (ATF2-ALL) E1a activation was reduced from about 20 fold with wild type ATF2 to about 3 fold with the mutant. Thus, phosphorylation of the activation domain in full length ATF2 appears to be essential for E1a activation. Since this post-translational modification is required for transcriptional activation by this domain, it seems that E1a acts only via transcriptionally active forms of ATF2. This parallels the results obtained with the truncated versions of ATF2 and E1a stimulation.

To assess the relative importance of the three sites, each site was mutated individually in ATF2 (figure 5.1B). When the threonine residues were mutated (ATF2-T1 and ATF2-T2) the fold E1a activation was similar to that of ATF2-ALL i.e. about 3 fold (figure 5.1A). Similarly, mutation of threonine residues in combination with either the serine or the other threonine (ATF2-T1S, ATF2-T2S or ATF2-T1T2) also gave E1a unresponsive variants of ATF2. Therefore, phosphorylation at the threonine residues is essential for E1a induction of transcription. Since, these are the critical phosphorylations for the activity of C2 this again shows the relationship between E1a activation and potentially active forms of ATF2.

Mutation of the serine residue in C2 reduced its basal level of transcription but did not render this domain transcriptionally inactive and did not affect its response to E1a. Such a mutation in the context of full length ATF2 (ATF2-S) also did not affect E1a activation with the fold induction with ATF2-S being the same as wild type i.e. approximately 20 fold (figure 5.1A). This further confirms the link between transcriptional activation by ATF2 and its E1a response and also suggests that ATF2 makes an active contribution during such a response.

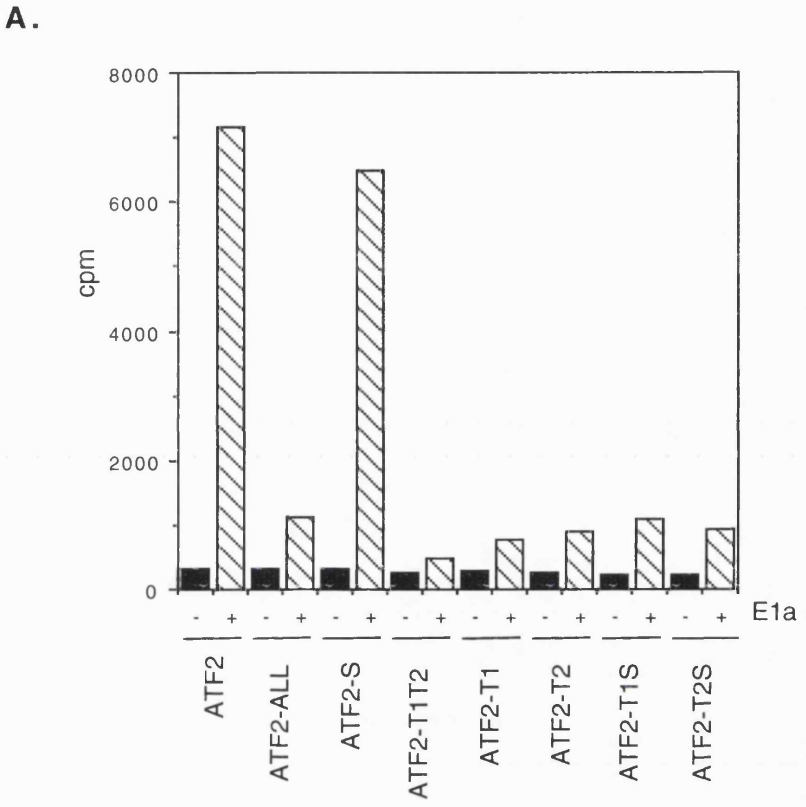
The requirement for the zinc finger in E1a activation was also determined. This has been measured previously (Flint and Jones, 1991) but it was important to repeat the transfections so that a comparison between the zinc finger and the phosphorylation sites could be drawn with regard to E1a activation of full length ATF2. Mutating the zinc finger, ATF2pmZn (figure 5.2B), resulted in a mutant that was unresponsive to E1a (figure 5.2A). It had the same fold induction, 3 fold, as ATF2-ALL and ATF2-T1T2 and therefore is as critical as phosphorylation for E1a activation of full length ATF2. However, in the context of the isolated activation domain, the zinc finger was found not to be as critical for basal transcription as the TPTP sequence. This suggests that the role the zinc finger plays in E1a activation is different to that played by phosphorylation. This is further substantiated by comparing ATF2pmZn with ATF2-S. Both mutations reduce the level of basal transcription in

Figure 5.1 Basal and E1a Activation of Full Length ATF2 Phosphorylation Mutants

CHO cells were transfected with 5μg GAL fusion, 2μg of the reporter G5E4CAT and where indicated 5μg of pCE.

A. Graph of transcription levels measured from the reporter.

B. Diagram of the ATF2 phosphorylation mutants used in the transfections. Mutated residues are highlighted.



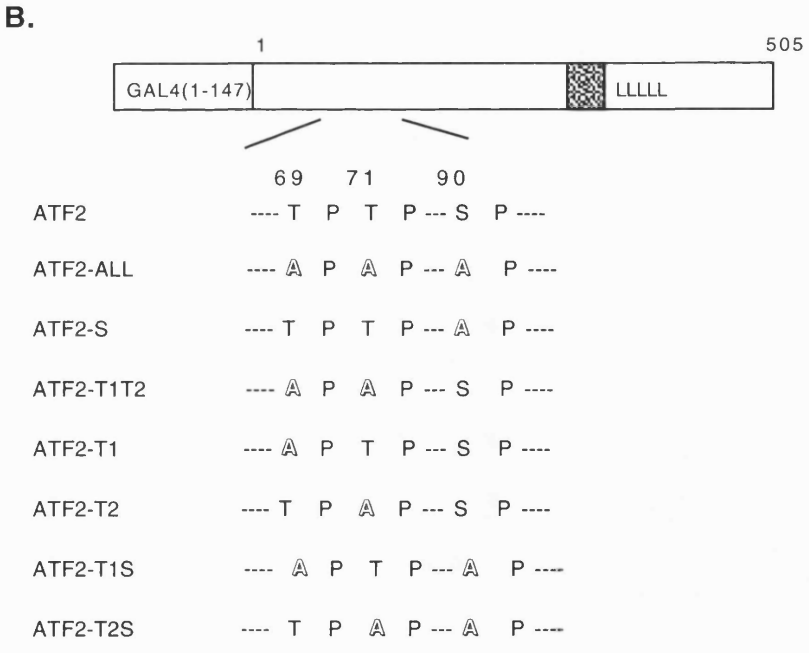


Figure 5.1 Basal and E1a activation of full length ATF2 phosphorylation mutants

Figure 5.2 Effect of Mutating the Zinc Finger in Full Length ATF2 Upon E1a Activation

CHO cells were transfected with 5μg GAL fusion, 2μg of the reporter G5E4CAT and where indicated 5μg of pCE.

A. Graph of transcription levels measured from the reporter.

B. Diagram of the zinc finger mutantion used in the transfection. The phosphorylation mutants used are described in figure 5.1B.

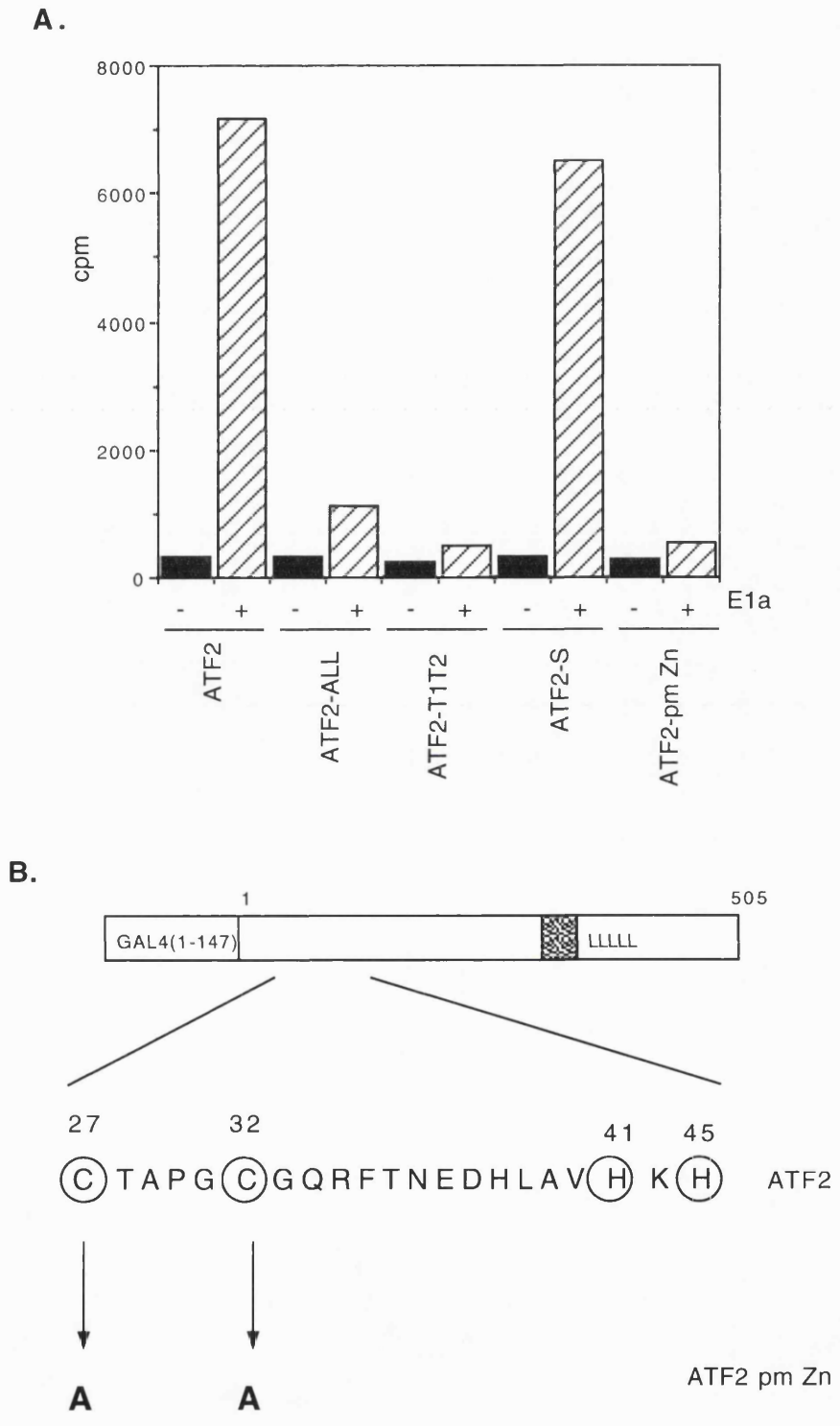


Figure 5.2 Effect of mutating the zinc finger in full length ATF2 upon E1a activation

the context of C2 by a similar amount (they do not destroy it) but in the full length protein (and also the short forms) only the zinc finger mutant is E1a unresponsive.

From this data it seems that phosphorylation is required for E1a induction by virtue of its involvement in transcriptional activation. However, the zinc finger appears to be involved in a different aspect of the activation process. This will be further discussed in relation to a model for E1a activation (section 6.6).

5.2.2 Requirement for the activation domain of E1a

The previous section has shown that the activation domain of ATF2 is required and involved in E1a stimulation of gene expression via ATF2. However, the role of E1a itself during such activation has not been addressed. It is known that the activation domain of E1a (CR3) is essential since only the 13s form of E1a can stimulate transcription via full length ATF2 (Flint and Jones, 1991). However, the role of E1a with respect to activation of truncated versions of ATF2 had not been investigated. Therefore, transient transfections were performed in CHO cells with GAL fusions along with either a 13s or 12s E1a expressing plasmid. Transcription was determined by co-transfection of a GAL-CAT reporter construct. As can be seen from figure 5.3A activation via full length ATF2 was 13s specific as expected. Δ4 (ATF2 lacking amino acids 8-94) was both 12s and 13s E1a unresponsive in agreement with published data (Flint and Jones, 1991). It was found that activation via Δ9 and C2, which only contain the N-terminus of ATF2, was 13s specific (figure 5.3B). Therefore, stimulation via the activation domain of ATF2 requires the activation domain of E1a. Thus it seems that the activation domains of both ATF2 and E1a are involved in transactivation.

To further establish the involvement of E1a's activation domain, transfections were performed using two mutants of E1a which are defective in transactivation. The activation domain of E1a consists of two regions as shown in figure 5.4. The C-terminal region (amino acids 183-188) contains the promoter targeting region required to tether E1a to promoters presumably via interaction with DNA bound transcription factors such as ATF2. The N-terminal region (amino acids 147-177) contains the activation domain and is thought to function by interacting with a component of the basal transcription machinery (Webster and Ricciardi, 1991). The mutants used in the transfections had single amino acid substitutions in critical residues only in this N-terminal activation region of CR3. Therefore, they should still be able to interact with ATF2 but they would not be able to directly activate transcription once they were brought to the promoter. However it was possible, depending on the mechanism by which E1a activates transcription, that these mutants could indirectly stimulate transcription. For example, the interaction of E1a with ATF2 may lead to a conformational change in ATF2. This could unmask the activation domain of ATF2 enabling it to interact with the transcription machinery and stimulate trancription. The level of activation in this situation would be expected to be lower than that obtained with wild type E1a because the activation domain of E1a no longer itself directly contributes to the activation process.

Figure 5.3 E1a Activation of ATF2 and its N-terminal Domain is 13s Specific

CHO cells were transfected with 5µg of GAL fusion, 2µg of the reporter G5E4CAT and where indicated 5µg of an E1a expressing plasmid

Key:

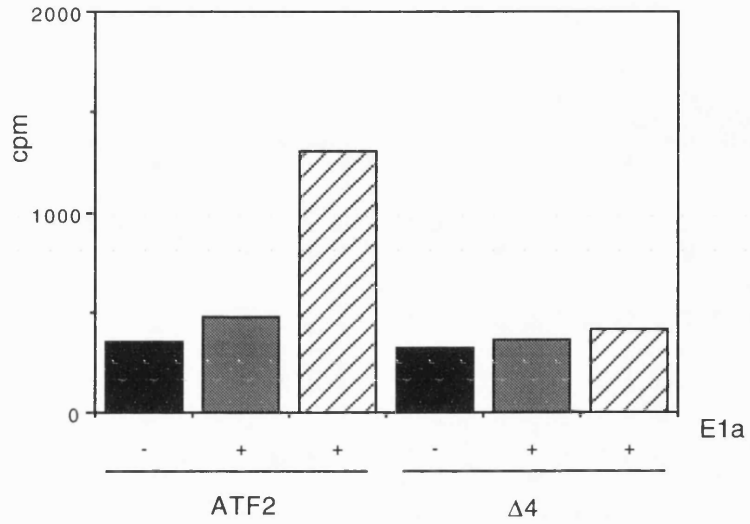
GAL fusion alone

GAL fusion plus 12s E1a expression plasmid (JF12)

GAL fusion plus 13s E1a expression plasmid (JN20)

A & B. Graphs of transcription measured from the reporter. Labels below the grouped lanes indicate the GAL fusion used in the transfection.

A.



B.

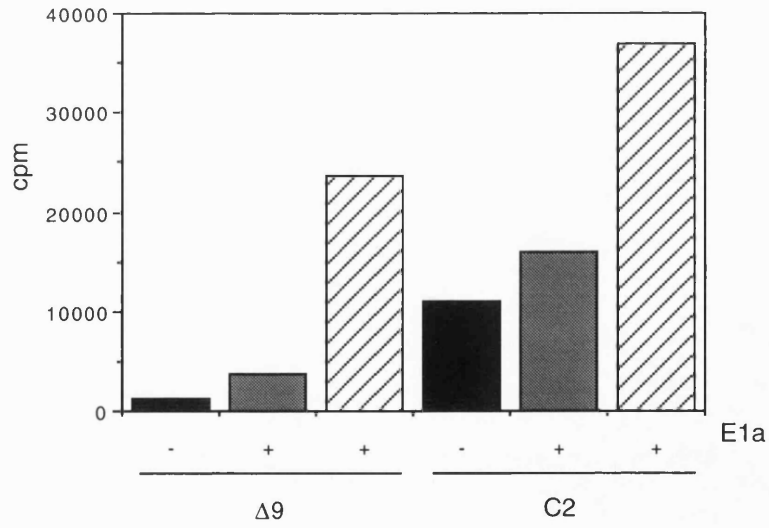


Figure 5.3 E1a activation of ATF2 and its N-terminal domain is 13s specific

Figure 5.4 Diagram Depicting the Sub Domains of E1a CR3 (Adapted from Webster and Ricciardi, 1991)

The activation domain (CR3) of 13s E1a comprises amino acids 140-188 and is lacking in the 12s form of the protein. Mutational analysis has been used to define two sub-domains within this region:

(i) N-terminal zinc finger region, amino acids 147-177, responsible for transactivation by 13s E1a

(ii) C-terminal promoter targeting region, amino acids 183-188, responsible for tethering E1a to promoters via protein-protein interactions with bound transcription factors.

Indicated are the four cysteine residues (cys-154, -157, -171, -174) which coordinate zinc in the activation region.

The two mutants used in the experiment presented in figure 5.5 have the following residues mutated:

- amino acid 147 V to L (mutant 147VL)
- amino acid 150 P to G (mutant 150PG)

These residues lie in the activation domain (highlighted in the diagram) and are known to be critical for E1a transactivation.

