Transcriptional Regulation of the *Evi-1* gene

Jeremy Leigh Fordham

PhD

University of Glasgow

Beatson Institute for Cancer Research

May 1996

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

for Roselyn and Nicholas. I adore you both.

Acknowledgements

Firstly, many thanks to my supervisor Dr Chris Bartholomew for his constant encouragement, scientific advice and insight into parenthood. In addition, I wish to thank my tutor Dr Dave Gillespie and Dr Paul Harrison for their valuable advice during my PhD. Thanks also to Kevin, Jerry, Rob, Vivien and Nick and all members of R4 for there constant enthusiasm for science and good beer. I wish to thank my parents and sisters, Siobhan and Emma, for their constant support throughout my academic career. Finally, I wish to thank my wife, Roselyn for her patience during the writing of this thesis, her wise advise, and the birth of our son, Nicholas Paul.

Abstract

Evi-1 shows a temporally and spatially restricted pattern of expression in murine embryonic development and is expressed predominantly in the kidney, lung and developing oocytes of adult mice. The positions of DNAse1 hypersensitive sites (DHS) within an 18kb region of the 5' Evi-1 locus have been examined to identify putative *Evi-1* gene regulatory regions, in murine kidney and spleen tissues. This analysis identified two DHS sites designated DHS I and DHS II. DHS I is located approximately 2kb upstream of the transcription initiation sites whereas DHS II maps over exon I.

The transcriptional activity of the *Evi-1* promoter has been investigated by inserting the 5' region of the gene into a luciferase construct and the activity examined by transient transfection of cells which express low levels of *Evi-1*. A substantial induction of luciferase activity is observed with a 5kb fragment of the 5' Evi-1 locus containing exon I, intron I and exon II which includes DHS I and DHS II.

Subsequent deletion mutagenesis has identfied two regions within DHS II, located between -338 to -284 and -284 to -254, which upon removal result in a substantial reduction of promoter activity. One of these, located between -338 to -284, binds several proteins when examined by footprinting and electrophoretic mobility shift assays. Interestingly, the most abundant factor, designated EvBP1, has been shown to bind a 14bp imperfect palindromic sequence, tttccctggggaaa, which is absolutely conserved in the human *Evi-1* promoter sequence. This sequence contains homology with putative binding sites for, AP3, AP2 and C/EBP. However, competition studies in EMSA with consensus binding site oligonucleotides failed to identify the components of EvBP1. EvBP1 might be a novel ubiquitous transcription factor which is required for regulation of the *Evi-1* promoter. Furthermore, EMSA analysis of the second deleted region between -284 to -254 has identified a CCAAT binding protein, possibly CP1, which may also be important in basal promoter activity.

Functional assays have failed to identify either promoter or enhancer activity for DHS I. Since there are no known appropriate high *Evi-1* expressing cell lines we have established *Evi-1* expressing kidney cultures as a system to examine tissue specific expression. This allowed the identification of a 3kb region necessary for higher activity in the cultures. This activity might correlate with a DHS site which is part of DHS II complex.

Contents

Part 4: Discussion

Part 5: References

List of Figures

List of Tables

Abbreviations

A	:adenine
ad-MLP	:adenovirus major-late promoter
ATP	:adenosine triphosphate
DCR	:dominant control region
DHS	:DNAse1 hypersensitive site
DNA	:deoxynucleic acid
DTT	:dithiothrietol
EMSA	:electrophoretic mobility shift assay
Evi-1	:ecotropic viral integration site-1
Fp	:DNAse1 footprint
G	:guanine
HLH	:helix-loop-helix
HEPES	:N-2-hydroxy-ethylpiperazine-N'-2-ethanesulphonic acid
LTR	:long terminal repeat
LCR	:locus control region
μ	:micro
m	:milli
mm	:millimetre
M	:molar
n	:nano
p	:pico
PAGE	:polyacrylamide gel electrophoresis
PBS	:phosphate buffered saline
PCR	:polymerase chain reaction
PEV	:Positional effect variegation
Pol II	:RNA polymerase II enzyme
RNA	:ribonucleic acid
mRNA	:messenger ribonucleic acid
RT	:room temperature
SAR	:Scaffold associated regions
SDS	:sodium dodecyl sulphate
ts	:temperature sensitive
T	:SV40 large T-antigen
TAFs	:TBP associated factors
TBP	:TATA binding protein
UV	:ultra violet

Part 1: Introduction

Chapter 1. Eukaryotic gene transcription

1.1 Transcription control sequences in eukaryotes

Transcription in eukaryotes is dependent on three different RNA polymerases; RNA polymerase I which transcribes large ribosomal RNAs (rRNA); RNA polymerase II; which transcribes messenger RNA (mRNA) and RNA polymerase III; which transcribes transfer RNAs (tRNA) and small nuclear RNAs (snRNA). Although all the polymerases are capable of transcribing RNA synthesis from a DNA template they require accessory protein factors to initiate transcription from the appropriate start site in specific promoters. Since the emphasis of this thesis is on transcriptional regulation of a polymerase II gene, which is also responsible for transcribing the majority of genes, it will be discussed below. However, many of the principles for RNA polymerase II are the same for the other polymerases.

Promoters:TATA box and the basal transcription machinery

Basal transcription is dependent on the regulated assembly of multi-protein complexes to *cis*-acting sites within the promoter regions of genes. The pivotal component, and the DNA binding activity, of the transcription machinery was originally identified as a general initiation factor for RNA polymerase II (pol II), designated TFIID (Matsui *et al.*, 1980). TFIID binds to a conserved *cis*-acting element containing a TATAAAA or related sequence (TATA box) (Sowadogo & Roeder, 1985b; Nakajima *et al.*, 1988) which in higher eukaryotes is located about 25-30 nucleotides upstream of the transcription start site in the majority of gene promoters transcribed by RNA polymerase II (class II promoters) (Breathnach & Chambon, 1981). Interestingly, TFIID is itself a multi-protein complex composed of a TATA binding protein (TBP) which is

responsible for binding to the TATA box and TBP associated factors (TAFs) which are important for activator-responsive transcription (reviewed by Hernandez, 1993). Although initiation complex formation and subsequent basal transcription by RNA polymerase II may be mediated by TBP alone in a reconstituted system, TAFs are necessary for response to activators and is probably more reflective of the *in vivo* situation.

In the presence of activators, the initiation complex is formed by the binding of TFIID to the TATA box. TBP binding to the TATA box results in the DNA being severely bent to accommodate the convex undersurface of the TBP saddle. As a consequence binding sites for two other general transcription factors, TFIIA and TFIIB, are more accessible and therefore bind TFIID to stabilises DNA bending within the promoter. This sub-complex is then further stabilised by the recruitment of pol II and TFIIF to form a minimal initiation complex. Subsequently, two other general factors, TFIIE and TFIIH, bind to form a complete initiation complex (reviewed by Zawel & Reinberg, 1993; Buratowski, 1994). Transcription is then initiated in a final ATP-dependent step mediated by the helicase activity of TFIIH, which allows conversion of the initiation complex to an elongation complex thereby facilitating promoter clearance (Goodrich & Tjian, 1994).

Despite no recognisable TATA box sequence the TFIID complex has also been shown to be important for the transcription of TATA-less class II promoters (Pugh & Tjian, 1990, 1991; Smale *et al.*, 1990). In addition, TFIID is also an important component for transcription of RNA polymerase I and III (class I and III respectively) (reviewed by Gill, 1992). Interestingly, an initiator (Inr) element, distinct from the TATA box, has also been described which directs initiation complex formation from TATA-less promoters. The Inr element is also found in TATA containing promoters where it may act in isolation or synergistically with the TATA box (Roeder, 1991). A number of proteins have been shown to recognise the Inr element including TFII-I (Roy

et al., 1993), YY1 (Seto *et al.*, 1991) and TAF150 (Verrijzer *et al.*, 1994). The mechanism of complex formation by the Inr element is not well understood as for the TATA box. However, *in vivo* studies have suggested that the Inr element is dominant over the TATA box in the adenovirus major-late promoter (ad-MLP) which contains both elements (Carcamo *et al.*, 1991). It will obviously require further analysis to assess if this observation is a general phenomena.

Although TFIID has been shown to be rate limiting in transcription initiation presumably other steps in initiation complex assembly may also have a regulatory role depending on the local cellular environment, for example, the availability of ATP for promoter clearance. Consistent with this, a number of transcriptional regulators bind basal transcription factors suggesting that they are targets for activator/repressor proteins. It is believed that activators increase the number of initiation complexes, rather than their rate of formation (White *et al.*, 1991; 1992; Choy & Green, 1993). Although the precise mechanism of *trans*-activator mediated transcriptional induction is unknown, activators have been shown to directly contact the basal initiation complex presumably stabilising complex formation.

As previously mentioned the TAFs are vital for the response of the initiation complex to activators. It is therefore significant that an increasing number of TAFs have been identified which interact directly with *trans*-activators (reviewed by Gill & Tijan, 1992). Well studied examples include the 100kD TAF (TAF110) which interacts with the Sp1 transcription factor (Weinzierl *et al.*, 1993; Hoey *et al.*, 1993) and TAF40 which binds to the herpes simplex virus *trans*-activator VP16 (Goodrich *et al.*, 1993). In addition, activators may directly interact with the basal transcription factors. For example, the activators including the USF (Sowadoga & Roeder, 1985a) and ATF (Horikoshi *et al.*, 1988) interact with TFIID, while VP16 may interact with both TFIID (Stringer *et al.*, 1990) and TFIIB (Lin & Green, 1991; Lin *et al.*, 1991; Goodrich *et al.*, 1993). Interestingly, in the absence of a discernible TATA box, upstream activators

such as Sp1 have been implicated in directing the TFIID complex to the appropriate region of DNA (Dynan, 1986).