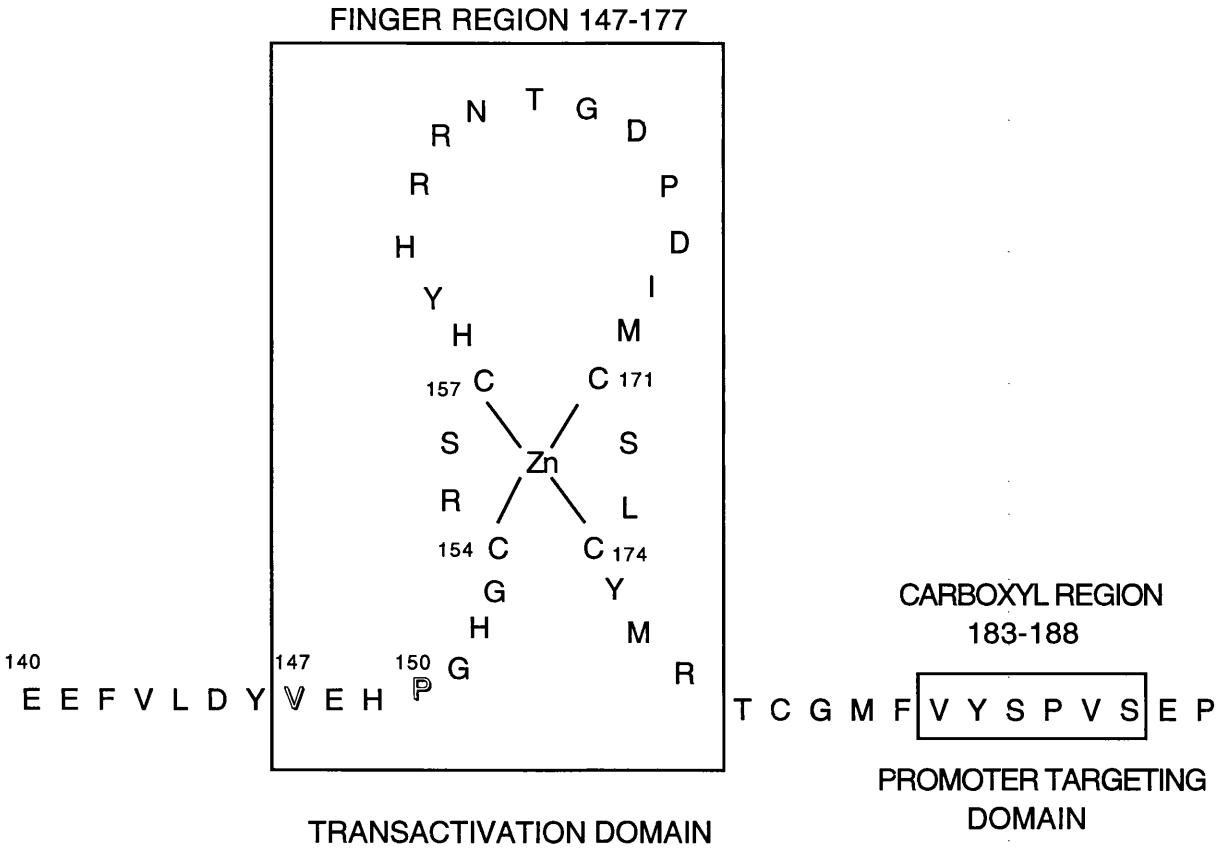


Figure 5.4 Diagram depicting the sub-domains of E1a CR3

When these mutants were transfected into CHO cells along with GAL-ATF2 no stimulation of transcription of the GAL-CAT reporter was observed (figure 5.5) whereas activation by wild type 13s E1a resulted in transcription levels about 20 fold above the level of basal transcription with GAL-ATF2 alone. Thus it seems that the proposed interaction between ATF2 and E1a is not sufficient to convert ATF2 into a transcriptional activator if the activation region of E1a is non-functional. This suggests that the activation region of E1a plays an important part in the conversion of ATF2 from an inactive to an active form (see section 5.3 and 6.4).

5.2.3 Physical interactions between ATF2 and E1a

E1a does not bind to DNA in a sequence specific manner (Chatterjee et al., 1988). Therefore, it is proposed to be brought to the promoter via its association with DNA bound transcription factors such as ATF2. However, at the time of starting this work a direct physical interaction between E1a and ATF2 had not been demonstrated. The standard methods for detecting interactions between proteins such as co-immunoprecipitation and EMSA with *in vitro* translated E1a and ATF2 had already been tried in the laboratory but without success. This suggests that either the interaction between these two proteins is weak and hence not detectable by conventional methods, that the correct binding conditions had not been utilized or that direct binding between these proteins does not occur. Therefore, it was decided to employ a more sensitive binding assay to see if a direct interaction, even a relatively weak one, could be detected.

The system chosen for the binding studies utilized the GST expression system for the production of proteins (Smith and Johnson, 1988). This has the advantage of being able to produce large amounts of GST fusion protein in bacteria which can then be recovered by adsorption to glutathione sepharose beads and washed to remove any contaminating bacterial proteins thus resulting in a pure protein preparation. The adsorbed proteins can then be used for column chromatography studies or Far-Western blotting assays. Both of these methods are more sensitive than co-immunoprecipitation or EMSA.

Column chromatography was the first method used and the particular pGEX expression vector used for the expression of proteins, pGEX-2TK, contained a protein kinase A (PKA) site between the GST moiety and the cloned protein of interest (Kaelin et al., 1992). This meant that ^{32}P-labelled proteins could be prepared by phosphorylating the purified GST fusions *in vitro* with γ^{32}P-ATP and PKA. Since this vector also contains a thrombin cleavage site, the GST portion can then be cleaved off by this enzyme and removed by re-adsorption to glutathione beads thus leaving pure labelled protein. However, since the GST portion is not labelled, any studies which look for bound proteins by autoradiography as the method of detection means that it is not always necessary to remove the GST portion after cleavage.

GST-ATF2, -E1a, -pX and also GST alone were prepared from bacteria by a standard protocol used for GST studies (see materials and methods). pX is a viral transactivator

Figure 5.5 **A Functional Activation Domain of E1a is Required for Transactivation via GAL-ATF2**

CHO cells were transfected with 5μg of GAL-ATF2, 2μg of the reporter G5E4CAT and where indicated 5μg of wild type or mutant E1a as described in the key and figure 5.4.

Key:


GAL-ATF2


GAL-ATF2 + pCE (wild type E1a)


GAL-ATF2 + 147VL E1a mutant


GAL-ATF2 + 150PG E1a mutant

Graph shows the level of transcription measured from the reporter.

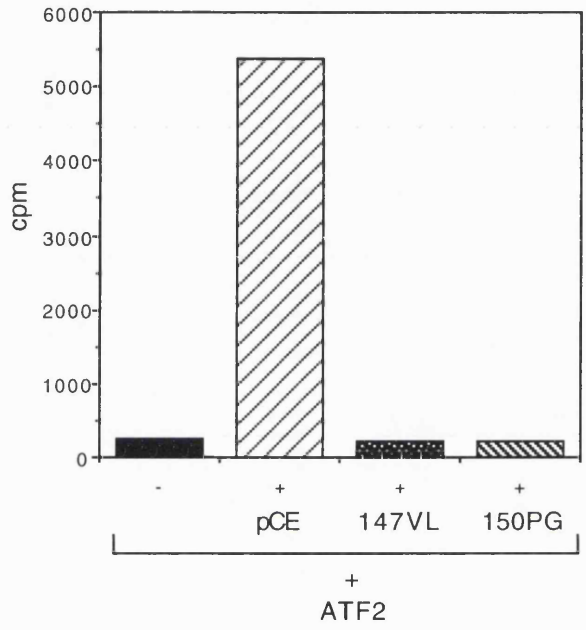


Figure 5.5 **A functional activation domain of E1a is required for transactivation via GAL-ATF2**

protein expressed by the hepatitis B virus (HBV) (Spandau and Lee, 1988; Maguire et al., 1991) and one report has shown it to function in a similar manner to E1a i.e. it requires ATF2 in order to bind to DNA in a sequence specific manner and thus be brought to the promoter to stimulate transcription. However, the situation is somewhat different to that of E1a since pX stimulates binding of ATF2 to CRE-like sequences which it would otherwise not recognize whereas E1a does not alter the DNA binding specificty of ATF2. This report showed that pX bound directly to ATF2 by using EMSA and Far-Western blotting techniques. Therefore, GST-pX was included in this experiment as a positive control for ATF2 binding.

The expressed GST fusions once bound to the glutathione sepharose beads were used to make columns of 50µl in size and ^{32}P-ATF2 prepared from bacteria was used as the substrate for binding to these columns. It was necessary to disrupt any ATF2 homodimers before performing the binding because the dimeric form may be stable enough, in the absence of a sufficient driving force, to prevent formation of monomeric ATF2 and subsequent interaction with proteins on the columns. This is not a problem if ATF2 interacts with the proteins in a dimeric form but if disruption of the dimers is required, as would be the case for interaction of free ATF2 with bound ATF2, dissociation of the ATF2 dimers is essential. Therefore, ^{32}P-ATF2 in BLOTTO loading buffer containing 0.1M NaCl was heated to 55^{o}C for 7 minutes (based on Benbrook and Jones, 1990) and then immediately chilled on ice before being passed over the columns. After recycling the ATF2 through the columns 5 times any unbound proteins were washed off with loading buffer with 0.1M NaCl. Bound proteins were eluted from the columns by washing with loading buffer containing a higher salt concentration, firstly 0.5M NaCl and then 1M NaCl. The washes from the columns were run on SDS-PAGE and any labelled ATF2 present in the samples detected by autoradiography as shown in figure 5.6.

Labelled full length ATF2 (apparent MW ~66kD) was found to bind to the GST-ATF2 column as shown by its elution from the column with a 0.5M NaCl wash. This is expected since the free ATF2 should be able to dimerize through its leucine zipper with the ATF2 on the column. Also eluted in this 0.5M NaCl wash were several labelled short forms of ATF2 (apparent MWs 30-43kD). However, these were present at a much lower level (less than 0.5% of their input level) than full length ATF2 (present at approximately 10% of its input level). These shortened products are thought to represent either degraded or truncated variants of ATF2 which, based on their size and the fact that ATF2 is labelled at its N-terminus, are presumed to lack the leucine zipper region of ATF2. Thus explaining why they are only very weakly retained by the GST-ATF2 column. In fact it may be that these forms are not specifically bound by the column at all but instead represent degradation of the bound full length ATF2.

There was no interaction of ATF2 with the GST column which shows that the observed binding of full length ATF2 to the GST-ATF2 column was through ATF2 and not the GST moiety. In addition, there was no interaction of ATF2 with either GST-E1a or GST-pX since

Figure 5.6 Labelled ATF2 is retained on a GST-ATF2, but not a GST-E1a, Column

10µl of ^{32}P-labelled ATF2 was loaded onto 50µl GST columns and chromatography performed as described in materials and methods. 5µl (1/20th) of the flow through, 10µl (1/10th) of the first wash and 50µl (whole sample) of fractions 2 and 3 of the 0.5M NaCl elution step were run on 10% SDS-PAGE. The 1M NaCl elution stage is not shown. Any ^{32}P-ATF2 present in the samples was detected by autoradiography. Labels above the lanes indicate the GST column used.

Full length ATF2 migrates with an apparent molecular weight of approximately 66kD which corresponds to the uppermost protein band in each sample lane. The smaller protein bands visible (apparent molecular weights 30-43kD) correspond to truncated/degraded forms of ATF2.

M=molecular weight markers, the sizes of which (kD) are indicated on the right of the gel.

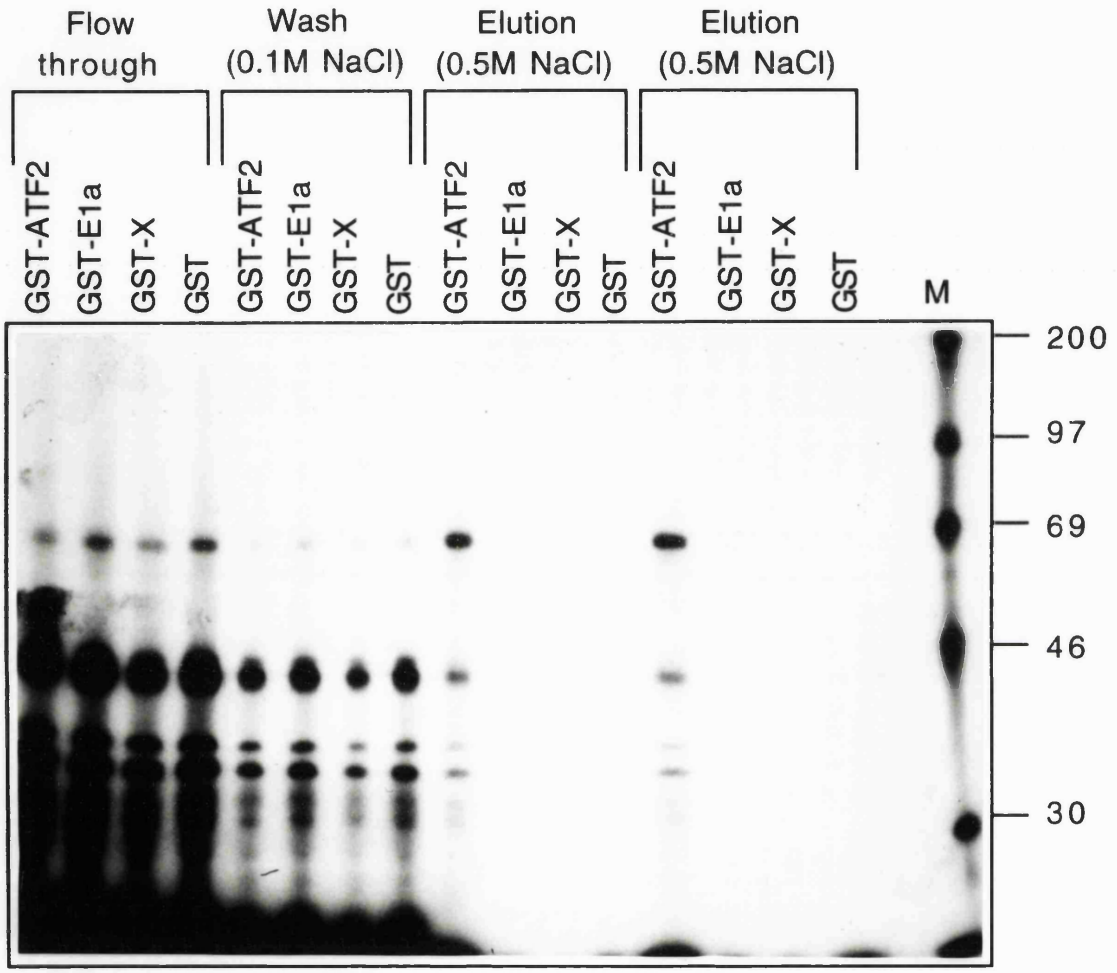


Figure 5.6 Labelled ATF2 is retained on a GST-ATF2, but not a GST-E1a, column

ATF2 was not eluted from these columns at either 0.5M (figure 5.6) or 1M (data not shown) NaCl concentrations. The result with pX was surprising given that an interaction between pX and ATF2 has been reported. However, in that report the binding assays used only partially purified proteins so a direct interaction between ATF2 and pX may not occur. Furthermore other groups have also been unable to reproduce the binding data so the result remains controversial.

The lack of binding of ATF2 to E1a suggests that there is not a direct interaction between these proteins. However, it is possible that the correct conditions for the binding have not been used in this experiment. Although there is binding of ATF2 to the GST-ATF2 column, it does not necessarily mean that the binding conditions are suitable for ATF2 and E1a since the ATF2-ATF2 interaction occurs via the leucine zipper motif and is a strong interaction whereas that of ATF2-E1a is thought to be relatively weak.

Far-Western blotting with the GST proteins was also carried out. For this GST-ATF2, GST-E1a and GST proteins were eluted from glutathione sepharose beads and cleaved with thrombin. The cleaved proteins were then run on SDS-PAGE, transferred to nitrocellulose and probed with ^{32}P-ATF2. Any bound ^{32}P-ATF2 was detected by autoradiography as shown in figure 5.7. Using this method an interaction between free ATF2 and that immobilized on the filter was detected. Although the proteins on the filter had been cleaved to remove the GST portion, the cleavage reaction seems to have been incomplete as GST-ATF2 is detected in the blot as well as ATF2. However, the binding to GST-ATF2 occurs through the ATF2 moiety since there was no binding to GST alone. Truncated/degraded forms of ATF2 were also detected on the filter (the aserisk labelled bands in figure 5.7). Since these shortened forms are not labelled, it cannot be determined which region of the full length protein they lack. However, since they were found to hybridize as efficiently as full length ATF2 with the ^{32}P-ATF2 probe, they presumably still possess the leucine zipper region. There was found to be no binding of ATF2 to E1a which is in agreement with the column data. A number of different conditions were tested for blotting e.g. the proteins on the filters were denatured/renatured before probing, the blots were probed at room temperature or 4^{o}C and the blots were also probed with ATF2 in the presence of nuclear extract. These various conditions had differing effects on the level of background and non-specific binding of ATF2 to the filter e.g. binding of ATF2 to GST was observed when the binding was carried out at room temperature but in no instance was binding of ATF2 to E1a detected (data not shown).

The one major problem with the above studies was that there was not a suitable positive control for ATF2 binding. As explained earlier, although homodimerization between ATF2 proteins can be detected, this interaction is a reasonably strong one. This is evidenced by the fact that although ATF2 can heterodimerize with c-Jun it only does so if both proteins are synthesized at the same time, mixing ATF2 and c-Jun after synthesis does not result in heterodimers due to the stability of the homodimers (Benbrook, D. M. pers. comm.).

Figure 5.7 Far-Western Blot Probed with Labelled ATF2

Proteins purified by the GST system were run on 15% SDS-PAGE and transferred to nitrocellulose. The filter was probed with ^{32}-P ATF2 and its binding detected by autoradiography as shown by the arrows. Labels above the gel indicate the proteins transferred to nitrocellulose. The asterisk indicates truncated/degraded forms of ATF2 detected on the filter.

M=molecular weight markers, the sizes of which (kD) are indicated on the left of the gel.

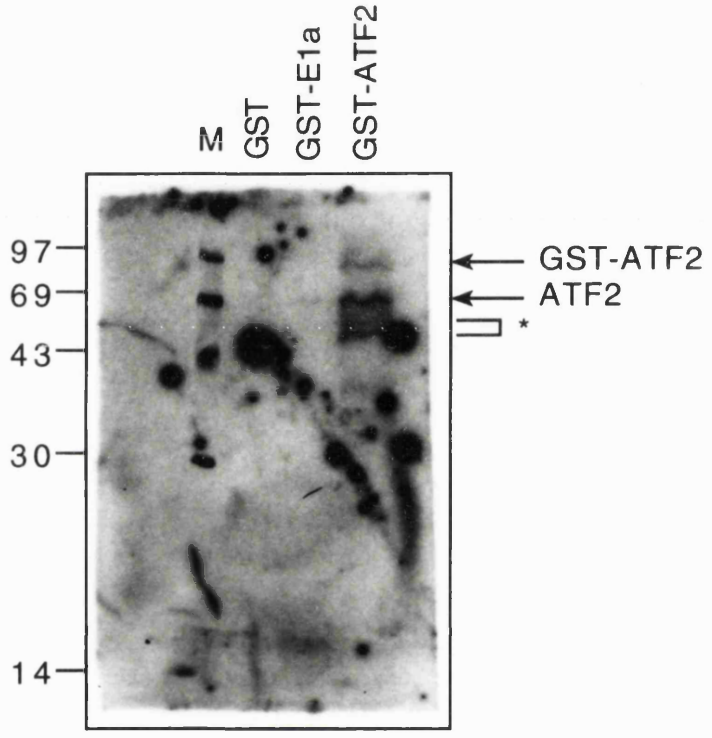


Figure 5.7 Far Western blot probed with labelled ATF2

Therefore, I decided to try the binding studies the other way around i.e. bind labelled E1a to immobilized ATF2 and use as the positive control binding of E1a to TBP.