Recently inhibitors of initiation complex formation have been identified, Dr1 and Dr2 (Inostrozola *et al.*, 1992; Merino *et al.*, 1993), which bind to TBP and hinder access of the basal transcription factors. Interestingly, the TFIIA and GAL4-derived activators have been shown to antagonise the action of the Dr2 (found to be topoisomerase I) (Merino *et al.*, 1993) by blocking their binding to the TFIID complex.

It is clear that the induction and maintenance of transcription is dependent on both the assembly of the basal transcription machinery and co-operation with trans-activator proteins. Consequently, *cis*-binding sites for *trans*-activator proteins are consistently present within class II promoters (Guerente & Bermingham-McDonogh, 1992). A number of conserved *trans*-activator binding sites have been identified which are frequently found in promoter regions (a few are discussed below). In addition, there are a diversity of other less common sequences which play important roles in promoter regulation which are not mentioned (Faisst & Meyer, 1992).

GC boxes

The GC boxes are GC rich regions of DNA which usually bind the ubiquitous transcription factor Sp1, although related sites may also be bound by other activators including EGR-1 (Christy & Nathans, 1989) and WT1 (Rauscher III *et al.*, 1990). GC boxes are found upstream of a number of TATA-containing promoters and act as upstream activator sequences to induce transcription. In some genes with TATA-less promoters containing multiple GC boxes, Sp1 binding is crucial for transcription initiation (Pugh and Tjian, 1990, 1990). Since TATA-less promoters generally have multiple transcription initiation sites this might suggest that the fidelity of this interaction may not be as great as for TATA promoters.

CAAT Boxes

The CAAT box motif is defined by the pentanucleotide sequence CCAAT and is commonly found 50-100 nucleotides upstream of the transcriptional initiation site in class II promoters. The first CCAAT binding proteins identified were C/EBP (Graves *et al.*, 1986) and CTF/NF1 (Jones *et al.*, 1985). The two proteins have different properties, C/EBP was found to be heat stabile and would bind a GCAAT mutant whereas CTF/NF1 was heat labile and was unable to bind this mutant. C/EBP and CTF/NF1 are not related in their DNA binding motifs but may bind the same CCAAT binding site in the HSV tk promoter. C/EBP is a member of the basic leucine zipper family whereas CTF/NF1 has no structural homology with previously identified structural motifs. Subsequent analysis has identified other proteins which bind the CCAAT motif, designated CP1 and CP2. CP1 is a ubiquitous transcription factor and is not heat stabile, in contrast to the C/EBP family. Presumably the adjacent sequences outside the core site, and their profile of expression, contributes to the specificity of each of these proteins to particular CCAAT sites.

Octamer

The octamer sequence, defined by ATGCAAAT, was originally identified in the immunoglobulin heavy chain (IgH) enhancer and promoter and kappa light chain (IgK) promoter (Falkner & Zachau, 1984; Parslow *et al.*, 1984) and was shown to be important in lymphoid cell specific expression. The motif was also found in non-lymphoid restricted promoters, such as the histone H2B promoter (Harvey *et al.*, 1982). Two binding activities were originally identified for the octamer motif designated Oct1 and Oct2 (Singh *et al.*, 1986; Staudt *et al.*, 1986). The Oct1 binding activity was

observed to be ubiquitous whereas the Oct2 protein was restricted to lymphoid cells (Staudt *et al.*, 1986). Separate studies have identified multiple other binding activities for the octamer motif suggesting that a range of transcription factors may bind to this site (Rosales *et al.*, 1987; Gerster *et al.*, 1987).

1.2 Modularity of promoters

Although the sites described previously have been discussed as separate modules the regulation of transcription of class II promoters is more complex being dependent on multiple protein/DNA interactions which all contribute to overall activity. This allows greater control and flexibility of the promoter in different environments such as during development, in cell specificity and in response to hormonal, heat or stress induction.

A number of well characterised promoters including SV40 and HSV tk possess activator binding sequences, such as CCAAT and GC boxes, which have intrinsic activity in isolation but are collectively essential for maximal activity (Jones *et al.*, 1985). Interestingly, the binding of multiple activators to a promoter region can lead to a synergistic activation, which is characterised by an increase that is greater than the activity expected from the sum of the individual binding sites alone. Synergy is observed with a number of different transcription factors and probably reflects a common theme in eukaryotic gene regulation. The chicken *mim-1* gene has been shown to be synergistically activated by v-myb and C/EBP transcription factors (Burk *et al.*, 1993). Furthermore, synergistic activation of transcription has been observed between members of the hormone receptor family. For example, progesterone and glucocorticoid may synergistically activate an artificial promoter containing upstream GRE and PRE response elements (Tsai *et al.*, 1989).

In addition to synergy, the context of a binding site within a promoter may also be important, for example, the mouse major histocompatibility (MHC) class 1 genes are

responsive to TNFα induction. The promoters contain two binding sites for AP2 and KBF1 which are occupied by an AP2-like factor and KBF1 during constitutive levels of expression. However, upon induction these factors are displaced by the binding of two NF-κB dimers (Israel *et al.*, 1989). This illustrates that the contribution of any given binding site to overall activity is dependent on (1) surrounding sites, which may be overlapping and (2) the availability, concentration and specificity of different *trans*-acting factors. An illustration of altered transcriptional regulatory properties in different promoter contexts is given by the GATA-1 transcription factor. The GATA-1 protein has been characterised as both a transcriptional activator and a positive regulator of erythroid differentiation (Martin *et al.*, 1990; Orkin *et al.*, 1992). By contrast, the protein may also act as a developmental repressor in the ε-globin gene (Raich *et al.*, 1995). Transcription factors may therefore act as activators or repressors depending on there immediate environment.

1.3 DNA is associated with chromatin

In higher organisms, DNA is generally tightly associated as a nucleoprotein complex in chromatin. Chromatin associated DNA is organised into nucleosomes: regular repeating structural units consisting of 166bp repeats of DNA wound around a histone octamer approximately 1.8 times to form the 'beads-on-a- string' 10nm fibre. This basic nucleosomal core is composed of the histones H2, H3, H4, H5 and undergoes further condensation, which is mediated by histone H1 in the nucleosomal linker to form the 30nm fibre. Further condensation then proceeds to form higher order condensed structures such as is seen for heterochromatin. However, the molecular framework of such structures is not well understood.

Gene transcription *in vivo* is dependent on the interaction of transcription factor proteins with chromatin associated DNA, in contrast to *in vitro* naked DNA

transcription from linear templates, which are generally not chromatin associated. In this state the DNA is less accessible to transcription factor protein interaction and, as a consequence, unable to initiate transcription. In contrast to the bulk of chromatin, transcriptionally active regions of DNA acquire a more relaxed, or unwound, structure with the chromatin framework.

It has been shown that structural relaxation surrounding transcriptionally active genes also correlates with a general increase in the sensitivity to nuclease digestion and that small 15-200bp regions exhibit extreme nuclease sensitivity. The most commonly used method of mapping nuclease hypersensitive sites (DHS) is DNAse1 digestion of nuclei followed by indirect labelling of the resulting purified double-stranded DNA. These DHS sites (Wu *et al.*, 1979a & b) are generally associated with active promoter and enhancer elements (reviewed by Gross and Garrard, 1988). Consequently, mapping DNAse1 sensitivity is a useful method of identifying gaps in the nucleosomal array and therefore the location of potentially transcriptionally active genes. However, the resolution of the technique, particularly with fragments larger than 1kb, is less than 50bp. A more accurate method for mapping the fine structure of hypersensitivity sites is offered by DNAse1 footprinting. This technique involves running the footprinting reactions alongside DNA sequencing tracts of the footprinted region which allows the mapping of individual proteins or protein complexes to specific recognition sequences (recently reviewed by Plumb and Goodwin, 1991).

1.4 Biology of Hypersensitivity Sites

DHS sites are ubiquitous among eukaryotes, being found in plants, fungi, viral and episomal genomes and animals. DHS sites may be conveniently grouped into two classes, constitutive or inducible.

Constitutive DHS sites

These sites are often associated with the promoter regions of low basal expressing genes such as housekeeping genes. These DHS sites are also found in promoter regions poised for transcriptional activation, such as the mouse metallothionein gene (Senear & Palmitter, 1983) and the heat shock genes (Wu, 1980; Elgin, 1988). The presence of the constitutive DHS sites are independent of gene expression.

Inducible DHS sites

Inducible DHS sites precede gene expression and may persist long after the initial stimulus. The nucleosomal displacement associated with induction of a particular DHS site is usually concomitant with the activation of a linked gene. Inducible DHS sites are classically associated with steroid inducible genes (Jantzen *et al.*, 1987; Kaye *et al.*, 1986) but a variety of other examples exist including, lipopolysaccharide-inducible DHS sites in κ immunoglobulin gene enhancer of pre-B-cells (Parslow & Granner, 1983) and light and oxygen tension-inducible DHS sites in plant genes (Paul *et al.*, 1987; Kaufman *et al.*, 1987).