For these studies GST-ATF2 was prepared from bacteria, bound to glutathione beads and used for column chromatography. Since the region of ATF2 required for an E1a response has been delineated to the N-terminus, I decided to see if this region (Δ9 amino acids 1-112) could bind E1a i.e. was there a correlation between the region of ATF2 required for *in vivo* E1a activation and *in vitro* E1a binding. Conversely, an ATF2 construct lacking the N-terminus has been shown to be E1a unresponsive (see figure 5.3A) so it was important to test if such a GST fusion could bind E1a. Control columns consisting of GST alone or GST-TBP were also prepared. The first control is to check that no interaction between E1a and GST occurs and the second is the positive control for the binding reaction (Lee et al., 1991).

^{35}S-E1a was synthesized by *in vitro* translation in the presence of ^{35}S methionine. This was then passed over the GST columns in MJ loading buffer 5 times before the columns were washed in the same buffer. The beads were extracted from the columns and boiled in Laemmli buffer to release the bound proteins. These were then run on SDS-PAGE and any labelled E1a present was detected by autoradiography. From figure 5.8 it can be seen that E1a binds well to the GST-TBP column with greater than 50% of the loaded E1a being bound. No such interaction was found to occur with the GST column. However, E1a was found to bind to the GST-ATF2 column although far more weakly than for binding to TBP; only about 2% of the input was retained on the column. Thus, in this assay a direct interaction between E1a and ATF2 was detected whereas in the studies above using GST-E1a no such interaction was observed. This discrepancy could be explained in a number of ways. Firstly, different buffers were used for loading the samples and performing the binding reactions in the two sets of experiments (BLOTTO versus MJ). Secondly, in the later experiment E1a was synthesized in reticulocyte lysate as opposed to bacteria. This means that potentially it will be post-translationally modified e.g. phosphorylated and this may be necessary for binding to occur especially given that E1a is a phosphoprotein. However, it should be noted that E1a phosphorylation has not been found to be relevant for its function (Tremblay et al., 1989). Finally, it may also be that synthesis in reticulocyte lysate leads to the correct folding of E1a whereas this is not the case for over-production in bacteria. Hence the interaction is only seen with *in vitro* translated E1a. Whatever the explanation it is clear that a direct interaction between ATF2 and E1a can occur.

Binding of E1a to mutants of ATF2 lacking either the amino or carboxy regions was also found to occur. However, the binding was decreased compared to that of full length ATF2. Binding to ATF2 lacking the N-terminus was very unexpected since this construct is not activated by E1a. This will be discussed later (section 5.3). Thus it seems that E1a can bind to either the N- or C-terminus of ATF2 but binding is most efficient with the full length protein.

Figure 5.8 Labelled E1a is retained on a GST-ATF2 column

10µl of *in vitro* translated ^{35}S-E1a was loaded onto 50µl GST columns and chromatography performed as described in materials and methods. 50µl (1/2) of the flow through, 50µl (all) the last wash and 50µl (all) the beads from the column were run on 10% SDS-PAGE. Any ^{35}S-E1a present in the samples was detected by autoradiography. Labels above the lanes indicate the GST column used.

M=molecular weight markers, the sizes of which (kD) are indicated on the left of the gel.

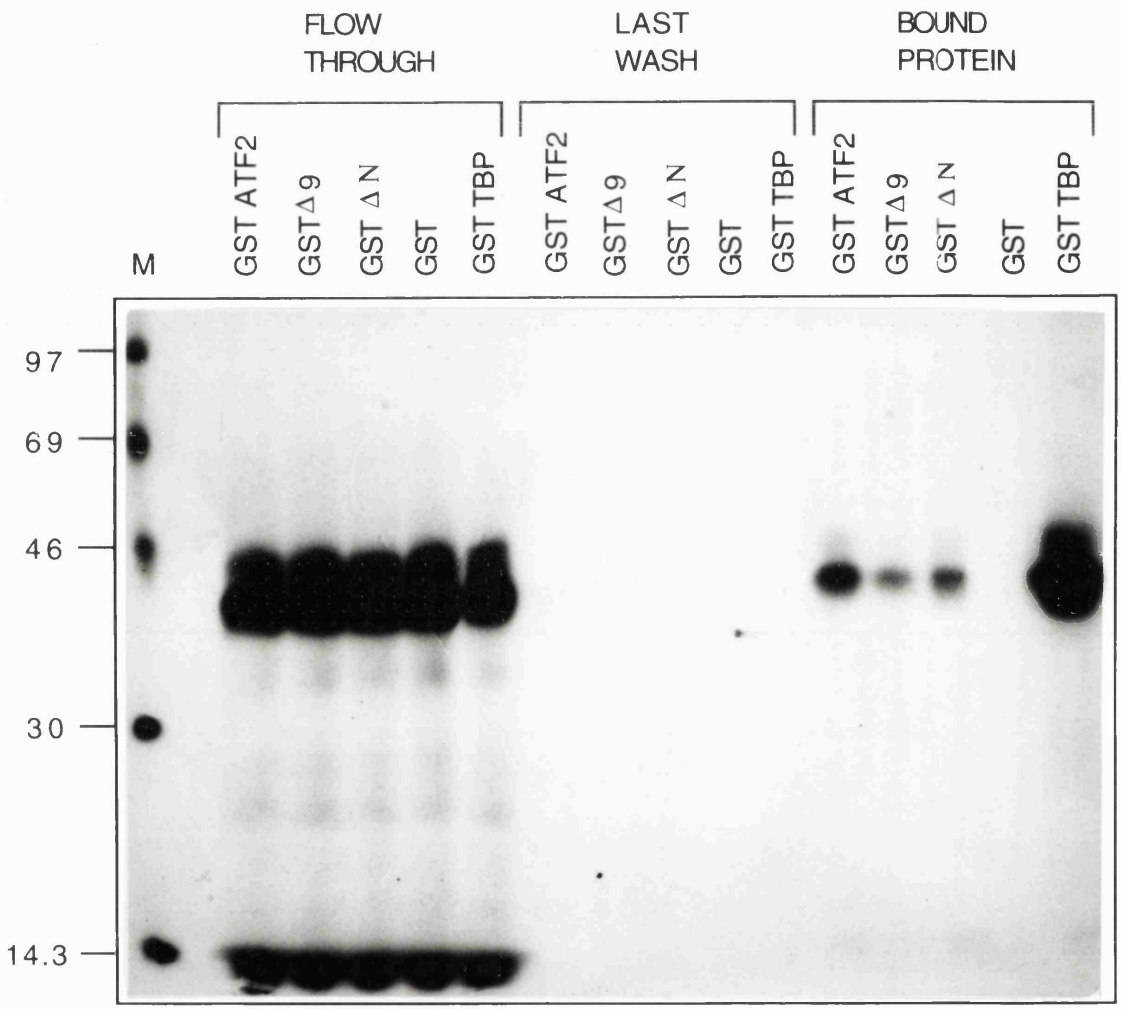


Figure 5.8 Labelled E1a is retained on a GST-ATF2 column

5.3 DISCUSSION

In this chapter the mechanism of E1a activation via full length ATF2 has been addressed. Firstly, the role of the activation domain in ATF2 was investigated and it was found that this domain had to be in a moderately active state for an E1a reponse to be obtained. This was shown by introducing mutations known to compromise the activity of truncated versions of ATF2 into the context of the full length protein and assessing their effect on E1a inducibility. Phosphorylation of the TPTP motif which is critical for transcriptional activation by the N-terminus of ATF2 was found to be essential for E1a activation of full length ATF2. Phosphorylation of the SP motif which is not as critical as the TPTP sequence for basal transcription was found not to be required for E1a activation. Thus it seems that if ATF2 is capable of activating transcription via its N-terminal domain it is also capable of supporting an E1a response. The mutants suggest that the activation domain of ATF2 does not have to be fully active for an E1a response to occur but that phosphorylation of the TPTP motif is absolutely required.

The zinc finger in the activation domain was also investigated. As with the SP phosphorylation, the zinc finger is not essential for basal transcription by the N-terminus of ATF2. However, in contrast to the serine phosphorylation, the zinc finger is critical for E1a activation. The role the zinc finger plays in E1a activation is unclear at present. However, from the data presented there appear to be two main possibilities. Firstly, the zinc finger could be involved in E1a transactivation via the TPTP motif (see sections 6.3 and 6.6 for details). Secondly, the zinc finger may be required for the interaction of E1a with ATF2 (discussed below). Zinc fingers are one of the major structural motifs involved in protein-DNA interactions (see section 1.1.4 (i)), however, instances of protein-protein interactions via zinc fingers have also been determined e.g. the zinc finger transcription factors YY1 and Sp1 interact with each other (Lee et al., 1993). The region of YY1 essential and sufficient for this interaction consists of the first one and a half zinc fingers of its four zinc fingers present at the C-terminus (the four zinc fingers together also constitute the DNA binding domain). The region of Sp1 involved consists of its transcriptional activation domain D and three zinc fingers which form its DNA binding domain. Precise characterization to determine whether the zinc fingers of Sp1 are also involved in the interaction has not yet been carried out. However, it is interesting to note that both YY1 and Sp1 contain zinc fingers since E1a and ATF2 are also both zinc finger proteins and the regions of E1a and ATF2 required for activation are those containing the zinc finger.

The role of the activation domain of E1a in stimulation of transcription via full length ATF2 was also investigated. Firstly, CR3 of E1a was found to be essential to stimulate transcription via the activation domain of ATF2. Secondly, by using mutants of E1a which were defective in activation but still capable of being targeted to ATF2, it was found that an interaction between ATF2 and E1a was not sufficient to stimulate ATF2 activity. This could suggest that ATF2 is not converted into a transcriptional activator by E1a and that ATF2 does

merely act as a scaffold for E1a to stimulate transcription. However, work on the ATF2 mutants has already shown that the activation domain of ATF2 is essential for the E1a response. Therefore, it may be that E1a has to be transcriptionally active in order that conversion of ATF2 from its inactive to an active state occurs i.e. the activation domain of E1a may play a dynamic role in the conversion process.

With regard to the role E1a plays in ATF2 activation, the following mechanism can be envisioned based on the experimental data I have presented. Firstly, the interaction of E1a with ATF2 may well cause the unmasking and exposure of ATF2's activation domain by the induction of conformational changes. However, by itself this process may be insufficient to convert ATF2 into a transcriptional activator. Thus the activation domain of ATF2 then needs to either co-operate with another activator or be modified e.g. phosphorylated in order to be fully functional. E1a could participate in both these secondary steps; firstly E1a has its own activation domain with which ATF2 could co-operate and secondly E1a could stimulate pathways which lead to post-translational modification of ATF2. These models will be discussed later (see section 6.4). However, it is clear that the activation domains of both E1a and ATF2 are required in order to stimulate transcription via full length ATF2.

Finally, the nature of the physical interaction between E1a and ATF2 was investigated. This was achieved by performing column chromatography and Far-Western blotting studies with purified E1a and ATF2 proteins to determine if they could bind directly to each other. When both components were prepared in bacteria and ATF2 was used as the "probe", no interaction was observed. However, when E1a was synthesized in reticulocyte lysate and used as the probe, a weak interaction with bacterially produced ATF2 was detected. The buffers used in the two sets of experiments were different so no direct comparisons can be made. In order to make such a comparison, the experiment which uses E1a as the probe should be repeated but using the alternative buffer. If an interaction still occurs then the site of synthesis of E1a would seem to be important (for reasons discussed in section 5.2.3). However, it may be that reticulocyte lysate contains proteins which stabilize or mediate the interaction between E1a and ATF2 and that this accounts for the different results obtained.

During the course of this work, two other groups published data concerning the physical association between ATF2 and E1a. Abdel-Hafiz et al (1993) performed co-immunoprecipitation studies using bacterially produced ATF2 and E1a proteins. The E1a was expressed as a polyhistidine-fusion protein and was purified by retention on a nickel column followed by elution under denaturing conditions (8M urea) by changing the buffer pH from 8.0 to 4.5. The ATF2 was overproduced in bacteria and purified to 85% homogeneity by a protocol based on its heat stability. The two proteins (one of which was radioactively labelled) were mixed, recovered from solution by immunoprecipitation with an antibody specific for the unlabelled component and any association of the labelled protein was detected by autoradiography. Liu and Green (1994), on the other hand, performed GST batch binding assays; GST-ATF2 was purified from bacteria by adsorption to glutathione-agarose beads

whereas E1a was synthesized by *in vitro* translation in the presence of ^{35}S-methionine, $ZnCl_2$ and β-mercaptoethanol. The two proteins were then incubated together and any E1a bound to the GST-ATF2 beads after several washes was detected by autoradiography. In both studies a direct interaction between full length ATF2 and 13s E1a was detected which is consistent with my data. Chatton et al (1993) also published data on E1a interactions but this time using ATFa1, another member of the CREB family which is highly related to ATF2 and which is E1a responsive. They performed co-immunoprecipitations with *in vitro* translated ATFa1 and E1a and found that a direct interaction between these two proteins also occcurred. Thus the binding of ATF2 and ATFa1 to E1a appears to occur in an identical manner i.e. by direct association which is expected given the similarity between ATFa and ATF2.

The results presented in section 5.2.3 showed that E1a was not only capable of interacting with full length ATF2 but also of interacting with either the N-terminus (amino acids 1-106) or the C-terminus (amino acids 109-505) of ATF2 although binding to these truncated proteins was decreased in comparison with the binding to full length ATF2. This conflicts with the results of Liu and Green (1994) who found that E1a could only interact with the C-terminus of ATF2, mapping the site of interaction to amino acids 350-415 which comprises the bzip DNA binding domain of ATF2. Furthermore, my results also conflict with other binding data obtained in our laboratory using GST-ATF2 and E1a purified from baculovirus infected insect cells. These studies found, in agreement with Liu and Green's data, that E1a could only interact with full length ATF2 or a C-terminal region of ATF2 but not the N-terminal portion (G. Patel unpubl. obs.). The discrepancy between results may be due to the different methods used to synthesize E1a. In my studies and also those of Liu and Green, E1a was synthesized in rabbit reticulocyte lysate but in the latter case translation was carried out in the presence of $ZnCl_2$ and β-mercaptoethanol and in the case of G.Patel's studies E1a was synthesized in insect cells. These different methods of production may mean that E1a has a slightly altered conformation which precludes its interaction with the amino terminus of ATF2 whilst not affecting its ability to interact with either full length or the carboxy terminal region of ATF2. Alternatively, the different binding buffers used in the various experiments may destabilize certain interactions such as binding to the amino terminus whilst not affecting others. This may arise due to the different nature of the forces involved in the interactions e.g. binding to one region may involve electrostatic forces whereas binding to another may rely on the hydrophobicity of the interacting surfaces.

My results do, however, agree with those of Abdel-Hafiz et al (1993) who find that E1a can interact with the N-terminal 200 amino acids of ATF2 (an interaction with the C-terminus of ATF2 was not tested). Although their studies used bacterially produced proteins and I was unable to see an interaction using such a system, this can be explained since the type of fusion protein used for purification in the two experiments differed; I used a GST-E1a fusion in contrast to their histidine tagged E1a. As the histidine tag is less bulky than GST (30 amino

acid leader as compared to 218 amino acids of GST) it can be predicted that it would not interfere with either the correct folding of E1a or the binding of E1a to ATF2 so explaining the discrepancy. Furthermore, the ATF2 used in their binding asay was not 100% pure and the observed interaction may therefore be stabilized by contaminating bacterial proteins although this is unlikely to be the case. It is interesting to note that in their study the zinc finger in the amino terminus was shown to be important for E1a binding since if the cysteine residue at position 32 is mutated to a serine full length ATF2 is no longer able to bind E1a. This mutation was also shown to lose the ability to bind Zn. Thus, the amino terminus of ATF2 and the zinc finger in particular appears to be important for the interaction of ATF2 with E1a.

My results are also in agreement with those of Chatton et al (1993) concerning E1a and ATFa1 interactions. They find that E1a can bind to either the amino (amino acids 1-263) or carboxy (amino acids 325-462) terminal region of ATFa1 and that binding to either section is as strong as binding to the full length protein. Although I find binding to similar regions in ATF2, I find that the binding is weaker than to the full length protein. This can be explained in two ways. Firstly, the sizes of the deletions in the two cases are not identical which may influence the strength of any interaction seen with E1a and also we are not looking at identical proteins. Secondly, in my studies in order to get equivalent amounts of proteins for binding the total amount of protein used in the case of full length ATF2 was greater than for the other forms due to the presence of breakdown products which still retained the GST moiety and thus were retained on the beads. If these breakdown products were still capable of binding E1a, as might be expected given that both the N- and C-terminal portions seem to be capable of binding E1a, then the increase in binding may merely reflect the presence of these degraded forms. However, what does seem to be clear from these studies is that ATFa1 and ATF2 seem to interact with E1a in a similar way.

The importance of the zinc finger of ATFa1 for interaction with E1a was also assessed (Chatton et al., 1993). Mutation of histidine 27 to an asparagine residue in the context of full length ATFa1 did not affect binding presumably since E1a can still bind to the C-terminus of ATFa1. This conclusion is supported by the fact that mutation of the zinc finger in conjunction with a deletion of the C-terminus results in a form of ATFa1 unable to bind E1a. Thus as with ATF2, the zinc finger of ATFa1 seems to be important for binding E1a.

I did not investigate binding of the various E1a forms to ATF2. However, GST batch binding assays performed by G.Patel did assess this aspect of the E1a interaction. All the proteins were prepared from baculovirus infected insect cells and it was found that 13s E1a interacted more strongly with GST-ATF2 than did the 12s form. This parallels the *in vivo* data which finds that it is the 13s form of E1a which activates ATF2. It is also in partial agreement with the results of Liu and Green (1994) who find that only 13s E1a is able to interact with ATF2 although no interaction with 12s E1a was observed in this case. Finally, it is also in agreement with the data on ATFa1 (Chatton et al., 1993) which has been shown to interact more strongly with the 13s form of E1a.

Therefore, in conclusion, both my studies and the published data have shown that ATF2 and also ATFa1 are able to interact with 13s E1a. Although the indications are that this interaction is direct, it is possible that stabilizing/mediator proteins present in the reticulocyte lysate used to express ATF2/E1a are also involved. In the case of the studies using either bacterially or insect cell produced ATF2/E1a these proteins were only partially purified again allowing the possibility of an indirect or stabilized E1a-ATF2 interaction. However, the regions of ATF2 involved in the interaction remains unclear. All the studies indicate an interaction (when tested) with the C-terminal region of ATF2 and in some cases this interaction was found to be exclusive to the C-terminus. Some, but not all, studies indicate an additional interaction with the N-terminus which requires the zinc finger region. Therefore, further work is necessary in order to clarify the situation. It may be possible to gain some insight by performing experiments other than physical binding studies e.g. one way to validate the interaction of E1a with the C-terminal region of ATF2 *in vivo* would be to perform a two hybrid assay in mammalian cells similar to the ones used in yeast and similar to the original two hybrid experiments used in mammalian cells to demonstrate that ATF2 could mediate an E1a response (Liu and Green, 1990). It is not possible to use yeast for such studies because work by A. Eccleston has shown that E1a does not activate ATF2 in *S. Cerevisiae* nor does E1a interact with ATF2 in such a two hybrid system.