In addition to inducible DHS sites within a given cell type, DHS sites may also be induced in a tissue specific manner. An example of this is the chicken lysozyme gene locus. Interestingly this gene locus contains 10 DHS sites, two of which are macrophage specific and one is oviduct specific and hormone inducible (Theisen *et al.*, 1986). The functions of the other sites vary but include constitutive and inducible DHS associated sites (Theisen *et al.*, 1986; Steiner *et al.*, 1987; Baniahmad *et al.*, 1987). This illustrates that DHS sites may be induced in a temporal or spatial manner within the same gene locus. In addition to the examples given above, tissue specific

DHS sites also appear transiently in promoters during differing stages of development as in the globin genes (reviewed by Evans *et al.*, 1990).

Presumably the reorganisation in nucleosomal positioning, as assessed by an increase in DNA hypersensitivity is necessary, but not sufficient, for transcription activation which may also be reliant on limiting transcription factors to achieve optimal expression. The functional relevance of DHS sites to transcriptional regulation is discussed below. In addition to there putative involvement in transcriptional regulation they have been associated with other functions including chromosomal replication, recombination and segregation (DePamphilis, 1988 & 1993; Wolffe & Pruss, 1996).

1.5 The inhibitory role of nucleosomes on transcriptional regulation

The 'Pre-emptive' Model

A conceptually simple model for the repressing effect of histones on transcription has been proposed which involves the direct blocking of access to *cis*-acting sites within promoter and enhancer regions by histone octamers in resting cells. This may be by nucleosomal positioning or a higher-order structure, masking sites which would otherwise be exposed to *trans*-acting transcription factors. Chromatin reassembly, or gene activation, would therefore occur during replication when the promoter and enhancer regions are partially exposed by the replication fork, and allowing direct competition of *trans*-acting factors and nucleosomes. In this way the *trans*-acting factors must co-operate to create a sufficiently stable complex which prevents nucleosome assembly.

The factors vital for stabilising an active transcriptional complex may be differentiated from those required for transcription initiation itself. The maintenance of nucleosomal free regions is not fully understood, although it has been proposed that

specific *cis*-acting structures within DNA aid in this process. For example, AT-rich *cis*-acting scaffold associated regions (SAR) have also been identified as important determinants in the maintenance of chromatin structural loops and transcriptionally active domains (Laemmli *et al.*, 1992). These sites may be important in the binding of proteins which associate with histone H1 and maybe other histone proteins to control nucleosome assembly or alternatively may represent sites for topoisomerase enzymes which can alter torsion stress of the DNA over large distances, and therefore chromatin structure. These mechanisms are not discussed in detail here but are reviewed in Laemmli *et al.*, 1992.

Alternatively, the folding of DNA during nucleosomal assembly may bring together two distant regulatory elements closer together in space (reviewed by van der Vliet & Verrijzer, 1993). Such DNA looping(bending) may occur through homo- or hetero- dimerisation of widely spaced *trans*-acting factors, such as Sp1 (Su *et al.*, 1991), progesterone receptor (Theveny *et al.*, 1987) and GATA (Fong & Emerson, 1992). This may restrict the formation of repressive histone/DNA contacts and potentiates the transcription process, as with the *Xenopus* vitellogenin B1 gene promoter, where DNA bending allows the interaction of the promoter with oestrogen receptor binding sites (Schild *et al.*, 1993) .

Support for this model has come from a number of chromatin assembly experiments. Pre-initiation complex formation within the adenovirus type-2 major late promoter (Ad2-MLP) only occurs before chromatin assembly (Knezetic *et al.*, 1988). Furthermore, pre-incubation of Ad2-MLP with TFIID, a component of the pre-initiation complex, protects the promoter from nucleosomal assembly (Workman & Roeder, 1987). Interestingly, if TFIID is added simultaneously with histone proteins in the presence of the upstream activator factor (USF) protein a transcriptionally active complex is formed (Workman *et al.*, 1988; Workman *et al.*, 1990). The mechanism of this interaction is unknown but may reflect an increase in overall stability mediated by

the presence of the USF protein. Similarly experiments have also shown that *trans*-acting tissue specific transcription factors bound close to the promoter region of the β-globin gene are able to exclude nucleosomes so long as the nucleosomes were not added prior to the transcription factors (Emerson & Felsenfeld., 1984). This suggests that the primary role of *trans*-activators may be to stabilise the basal transcription machinery, resulting in elevated levels of gene expression, which is mediated by the interaction of *trans*-activators with both the transcription machinery and nucleosomes.

Dynamic Model

In the 'dynamic' model *trans*-acting factors can directly displace/alter nucleosomal positioning and consequently is not dependent on DNA replication. An example of the 'dynamic' model is illustrated by nucleosomal reorganisation following hormone induction of a glucocorticoid-response element (GRE) within the mouse mammary tumour virus (MMTV) long terminal repeat (LTR). Prior to induction nucleosomes are precisely located along the LTR. Upon induction, the GRE region (-50 to -250) is bound by its hormone activated receptor, glucocorticoid receptor (GR), which displaces a single nucleosome (Richard-Foy & Hager, 1987). This in turn allows access of another transcription factor, nuclear factor 1 (NF1) and possibly other factors, to this region and subsequent activation of the promoter (Cordingley *et al.*, 1987). This response is very rapid with transcriptional activation taking place within a minute of exposure to the hormone and the entire process is reversed upon removal of the hormone (Zaret & Yamamoto, 1984). In view of this transient nature, presumably displacement of the original inducing *trans*-acting factor (GR) is functionally important in controlling reversion of the local chromatin structure. This displacement may be mediated by the binding of other higher affinity *trans*-factors to the promoter, such as NF1.

Similar mechanisms appear to be important in the induction of the Saccharomyces cerevisae *PHO5* gene promoter. The *PHO5* gene promoter is activated in low phosphate conditions and contains functionally important upstream activator sequences (UAS) adjacent to the promoter. The nucleosomal positioning over the promoter and UAS regions have been determined (Almer & Horz, 1986). One of the UAS sequences contain a DHS site in both inducing and non-inducing conditions. In low phosphate conditions the product of the PHO4 gene and PHO2 gene bind to the UAS and displace four nucleosomes exposing the TATA box and a PHO4 binding site (Almer *et al.*, 1986; Almer & Horz., 1986). Overexpression studies revealed that displacement of the UAS is dependent on both PHO2 and PHO4 proteins, since the presence of just one protein was insufficient (Fascher *et al.*, 1990).

In contrast to the assumption that all DHS regions are free of nucleosomes, the studies with the MMTV LTR also suggest that increased transcription factor binding is due to less stringent binding rather than total displacement of the nucleosomal structure. This is supported by experiments cross-linking the DHS region induced by GRE/GR complex formation in the MMTV LTR and finding it still contains nucleosomal histones (Bresnick *et al.*, 1992). However in contradiction to these observation, no histone proteins could be cross-linked from the 5' DHS sites of the *Drosophila hsp70* genes (Karpov *et al.*, 1984) or the 5' DHS site of the chicken β-globin gene (McGhee *et al.*, 1981). Consequently, there may be more than one mechanism controlling the access of transcription factors to DHS regions. In this respect it seems probable that the mechanisms of repression mediated by chromatin structure may involve aspects of both the pre-emptive and dynamic models, depending on the promoter and the local environment.

Genetic analysis carried out in yeast has illustrated the importance of chromatin structure as a general repressor of transcription. A single copy of the yeast H4 gene was placed under the control of a galactose-inducible promoter and the cells grown on

glucose media. The subsequent depletion of H4 led to an activation of repressed genes including *PHO5*, *CYC1* and *GAL1* (Han *et al.*, 1988; Han & Grunstein., 1988). The induction was independent of upstream activator sequences. However, the level of induction was significantly lower then following normal glucose mediated induction of the gene. Consequently, the data show that nucleosomes can block transcription from inducible promoters, however, full promoter activation requires other factors in addition to nucleosomal de-repression. Promoters not affected by nucleosomal loss may represent promoter regions constitutively free of nucleosomes, possibly because of the presence of activators.

Furthermore genetic analysis, again in yeast, has identified non-histone proteins which appear to influence transcription by altering chromatin structure. Mutations in the negative regulators SIN1 (a member of the high mobility group (HMG) 1 family) and SIN2 (histone H3) were shown to alleviate transcriptional constraints previously mediated by mutations in the positive regulating SWI proteins (SWI1, SWI2 and SWI3). It has been suggested that the SWI proteins may act to relieve inhibitory effects of chromatin structure (Peterson & Hershkowizt, 1992; Hirschhorn *et al.*, 1992)). Interestingly, a recent study has shown that a SWI/SNF complex may bind DNA and introduce positive supercoils, consistent with a role in chromatin structure remodelling and transcriptional activation (Quinn *et al.*, 1996). The HMG family are non-histone chromosomal associated proteins, which bind to DNA via a HMG-box domain, and appear to be important in transcriptional control, although there functions are not well understood (reviewed by Landsman and Bustin, 1993).

1.6 General properties of active chromatin

Active chromatin is characterised by a general depletion of histone H1, hypomethylation and histone acetylation. The importance of these three characteristics are discussed below;

Histone H1

The linker histone H1 can repress transcription by all three RNA polymerases. Crosslinking experiments have demonstrated that H1 is depleted in active as opposed to inactive genes (Nacheva *et al.*, 1989; Bresnick *et al.*, 1992). H1 has been implicated in the reorganisation of nucleosomal positioning and/or the formation of higher-order chromatin structures (Felsenfeld & McGhee, 1986), although the mechanism of action is unclear.

Methylation

In vertebrates there is an uneven distribution of the dinucleotide CpG. The bulk of DNA is depleted of CpGs with the cytosines within this dinucleotide being methylated (Bird, 1987). By contrast, CpG rich stretches of DNA or "CpG islands" are often associated with the 5' region of housekeeping and tissue specific genes which are generally hypomethylated (Bird, 1986; Gardiner-Garden & Frommer, 1987) although this general rule is not true for the genes on the inactivated X-chromosome (Wolf & Migeon, 1985).