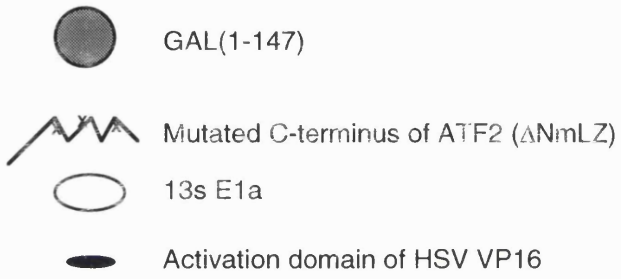
The outline of such an experiment is shown diagramatically in figure 5.9. GAL-ΔNmLZ is a GAL-ATF2 fusion lacking the amino terminus of ATF2 and containing a theoretical mutated leucine zipper. This mutation would prevent the ATF2 portion of the fusion from heterodimerizing with endogenous ATF2 in cells but would still allow homodimerization with the mutated zipper of a second GAL-ΔNmLZ molecule. This construct is transfected into CHO cells along with a GAL-CAT reporter construct in order to measure the level of transcription stimulated by the fusion. An E1a plasmid is also co-transfected but since it has been shown that GAL-ΔN is E1a unresponsive, E1a must be altered so that it can stimulate transcription via ATF2 even in the absence of ATF2's activation domain. One way to achieve this is to fuse the acidic activator VP16 of the herpes simplex virus to E1a. This activator when tethered to the promoter strongly activates transcription in an ATF2 independent manner. Thus if E1a does interact *in vivo* with the C-terminus of ATF2, E1a-VP16 should be brought to the promoter and so activate transcription.

The same experiment is not feasible with the N-terminus of ATF2 (GAL-Δ9) since this domain is E1a responsive and therefore any activation seen may be due to activation of this N-terminal region by E1a in an indirect manner (see below and section 6.4 (iii) for details) and not necessarily a result of transactivation by VP16 functioning at the promoter. However, if the activation domain of E1a was mutated then activation due to VP16 function could be examined. In this situation, only if E1a interacted with the amino-terminus of ATF2 would

Figure 5.9 Two Hybrid System To Investigate E1a-ATF2 Interactions *In Vivo*

Schematic diagram of a two hybrid system which could be used in mammalian cells to detect E1a and ATF2 interactions.

Key:



The crosses on the C-terminus of ATF2 indicate mutations designed to prevent heterodimerization with wild type ATF2 in the cell. These mutations allow homodimerization to occur to between mutated proteins.

A. GAL-ΔNmLZ transfected into CHO cells along with a GAL-CAT reporter does not stimulate CAT transcription

B. E1a-VP16 should not activate transcription of the minimal GAL-CAT reporter

C. Both GAL-ΔNmLZ and E1a-VP16 are transfected into CHO cells along with the reporter. If E1a interacts with the C-terminus of ATF2, E1a-VP16 will be brought to the promoter and VP16 will activate CAT transcription.

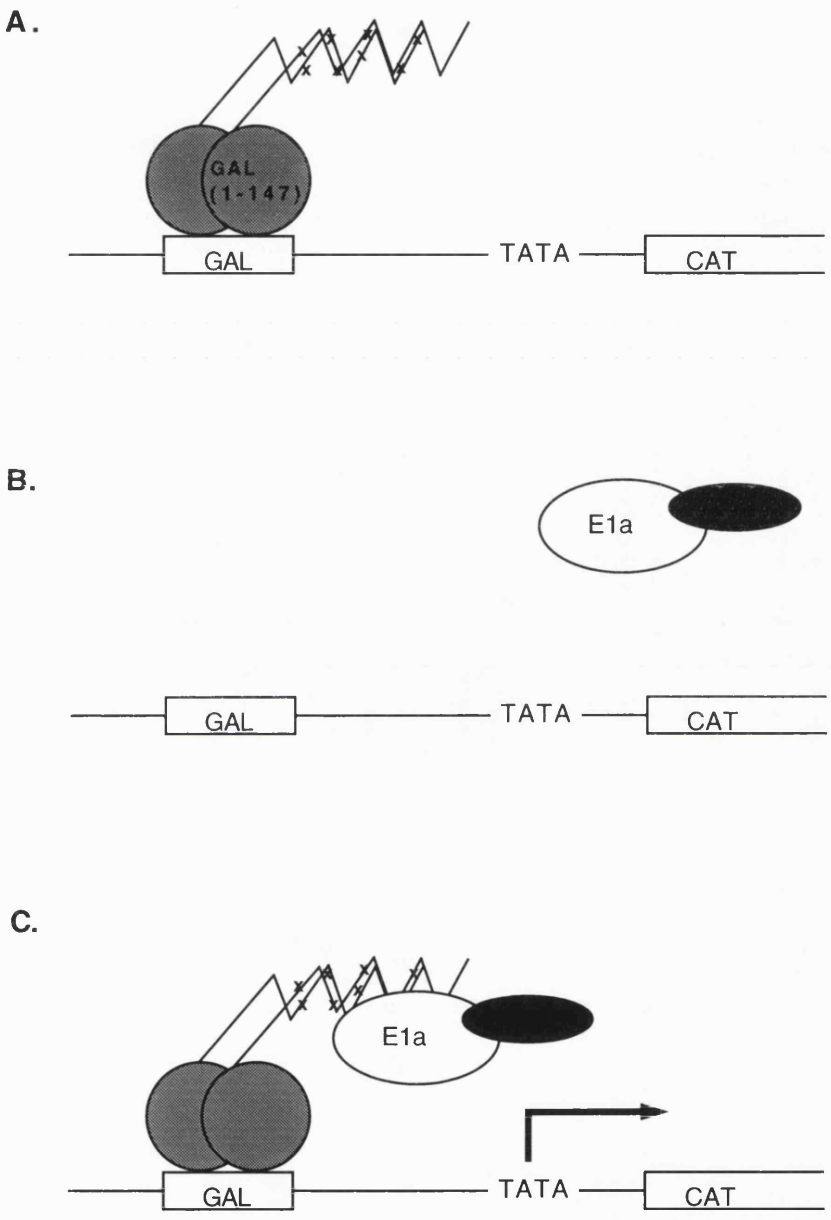


Figure 5.9 Two hybrid system to investigate E1a-ATF2 interactions *in vivo*

VP16 be brought to the promoter to stimulate transcription. If such an interaction did occur, the experiment could be repeated but using the N-terminus of ATF2 containing a mutated zinc finger to assess the contribution of this structure to the interaction.

In the above two hybrid experiments, one component of the system has been mutated; ATF2 in the case of the carboxy terminus and E1a in the case of the amino terminus. If these mutations are in residues critical for an interaction between ATF2 and E1a then the experiment will be invalid. However, with regard to the ATF2 mutation, E1a has been shown to interact with a variety of bzip and other DNA binding domains (Liu and Green, 1994) and so the sequence of the binding site in ATF2 does not appear to be critical. In the case of E1a mutations, the activation and targeting regions in CR3 of E1a have been delineated (Webster and Ricciardi, 1991) and so providing mutations are made only in the activation region no affect on binding should occur. The effect of the mutations in both proteins can be checked by *in vitro* binding assays using mutant versions of E1a and full length ATF2.

One final way of investigating the interaction between E1a and the N-terminus of ATF2 *in vivo* involves mutating the targeting region of CR3 whilst leaving the activation region wild type. If E1a functions by physically associating with the N-terminus of ATF2, then such a mutation should prevent activation. However, if E1a acts indirectly e.g. by altering the phosphorylation state of the N-terminus by stimulating certain kinases (see section 6.4 (iii)) then activation will still be observed. This experiment would thus cast light on the mechanism of E1a activation *in vivo* as well as addressing the regions of ATF2 involved in the interaction with E1a. It should be noted that these experiments do not answer the question as to whether the interaction is direct or not since cellular proteins are available for binding to E1a and ATF2. However, they do address whether the physical interactions determined so far *in vitro* are also relevant *in vivo* and also aid in understanding the role of the zinc finger in E1a activation.

CHAPTER 6:
DISCUSSION

DISCUSSION

The transcription factor ATF2 is a member of the ATF/CREB family of bzip proteins which bind to the consensus sequence TGACGTCA (Hai et al., 1989) and regulate transcription in response to a wide variety of stimuli. Attention was originally focused on ATF2 because it was the only family member found to be involved in the adenovirus 13s E1a stimulation of transcription via ATF/CRE sites (Flint and Jones, 1991). Other members have since also been shown to be E1a responsive (Chatton et al., 1993). I have studied ATF2 in greater detail and have found that it contains a cryptic transcriptional activation domain in its amino terminus. This domain is regulated by phosphorylation in response to serum and UV light and is essential for the E1a response by ATF2.

6.1 The Activation Domain Of ATF2

Amino acids 19-104 of ATF2 function as a strong transcriptional activator when fused to a heterologous DNA binding domain (section 3.2.1). This region does not possess any of the obvious characteristics typically ascribed to transcriptional activators i.e. a predominance of either glutamine, proline, isoleucine or acidic residues (Hahn, 1993) since the percentage composition of these respective residues is only 2.6%, 7.8%, 1.3% and 15.6%. Other examples of activation domains which do not fit into this classification have been described e.g. the HOB1 and HOB2 activation domains of Fos and c-Jun. These domains were identified by their homology and potential to form a helical structure (Sutherland et al., 1992). In this regard it is interesting to note that the secondary structure predictions for the N-terminus of ATF2 also include a large amount of helical structure (data not shown). This agrees with the proposal that structure rather than primary sequence is a key feature of activation domains, section 1.1.4 (ii).

6.2 Mechanism of Transcriptional Activation by C2

Transcriptional activation by the N-terminus of ATF2 (C2) was shown by mutational analysis to be dependent upon two critical threonine residues at positions 69 and 71 and to be somewhat less dependent upon a serine at position 90 (section 4.2.2). Each of these amino acids has been shown to be phosphorylated *in vitro* and one or more are known to be phosphorylated *in vivo*. This suggests that it is this post-translational modification which is important for the activity of C2. A second feature, the zinc finger motif between residues 25 and 49, is also required for transcriptional activation but like the serine residue is not essential for activation (section 4.2.1).

How do these features enable C2 to stimulate transcription? It is highly likely that the TPTP motif is a point of contact between C2 and a component of the transcription machinery such as a general factor, TAF or coactivator and that this interaction requires phosphorylation of the threonine residues in order to either take place or to activate transcription i.e. for the interaction to be productive. It also seems reasonable to assume that the phosphorylated

serine assists this interaction either by cooperating with the TPTP motif in binding, by influencing the conformation of the interacting region or by binding to a separate site on the target protein. The potential role of the zinc finger in this situation may be more complicated. It may be that the zinc finger, as proposed for the serine, influences the efficiency of the TPTP interaction with the transcription machinery e.g. by affecting the conformation of the activation domain or by affecting the activity of the kinases responsible for phosphorylating the TPTP motif (see section 6.3). Alternatively, the zinc finger may function independently of this motif and target a component of the transcription machinery itself. This interaction cannot substitute for the function of the TPTP motif but rather functions alongside it. These possible models can only be tested when the transcriptional target(s) of C2 are known.

6.3 Regulation of C2's Activity By Phosphorylation

The three phosphorylation sites described in the previous section are potential substrates for the proline-directed MAPKs and I have shown that serum and UV light which stimulate two of the MAPK subfamilies, the Erks and SAPKs, are able to potentiate C2's activity via these sites (section 4.2.3). However, due to the overlap known to exist between the serum and UV stimulation pathways e.g. both involve Ras, the actual kinase (Erk or SAPK) responsible for the regulation *in vivo* remains unclear. Therefore, *in vitro* studies have been used in the laboratory to try and clarify the situation.

I have shown that partially purified Erk1/2 is capable of phosphorylating C2 at the three identified sites *in vitro* (section 4.2.5). However, the level of C2 phosphorylation obtained was low in comparison to that of a known Erk substrate, Elk. Similarly, recombinant Erk2, activated by MEK immunoprecipitated from TPA treated cells, is also able to phosphorylate C2 at the three sites although again the level of phosphorylation was found to be poor. On the other hand SAPKs, which had been immunopurified from UV treated cells, were found to phosphorylate C2 efficiently since the level of C2's phosphorylation was only 2-3 fold lower than for the known SAPK substrate cJun. Phosphorylation again occurred at the three T/S P sites. In addition, C2 was found to bind and be phosphorylated by a kinase present in extracts from UV treated cells. The C2-bound kinase was also capable of phosphorylating cJun. When the kinase binding site in C2 was mapped, amino acids 48-65 of ATF2 were found to be essential. This implies that the kinase binds to a region wholly or partially contained between these residues, a region distinct from the sites of phosphorylation in C2. These are properties characteristic of SAPKγ (JNK1); it is activated by UV light and has been shown to bind to a region of cJun separate from that which it phosphorylates (Derijard et al., 1994). This suggests that the kinase bound by C2 is a SAPK. Therefore it appears that C2, at least *in vitro*, is preferentially phosphorylated by the SAPKs (G. Patel unpubl.).

The proposed SAPK binding site identified in C2 is known to be essential for the transcriptional activity of ATF2's N-terminal activation domain (construct Δ34 section 3.2.1). This therefore indicates a link between the docking of the kinase to ATF2 and ATF2 activation

and thus implies that the observed SAPK phosphorylation of ATF2 *in vitro* is physiologically relevant. This is also suggested by a transfection experiment performed in the laboratory using constructs known to specifically activate one of the MAPK pathways (as discussed in section 4.3). In this experiment it was demonstrated that activated Raf, which specifically stimulates Erk as opposed to SAPK activity (Minden et al., 1994a), does not activate ATF2. However, overexpression of MEKK1, which stimulates both Erk and SAPK activity, leads to ATF2 activation (A. Clarke, unpubl. obs.). Therefore, both the *in vitro* and *in vivo* data suggest that ATF2 is primarily phosphorylated by the SAPKs.

This result has important consequences for the role of the zinc finger in activation by C2. Serum stimulation experiments suggested that the zinc finger did not influence phosphorylation and thus activity of C2 (section 4.2.4). However, these experiments assessed its role with respect to stimulation predominantly by the Erks, not the SAPKs. It may be that in the latter case, the zinc finger is able to affect kinase function by aiding kinase binding. This is feasible because the SAPK binding site so far defined in C2 (residues 48-65) is known to overlap with the zinc finger motif (residues 25-49). It may also be that the zinc finger directly contributes to the binding site although this has not yet been determined. However, in either situation mutating the finger will alter the conformation of the binding site so preventing kinase binding and hence phosphorylation. Conformation is assumed to be an important determinant of the SAPK site because there is no obvious sequence homology between the kinase binding sites of ATF2 and cJun although this may reflect the possibility that ATF2 and cJun bind different SAPK isoforms.

From this discussion, it is feasible that the role the zinc finger plays in transcriptional activation by the N-terminus of ATF2 is to ensure phosphorylation of the critical TPTP motif (section 6.2). This can be tested in two ways. Firstly, UV stimulation experiments can be performed using the zinc finger mutant of C2 (C2pmZn). If the zinc finger is indeed required for SAPK phosphorylation of C2, then the UV response of C2pmZn would be expected to be diminished compared to that of wild type. Secondly, it could be determined directly whether the zinc finger mutant could bind SAPK by performing *in vitro* binding assays as were originally used to demonstrate that SAPK bound and phosphorylated C2.

6.4 Mechanism of E1a Activation

Sections 5.2.1 and 5.2.2 demonstrated that in order for E1a to stimulate transcription via ATF2, both proteins must possess active activation domains. It also appears that the proposed interaction between ATF2 and E1a does not convert ATF2 into a transcriptional activator if the activation region of E1a is non-functional. This suggests that this region of E1a plays an active role in the conversion process. Several mechanisms for such activation have already been mentioned (section 5.3). These are discussed here in greater detail:

(i) Induction of a Conformational Change

Physical association of E1a with ATF2 by itself may be insufficient to cause the conformational change in ATF2 required to expose its activation domain and stimulate transcription. E1a may also have to contact the transcription machinery e.g. via its interaction with TBP/TFIID (Lee et al., 1991) in order to be in the correct orientation/conformation to influence ATF2 structure. The E1a mutants used in the activation experiments (section 5.2.2) are known to be defective in TBP binding (Geisberg et al., 1994) which could therefore explain why, according to this scheme, they no longer activate transcription via ATF2.

A conformational change during transcriptional activation has been demonstrated for the transcription factor TFIIB (Roberts and Green, 1994). In its native (uninduced) state, TFIIB is proposed to be in a closed conformation arising from the intramolecular association of its N- and C-terminal domains. This conformation inhibits TFIIF and RNAPII binding and thus prevents initiation complex assembly. Binding of the acidic activator VP16 to TFIIB is thought to disrupt the intramolecular interaction so exposing the TFIIF and RNAPII binding sites and thereby encouraging transcription initiation. Evidence for a VP16 induced conformational change in TFIIB was obtained by performing V8 partial proteolysis on TFIIB. Addition of VP16 to the assay resulted in the appearance of a novel cleavage fragment, indicative of a conformational change. This fragment was not induced by a transcriptionally inactive VP16 mutant which has a decreased affinity for TFIIB. This implies that the structural change in TFIIB only occurs with transcriptionally active VP16, as proposed for the E1a induced conformational change of ATF2.

With regard to this mechanism of TFIIB activation, it is interesting to note that an interaction between the N- and C-terminus of ATF2 has previously been inferred based on the results of peptide competition studies (Abdel-Hafiz et al., 1993). Although this interaction has not been directly demonstrated, it does indicate that the proteolysis experiments used to study the VP16 induced TFIIB conformational changes may be useful to investigate proposed E1a induced alterations in ATF2 structure, especially now that the E1a-ATF2 binding conditions are known.

(ii) Overcoming Inhibition

Unmasking the activation domain of ATF2 in the context of the full length protein may not convert ATF2 into a transcriptional activator because of the influence of the proposed inhibitory sequences within ATF2. Therefore, E1a may play a role in overcoming this repression during its activation of ATF2. How E1a achieves this presumably reflects the nature of the repression involved. For example, repression could result from the binding of an inhibitor to ATF2 which prevents the activation domain of ATF2 from contacting its transcriptional target. In this case, E1a could act by displacing the repressor. Alternatively, repression may be caused by the inhibitory region of ATF2 binding to and blocking access of ATF2's activation domain to the transcription machinery. E1a could then function by

disrupting this non-productive interaction. In both cases E1a indirectly facilitates an interaction of ATF2's activation domain with its target in the transcription machinery so enabling ATF2 to stimulate transcription. E1a may also directly facilitate the interaction by influencing the conformation and accessibility of the basal machinery to ATF2 via its interaction with TBP.

(iii) Stimulating Phosphorylation

E1a may also stimulate transcription via ATF2 by increasing phosphorylation of ATF2. As discussed in section 6.3 phosphorylation of the TPTP motif within ATF2's activation domain is critical for its ability to transactivate. Therefore, it is feasible that E1a could stimulate the kinases responsible for phosphorylating these residues and so increase the activation potential of ATF2. Evidence for this proposal is given by the fact that in E1a-transformed cells cJun is found to be hyperphosphorylated (Hagmeyer et al., 1993). The sites of increased phosphorylation were mapped to ser-63 and -73, residues known to be phosphorylated by the SAPKs (the kinases thought to regulate ATF2). As with ATF2, these phosphorylated serines are required for the transactivation function of cJun (Smeal et al., 1991). Thus E1a appears to stimulate cJun and ATF2 by similar mechanisms. However, there is one important difference in the mechanisms i.e. cJun can be activated by 12s E1a (Hagmeyer et al., 1993) whereas ATF2 activation requires 13s E1a (section 5.2.2). This suggests that phosphorylation of ATF2 by itself is not sufficient for activation but requires other functions of E1a provided by CR3. This is also suggested by phosphorylation studies of ATF2. Full length ATF2 has been shown to be phosphorylated weakly by the Erks and efficiently by the SAPKs *in vitro* (G. Patel unpubl.) implying that the relevant phosphorylation sites are accessible to the kinases in the full length protein. However, *in vivo*, stimuli which lead to an increase in activity of these kinases do not produce an increase in ATF2 activity (section 4.2.3).

Stimulation of cellular kinase activity has been shown to be a function of another viral activator namely HBV pX. This protein appears to activate signal transduction via Ras and Raf leading to transcriptional activation of cJun (Natoli et al., 1994). The activation is dependent on the phosphorylation sites, ser-63 and ser-73, in cJun. pX has been shown to stimulate Erk1 and 2. However, this function of pX is probably not involved in activation of cJun but rather reflects the crossover thought to exist between the Erk and SAPK pathways. Whether pX stimulates the SAPKs as well as the Erks has not been reported. pX is proposed to stimulate these kinase pathways by acting on cytoplasmic factors that regulate or activate Ras-GTP complex formation and it is possible that E1a functions in a similar manner. However, it is equally likely that E1a increases phosphorylation of cJun and ATF2 by inhibiting the action of a phosphatase. However, no evidence for such a mechanism exists at present.

The above three proposals all entail the cooperation of two activation domains in order to stimulate transcription i.e. E1a and ATF2 appear to synergize with each other during transactivation. ATF2 needs to cooperate with a second activator because in the context of the full length protein its activation domain is inert (section 3.2.1). Similarly, it has been shown that full length E1a is a very poor activator despite the presence of a powerful transactivation domain within CR3 (Liu and Green, 1994). This suggests that inhibitory sequences in E1a act upon this domain to prevent it functioning.

Based on all these observations, the following mechanism of cooperation could be imagined to occur between E1a and ATF2. Firstly, interaction of E1a with ATF2 would be predicted to overcome the effect of the negative acting sequences in E1a by either directly blocking their actions or by altering the conformation of E1a and thereby destroying their capacity to repress. This elimination of repression enables E1a CR3 to interact with the transcription machinery. However, this does not result in the efficient stimulation of transcription in contrast to the situation with GAL-CR3 which strongly activates transcription (Lillie and Green, 1989). The discrepancy can be explained by the different nature of CR3 in the two experiments; in the ATF2 experiment CR3 is present as a monomer but as a GAL fusion it is dimeric. CR3 is therefore thought to act only in conjunction with a second activator (CR3 in the case of the GAL fusion). Since the activation domain of ATF2 is not functional at this stage, cooperation cannot occur and hence there is proposed to be little or no stimulation of transcription. However, the interaction of CR3 with the transcription machinery changes this by converting the N-terminal domain of ATF2 from a latent to an active state by the mechanisms described above. Therefore, there are now two fully functional activation domains and their combined actions on the transcription machinery results in the stimulation of transcription. According to this scheme ATF2 plays an active role during E1a transactivation and is no longer visualized as the passive player originally proposed in section 1.3.2. The main details of this mechanism are outlined in figure 6.1.

This mechanism provides an alternative explanation for the E1a mutants proposed by Geisberg et al (1994) to be defective in binding a limiting cellular factor (section 1.3.2). Such mutants are capable of interacting with TBP and ATF2 but do not activate transcription. Geisberg et al suggested that this was due to failure to bind e.g. a TAF or coactivator necessary for activated transcription. However, it is also possible that these residues are critical for E1a to convert ATF2 into an activator e.g. these residues may be involved in the stimulation of kinase activity. Mutation of these residues would therefore prevent cooperation between ATF2 and E1a and concomitantly prevent transactivation. A combination of these two mechanisms is also feasible in that binding of a limiting cellular factor may be a prerequisite for E1a to activate ATF2 e.g. E1a may have to interact with a TAF as well as TBP in order to influence ATF2 structure or repression.

The mechanism also explains why E1a fails to transactivate other members of the ATF/CREB family with which it can physically interact (section 1.3.2). These proteins are

Figure 6.1 Revised model of E1a activation via ATF2

The interaction of E1a with ATF2 bound at the promoter is proposed to cause a conformational change in E1a which enables it to interact with the basic machinery. E1a in turn stimulates three changes in ATF2:

(i) a conformational change

(ii) repression of inhibition

(iii) phosphorylation of the activation domain

These changes activate ATF2 which can therefore stimulate transcription.

Key:

Proposed repression domain in ATF2

In this model ATF2 is no longer seen as a passive player during E1a transactivation which contrasts to the original model presented in figure 1.3.

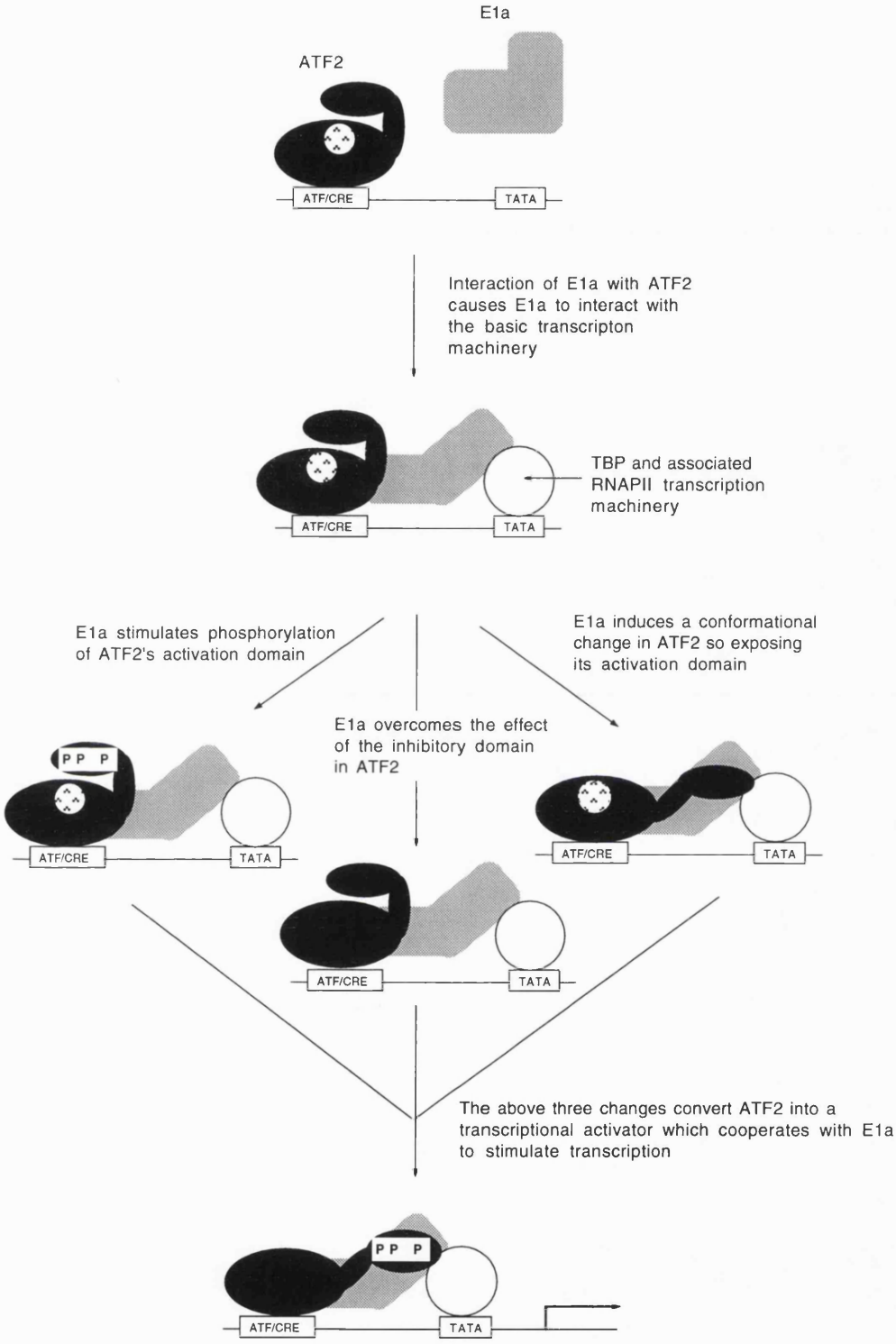


Figure 6.1 Revised model of E1a activation

presumably not able to either overcome the inhibitory sequences in E1a or to synergize with E1a by the mechanisms outlined above i.e. they can be regarded as passive players whereas E1a requires an active partner in order to stimulate transcription via ATF/CRE sites (see also section 6.6).

6.5 Synergy

The above mechanism details how E1a and ATF2 synergize with each other to activate transcription. Synergy is a common feature of transcriptional activators and well documented examples of activation domains which need to cooperate include the HOB motifs of Fos and Jun (Sutherland et al., 1992) and the glutamine (Q) and proline (P) rich N- and C-terminal domains of Oct2 (Tanaka et al., 1994; Tanaka and Herr, 1994). Each of these domains on its own when fused to a heterologous DNA binding domain (DBD) is incapable of activating transcription to any great extent. However, when these motifs are duplicated either with themselves or with a corresponding activation domain e.g. P-POU DBD-P or P-POU DBD-Q; GAL DBD-HOB1HOB1 or GAL DBD-HOB1HOB2, strong transactivation results. Furthermore, these motifs have been found to cooperate with like domains from other transcription factors e.g. CCAAT/enhancer binding protein (C/EBP) contains a HOB2 motif which can synergize with the Jun HOB1 motif and the Sp1 glutamine-rich domain B can cooperate with the Oct2 C-terminal region. This gives credence to the proposal that CR3 of E1a can cooperate with itself and also with the activation domain of ATF2. Whether the activation domain of ATF2 can cooperate with itself has not yet been assessed. Thus further work is required to determine the precise interplay between these activation domains. For example, it needs to be confirmed that monomeric CR3 is inactive in the absence of cooperation with other transcription factors. This can be achieved by making a Myb DBD-CR3 expression construct and measuring the level of transcription from a reporter plasmid bearing Myb binding sites upstream of the CAT gene since the Myb DBD binds DNA as a monomer (Frampton et al., 1991). This strategy can also be applied to investigate synergy involving the activation domain of ATF2. For example it can be tested whether a Myb DBD-C2 fusion (monomeric) is less active than a GAL DBD-C2 fusion (dimeric). It can also be determined whether E1a cooperates with the monomeric as well as the dimeric physiologically relevant version of ATF2. A second approach to studying cooperation is to vary the number of binding sites present in the reporter construct. This assesses whether cooperation can take place in trans i.e. between activation domains of different proteins e.g. does GAL-C2 activate transcription 2 fold higher on a $(GAL)_2$-CAT reporter than a $(GAL)_1$-CAT reporter or does synergism occur so resulting in transcription levels greater than 2 fold. Finally, if repression sequences are truly identified in either E1a or ATF2 their effect on both activation and synergism need to be determined using similar approaches e.g. fusions containing different combinations of activators and repressors, reporters which bind both activators and repressors. In this way the requirements of activators needed to overcome repression may be determined.

A mechanism by which identical activators can synergize with each other to stimulate transcription has been suggested by work carried out on the TAFs. These studies have shown that a single activation domain may interact with multiple TAFs e.g. NTF-1 interacts with $TAF_{II}150$ and $TAF_{II}60$ (Chen et al., 1994). Thus it is possible that this domain stimulates transcription by two separate pathways involved in initiation complex formation, each pathway being linked to one of the TAFs contacted by the activation domain. Only when both pathways are activated is transcription efficiently stimulated and this only occurs when multiple activation domains are available for interaction. The same mechanism presumably applies for synergism between different activators i.e. the activators contact their respective TAF targets so triggering distinct pathways involved in initiation complex formation resulting in high levels of transcription. However, cooperation does not occur with all activators suggesting that either some activators stimulate the same pathways or that the pathways stimulated give rise to the same effect rather than two stimulatory effects. The direct mechanisms of synergism given here will most likely contribute to the cooperation proposed to exist between E1a and ATF2 during transactivation in addition to the more indirect mechanisms described in section 6.4.

6.6 ATF Family Members and E1a responsiveness

Several ATF2 related proteins have been identified; ATFa1, 2 and 3 (Chatton et al., 1993) and CRE-BPaα, β, γ and δ (Zu et al., 1993). These proteins have several regions of extremely high homology when compared with the ATF2 sequence as shown in figure 6.2. It is interesting to note that the N-terminus of the ATFa proteins and CRE-BPaα and β is conserved (this region is partially or wholly absent in the CRE-BPa splice variants γ and δ) because this is the domain necessary and sufficient for E1a transactivation of ATF2. It was therefore expected that all these proteins would be E1a responsive. However, when CRE-BPaα and the ATFa proteins were tested for their ability to mediate E1a transactivation, only the latter were found to be functional (Chatton et al., 1993; Nomura et al., 1993). Since all three of the ATFa species were equally responsive, ATFa1 was used as the model to investigate this family's response further.

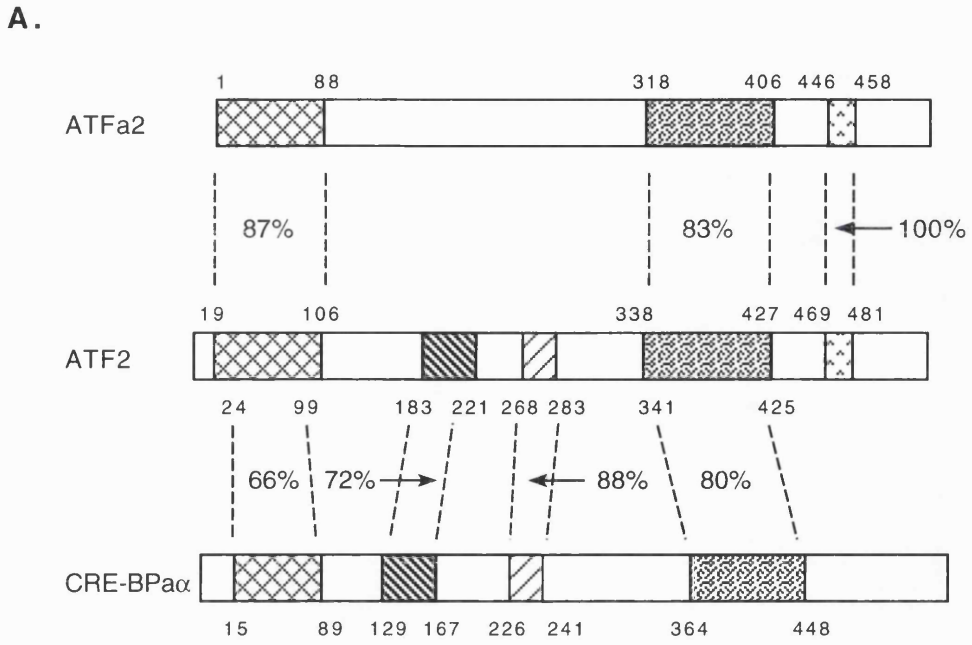
It was found that the region of ATFa1 sufficient for E1a transactivation mapped to its conserved N-terminus, amino acids 2-82 (equivalent to amino acids 21-100 of ATF2). In addition, the zinc finger located in this domain was found to be critical for ATFa1 to respond to E1a. This mirrors the situation found for ATF2 and suggests that E1a stimulates transcription via ATFa1 by the same mechanisms as via ATF2. This is also suggested by the fact that the N-terminus of ATFa1, like ATF2, contains an activation domain which is inactive (masked) in the full length protein (Chatton et al., 1994). The activity of this domain is dependent on the integrity of its zinc finger and, although its importance has not been assessed, a TPTP phosphorylation motif is also present. Therefore, E1a may well activate ATFa1 by stimulating the unmasking of ATFa1's activation domain and by augmenting its

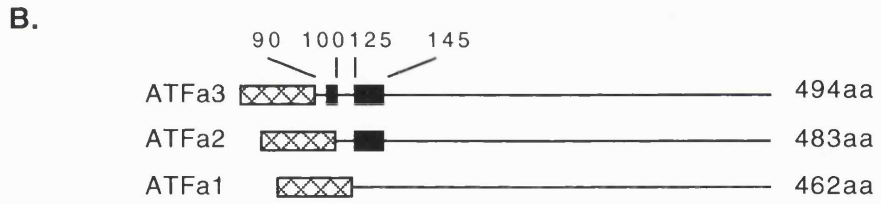
Figure 6.2 Homolgy between ATF2-like proteins (from Chatton et al., 1993; Zu et al., 1993)

A. The regions of homology between ATFa2, ATF2 and CRE-BPaα are depicted by the shaded boxes. The degree of homology compared to ATF2 in each region are shown between the proteins.
Numbers above and below the proteins correspond to the amino acid positions.

B. Structure of the ATFa1, 2 and 3 proteins. The black boxes correspond to the extra amino acid sequences present in ATFa2 and 3. All three forms possess the N-terminal ATF2 homology domain as indicated by the hatched boxes.

C. Structure of the CRE-BPa subfamily. Only CRE-BPaα and β contain the complete N-terminal ATF2 homology domain as indicated by the hatched box.





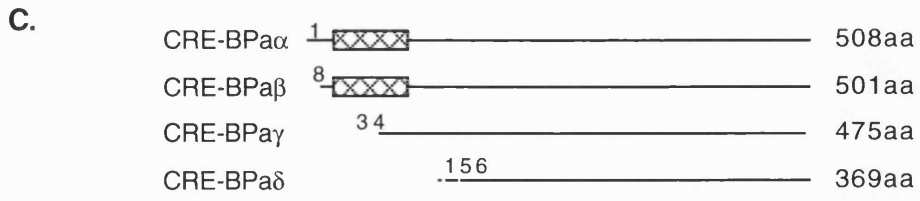


Figure 6.2 Homology between ATF2-like proteins

activation potential by increasing the degree of phosphorylation at the threonine residues. If this is the case, why does E1a not activate CRE-BPaα since this protein also contains a zinc finger as well as the critical TPTP phosphorylation motif within its N-terminus?