DNA methylation is generally associated with inactive chromatin, and as a consequence has been implicated as a epigenetic element in the repression of gene activity. Pre-established and *de novo* DNA methylation patterns are maintained by the

DNA methyltransferase (MTase) enzyme. In contrast to *in vivo* some autosomal genes are methylated and consequently inactivated in cell culture including the thymidine kinase (Wise & Harris, 1988) and metallothionein genes (Compere & Palmitter, 1981). The importance of DNA methylation in transcriptional repression has been questioned due to the lack of appreciable methylation in either the *Drosophila* or *Caenorhabditis elegans* genome. However, homozygous MTase null mice do not survive past mid-gestation (Li *et al.*, 1992) suggesting an importance of methylation in higher eukaryotes.

The mechanism by which DNA methylation results in transcriptional repression is poorly understood. It has been shown that some transcription factors are unable to bind to methylated DNA, and it has therefore been proposed that direct methylation of their appropriate binding sites would be a mechanism of repressing transcription (reviewed by Tate & Bird, 1993). However, transcriptional repression mediated through methylation, of some promoters do not appear to rely on methylation of particular binding sites, an example being the γ-globin promoter (Murray & Grosveld, 1987). A number of proteins have been identified which specifically bind to methylated DNA, including MeCP1, MeCP2, MDBP1, MDBP2, which presumably may compete with *trans*-activators and thereby offer another mechanism of transcriptional repression (reviewed by Tate & Bird, 1993). The function of these proteins is unclear although MeCP1 has been shown to repress methylated promoters *in vitro* (Boyes & Bird, 1991).

Since greater than 80% of methylated CpG residues are packaged into H1-containing nucleosomes (Ball *et al.*, 1983), methylation of DNA might increase the binding affinity of histone H1 for this region possibly resulting in reorganised nucleosomal positioning and/or the formation of higher-order chromatin structures mediated by H1 (Felsenfeld & McGhee, 1986). It has been observed that methylated plasmid DNA is less sensitive to nucleases than its hypomethylated counterpart if pre-incubated with histone H1 (Higurashi & Cole, 1992). Furthermore, *in vitro*

transcription experiments have identified a methylated CpG binding protein, believed to be histone H1, responsible for repression of the vitallogenin promoter (Jost *et al.*, 1991).

Acetylation

The core histone proteins (H2, H3, H4 and H5) contain long extended N-terminal hydrophilic tails which are not thought to be important in histone core particle assembly (Whitlock & Simpson, 1977; Ausio *et al.*, 1989) but rather in histone/protein (Johnson *et al.*, 1990) or histone/DNA (Hill & Thomas, 1990) interactions. The positive charges present in the N-terminal tails may be neutralised by the addition of acetyl groups to the ε-amino group of specific N-terminal lysine residues (hyper-acetylation). It has been proposed that histone acetylation weakens N-terminal interaction with the negatively charged DNA backbone thereby mediating unfolding of chromatin fibers (McGhee & Felsenfeld, 1980) and there is also evidence suggesting that higher order chromatin folding is also affected (Annunziato *et al.*, 1988). Furthermore, histone hyper-acetylation has been shown to allow greater transcription factor access to nucleosomal DNA (Lee *et al.*, 1993). All these studies are consistent with acetylation being a positive regulator of transcription.

Hyper-acetylation of H3 and H4 (Cahal *et al.*, 1980; Waterborg & Matthews, 1984) has been associated with induced levels of transcriptional activity (Johnson *et al.*, 1987; Chan *et al.*, 1988; Hebbes *et al.*, 1988). On the strength of these studies, is has been postulated that transcriptionally active genes are located in hyperacetylated chromatin regions and that the acetylation sites are likely to be used in non-random fashion (Cahal *et al.*, 1980; Waterborg & Matthews, 1984; Chambers & Shaw, 1984; Chicoine *et al*, 1987). Interestingly, deletion of residues 4-28 in H4 results in repressing transcription from inducible promoters while deletion of 4-30 in H3 has the

opposite effect and increases expression (Durrin *et al*, 1991). This is consistent with histone acetylation being important in the control of specific gene expression.

1.7 Enhancers

In addition to *cis*-binding sequences within the promoter of genes transcription factor binding sites may also be found up- or downstream to the promoter region. These distal regions are often termed enhancer or silencer (repressor) elements depending on their effect on transcription. In a similar manner to DHS sites enhancers may be constitutive, tissue specific or inducible in nature. Classically, enhancers are characterised by their ability to activate transcription in an orientation independent manner, and may be found considerable distances upstream of the transcriptional initiation site, as illustrated by the locus control region (LCR) in controlling globin expression (reviewed by Evans *et al.*, 1990). In addition, enhancers may also be found downstream of genes (O'Prey *et al.*, 1993), within introns (Haung *et al.*, 1993) and within coding exons (Tognoni *et al.*, 1985).

There have been two proposed mechanisms for how enhancers may function; the first more classical model results in increasing the rate of transcription of a linked gene, and has been termed the graded model (Lewin, 1990, 1994). This might be achieved by the interaction of the enhancer directly with the initiation complex, the intervening DNA looping out to allow the interaction. In support of this model the Sp1 transcription factor has been demonstrated to form multimers and stabilise DNA loops (Su *et al.*, 1991).

The second model, designated the binary model (Walters *et al.*, 1995), increases the probability that a promoter will achieve and maintain an active state and consequently an observed increase in transcription is due to an increase in the proportion of cells expressing as opposed to an increase within cells. The model proposes that

enhancers act to suppress position effect variegation (PEV) (Walters *et al.*, 1996) which is the suppression of genes due to the surrounding heterochromatin and correlates with an loss in DNAse1 hypersensitivity and an increase in methylation (PEV is reviewed by Karpen, 1994). DNA looping may also be important in this function to bring active genes under the control of heterochromatin regions. Enhancers may also act as origins of DNA replication (DePamphilis, 1988), where the activity is mediated by interaction of transcription and replication factors (DePamphilis, 1993). Enhancers elements are also thought to be important during murine development after the 2-cell embryo stage (reviewed by Majumder & DePamphilis, 1995).

1.8 Positional effects and higher order chromatin structure

Interestingly, enhancers may act over great distances yet function on specific promoters, in view of their orientation and position independent manner it might be expected that they would act promiscuously on surrounding promoters. Since this is not usually observed, except with chromosomal translocations involved in oncogenesis, there are presumably mechanisms which usually repress distant enhancers. A possible mechanism for this might be the nature of the surrounding chromatin structure. In higher eukaryotes, chromatin is organised into discrete active (euchromatin) and inactive (heterochromatin) regions which are believed to exist as DNA loops anchored to the nuclear scaffold (reviewed by Gasser and Laemmli, 1986). It has been suggested that genes brought into heterochromatin regions may be transcriptionally suppressed, by PEV (reviewed by Karpen, 1994), although the precise mechanism of this repression is unknown.

Analyses of viral gene expression following Moloney leukaemia proviral insertions in the mouse identified positional effects in expression mediated by silencer and enhancer activities at the site of integration (Jaenish *et al.*, 1981), although the

silencing effects may also be due to the induction of *de novo* methylation. By contrast, subsequent studies have been used to identify and investigate the sequences important in locus control regions (LCRs) which show position independent expression (PIE), a well studied example being the human β-globin gene locus (Grosveld *et al.*, 1987; Talbot *et al.*, 1989).

The DNA loop regions are believed to be flanked with nuclear targeting sequences known as scaffold associated regions (SARs) or matrix associated regions (MARs), which were identified as regions of DNA which preferentially associated with detergent-extracted nuclear remnants (termed nuclear matrices or scaffolds) (Mirkovitch *et al.*, 1984). SARs are believed to protect defined regulatory domains from repressive/enhancing effects mediated by neighbouring chromatin and are found in a number of gene loci. The human β-globin cluster has been shown to contain eight SARs (Jarman & Higgs, 1988) while the chicken lysozyme gene is flanked 5' and 3' by SARs which correspond to the DNAse1 hypersensitive region (Phi-Van & Stratling, 1988). SARs flanking both homologous and heterologous promoters have been shown to stimulate transcription by 10-20 fold (Klehr *et al.*, 1991). In addition, the 5' SAR of the chicken lysozyme gene has been shown to impose limited PIE in stably transformed cells (Stief *et al.*, 1989). SAR regions have been shown to have no effect in transient transfecting assays, consequently it is possible that the increase in transcription observed in stable transfection is mediated by some association with the chromatin framework. Consistent with this is that SAR regions are typically targets for topoisomerase I and II enzymes which are associated with controlling DNA torsion stress and chromosome condensation (review by Laemmli *et al.*, 1993). Similar SAR like structures, designated Scs elements, have also been identified in the *Drosophilia hsp*70 promoter (Udvardy *et al.*, 1985) which may confer PIE to integrated constructs (Kellum & Schedl, 1991).

1.9 Transcription factor families

Transcription factors are modular proteins containing domains for DNA binding, dimerisation and transcriptional activation or repression. Transcription factors have been characterised into 'families' by similarities in there structural domains. This is a convenient way of grouping proteins allowing a framework of transcription factors to be built which are related through their DNA recognition motif. This at least helps to rationalise what is an immensely complex interplay of multiple transcription factors directing the proliferation and differentiation of diverse cell types.

(I) Structural domains important in DNA recognition

Prokaryotic Helix-turn-helix (HTH) proteins

The first helix-turn-helix (HTH) structural motifs described were from the crystal structures of the bacteriophage λ Cro protein (Anderson *et al.*, 1981), λ repressor protein (McKay & Steitz, 1981) and the *Escherichia coli* CAP protein (Pabo & Lewis, 1982). These studies demonstrated a conserved recognition sequence consisting of an α-helix followed by a turn and then a second α-helix. Subsequent studies identified a large family of HTH DNA recognition motif prokaryotic transcription factors including Lac repressor, Trp repressor. A similar structural motif is observed in the bacterial LexA protein except that the turn region is larger than that seen in other HTH proteins.

In contrast to other structural motifs (Zinc-fingers, leucine zipper, etc) the HTH motifs are unable to function as an independent domain and require a larger DNA binding domain to function. The λ repressor-operator has been used extensively to characterise the specific contacts of the HTH domain made with the DNA recognition sequence. It should also be noted that contacts outside the HTH region are important in

specificity of DNA recognition (Jordan & Pabo, 1988). The HTH proteins bind as dimers with each monomer recognising a half-site.

Homeodomain

The homeodomain was identified as a structural motif present in proteins important in *Drosophila* development and shows similarities with the HTH family (McGinnis *et al.*, 1984; Scott & Weiner, 1984). Subsequently, homeodomains were found to be more widespread in eukaryotic gene regulation. The 60 residue homeodomain forms a folded structure which may bind DNA independently. The crystal structures of the *Drosophila* engrailed and yeast MAT α2 homeodomain-DNA complex have been determined. The engrailed homeodomain contains three α-helices, helix 1 and 2 are packed antiparallel to each other with helix 3 perpendicular to the first two helices (Kissinger *et al.*, 1990), a similar structure is also seen for MAT α2 (Wolberger *et al.*, 1991). Helix 2 and helix 3 represent the HTH unit, originally identified in prokaryotic proteins, although the helices are significantly longer. This difference is believed to alter the way the prokaryotic HTH and eukaryotic homeodomains line up and contact the DNA. The evolutionary significance of the differences in DNA recognition between prokaryotic HTH and eukaryotic homeodomain proteins is unclear.

The homeodomain may act as a separate domain, binding DNA in the absence of the rest of the protein. However, residues outside the homeodomain are also important in DNA recognition. A sub-family of homeodomain proteins which include pit-1 and Oct-1, contain a conserved 65-75 residue region, termed the POU-specific domain, located on the N-terminal side of the homeodomain (Herr *et al.*, 1988). The POU domain makes DNA contacts adjacent to those made by the homeodomain domain.

Winged Helix-turn-Helix

This motif is based on the structure of the hepatocyte nuclear factor (HNF)3/fork head DNA binding domain, where two loops stick out from either side of the HTH motif and make DNA contacts, primarily with the phosphate backbone (Clark *et al.*, 1993). Subsequent studies have identified similar motifs in the heat shock transcription factor family, although only one loop is observed (Harrison *et al.*, 1994). Similarly, the previously unclassified ETS domain DNA-binding motif has recently been shown to contain a winged HTH motif, in the Ets family members Fli-1 (Liang *et al.*, 1994) and Ets-1 (Donaldson *et al*, 1994).

Paired Box motif

The paired box Pax motif was originally identified in three Drosophila segmentation genes (Bopp *et al.*, 1986) but has subsequently been found in human, mouse, chicken and zebrafish (reviewed in Deutsch & Gruss, 1991). Sub-groups of the Pax motifs have been identified which contain a conserved octapeptide or homeodomain motifs outside the paired-box domain. There are an ever increasing family of Pax genes (designated Pax 1 to Pax 8 in mouse) which are important in development (reviewed in Gruss & Walther, 1992).

Zinc binding domains

Members of the zinc-finger family of transcription factors have been implicated in many roles of eukaryotic gene regulation including development, differentiation and proliferation signals, basal transcription regulation and oncogenesis. The family was originally formed by the identification of a 30 residue zinc-finger motif Cys-X_{2to4}-Cys-

X_{12}-His-X_{3-5}-His (Cys_2-His_2) from the Xenopus transcription factor III A (TFIIIA) protein (Miller *et al.*, 1985) which binds to the internal promoter of the 5S rRNA genes. Subsequent studies have identified other cyteine-rich domains which co-ordinate Zn atoms facilitating DNA recognition, which may divided into four structural classes. The first class is the TFIIIA like zinc-finger family. The structure of the TFIIIA zinc-finger domain has been determined by 2D NMR (Parraga *et al.*, 1988) showing that the motif contains an anti-parallel β-sheet and an α-helix. The two cysteines present within the β-sheet and the two histidines in the α-helix co-ordinate a central zinc ion holding the secondary structures together forming a compact globular domain.

The second class of zinc-fingers is the, Cys_2-Cys_2, structural motif originally identified in members of the GATA family of transcription factors. The proteins have a central core of two irregular antiparallel β sheets and a α-helix (Omichinski *et al.*, 1993) stabilised by four cysteines co-ordinated around a central zinc-ion .

A third class contains six cysteines and two Zn ions per binding element which has been termed a, binuclear cluster, identified in the yeast GAL4 protein (Marmorstein *et al*, 1992) and PPR1 (Marmorstein & Harrison, 1994).

The fourth class includes the steroid receptor proteins which act as second messengers for transducing hormonal signals including steroid, vitamin D, retinoids and thyroid hormones. The proteins have distinct DNA-binding, *trans*-activation and ligand binding domains. The DNA-binding domain possesses eight conserved cysteine residues which may co-ordinate with zinc to form a distinct structural motif from the (Cys_2-His_2) zinc-fingers described previously for TFIIIA (Frankel and Pabo., 1988). The crystal structure of a glucocorticoid receptor complex suggests that the protein binds as a dimer to the recognition sequence (Luisi *et al.*, 1991).

In addition, there are two other relatively recently identified Zn binding motifs which have not been demonstrated to bind specifically to DNA. Although the domains may be important in protein/protein interactions they are mentioned since the precise

function of the motifs still remains unclear. The LIM family of proteins are an interesting new family of proteins identified which possess novel cysteine rich zinc-binding domains (reviewed by Sanchez-Garcia & Rabbitts., 1994). The LIM domain was originally identified in the protein products of three genes: the *lin-11* (Freyd *et al.*, 1990) and *mec-3* (Way & Chalfie, 1988) genes from *C. elegans* and *ISL1* gene (Karlsson *et al*, 1990) from rat. Subsequently LIM proteins have been sub-divided into three groups; the first group, which include Mec3, Lin-11 and ISL3 possess LIM domains associated with a homeodomain (LIM-HD), the second, LIM-PK, contains LIM domains linked to a protein kinase domain, and the third possess just LIM domains and have therefore been designated LIM-only proteins (reviewed by Sanchez-Garcia & Rabbitts, 1994). LIM proteins have been identified in a range of eukaryotic organisms including rat, chicken, quail, Xenopus, *Drosophila*, mouse and human and have been implicated in developmental regulation.. Interestingly, LIM proteins have also been associated with human leukaemias. The human RBTN1 and RBTN2 LIM domain proteins have been implicated in the progression of some childhood human T-cell acute leukaemias involving chromosomal translocation within 11p15 and 11p13 receptively, which is where the human genes are located (Sanchez-Garcia & Rabbitts, 1993). Although the LIM domain motifs appears to be important in protein/protein interaction analysis of the primary sequence suggests that it might form two folds that are similar to GATA zinc-fingers, consequently it is possible that be important in DNA recognition under the appropriate conditions.

The second motif has been designated the ring-finger motif (Freemont, 1993). The NMR structure of the ring-finger has been identified (Barlow *et al.*, 1993). Interestingly, the breast and ovarian cancer susceptibility gene, BRCA-1 contains a zinc ring-finger motif (Miki *et al.*, 1994) although the function of this domain remains unclear.

The members of zinc-finger family of transcription factors are summarised in table 1.1 below;

Table 1.1: Selected families of zinc-finger domains

Zinc-domain type	Approximate consensus sequence	Function
Cys_2-His_2 (TFIIA)	C-X_{2-4}-C-X_{12}-H-X_{3-5}-H	Nucleic acid binding
Cys_2-Cys_2 (GATA-1)	C-X_2-C-X_{17}-C-X_2C	DNA binding
Cys_8 (Steroid thyroid receptor)	C-X_2-C-X_{13}-C-X_2C-X_{15}C-X_5-C-X_{12}-C-X_4-C	DNA binding, oligomerisation
Cys_6 (GAL4)	C-X_2-C-X_6-C-X_6-C-X_2-C-X_6-C	DNA binding
Cys_2HisCys$_5$ (LIM domain)	C-X_2-C-X_{17-19}-H-X_2-H-X_2C-X_2-C-X_{16-20}-C-X_{2-3}-C	Protein/Protein interaction, DNA binding ?
Cys_2HisCys$_4$ (Ring-finger)	C-X_2-C-X_{9-27}-C-X_{1-3}-H-X_{2-3}C-X_2-C-X_{4-48}-C-X_2-C	Protein/Protein interaction ? DNA binding ?

Adapted from Berg & Shi, 1996

Abbreviations: C; Cys, H; Hys, X; other amino acids.