The answer appears to lie with the mechanism of TPTP phosphorylation. As discussed in section 6.3 the SAPKs appear to be responsible for phosphorylating this motif and these kinases are known to bind to a site in ATF2 comprising some/all of the region between residues 48-65. Comparison of the sequences of ATF2, ATFa1 and CRE-BPaα in this region (figure 6.3) reveals extremely high conservation between ATF2 and ATFa1 (89% conservation) whereas CRE-BPaα is poorly conserved (50%) particularly in the area amino acids 48-57 (only 18% conservation) suggesting the possibility that CRE-BPaα is unable to bind and thus be phosphorylated by the SAPKs. This means that even upon stimulation of these kinases by E1a, there would be no effect on the level of CRE-BPaα phosphorylation and no transactivation. Since there appears to be a link between phosphorylation and E1a activation, the possibility that CRE-BPaα does not bind SAPK could therefore explain its inability to respond to E1a.

I have suggested that the zinc finger of ATF2 also contributes to SAPK binding (section 6.3). The sequence of this region in CRE-BPaα is much less homologous (69% conserved) than the corresponding sequence of ATFa1 (93% conserved) when compared to ATF2. Furthermore, when the zinc finger of ATF2 was replaced with that of CRE-BPaα, the resultant mutant was unable to be activated by E1a (section 4.3). These data together suggest that the lack of sequence conservation in the zinc finger may contribute to CRE-BPaα's lack of E1a inducibility by again preventing SAPK binding. However, it is also possible that the zinc finger functions independently of SAPK binding during transactivation (section 6.2) e.g. it may function by targeting a coactivator, a role which is presumably required for E1a induction. In this situation, the less conserved zinc finger of CRE-BPaα is presumably unable to target such a coactivator and thus is unable to respond to E1a. This latter proposal can only be addressed when the transcriptional target(s) of ATF2 are known (see section 3.3 for experiments which identify such targets).

The proposed contribution of SAPK binding towards E1a inducibility, however, can readily be tested by determining whether CRE-BPaα is capable of binding the SAPKs or not. If it is incapable of binding these kinases then chimeras between itself and ATF2 can be constructed to assess the contribution of the zinc finger in binding. Such constructs can also be used to precisely define the SAPK binding site within ATF2. It is highly likely that CRE-BPaα is phosphorylated by the Erks on the TPTP motif because the activity of CRE-BPaα is increased by the phorbol ester TPA, a known stimulus of these kinases (Chatton et al., 1994). If this is the case, then it would appear that E1a is incapable of activating this subfamily of MAPKs and may provide some insight/tools as to the mechanism of kinase stimulation by E1a. This ATF family member may therefore prove to be a useful tool with which to study the mechanism of E1a transactivation.

Figure 6.3 Sequence comparison of ATF2, ATFa1/2/3 and CRE-BPaα

The amino acid sequence of the N-terminal homology regions of ATF2, ATFa1/2/3 and CRE-BPaα are shown. Amino acids which differ from those of ATF2 are highlighted with asterisks. The residues highlighted in the ATF2 sequence are those cys and his residues which compose the zinc finger motif. Numbers indicate amino acid position.

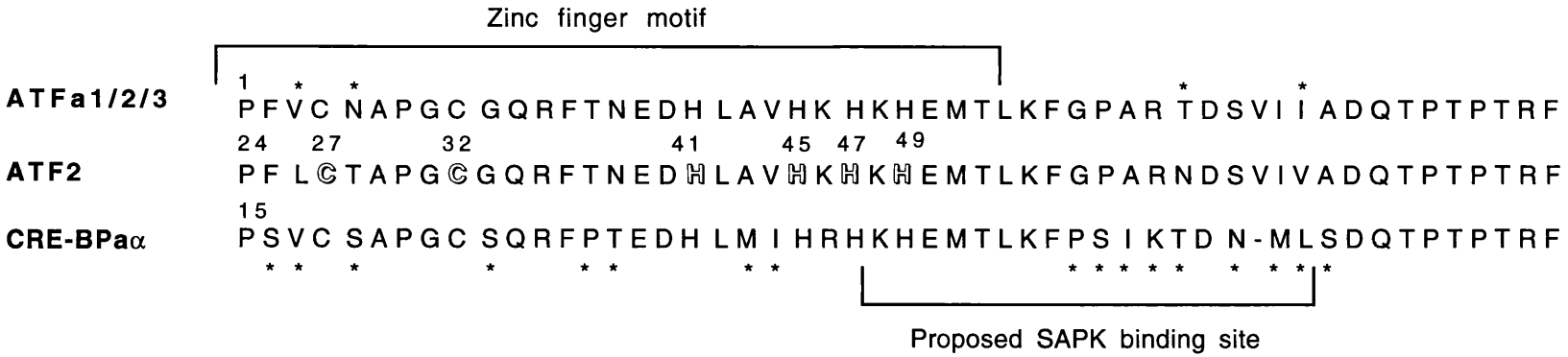


Figure 6.3 Sequence comparison of ATF2, ATFa1/2/3 and CRE-BPaα

6.7 ATF2 and Signal Transduction

The physiological role of only one CREB/ATF family member has been well established i.e. the regulation of gene expression by CREB in response to extracellular stimuli (section 1.2.2). Roles for the other family members have remained obscure. However, with the identification and cloning of the MAPK and related signalling pathways, it has become apparent that the CREB/ATF family is more generally involved in signal transduction than originally thought. In this thesis, I have shown that ATF2 contains a transcriptional activation domain whose activity is stimulated in response to UV light and serum via the action of the SAPKs and Erks. Thus ATF2 appears to be a regulator of gene expression via ATF/CRE sites in response to the MAPK or MAPK-related pathways. Not all promoters containing ATF/CREs are ATF2 responsive due to the regulatory mechanisms described in sections 1.2.1-1.2.5 which control ATF/CREB family activity. However, several promoters have been identified which appear to be influenced by ATF2. These promoters are described below along with the signalling pathways thought to converge on ATF2 at these promoters:

(i) cJun Promoter

cJun transcription is induced by the cytokines interleukin-1 (IL-1) (Muegge et al., 1993) and tumour necrosis factor α (TNFα) (Westwick et al., 1994), the alkylating agents N-methyl-N'-nitro-N-nitroso-guanidine (MNNG) and methylmethane-sulphonate (MMS), and short wave ultraviolet light (UV-C) (van Dam et al., 1995). These signals mediate their effects through two ATF-like AP-1 sites in the cJun promoter, jun1 (TGACATCA) at position -71 and jun2 (TTACCTCA) at position -191. These sites have been shown to bind cJun/cJun homodimers and cJun/ATF2 heterodimers in cells containing these transcription factors (van Dam et al., 1993) and to bind ATF2/ATF2 homodimers in cells lacking cJun (van Dam et al., 1995). These sites are constantly occupied, even under conditions of low promoter activity, which suggests that regulation of transcription via these sites arises by post-translational modification of the bound ATF2 and cJun (Herr et al., 1994). Consistent with this proposal, the transactivation potential of cJun has been shown to be increased by phosphorylation of its activation domain in response to UV-C (Hibi et al., 1993). I have shown that ATF2 is similarly regulated (chapter 4) and therefore also presumably participates in the regulation of cJun expression. This has been confirmed by demonstrating that in cells lacking cJun, UV-C is still able to stimulate transcription controlled by jun1 and 2. In fact, ATF2 may be the major regulator of UV-C induced cJun transcription since this transcription factor is more strongly activated than cJun by this stimulus (van Dam et al., 1995).

The other known stimuli of cJun transcription also seem to act by increasing the transactivation potential of cJun and ATF2 via post-translational modification since IL-1 (Sluss et al., 1994), TNFα (Westwick et al., 1994), MNNG and MMS (van Dam et al., 1995), like UV-C, all activate the SAPKs which are responsible for phosphorylating ATF2 and cJun. It is possible that each signal preferentially stimulates a different SAPK isoform or related kinase

whilst only weakly stimulating the other isoforms/kinases e.g. IL-1 may stimulate the MAPK homolog, RK (Freshney et al., 1994), to a much greater extent than it stimulates any of the SAPKs. This in turn would lead to the differential activation of ATF2 and cJun, provided that the various kinases have differing activities towards these substrates, and thus could provide a mechanism for precisely regulating cJun transcription. Whether such a complex control mechanism operates can only be addressed once all the relevant kinases have been cloned and their properties established.

(ii) E-selectin Promoter

E-selectin (ELAM-1) is a member of the selectin family of endothelial cell adhesion proteins which recognize carbohydrate ligands on circulating immune cells. Its expression is both cell specific and inducible since it is expressed only on endothelial cells in response to induction by the cytokines IL-1 and TNFα. Induction is controlled at the level of transcription and several sites within the E-selectin promoter are known to be involved; an ATF/CRE site, three NF-κB sites and an HMG I(Y) site (Whitley et al., 1994). The CRE has the same sequence as the jun1 regulatory site i.e. TGACATCA and *in vitro* binds ATF2, ATFa and ATF3 homodimers and ATF2/cJun, ATFa/cJun and ATF3/cJun heterodimers specifically and with high affinity. Furthermore, ATFa or an ATFa-like protein has been shown to be part of the complex bound to the site *in vivo* (Kaszubska et al., 1993). Given the homology of ATFa with ATF2 and the sequence of the CRE, it is highly likely that ATF2 and hence cJun are also bound to this site *in vivo*. Thus similar, if not identical, factors appear to regulate transcription via the CRE of the E-selectin promoter as are known to regulate transcription via the jun1 site of the cJun promoter.

The mechanism of regulation via these two promoter elements may also be identical because (i) the same signals induce transcription via these sites and (ii) post-translational modification of bound transcription factors appears to be involved in both cases (van Huijsduijnen et al., 1992; Herr et al., 1994). Therefore, cytokine induction of the E-selectin gene is likely to be mediated by kinases acting on ATF2, ATFa and cJun so increasing their transactivation potential. The involvement of ATFa in such a scheme is not unexpected since this factor is proposed to function in an analogous manner to ATF2 (section 6.6). If this is the case, it is feasible that ATFa also plays a role in regulating cJun transcription.

This regulatory mechanism, although necessary for cytokine induction of E-selectin transcription, is not sufficient since multimers of the E-selectin CRE are unable to confer cytokine inducibility upon a heterologous promoter (van Huijsduijnen et al., 1992). It has been shown that NF-κB is also required because mutations which decrease binding *in vitro* of NF-κB to any of the NF-κB sites in the E-selectin promoter abolish cytokine-induced E-selectin gene expression *in vivo* (Whitley et al., 1994). NF-κB has been shown to require the CRE site for its action and furthermore it has been shown to interact directly with ATF2 and ATFa *in vitro* (Kaszubska et al., 1993). Thus one way in which NF-κB may act to induce E-selectin

expression is by activating ATF2 via direct physical association e.g. the interaction may stimulate the unmasking of ATF2's activation domain by inducing a conformational change in the protein as suggested for ATF2 activation by 13s E1a (section 6.4). Since E1a and NF-κB both bind to the bzip region of ATF2 (Kaszubska et al., 1993; Liu and Green, 1994), it is likely that these proteins produce the same or similar structural changes in ATF2 which lead to its activation. In the case of NF-κB, however, activation will only occur upon cytokine induction because only then is NF-κB released from its association with IκB in the cytoplasm and is therefore free to interact with ATF2 in the nucleus (Liou and Baltimore, 1993; Baeuerle and Henkel, 1994).

The requirement for NF-κB suggests that post-translational modification of ATF2 alone is insufficient for its activation as discussed earlier (section 6.4 (iii)). Thus cooperation with other transcription factors is necessary. However, it appears that this is not true in all situations. This is evidenced by comparing transactivation via the E-selectin CRE and jun1 with transactivation via jun2. The first two regulatory elements have the same sequence and are essential for induction of their promoters but when multimerized in front of a heterologous promoter are unable, or are only very weakly able, to confer a transcriptional response to known regulatory signals e.g. cytokines for the E-selectin CRE (van Huijsduijnen et al., 1992) or UV irradiation/12s E1a for jun1 (van Dam et al., 1993; van Dam et al., 1995). Thus jun1, like the E-selectin CRE, appears to function only in conjunction with other regulatory elements in the promoter. Jun2, on the other hand, is able to bestow UV responsiveness onto a heterologous promoter (even in cells lacking cJun) and thus does not need to cooperate with additional regulatory elements (van Dam et al., 1995). This suggests that post-translational modification of ATF2 in this instance is sufficient for activation. However, it may be that the binding of ATF2 to this sequence induces a conformational change in ATF2 which obviates the requirement for cooperation; binding may induce the exposure of ATF2's activation domain. Accordingly, this conformational change is not produced upon binding to jun1 or to the E-selectin CRE. These sequences differ from jun2 by only two nucleotides (TGACATCA versus TTACCTCA) but presumably the contacts these bases make with ATF2 are of sufficient importance as to be able to influence its structure. A similar scheme has been proposed for regulation of the transcription factor SRF (Hill et al., 1994); DNA binding in this case is thought to be associated with a conformational change resulting in the production of a serum regulated activity within SRF. To test whether ATF2 alters in conformation upon binding to specific CRE sequences, partial proteolysis experiments (see also section 6.4 (i)) could be performed on ATF2 bound to either jun1 or jun2. Comparison of the resultant cleavage patterns should indicate if the structure of ATF2 differs in the two cases.

The role of DNA binding in ATF2 activation could explain why GAL-ATF2 is not stimulated by serum as measured by the reporter, G5E4CAT (section 4.2.3). In such an experiment ATF2 binds to the reporter via the GAL DNA binding moiety, not its own, and therefore presumably does not adopt an active conformation at the promoter. Thus, even if its

activation domain is phosphorylated and hence activated in response to serum, its activity will remain masked and transcription will remain at basal levels. Since the reporter does not possess additional regulatory sites upstream of the CAT gene, cooperation between ATF2 and other transcription factors cannot compensate for the lack of ATF2 DNA binding. Such a reporter provides a system for investigating ATF2 cooperation with transcription factors; by sub-cloning different regulatory elements into the reporter the effect of various transcription factors on transactivation by GAL-ATF2 can be assessed.

It is difficult to study the role sequence specific DNA binding plays in ATF2 activation because this requires the use *in vivo* of reporters driven by CRE sequences and these sites may then be occupied/regulated by other ATF/CREB or related proteins. However, by performing transfections in e.g. F9 cells which lack detectable Jun and Fos protein, by overexpressing ATF2 in conjunction with a variety of CRE driven reporter constructs, it may be possible to precisely define the sequences required to stimulate the proposed unmasking of ATF2's activation domain. In addition, methylation interference assays and DNA footprinting may be used to identify nucleotides which contact ATF2.

A third protein, HMG I(Y), is also involved in the regulation of E-selectin expression because mutations which interfere with its binding to the E-selectin promoter decrease the level of cytokine-induced expression (Whitley et al., 1994). This reflects the role that HMG I(Y), a low molecular weight non-histone chromosomal protein, plays in transcriptional activation i.e. structural activation (Tjian and Maniatis, 1994 and section 1.1.5 (iv)). HMG I(Y) facilitates binding and interaction of transcription factors at promoters into stereospecific complexes which interact efficiently with and thus stimulate the basal transcription machinery. Since HMG I(Y) is known to interact directly with NF-κB and ATF2 (Du et al., 1993), it can thus facilitate cooperation between these proteins and thereby activate E-selectin transcription via these transcription factors.

(iii) β-Interferon (β-IFN) Promoter

The human β-IFN and E-selectin promoters share a highly similar arrangement of regulatory elements (Whitley et al., 1994) as shown in figure 6.4. Furthermore induction of β-IFN expression by viruses is mediated by the same proteins as required for the cytokine induction of the E-selectin gene i.e. ATF2, NF-κB and HMG I(Y) (Du et al., 1993). The mechanisms controlling ATF2 activity at these promoters are similar, if not identical. Thus cooperation between ATF2 and NF-κB and enhancement of this cooperation by HMG I(Y), as described above, takes place at the β-IFN promoter. In addition HMG I(Y) appears to activate ATF2 and this presumably occurs by a direct association generating the active conformation of ATF2. HMG I(Y) also stimulates the binding of ATF2 to the β-IFN promoter *in vitro*. However, the physiological importance of this stimulation not been addressed (it may be that *in vivo* ATF2 is constitutively bound to the β-IFN promoter).

Figure 6.4 Comparison of the human E-selectin and β-interferon promoters (Adapted from Whitley et al., 1994)

The human E-selectin and β-interferon promoters both contain four regulatory domains; PDI-IV in the E-selectin promoter and PRDI-IV in the β-interferon promoter. Transcription factors bound to these domains are:

PDII; ATF2 and HMG I(Y)

PDIV; NF-κB and HMG I(Y)

PDIII; " "

PDI; " "

PRDIV; ATF2 and HMG I(Y)

PRDIII; Interferon regulatory factor (IRF) and HMG I(Y)

PRDI; IRF

PRDII; NF-κB and HMG I(Y)

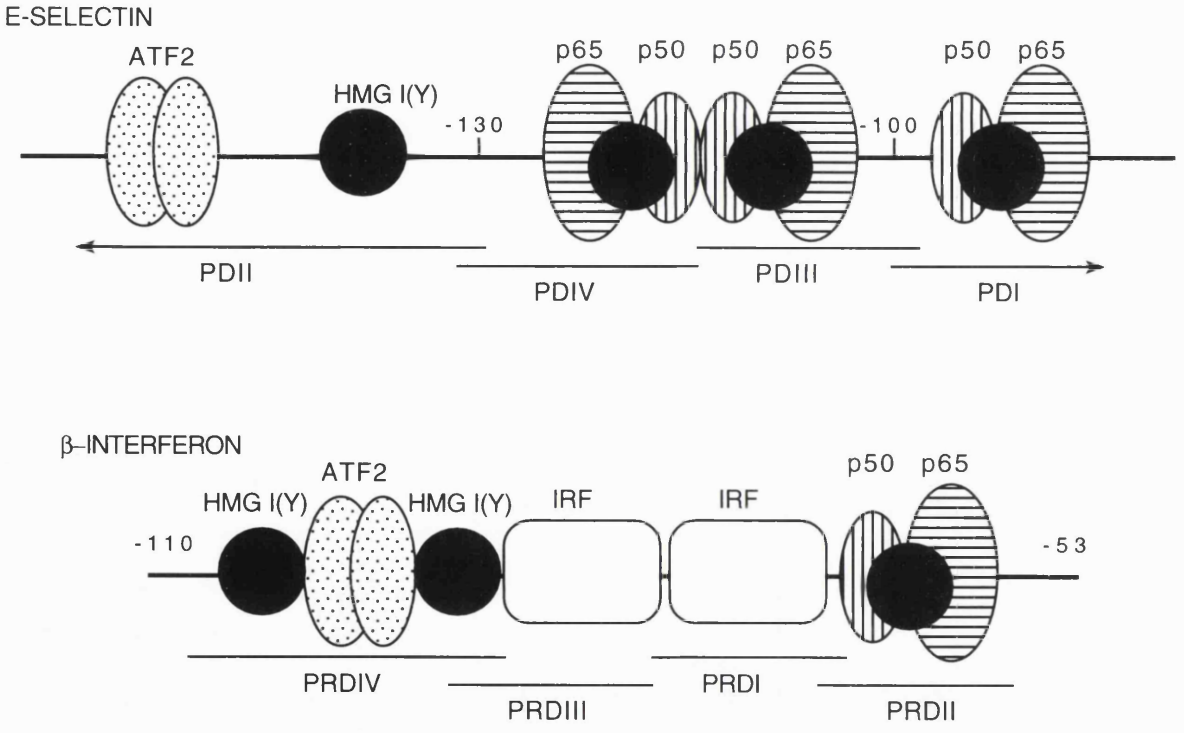


Figure 6.4 Comparison of the human E-selectin and β-interferon promoters

(iv) Prostaglandin Synthase2 (PGS2) Promoter

The PGS2 gene is stimulated by a variety of agents including PDGF, CSF, serum, endotoxin and ILs. Several of these stimuli e.g. PDGF, CSF and IL-2 are known to act via the Src family of tyrosine kinases and accordingly v-Src itself is able to stimulate PGS2 transcription. Deletion and mutational analysis determined that v-Src exerts its effect via a CRE site in the PGS2 promoter (Xie et al., 1994). This site is constitutively occupied and v-Src therefore acts by influencing post-translational modification of factor(s) bound at this site. Since v-Src increases cJun phosphorylation in response to UV-C irradiation (Devary et al., 1992), it is feasible that a similar mode of regulation operates at the PGS2 CRE. This is further suggested by the involvement of Ras in v-Src induction of PGS2 expression (Xie et al., 1994). The factor(s) phosphorylated at the PGS2 CRE have not been identified but an obvious candidate is ATF2 due to its similarities to cJun in regard to its regulation. Thus induction of PGS2 transcription may occur as described in (i).