Leucine zipper and helix-loop-helix

The leucine zipper motif was first identified as a dimerisation motif in the CCAAT/enhancer binding protein (C/EBP) (Landschulz *et al.*, 1988). Dimerisation forms two parallel α-helices in a coiled-coil arrangement (O'Shea *et al.*, 1991). Leucine zippers are characterised by a heptad repeat of leucines over a 30-40 residue segment with a positively charged 'basic region', of 30 residues, located at its N-terminus important in DNA binding (Landschulz *et al.*, 1988), the structural motif has therefore been termed as basic leucine zippers, or bZIP (Shuman *et al.*, 1990). Although the basic region is rich in arginine and lysines it also contains other residues which appear to influence the specificity of binding to the DNA-recognition sites (Agre *et al.*, 1989).

Members of the bZip family include the C/EBP family, AP-1 family members, Fos, Jun, ATF/CREB family and the yeast factor GCN4 (Harrison, 1991).

A related structural motif to the leucine zipper is the helix-loop-helix motif (HLH) which forms a α-helix followed by a loop and then another α-helix. Similarly to the bZip proteins the HLH sub-family have a basic region adjacent to the dimerisation motif which contacts DNA. Members of this sub-family include the Myc and Max proteins (Blackwood and Eisenman., 1991).

High Mobility Group (HMG) proteins

The chromatin associated HMG non-histone proteins are divided into three unrelated subclasses HMG1/2, HMG14/17 and HMG-I/Y. The HMG1/2 proteins contain a HMG1 box motif important in DNA recognition which have been found in a number of proteins. Although the biological functions of the HMG proteins are largely unknown it has been suggested that the HMG1/2 proteins are important in transcription, DNA replication and recombination. The properties of the HMG protein are reviewed by Landsman & Bustin, 1993 and Lilley, 1992.

Other DNA-binding structural motifs

There are a number of other proteins which bind to DNA in a sequence specific manner but do not fit into one of the families mentioned above, these include the myb, CTF/NF1, AP2. This suggests that there may be other structural motifs which have yet to be identified. Nevertheless, the vast majority of transcription factors identified may be grouped into the above families suggesting that evolutionary selection has resulted in a limited number of structural DNA-recognition motifs.

(II) *Trans*-activation domains

Gene-specific regulation is controlled by a diversity of transcriptional activator and repressors which have separable domains for directing the protein to its appropriate binding site, by contrast the structures of transcriptional activator and repressor domains are poorly understood.

Many DNA binding factors act as transcriptional activators containing domains of approximately 30-100 amino acids important in mediating this function. The first activation domains to be characterised were from the yeast DNA binding proteins GAL4 (Ma & Ptsashne, 1987) and GCN4 (Hope & Struhl, 1986). These studies showed the regions to be similar in their general negative charge (acidic) and ability to form amphipathic α-helical structures. Acidic domains or "acid blobs" have been identified in other *trans*-acting DNA binding proteins including *c-Jun* (Bohmann *et al.*, 1987), *RelA* (Blair *et al.*, 1994), glucocorticoid hormone receptor (Hollenberg & Evans, 1988) and Sp1 (Courey & Tjian, 1991). Similarly, other structural motifs have been isolated which may act as activator domains, these include proline rich domains, CTF/NF1 (Kim & Roeder, 1993), Oct-2 (Tanaka *et al.*, 1994a, 1994b) and glutamine-rich domains, Sp1 (Gill *et al.*, 1994) and Oct-2 (Tanaka *et al.*, 1994a, 1994b). *Trans*-activator proteins may possess more than one *trans*-activation domain which may be 1) of the same type, the yeast GAL4 protein has two acidic domains or 2) of different structure, Oct-2 has a proline and glutamine domain. In some cases activation domains may be masked requiring an induction event, such as hormone binding or phosphorylation, to alter the conformation of the protein (Tasset *et al.*, 1990).

Consistent with the modular nature of proteins chimeric proteins may be constructed containing heterologous DNA binding or *trans*-activation domains which maintain the appropriate function, an example is the fusion of the activation domain of VP16 to the DNA binding domain of GAL4 to create a GAL4-VP16 fusion protein.

The GAL4 is still able bind to its upstream activation sequence (UAS) binding site and activate transcription due to the VPI6 activation domain (Cousens *et al.*, 1989).

Trans-activators are believed to mediate their effects by binding to the basal transcription machinery, resulting in either an increase in rate of initiation or stabilising the pre-initiation complex once formed. Activators may bind directly to basal transcription factors including TBP, (Xu *et al.*, 1993) (Emili *et al.*, 1994), TFIIB (Choy & Green, 1994), TFIIA (Ozer *et al.*, 1994) and TFIIH (Xiao *et al*, 1994). Despite these direct interactions transcriptional activation also requires co-activators (mediators or adapters) which include some of the TAFs of the TFIID complex (reviewed by Gill and Tjian, 1992).

(III) Repression domains

In addition to the negative effects of the chromatin framework transcription repression may also be mediated by *trans*-acting transcription factors. The mechanism of repression by *trans*-acting factors may be active, where the repressors directly downregulates transcription. Examples of these type of proteins include WT1 (Madden *et al.*, 1991), *Drosophila even-skipped* (Ham *et al.*, 1992), Kruppel (Licht *et al.*, 1993) and the human Kruppel related protein, YY1 (Shi *et al.*, 1991). Although there are clear amino acid similarities between the repression domains of active repressors, they are generally rich in alanine, glutamine and/or proline and depleted of charged amino acids. Since these properties are reminiscent of the modular nature of activation domains observed within *trans*-activator proteins, it is perhaps not surprising that active repressors are believed to function by directly contacting the basal transcription machinery. For example, the *even-skipped* repressor prevents TFIID binding to the promoter (Austin & Biggin, 1995). Presumably, in addition to the general transcription

factors such as TFIID, TFIIB, TFIIA etc, the TAFs and other co-activators may be targets for active *trans*-acting repressors.

Alternatively, repression may occur by a passive mechanism, where negative regulators either block binding of *trans*-activator or basal factors or form inactive heterodimers. An example of direct competition is the repression of retinoic acid induced transcription of the human ostoecalcin gene by activator protein 1 (AP1) through direct competition with the retinoic acid receptor for an overlapping DNA binding sites. Altered dimerisation may also result in repression, an example is the dimerisation of the bZip proteins, C/EBP and the C/EBP homologous protein (CHOP). CHOP has proline substitutions in its basic domain which prevent it binding DNA. As a consequence overexpression of CHOP during adipocyte maturation inhibits C/EBP activity (Ron & Habener, 1992).

Interestingly, a recent study has shown that the negative regulator adipocyte lipid-binding protein1, ALBP1 contains a domain which contains carboxypeptidase (CP) activity, vital for the repression function of the protein (He *et al.*, 1995). The authors suggests a novel mechanism of repression exists where the CP activity enzymatically cleaves the proteins involved in transcription. Consequently, there appears to exist a plethora of positive and negative acting transcription factors which act co-ordinately to modulate the level of transcriptional regulation.

1.10 Mechanisms for modulation of transcription factor activity

Dimerisation

Dimerisation or higher complex formation is a common theme between transcription factors and gives another level of complexity to transcriptional control. A number of transcription factor families have been defined by their characteristic

dimerisation interfaces. These include helix-loop-helix, (myc, max, MyoD) (Amati & Land, 1994), leucine zipper, (C/EBP, Fos, Jun) (Lamb & McKnight, 1991) and Rel (NF-kB and dorsal) (Blank *et al.*, 1992). A consequence of heterodimerisation is that the protein components may have altered DNA binding and activation properties which, depending on the dimerisation partner, may also target the protein complex to a different set of promoters.

In addition to dimerisation between families proteins may also form complexes with unrelated proteins. For example the c-Myc protein may also bind to the zinc-finger protein YY1. However, dimerisation between proteins is not promiscuous since the bZip protein c-Fos can form heterodimers with the AP1 family member c-Jun protein but not form homodimers or dimers with the C/EBP bZip containing family.

The simplest scenario resulting from different heterodimers is altered DNA binding specificity. The nuclear hormone receptors contain a central DNA-binding domain which targets the receptor to hormone response elements (reviewed by Yamamoto, 1985). The 9-cis retinoic acid X receptor (RXR) may bind a variety of other monomers including the all-trans retinoic acid receptor (RAR), the thyroid hormone receptor (T_3R), the vitamin D3 receptor (VDR) and fatty acid/peroxisome proliferator-activated receptor (PPAR) (Bugge *et al.*, 1992; Leid *et al.*, 1992; Issemanm *et al.*, 1993). The resulting heterodimers target different hormone response elements (HREs) which may possess altered ligand responses altering their specificity for the HRE and *trans*-activation potential (Forman *et al.*, 1995).

The dimerisation observed between the Max, Myc, and Mad proteins, which are members of the basic helix-loop-helix/leucine zipper family, results in different *trans*-activation potentials. Myc *trans*-activation is dependent on its heterodimerisation to the Max protein in order to bind to the recognition site CACGTG (E box) (Blackwood & Eisenman, 1991). In addition to binding to Myc, the Max protein can also form

homodimers or heterodimers with Mad related proteins such as Mxi1, which compete with Max/Myc heterodimers for the E box sequence (Ayer *et al.*, 1993). The Myc/Max target gene, human ornithine decarboxylase (ODC) gene, is induced due to the presence of an E Box sequence within its promoter (Pena *et al.*, 1993, 1995). Recent data indicates that the Myc/Max complex may be inhibited by overexpression of the Mxi1 protein, forming Max/Mxi1 heterodimers which bind to the E box and prevents Myc *trans*-activation (Wu *et al.*, 1996).

Phosphorylation

Another important mode of control of transcription factor activity is phosphorylation. This is underlined by the observation that the majority of transcription factor proteins identified exist as phosphoproteins. Phosphorylation of transcription factors may affect the protein's, sub-cellular localisation, *trans*-activation activity or DNA binding activity resulting ultimately in altered promoter regulation.

The control of transcription factor subcellular localisation has the obvious consequence of preventing transcription factors from entering the nucleus and thereby binding there *cis*-acting sites. Nuclear proteins are synthesised in the cytoplasm and are transported to the nucleus by the presence of a nuclear localisation signal (NLS) sequence in the protein. Regulation by direct phosphorylation of the NLS was first observed in the SV40 T antigen and is believed to be mediated by a phosphorylation site for Caesin kinase II (CKII) within the NLS and a $p34^{cdc2}$ site adjacent to the NLS (Jans *et al.*, 1991). Other nuclear proteins such as mouse c-Myc, c-Abl and human p53 also contain potential CKII or $p34^{cdc2}$ phosphorylation sites within the NLS (Jans *et al.*, 1991). It has been suggested that both CKII and $p34^{cdc2}$ might have a general role in protein import into the nucleus. Phosphorylation may also aid in the nuclear localisation of some members of the Rel family of transcription factor, by a different

mechanism. In this case NF-κB is held in an inactive complex with IκB, which is released by the phosphorylation of the IκB. Both protein kinase A and C (PKA & PKC) may phosphorylate IκB *in vitro* (Shirakawa & Mizel, 1989; Ghosh & Baltimore, 1990).

The control of transcription factor *trans*-activation mediated by phosphorylation has been extensively studied in the cAMP response element binding/activating transcription factor protein (CREB/ATF) family. The glutamine-rich activation domain of CREB contains a phosphorylation box (P-box) which contains consensus phosphorylation sites for a number of kinases (de Groot *et al.*, 1993). Upon induction by cAMP CREB is phosphorylated on serine 133 within the P-Box by PKA and as a consequence is converted to a potent transcriptional activator (Gonzalez & Montminy, 1989). Other members of the family may be phosphorylated by different kinases at alternative sites (de Groot *et al.*, 1993). In vitro, the repression of CREB following stimulation is dependent on dephosphorylation by protein phosphatase-1 (PP-1) (Hagiwara *et al.*, 1992) and CREB may also be dephosphorylated by PP-2A (Nichols *et al.*, 1992).

Phosphorylation may also affect DNA binding, an interesting example is the c-Jun protein which shows altered binding activity or *trans*-activation function depending on the sites phosphorylated. For example, phosphorylation at one set of sites towards the C-terminal end close to the basic domain (Ser226, Ser232 and Thr214) decreases DNA binding activity (Boyle *et al.*, 1991). Alternatively phosphorylation at other sites close to the activation domain, (Ser63 and Ser73) results in an increase in transcriptional activity (Binetruy *et al.*, 1991; Pulverer *et al*, 1991). A single stimulus, for example TPA, can induce dephophorylation at inhibitory sites while increasing phosphorylation at positive regulating sites (Pulverer *et al.*, 1992).

The examples here are just a fraction of the phosphorylation events that occur in transcription regulation. It is likely that phosphorylation plays a central role in controlling protein conformation and activity in a number of cellular events, as appears

to be the case in the signalling pathways, and these events are regulated by kinases and the phosphatases.

The points discussed previously have illustrated the immense complexity of eukaryotic gene regulation to create a multitude of different regulatory options for achieving appropriate tissue specific and developmental gene regulation. Consequently some of these aspects of gene regulation are discussed below with regard to tissue specific and developmental gene regulation.

1.11 Globin gene expression: a model for tissue specific and developmental gene regulation

An intriguing question in eukaryotic gene regulation is how transcriptional cascades mediate developmental and tissue specific pathways. A variety of developmental pathways (haemopoiesis, liver, kidney) have been used to assess how a set of transcription factors specify a particular cell lineage. A particularly good example is globin gene expression which has been used as an important paradigm for studying tissue specific and developmental transcription regulation. Erythroid cells undergo a series of morphological and biosynthetic alterations during erythropoiesis. The human β-globin locus contains five genes (ε,Gγ, Aγ, δ and β) successively expressed as erythropoieses progresses during development. The ε gene is transcriptionally active in the yolk sac, the two γ genes are expressed in the foetal liver followed by expression of δ and β in adult bone marrow. The promoters of all five β-globin genes are also dependant on distant enhancer elements, termed the β-locus control region (LCR), 11-60kb upstream.

Regulatory elements within the β-globin promoter

The β-globin promoter contains three conserved positive *cis*-acting sites a TATA box, a CCAAT and a CACCC motif which are necessary for full activity in the mouse β-globin gene. The TATA box is important for the binding of the basal transcription factor TFIID which recruits other factors of basal transcription machinery into the pre-initiation complex, as discussed previously (Chapter 1.1).

The CCAAT box may bind at least five different proteins but evidence suggest that the ubiquitous CCAAT protein 1 (CP1) protein binds to the β-globin CCAAT element (Chodosh, *et al.*, 1988; de Boer et al., 1988, Gumucio *et al.*, 1988) and contributes to increase expression of the gene. Alternatively, a CTF/NF1-like factor, probably the CCAAT displacement factor (CDP) first identified in sea urchin (Barberis *et al.*, 1987), downregulates the human β-globin gene expression by binding adjacently to the CCAAT site and inhibiting binding of CP1 (deBoer *et al*, 1988). Interestingly, similar activities have also been observed in the chicken β-globin promoter (Plumb *et al.*, 1988) and the CDP factor binds to a duplicated CCAAT element in the γ-globin promoter preventing CP1 binding (Superti-Furga *et al.*, 1988). This suggests that this antagonistic relationship may be a common control mechanism in globin gene expression.

The CACCC box motif is essential for erythroid development, but may be found in non-erythroid promoters (Xiao *et al.*, 1987; Schule *et al.*, 1988). The motif may be bound by ubiquitous proteins such as Sp1 (Barnhart *et al.*, 1988; Jackson *et al.*, 1989) or more cell restricted proteins such as, TEF-2 (Xiao *et al.*, 1987), erythroid kruppel like factor (EKLF) (Miller & Bieker, 1993) or the newly identified lung erythroid kruppel-like factor (LEKF) (Anderson *et al.*, 1995).

The EKLF zinc-finger protein is expressed exclusively in erythroid cells and acts as a powerful *trans*-activator through the CACCC site in transient expression assays

(Miller & Bieker, 1993; Donze *et al.*, 1995). The EKLF protein is essential for cell-specific expression of the β-globin promoter (Bieker & Southwood,1995). Interestingly, the EKLF protein binds 8-fold more efficiently to the β-globin promoter relative to the γ-globin promoter, suggesting that EKLF is important in the transition from human γ- to β-globin gene switching (Donze *et al.*, 1995). Consistent with this, CACCC box mutations which cause β-thalassemias inhibit EKLF binding (Feng *et al.*, 1994). It is also interesting to note that the EKLF promoter is regulated by the GATA-1 zinc finger protein, another cell-restricted gene which is essential in erythroid development. This may, intriguingly suggest that essential erythroid transcription factors may regulate each other (Crossley *et al.*, 1994).

The β-globin promoter requires the LCR for tissue specific and developmental regulation

Although, some of the β-globin genes may function in a partial tissue- and stage-specific manner in the absence of the LCR, (Rutherford & Nienhuis, 1987), the LCR is essential for efficient transcription of the β-globin genes. The β-LCR is divided into four subdomains which correspond to specific DHS sites in erythroid cells. Surprisingly DNAse1 footprinting analysis only consistently identified three binding motifs within the LCR, being GATA, AP-1 and CACC/GGTG (Philipsen *et al.*, 1990; Talbot *et al.*, 1990; Prunzina *et al.*, 1991; Reddy & Shen., 1991; Straus & Orkin, 1992). The GATA and AP-1 motif were shown to bind the cell-restricted GATA-1 and NF-E2 proteins respectively, while the CACC/GGTG sequence binds Sp1 or Sp1-like factors (Philipsen *et al.*, 1990; Talbot *et al.*, 1990; Kingsley & Winoto, 1992). The model for developmental associated switching of the β-globin genes is believed to occur due to competition for a shared enhancer. This may reflect differences in the diversity of *trans*-activators/repressors able to interact with the LCR DHS sites or gene promoters. In

support of this, a stage specific factor, NF-E4, has been suggested to mediate direct association of the β-promoter with the enhancer. Mutations within the NF-E4 site, which abolish binding, have been shown to result in reciprocal changes between β and ε gene expression (Foley & Engel, 1992). Interaction of the promoter region with the enhancer would presumably be dependent on the DNA looping. It is therefore interesting to note that there are a number of Sp1 sites in the LCR and globin promoters since the Sp1 protein has been shown to stabilise DNA loops (Su *et al.*, 1991). Similarly, there is also evidence that the GATA-1 sites within globin promoters and the LCR are required for enhancer-dependent transcription (Fong & Emerson, 1992; Gong and Dean, 1993). Since the GATA-1, Sp1 and EKLF proteins may all physically interact, protein/protein interaction may be important in stabilising DNA looping (Merika & Orkin, 1995).

An alternative model may be that there is a shift in location of the gene relative to the LCR, altering its accessibility. Interestingly, the four DHS sites of the LCR appear to have different specificities for the gene promoters, with DHS-4 preferring β-expression (Fraser *et al.*, 1993). In transgenic mice studies the β-LCR was shown to act as dominant control region showing position independent expression (Grosveld *et al.*, 1987) and it is possible that the LCR may impart some of its activity by altering the surrounding chromatin structure. This form of enhancer control may be similar in manner to the PEV effect described previously (chapter 1.7). It is possible that aspects of both models may be important in controlling β-globin gene expression.