(v) Cyclin D Promoter

The cyclin D1 promoter has been cloned and the regulatory elements within it analyzed by deletion analysis (Herber et al., 1994). Three key elements have been identified including a CRE site at position -52. This site is required for induction of the cyclin D1 gene by cJun and to a lesser extent by serum. As with the other promoters described, the cyclin D1 CRE is constitutively bound and therefore regulation probably occurs by post-translational modification of bound factors. The sequence of this CRE is known to be a preferred binding site for ATF2 (Benbrook and Jones, 1994), thus it is likely that ATF2-cJun heterodimers are responsible for regulating transcription from this site via their phosphorylation-dependent transactivation. Such a mechanism ties in with the observed induction of cyclin D1 transcription by Ras (Filmus et al., 1994).

From the above section, it can be seen that ATF2 is implicated in regulating a variety genes in response to a number of stimuli. A common theme emerges from identifying its regulatory roles namely that ATF2 induces the expression of genes in response to and which aid in overcoming environmental stress; ATF2 is activated by stress agents such as UV-C, alkylating compounds, pro-inflammatory cytokines and viruses, and it activates genes such as PGS2, E-selectin and β-IFN which are associated with inflammation, immune responses and anti-viral activity. ATF2 also stimulates genes involved in proliferation e.g. cJun and cyclin D1. This is thought to serve a protective function because by stimulating growth ATF2 induces the cell to synthesize new proteins thus replacing any cellular constituents damaged by the stress agents.

It can also be seen that ATF2 activity is regulated in two major ways; by phosphorylation and by cooperation with other factors. In addition, there is evidence to suggest that ATF2 may be regulated by conformational changes induced in the protein upon

binding to certain DNA sequences. This means that promoter variations in the type of CRE site present and in the number of binding sites for ATF2 cooperative factors will influence the activity of ATF2 and thus allow for differential regulation of transcription by ATF2. Since many of the mechanisms proposed to control ATF2 activity are interconnected e.g. v-Src and TNFα activate NF-κB and ATF2 which can then cooperate with each other, stimulation by a single signal can lead to a very rapid induction of transcription allowing the cell to adapt rapidly to stressful conditions. Thus ATF2 appears to lie at the end of important signalling pathways which allow the cell to survive in damaging conditions. Further work is now required to confirm the role ATF2 plays and to fully the define the pathway involved and the influences regulating ATF2 activity.

REFERENCES

REFERENCES

Abdel-Hafiz, H. A., Chen, C. Y., Marcell, T., Kroll, D. J. and Hoeffler, J. P. (1993). Oncogene **8:** 1161-74.

Abdel-Hafiz, H. A., Heasley, L. E., Kyriakis, J. M., Avruch, J., Kroll, D. J., Johnson, G. L. and Hoeffler, J. P. (1992). Mol Endocrinol **6:** 2079-89.

Adya, N., Zhao, L.-J., Huang, W., Boros, I. and Giam, C.-Z. (1994). Proc Natl Acad Sci U S A **91:** 5642-6.

Agoff, S. N. (1994). Oncogene **9:** 3707-3711.

Alberts, A. S., Montminy, M., Shenolikar, S. and Feramisco, J. R. (1994). Mol Cell Biol **14:** 4398-407.

Amati, B., Brooks, M. W., Levy, N., Littlewood, T. D., Evan, G. I. and Land, H. (1993). Cell **72:** 233-45.

Ammerer, G. (1994). Curr Opin Gen Dev **4:** 90-95.

Anderson, W. F., Ohlendorf, D. H., Takeda, Y. and Matthews, B. W. (1981). Nature **290:** 754-759.

Andrews, N. C. and Faller, D. V. (1991). Nucleic Acids Res **19:** 2499.

Archer, T. K., Cordingley, M. G., Wolford, R. G. and Hager, G. L. (1991). Mol Cell Biol **11:** 688-98.

Arias, J., Alberts, A. S., Brindle, P., Claret, F. X., Smeal, T., Karin, M., Feramisco, J. and Montminy, M. (1994). Nature **370:** 226-9.

Aronheim, A., Engelberg, D., Li, N., Al-Alawi, N., Schlessinger, J. and Karin, M. (1994). Cell **78:** 949-961.

Assa-Munt, N., Mortishire-Smith, R. J., Aurora, R., Herr, W. and Wright, P. E. (1993). Cell **73:** 193-205.

Attardi, L. D. and Tjian, R. (1993). Genes Dev **7:** 1341-53.

Auble, D. T. and Hahn, S. (1993). Genes Dev **7:** 844-56.

Aurora, R. and Herr, W. (1992). Mol Cell Biol **12:** 455-67.

Baeuerle, P. A. and Henkel, T. (1994). Annu Rev Immunol **12:** 141-79.

Baichwal, V. R. and Tjian, R. (1990). Cell **63:** 815-25.

Baichwal, V. R., Park, A. and Tjian, R. (1992). Genes Dev **6:** 1493-502.

Bandara, L. R., Buck, V. M., Zamanian, M., Johnston, L. H. and La Thangue, N. (1993). Embo J **12:** 4317-24.

Barbeau, D., Charbonneau, R., Whalen, S. G., Bayley, S. T. and Branton, P. E. (1994). Oncogene **9:** 359-73.

Bayley, S. T. and Mymyrk, J. S. (1994). Intl J Oncology **5:** 425-444.

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

Benbrook, D. M. and Jones, N. C. (1994). Nucleic Acids Res **22:** 1463-9.

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.

Berk, A. J. (1986). Annu Rev Genet **20:** 45-79.

Blumer, K. J. and Johnson, G. L. (1994). Trends Biochem Sci **19:** 236-40.

Boulton, T. G., Nye, S. H., Robbins, D. J., Ip, N. Y., Radziejewska, E., Morgenbesser, S. D., DePinho, R. A., Panayotatos, N., Cobb, M. H. and Yancopoulos, G. D. (1991). Cell **65:** 663-75.

Boyer, T. G. and Berk, A. J. (1993). Genes Dev **7:** 1810-23.

Breg, J. N., van Opheusden, J. H. J., Burgering, M. J., Boelens, R. and Kaptein, R. (1990). Nature **346:** 586-9.

Brent, R. and Ptashne, M. (1985). Cell **43:** 729-736.

Bresnick, E. H., Bustin, M., Marsaud, V., Richard, F. H. and Hager, G. L. (1992). Nucleic Acids Res **20:** 273-8.

Brindle, P., Linke, S. and Montminy, M. (1993). Nature **364:** 821-4.

Buratowski, S. (1994). Cell **77:** 1-3.

Caswell, R., Hagemeier, C., Chiou, C.-J., Hayward, G., Kouzarides, T. and Sinclair, J. (1993). J Gen Virol **74:** 2691-2698.

Chatterjee, P. K., Bruner, M., Flint, S. J. and Harter, M. L. (1988). Embo J **7:** 835-41.

Chatton, B., Bocco, J. L., Gaire, M., Hauss, C., Reimund, B., Goetz, J. and Kedinger, C. (1993). Mol Cell Biol **13:** 561-70.

Chatton, B., Bocco, J. L., Goetz, J., Gaire, M., Lutz, Y. and Kedinger, C. (1994). Oncogene **9:** 375-85.

Chellappan, S., Kraus, V. B., Kroger, B., Munger, K., Howley, P. M., Phelps, W. C. and Nevins, J. R. (1992). Proc Natl Acad Sci U S A **89:** 4549-53.

Chen, J-L., Attardi, L. D., Verrijzer, C. P., Yokomori, K. and Tjian, R. (1994). Cell **79:** 93-105.

Choy, B. and Green, M. R. (1993). Nature **366:** 531-6.

Chrivia, J. C., Kwok, R. P., Lamb, N., Hagiwara, M., Montminy, M. R. and Goodman, R. H. (1993). Nature **365:** 855-9.

Cobrinik, D., Whyte, P., Peeper, D. S., Jacks, T. and Weinberg, R. A. (1993). Genes Dev **7:** 2392-404.

Comai, L., Tanese, N. and Tjian, R. (1992). Cell **68:** 965-76.

Courey, A. J. and Tjian, R. (1988). Cell **55:** 887-98.

Cowell, I. G. (1994). Trends Biochem Sci **19:** 38-42.

Cress, W. D. and Triezenberg, S. J. (1991). Science **251:** 87-90.

Davis, R. J. (1993). J Biol Chem **268:** 14553-6.

Dekker, N., Cox, M., Boelens, R., Verrijzer, C. P., van der Vliet, P. C. and Kaptein, R. (1993). Nature **362:** 852-5.

Delmas, V., Laoide, B. M., Masquilier, D., de Groot, R. P., Foulkes, N. S. and Sassone-Corsi, P. (1992). Proc Natl Acad Sci U S A **89:** 4226-30.

Derijard, B., Hibi, M., Wu, I. H., Barrett, T., Su, B., Deng, T., Karin, M. and Davis, R. J. (1994). Cell **76:** 1025-37.

Devary, Y., Gottlieb, R. A., Smeal, T. and Karin, M. (1992). Cell **71:** 1081-91.

Drapkin, R. and Reinberg, D. (1994). Trends Biochem. **19:** 504-508.

Du, W., Thanos, D. and Maniatis, T. (1993). Cell **74:** 887-98.

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

Dynlacht, B. D., Hoey, T. and Tjian, R. (1991). Cell **66:** 563-76.

Dynlacht, B. D., Weinzierl, R. O., Admon, A. and Tjian, R. (1993). Nature **363:** 176-9.

Dyson, N., Dembski, M., Fattaey, A., Ngwu, C., Ewen, M. and Helin, K. (1993). J Virol **67:** 7641-7.

Eckner, R., Ewen, M. E., Newsome, D., Gerdes, M., DeCaprio, J. A., Lawrence, J. B. and Livingston, D. M. (1994). Genes Dev **8:** 869-884.

Ellenberger, T. E., Brandl, C. J., Struhl, K. and Harrison, S. C. (1992). Cell **71:** 1223-37.

Ewen, M. E., Xing, Y. G., Lawrence, J. B. and Livingston, D. M. (1991). Cell **66:** 1155-64.

Fattaey, A. R., Harlow, E. and Helin, K. (1993). Mol Cell Biol **13:** 7267-77.

Ferreri, K., Gill, G. and Montminy, M. (1994). Proc Natl Acad Sci U S A **91:** 1210-3.

Filmus, J., Robles, A. I., Shi, W., Wong, M. J., Colombo, L. L. and Conti, C. J. (1994). Oncogene **9:** 3627-3633.

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.

Foulkes, N. S., Mellstrom, B., Benusiglio, E. and Sassone-Corsi, P. (1992). Nature **355:** 80-4.

Frampton, J., Gibson, T. J., Ness, S. A., Doderlein, G. and Graf, T. (1991). Protein Eng **4:** 891-901.

Freshney, N. W., Rawlinson, L., Guesdon, F., Jones, E., Cowley, S., Hsuan, J. and Saklatvala, J. (1994). Cell **78:** 1039-1049.

Frost, J. A., Geppert, T. D., Cobb, M. H. and Feramisco, J. R. (1994). Proc Natl Acad Sci U S A **91:** 3844-8.

Ge, H. and Roeder, R. G. (1994). Cell **78:** 513-23.

Geisberg, J. V., Lee, W. S., Berk, A. J. and Ricciardi, R. P. (1994). Proc Natl Acad Sci U S A **91:** 2488-92.

Gill, G. (1994). Curr Biol **4:** 374-376.

Gill, G., Pascal, E., Tseng, Z. H. and Tjian, R. (1994). Proc Natl Acad Sci U S A **91:** 192-6.

Girling, R., Partridge, J. F., Bandara, L. R., Burden, N., Totty, N. F., Hsuan, J. J. and La Thangue, N. B. (1993). Nature **362:** 83-7.

Gonzalez, G. A. and Montminy, M. R. (1989). Cell **59:** 675-80.

Gonzalez, G. A., Menzel, P., Leonard, J., Fischer, W. H. and Montminy, M. R. (1991). Mol Cell Biol **11:** 1306-12.

Gonzalez, G. A., Yamamoto, K. K., Fischer, W. H., Karr, D., Menzel, P., Biggs, W., Vale, W. W. and Montminy, M. R. (1989). Nature **337:** 749-52.

Goodrich, J. A. (1994). Curr Opin Cell Biol **6:** 403-409.

Goodrich, J. A. and Tjian, R. (1994). Cell **77:** 145-56.

Goodrich, J. A., Hoey, T., Thut, C. J., Admon, A. and Tjian, R. (1993). Cell **75:** 519-30.

Graham, F. L., Smiley, J., Russell, W. C. and Nairn, R. (1977). J Gen Virol **36:** 59-74.

Gupta, S., Seth, A. and Davis, R. J. (1993). Proc. Natl. Acad. U. S. A. **90:** 3216-3220.

Hagemeier, C., Walker, S., Caswell, R., Kouzarides, T. and Sinclair, J. (1992). J Virol **66:** 4452-6.

Hagiwara, M., Brindle, P., Harootunian, A., Armstrong, R., Rivier, J., Vale, W., Tsien, R. and Montminy, M. R. (1993). Mol Cell Biol **13:** 4852-9.

Hagmeyer, B. M., Konig, H., Herr, I., Offringa, R., Zantema, A., van der Eb, A. J., Herrlich, P. and Angel, P. (1993). Embo J **12:** 3559-72.

Hahn, S. (1993). Cell **72:** 481-3.

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

Han, J., Lee, J.-D., Bibbs, L. and Ulevitch, R. J. (1994). Science **265:** 808-811.

Hannon, G. J., Demetrick, D. and Beach, D. (1993). Genes Dev **7:** 2378-91.

Hardy, S. and Shenk, T. (1989). Mol Cell Biol **9:** 4495-506.

Harlow, E., Whyte, P., Franza, B. J. and Schley, C. (1986). Mol Cell Biol **6:** 1579-89.

Helin, K., Harlow, E. and Fattaey, A. (1993). Mol Cell Biol **13:** 6501-8.

Helin, K., Lees, J. A., Vidal, M., Dyson, N., Harlow, E. and Fattaey, A. (1992). Cell **70:** 337-50.

Herber, B., Truss, M., Beato, M. and Müller, R. (1994). Oncogene **9:** 1295-1304.

Hernandez, N. (1993). Genes Dev **7:** 1291-308.

Herr, I., van Dam, H. and Angel, P. (1994). Carcinogenesis **15:** 1105-13.

Herr, W., Sturm, R. A., Clerc, R. G., Corcoran, L. M., Baltimore, D. M., Sharp, P. A., Ingraham, H. A., Rosenfeld, M. G., Finney, M., Rukun, G. and Horvitz, H. R. (1988). Genes Dev **2:** 1513-1516.

Hibi, M., Lin, A., Smeal, T., Minden, A. and Karin, M. (1993). Genes Dev **7:** 2135-48.

Hill, C. S., Wynne, J. and Treisman, R. (1994). EMBO J **13:** 5421-5432.

Hisatake, K., Hasegawa, S., Takada, R., Nakatani, Y., Horikoshi, M. and Roeder, R. G. (1993). Nature **362:** 179-81.

Hoeffler, J. P., Lustbader, J. W. and Chen, C-Y. (1991). Mol Endocrinol **5:** 256-66.

Hoeffler, J. P., Meyer, T. E., Yun, Y., Jameson, J. L. and Habener, J. F. (1988). Science **242:** 1430-3.

Hoey, T., Weinzierl, R. O., Gill, G., Chen, J-L., Dynlacht, B. D. and Tjian, R. (1993). Cell **72:** 247-60.

Hoopes, B. C., LeBlanc, J. F. and Hawley, D. K. (1992). J Biol Chem **267:** 11539-47.

Horikoshi, N., Maguire, K., Kralli, A., Maldonado, E., Reinberg, D. and Weinmann, R. (1991). Proc Natl Acad Sci U S A **88:** 5124-8.

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

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

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

Ivey-Hoyle, M., Conroy, R., Huber, H. E., Goodhart, P. J., Oliff, A. and Heimbrook, D. C. (1993). Mol Cell Biol **13:** 7802-12.

Jacq, X., Brou, C., Lutz, Y., Davidson, I., Chambon, P. and Tora, L. (1994). Cell **79:** 107-117.

Jainchall, J. L., Aaronson, S. A. and Todaro, G. J. (1969). J Virol **4:** 549-553.

Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B. and Smale, S. T. (1994). Mol Cell Biol **14:** 116-27.

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. A., Livingston, D. M. and Flemington, E. K. (1992). Cell **70:** 351-64.

Kaelin, W. G., Pallas, D. C., DeCaprio, J. A., Kaye, F. J. and Livingston, D. M. (1991). Cell **64:** 521-32.

Kassavetis, G. A., Joazeiro, C. A., Pisano, M., Geiduschek, E. P., Colbert, T., Hahn, S. and Blanco, J. A. (1992). Cell **71:** 1055-64.

Kaszubska, W., van Huijsduijnen, R. H., Ghersa, P., DeRaemy, S. A., Chen, B. P., Hai, T., DeLamarter, J. F. and Whelan, J. (1993). Mol Cell Biol **13:** 7180-90.

Keegan, L., Gill, G. and Ptashne, M. (1986). Science **231:** 699-704.

Kim, S.-J., Wagner, S., Liu, F., O'Reilly, M. A., Robbins, P. D. and Green, M. R. (1992). Nature **358:** 331-4.

Kimelman, D., Miller, J. S., Porter, D. and Roberts, B. E. (1985). J Virol **53:** 399-409.

Kingston, R. E. and Green, M. R. (1994). Curr Biol **4:** 325-332.

Kissinger, C. R., Liu, B., Martin-Blanco, E., Kornberg, T. B. and Pabo, C. O. (1990). Cell **63:** 579-590.

Kokubo, T., Gong, D-W., Roeder, R. G., Horikoshi, M. and Nakatani, Y. (1993). Proc Natl Acad Sci U S A **90:** 5896-900.

Kokubo, T., Gong, D. W., Wootton, J. C., Horikoshi, M., Roeder, R. G. and Nakatani, Y. (1994). Nature **367:** 484-7.

Kokubo, T., Gong, D. W., Yamashita, S., Horikoshi, M., Roeder, R. G. and Nakatani, Y. (1993a). Genes Dev **7:** 1033-46.

Kokubo, T., Gong, D. W., Yamashita, S., Takada, R., Roeder, R. G., Horikoshi, M. and Nakatani, Y. (1993b). Mol Cell Biol **13:** 7859-63.

Kretzschmar, M., Kaiser, K., Lottspeich, F. and Meisterernst, M. (1994). Cell **78:** 525-34.

Kretzschmar, M., Meisterernst, M. and Roeder, R. G. (1993). Proc Nati Acad Sci U S A **90:** 11508-12.

Kretzschmar, M., Stelzer, G., Roeder, R. G. and Meisterernst, M. (1994). Mol Cell Biol **14:** 3927-37.

Kwok, R. P., Lundblad, J. R., Chrivia, J. C., Richards, J. P., Bachinger, H. P., Brennan, R. G., Roberts, S. G., Green, M. R. and Goodman, R. H. (1994). Nature **370:** 223-6.

Kyriakis, J. M., App, H., Zhang, X-F., Banerjee, P., Brautigan, D. L., Rapp, U. R. and Avruch, J. (1992). Nature **358:** 417-21.

Kyriakis, J. M., Banerjee, P., Nikolakaki, E., Dai, T., Rubie, E. A., Ahmad, M. F., Avruch, J. and Woodgett, J. R. (1994). Nature **369:** 156-60.

La Thangue, N. B. (1994). Trends Biochem Sci **19:** 108-14.

Lalli, E. and Sassone-Corsi, P. (1994). J Bioi Chem **269:** 17359-62.

Landschulz, W. H., Johnson, P. F. and McKnight, S. L. (1988). Science **240:** 1759-64.

Lange-Carter, C. A., Pleiman, C. M., Gardner, A. M., Blumer, K. J. and Johnson, G. L. (1993). Science **260:** 315-319.

Laoide, B. M., Foulkes, N. S., Schlotter, F. and Sassone-Corsi, P. (1993). Embo J **12:** 1179-91.

Laurent, B. C., Treich, I. and Carison, M. (1993). Genes Dev **7:** 583-91.

Lee, J.-S., Galvin, K. M. and Shi, Y. (1993). Proc. Natl. Acad. Sci. U. S. A. **90:** 6145-6149.

Lee, K. A. and Green, M. R. (1987). Embo J **6:** 1345-53.

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

Lees, J. A., Saito, M., Vidal, M., Valentine, M., Look, T., Harlow, E., Dyson, N. and Helin, K. (1993). Mol Cell Biol **13:** 7813-25.

Leevers, S. J., Paterson, H. F. and Marshall, C. J. (1994). Nature **369:** 411-414.

Leong, K., Brunet, L. and Berk, A. J. (1988). Mol Cell Biol **8:** 1765-74.

Li, Q. and Wrange, O. (1993). Genes Dev **7:** 2471-82.

Li, Y., Graham, C., Lacy, S., Duncan, A. M. and Whyte, P. (1993). Genes Dev **7:** 2366-77.

Lieberman, P. M. and Berk, A. J. (1991). Genes Dev **5:** 2441-54.

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:** 1091-100.

Lin, A., Frost, J., Deng, T., Smeal, T., Al-Alawi, N., Kikkawa, U., Hunter, T., Brenner, D. and Karin, M. (1992). Cell **70:** 777-789.

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

Lin, Y-S., Ha, I., Maldonado, E., Reinberg, D. and Green, M. R. (1991). Nature **353:** 569-71.

Liou, H. C. and Baltimore, D. (1993). Curr Opin Cell Biol **5:** 477-87.

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

Liu, F. and Green, M. R. (1994). Nature **368:** 520-5.

Liu, F., Thompson, M. A., Wagner, S., Greenberg, M. E. and Green, M. R. (1993). J Biol Chem **268:** 6714-20.

Lobo, S. M., Tanaka, M., Sullivan, M. L. and Hernandez, N. (1992). Cell **71:** 1029-40.

Loriaux, M. M., Rehfuss, R. P., Brennan, R. G. and Goodman, R. H. (1993). Proc Natl Acad Sci U S A **90:** 9046-50.

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

Ma, J. and Ptashne, M. (1987). Cell **48:** 847-53.

Maekawa, T., Sakura, H., Kanei-Ishii, C., Sudo, T., Yoshimura, T., Fujisawa, J., Yoshida, M. and Ishii, S. (1989). Embo J **8:** 2023-8.

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

Manohar, C. F., Kratochvil, J. and Thimmapaya, B. (1990). J Virol **64:** 2457-66.

Marais, R., Wynne, J. and Treisman, R. (1993). Cell **73:** 381-93.

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

Martinez, E., Dusserre, Y., Wahli, W. and Mermod, N. (1991). Mol Cell Biol **11:** 2937-45.

McCormick, F. (1993). Nature **363:** 15-6.

McKay, D. B. and Steitz, T. A. (1981). Nature **290:** 744-749.

Meisterernst, M., Roy, A. L., Lieu, H. M. and Roeder, R. G. (1991). Cell **66:** 981-93.

Merino, A., Madden, K. R., Lane, W. S., Champoux, J. J. and Reinberg, D. (1993). Nature **365:** 227-32.

Mermod, N., O'Neill, E. A., Kelly, T. J. and Tjian, R. (1989). Cell **58:** 741-53.

Metz, R., Bannister, A. J., Sutherland, J. A., Hagemeier, C., O'Rourke, E. C., Cook, A., Bravo, R. and Kouzarides, T. (1994). Mol Cell Biol **14:** 6021-9.

Miller, J., McLachlan, A. D. and Klug, A. (1985). Embo J **4:** 1609-14.

Minden, A., Lin, A., McMahon, M., Lange-Carter, C., Derijard, B., Davis, R. J., Johnson, G. L. and Karin, M. (1994a). Science **266:** 1719-1723.

Minden, A., Lin, A., Smeal, T., Dérijard, B., Cobb, M., Davis, R. and Karin, M. (1994b). Mol Cell Biol **14:** 6683-6688.

Molina, C. A., Foulkes, N. S., Lalli, E. and Sassone-Corsi, P. (1993). Cell **75:** 875-86.

Moran, E. and Mathews, M. B. (1987). Cell **48:** 177-8.

Muegge, K., Vila, M., Gusella, G. L., Musso, T., Herrlich, P., Stein, B. and Durum, S. K. (1993). Proc Natl Acad Sci U S A **90:** 7054-8.

Murre, C., McCaw, P. S. and Baltimore, D. (1989). Cell **56:** 777-83.

Nakielny, S., Cohen, P., Wu, J. and Sturgill, T. (1992). EMBO J **11:** 2123-2129.

Natoli, G., Avantaggiati, M. L., Chirillo, P., Puri, P., Ianni, A., Balsano, C. and Levrero, M. (1994). Oncogene **9:** 2837-2843.

Nebreda, A. R., Hill, C., Gomez, N., Cohen, P. and Hunt, T. (1993). Febs Lett **333:** 183-187.

Nevins, J. R. (1987). Microbiol. Rev. **51:** 419-430.

Nevins, J. R. (1992). Science **258:** 424-9.

Nichols, M., Weih, F., Schmid, W., DeVack, C., Kowenz-Leutz, E., Luckow, B., Boshart, M. and Schutz, G. (1992). Embo J **11:** 3337-46.

Nomura, N., Zu, Y. L., Maekawa, T., Tabata, S., Akiyama, T. and Ishii, S. (1993). J Biol Chem **268:** 4259-66.

O'Shea, E. K., Klemm, J. D., Kim, P. S. and Alber, T. (1991). Science **254:** 539-44.

O'Shea-Greenfield, A. and Smale, S. T. (1992). J Biol Chem **267:** 1391-1401.

Ohlendorf, D. H., Anderson, W. F., Lewis, M., Pabo, C. O. and Matthews, B. W. (1983). J Mol Biol **169:** 757-69.

Ozer, J., Moore, P. A., Bolden, A. H., Lee, A., Rosen, C. A. and Lieberman, P. M. (1994). Genes Dev **8:** 2324-2335.

Pabo, C. O. and Lewis, M. (1982). Nature **298:** 443-7.

Pabo, C. O. and Sauer, R. T. (1992). Annu Rev Biochem **61:** 1053-95.

Pages, G., Brunet, A., L'Allemain, G. and Pouyssegur, J. (1994). Embo J **13:** 3003-10.

Perlmann, T. and Wrange, O. (1991). Mol Cell Biol **11:** 5259-65.

Peterson, M. G. and Tjian, R. (1992). Nature **358:** 620-1.

Pina, B., Bruggemeier, U. and Beato, M. (1990). Cell **60:** 719-31.

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

Pugh, B. F. and Tjian, R. (1990). Cell **61:** 1187-1197.

Quinn, P. G. (1993). J Biol Chem **268:** 16999-17009.

Regier, J. L., Shen, F. and Triezenberg, S. J. (1993). Proc Natl Acad Sci U S A **90:** 883-7.

Roberts, S. G. E. and Green, M. R. (1994). Nature **371:** 717-720.

Roberts, S. G., Ha, I., Maldonado, E., Reinberg, D. and Green, M. R. (1993). Nature **363:** 741-4.

Rouse, J., Cohen, P., Trigon, S., Morange, M., Alonso-Llamazares, A., Zamanillo, D., Hunt, T. and Nebreda, A. R. (1994). Cell **78:** 1027-1037.

Ruden, D. M. (1990). Mol Cell Biol **10:** 1432-8.

Ruppert, R., Cole, T. J., Boshart, M., Schmid, E. and Schütz, G. (1992). EMBO J. **11:** 1503-1512.

Ruppert, S., Wang, E. H. and Tjian, R. (1993). Nature **362:** 175-9.

Sadowski, C. L., Henry, R. W., Lobo, S. M. and Hernandez, N. (1993). Genes Dev **7:** 1535-48.

Sadowski, I. and Ptashne, M. (1989). Nucleic Acids Res **17:** 7539.

Sadowski, I., Ma, J., Triezenberg, S. and Ptashne, M. (1988). Nature **335:** 563-4.

Sanchez, I., Hughes, R. T., Mayer, B. J., Yee, K., Woodgett, J. R., Avruch, J., Kyriakis, J. M. and Zon, L. I. (1994). Nature **372:** 794-798.

Schneider, J. F., Fisher, F., Goding, C. R. and Jones, N. C. (1987). Embo J **6:** 2053-60.

Schwarz, J. K., Devoto, S. H., Smith, E. J., Chellappan, S. P., Jakoi, L. and Nevins, J. R. (1993). Embo J **12:** 1013-20.

Shirodkar, S., Ewen, M., DeCaprio, J. A., Morgan, J., Livingston, D. M. and Chittenden, T. (1992). Cell **68:** 157-66.

Simon, M. C., Fisch, T. M., Benecke, B. J., Nevins, J. R. and Heintz, N. (1988). Cell **52:** 723-9.

Slavicek, J. M., Jones, N. C. and Richter, J. D. (1988). Embo J **7:** 3171-80.

Sluss, H. K., Barrett, T., Dérijard, B. and Davis, R. J. (1994). Mol Cell Biol **14:** 8376-8384.

Smeal, T., Binetruy, B., Mercola, D. A., Birrer, M. and Karin, M. (1991). Nature **354:** 494-6.

Smith, D. B. and Johnson, K. S. (1988). Gene **67:** 31-40.

Southgate, C. D. and Green, M. R. (1991). Genes Dev **5:** 2496-507.

Spandau, D. F. and Lee, C. H. (1988). J Virol **62:** 427-34.

Srebrow, A., Muro, A. F., Werbajh, S., Sharp, P. A. and Kornblihtt, A. R. (1993). Febs Lett **327:** 25-8.

Steitz, T. A., Ohlendorf, D. H., McKay, D. B., Anderson, W. F. and Matthews, B. W. (1982). Proc Natl Acad Sci U S A **79:** 3097-100.

Stephens, C. and Harlow, E. (1987). Embo J **6:** 2027-35.

Stern, S. and Herr, W. (1991). Genes Dev **5:** 2555-66.

Stringer, K. F., Ingles, C. J. and Greenblatt, J. (1990). Nature **345:** 783-6.

Sun, X., Ma, D., Sheldon, M., Yeung, K. and Reinberg, D. (1994). Genes Dev **8:** 2336-2348.

Sutherland, J. A., Cook, A., Bannister, A. J. and Kouzarides, T. (1992). Genes Dev **6:** 1810-9.

Taggart, A. K., Fisher, T. S. and Pugh, B. F. (1992). Cell **71:** 1015-28.

Tanaka, M. and Herr, H. (1994). Mol Cell Biol **14:** 6056-6067.

Tanaka, M., Clouston, W. M. and Herr, W. (1994). Mol Cell Biol **14:** 6046-6055.

Tjian, R. and Maniatis, T. (1994). Cell **77:** 5-8.

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

Triezenberg, S. J., Kingsbury, R. C. and McKnight, S. L. (1988). Genes Dev **2:** 718-29.

Ullrich, A. and Schlessinger, J. (1990). Cell **61:** 203-12.

Usheva, A. and Shenk, T. (1994). Cell **76:** 1115-21.

Vallee, B. L., Coleman, J. E. and Auld, D. S. (1991). Proc Natl Acad Sci U S A **88:** 999-1003.

van Dam, H., Duyndam, M., Rottier, R., Bosch, A., de Vries-Smits, L., Herrlich, P., Zantema, A., Angel, P. and van der Eb, A. J. (1993). Embo J **12:** 479-87.

van Dam, H., Wilhelm, D., Herr, I., Steffen, A., Herrlich, P. and Angel, P. (1995). EMBO J. in press.

van Huijsduijnen, R. H., Whelan, J., Pescini, R., Becker-Andre, M., Schenk, A-M. and DeLamarter, J. F. (1992). J Biol Chem **267:** 22385-22391.

van Wijnen, A. J., Bidwell, J. P., Fey, E. G., Penman, S., Lian, J. B., Stein, J. L. and Stein, G. S. (1993). Biochemistry **32:** 8397-402.

Verrijzer, C. P., Yokomori, K., Chen, J. L. and Tjian, R. (1994). Science **264:** 933-41.

Voronova, A. and Baltimore, D. (1990). Proc Natl. Acad. Sci U. S. A. **87:** 4722-4726.

Waeber, G., Meyer, T. E., LeSieur, M., Hermann, H. L., Gerard, N. and Habener, J. H. (1991). Mol Endocrinol **5:** 1418-1430.

Wagner, S. and Green, M. R. (1993). Science **262:** 395-9.

Wang, H-G. H., Rikitake, Y., Carter, M. C., Yaciuk, P., Abraham, S. E., Zerler, B. and Moran, E. (1993). J Virol **67:** 476-88.

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

Weinzierl, R. O., Dynlacht, B. D. and Tjian, R. (1993a). Nature **362:** 511-7.

Weinzierl, R. O., Ruppert, S., Dynlacht, B. D., Tanese, N. and Tjian, R. (1993b). Embo J **12:** 5303-9.

Weiss, M. A., Ellenberger, T., Wobbe, C. R., Lee, J. P., Harrison, S. C. and Struhl, K. (1990). Nature **347:** 575-8.

Westwick, J. K., Weitzel, C., Minden, A., Karin, M. and Brenner, D. A. (1994). J Biol Chem **269:** 26396-26401.

Wheat, W. H., Roesler, W. J. and Klemm, D. J. (1994). Mol Cell Biol **14:** 5881-90.

Whelan, J., Ghersa, P., van Huijsduijnen, R. H., Gray, J., Chandra, G., Talabot, F. and DeLamarter, J. F. (1991). Nucleic Acids Res **19:** 2645-53.

White, R. J. and Jackson, S. P. (1992). Cell **71:** 1041-53.

Whitley, M. Z., Thanos, D., Read, M. A., Maniatis, T. and Collins, T. (1994). Mol Cell Biol **14:** 6464-6475.

Whyte, P., Buchkovich, K. J., Horowitz, J. M., Friend, S. H., Raybuck, M., Weinberg, R. A. and Harlow, E. (1988). Nature **334:** 124-9.

Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. and Pabo, C. O. (1991). Cell **67:** 517-28.

Wolffe, A. P. (1994a). Trends Biochem Sci **19:** 240-4.

Wolffe, A. P. (1994b). Curr Biol **4:** 525-528.

Wu, L., Rosser, D. S., Schmidt, M. C. and Berk, A. (1987). Nature **326:** 512-5.

Xie, W., Fletcher, B. S., Anderson, R. D. and Herschman, H. R. (1994). Mol Cell Biol **14:** 6531-6539.

Yan, M., Dai, T., Deak, J. C., Kyriakis, J. M., Zon, L. I., Woodgett, J. R. and Templeton, D. J. (1994). Nature **372:** 798-800.

Yee, S.-P. and Branton, P. E. (1985). Virology **147:** 142-53.

Yokomori, K., Chen, J-L., Admon, A., Zhou, S. and Tjian, R. (1993). Genes Dev **7:** 2587-97.

Yokomori, K., Zeidler, M. P., Chen, J-L., Verrijzer, C. P., Mlodzik, M. and Tjian, R. (1994). Genes Dev **8:** 2313-2323.

Yoshinaga, S. K., Peterson, C. L., Herskowitz, I. and Yamamoto, K. R. (1992). Science **258:** 1598-604.

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

Zawel, L. and Reinberg, D. (1993). Prog Nucleic Acid Res Mol Biol **44:** 67-108.

Zheng, C.-F. and Guan, K.-L. (1993). J Biol Chem **268:** 11435-11439.

Zhou, Q., Boyer, T. G. and Berk, A. J. (1993). Genes Dev **7:** 180-7.

Zhou, Q., Lieberman, P. M., Boyer, T. G. and Berk, A. J. (1992). Genes Dev **6:** 1964-74.

Ziff, E. B. (1990). Trends Genet **6:** 69-72.

Zu, Y.-L., Maekawa, T., Nomura, N., Nakata, T. and Ishii, S. (1993). Oncogene **8:** 2749-58.