1.12 Nuclear oncogenes may disrupt transcriptional cascades

Disruption of normal transcription factor cascades during development can lead to gross phenotypic alterations, which have been extensively studied in *Drosophila* development (Nusslein-Volhard *et al.*, 1987; Warren & Carrol, 1995). Alternatively,

aberrant expression patterns results in increased proliferation or altered differentiation which in some circumstances progress to oncogenesis. The association of transcription factors with cancer progression underlines the importance of transcription factors in development. Ironically, therefore, identification of cancer associated nuclear oncogenes/tumour supressors has co-ordinately led to the identity of many novel transcription factors which are vital for appropriate developmental regulation.

Cancer progression may result from the loss of function of a cellular protein, designated a tumour suppressor, or a dominant gain of function, termed a oncogene, which represents the most well studied group. Altered expression of proteins may induce cancer by abrogating or mimicking growth factors responses (reviewed by Cross & Dexter, 1991) or transcription factors (reviewed by Nichols & Nimer, 1992), directly affecting endogenous gene expression. However, cancer progression is believed to be a multi-step process reliant on the synergy of oncogenes for tumourgenicity (reviewed by Adams & Cory, 1992).

Many nuclear oncogenes have been identified in the haemopoietic system where chromosomal translocations and inversions commonly activate genes encoding nuclear DNA-binding proteins. As a consequence many of the examples given here are derived from the haemopoietic system. However, altered expression of transcription factors is also important in non-haemopoietic tissues, for example, fusions of paired box (PAX) PAX3 and PAX7 genes with the FKHR gene are associated with alveolar rhabdomyosarcoma (Shapiro *et al.*, 1993; Davies *et al.*, 1994) or the t(11;22) translocations create a *EWS/Fli-1* fusion gene associated with Ewing's sarcoma (Delattre *et al.*, 1992). The increased levels of nuclear oncogene products observed in tumours may be as a consequence of a variety of mechanisms discussed below.

Cellular oncogenes are targets for retroviral insertions

Oncogenes are often associated with common sites of retroviral integration either at the oncogene locus or at more distant sites resulting in aberrant expression. For example, *c-Myc* (Hayward *et al.*, 1981), *c-Myb* (Shen-Ong *et al.*, 1986), *Fli-1* (Ben-David *et al.*, 1991), *Spi-1* (Moreau-Gaschelin *et al.*, 1988) and *Evi-1* (Morishita *et al.*, 1988). The mechanisms of oncogene activation may be due to either promoter insertion or alternatively activation of the appropriate oncogene promoter by an enhancer, presumably the viral LTR.

Translocation resulting in enhancer activation

Chromosomal translocations and inversions may have the consequence of bringing a nuclear oncogene under the control of heterologous enhancer or promoter elements. Table 1.2 (adapted from Rabbits, 1994) shows a number of examples of non-fusion chromosome translocation events resulting in aberrant expression of the appropriate gene in haematopoietic tumours. For a more comprehensive review see Rabbits 1994 and references therein.

The predominant t(8;14) (q24;q32) translocation in B-cell acute leukaemia, rearranges one allele of the *c-Myc* oncogene into the heavy-chain Ig gene locus which brings the c-Myc gene under the control of the strong Ig enhancer (Dalla-Favera *et al.*, 1983; Taub *et al*, 1982; Adams *et al* 1983). Similarly, the t(10;14) (q24;q11) translocation in T-cell acute lymphoma brings the developmentally important *HOX11* gene under the control of the regulatory sequences present in the TCR locus (Dube *et al.*, 1991) and the inv (3)(q21;q26) in acute myeloid leukaemia brings the *Evi-1* oncogene under the influence of the enhancer elements of the Ribophorin I gene (Suzukama, *et al.*, 1994).

Table 1.2: Examples of non-fusions from haematopoietic tumours

Type	Affected gene	Disease	Rearranging gene
Basic-helix-loop-helix			
t(8; 14)(q24; q32)	*c-Myc* (8q24)	BL, BL-ALL	IgH, IgL
t(7; 19)(q35; p13)	*LYL1* (19p13)	T-ALL	TCR-β
t(1; 14)(p32; q11)	*TAL1/SCL* (1p32)	T-ALL	TCR-α
t(7; 9)(q35; q34)	*TAL2* (9q34)	T-ALL	TCR-β
LIM Proteins			
t(11; 14)(p15; q11)	*RBTN1/Ttg1* (11p15)	T-ALL	TCR- δ
t(11; 14)(p13; q11)	*RBTN1/Ttg1* (11p15)	T-ALL	TCR- δ /α / β
Homeobox proteins			
t(10; 14)(q24; q11)	*Hox11* (10q24)	T-ALL	TCR- α / β
Zinc-finger proteins			
t(3;14)(q27; q32)	*Laz3/BCL-6* (3q27)	NHL/DLCL	IgH
inv (3) (q21; q26)	*Evi-1* (3q26)	AML	
Others			
t(14;18)(q32; 21)	*BCL2* (18q21)	FL	IgH, IgL
t(14;19)(q32; q13.1)	*BCL3* (19q13.1)	B-CLL	IgH
t(3;14)(q27; q32)	*TAN1* (9q34.3)	T-ALL	TCR-β

Adapted from Rabitts 1994

T-ALL; T-cell acute lymphocytic leukaemia, BL; Burkitt's Lymphoma, FL; Folliculalr lymphoma, B-CLL; B-cell Chronic Lymphocytic leukaemia, AML; Acute Myeloid Leukaemia, TCR; T-cell receptor locus, Ig; immunoglobulin locus.

Gene amplification

Gene amplifications are common in many types of cancer and are generally associated with more advanced tumours (Schwab & Amler, 1990). Analysis of the human promyelocytic leukaemia cell line HL-60 showed a 10-fold amplification of the c-Myc gene on double minute chromosomes (Dalla-Favera *et al.*, 1982) and was the first gene to be found amplified in human neoplasms. Amplification of another myc family member, N-myc, is characteristic in the progression of neuroblastomas (Brodeur *et al.*,

1984) while the unrelated c-myb gene has also been found to be amplified, for example, in two cell lines from a colon carcinoma (Alitalo *et al.*, 1984). The mechanisms of mammalian gene amplification is not discussed here but is reviewed by Stark, 1993.

Chimeric proteins

Chromosomal translocations and inversions may result in the fusion of two genes, thereby creating a chimeric protein, with altered function (DNA-binding, *trans*-activation) relative to the wild type proteins. Table 1.3 (adapted from Rabbits, 1994) shows a number of examples of transcription factor fusion-proteins resulting from chromosome translocation events in haematopoietic tumours. For a more comprehensive review see Rabbits 1994 and references therein.

For example, the translocation t(1;19) (q23;p13), that is associated with about a quarter of human pre-B-cell acute lymphoid leukaemias, results in a *E2A-PBX1* chimera. The N-terminal *trans*-activation domain of the bHLH transcription factor *E2A* gene located on chromosome 19 fuses to the homeobox *PBX1* gene on chromosome 1, resulting in a number of E2A-PBX1 oncoproteins. Consequently, the PBX1 transcription factor acquires *trans*-activator function and still recognises it binding sites in lymphoid cells where the gene it not usually expressed, suggesting that the formation of chimeric proteins with novel activities may play an important part in the progression of cancer (reviewed by Look, 1995; Rabbits 1994).

Table 1.3: Examples of fusions-proteins from haematopoietic tumours

Type	Affected gene	Protein domain	Disease
t(1;19)(q23; q13.3)	*PBX1* (1q23)	Homeodomain	pre-B-ALL
	E2A (19p13.3)	AD-b-HLH	APL
t(15;17)(q21; q11-22)	*PML* (15q21)	Zinc-finger	
	RARα (17q21)	Zinc-finger	
t(3;21)(q26; q22)	*AML1* (21q22)	DNA binding/runt homolgy	AML
	EVI-1 (3q26)	Zinc-finger	AML
t(8;21)(q22; q22)	AML1 (21q22)	DNA binding/runt homology	AML
	ETO (8q22)	Zinc-finger	
t(3;21)(q26; q22)	AML1 (21q22)	DNA binding/runt homology	AML
	EAP (3q26)	Sn protein	
t(3;8)(q26; q22)	EVI-1 (3q26)	Zinc-finger	AML
	ETO (8q22)	Zinc-finger	
t(16;21)(p11; q22)	FUS (16p11)	RNA binding	Myeloid
	ERG (21q22)	Ets protein	
t(6;9)(p23; q34)	DEK (6p23)	?	AML
	CAN (9q34)	Leucine zipper	

Adapted from Rabitts 1994

Abbreviations: pre-B-ALL; pre-B-cell acute lymphocytic leukaemia, APL; Acute Promyelocytic Leukaemia, AML; Acute Myeloid Leukaemia. AD-b-HLH; Activation Domain-basic-Helix-Loop-Helix.

It is clear that nuclear oncogenes are not only important in cancer progression but are also vital genes in normal cell proliferation and differentiation, often being pivotal genes in developmental and cell cycle regulation. Interestingly, a specific nuclear oncogene, *Evi-1*, encodes a protein of the zinc-finger family which is an essential transcription factor for appropriate murine and human development and is also associated in some leukaemias, following ectopic expression. The rest of this chapter discusses the properties of the *Evi-1* gene.

1.13 Identification of *Evi-1* as a proto-oncogene

Evi-1 was originally identified as a common site of viral integration in the DNA of AKXD murine myeloid tumours (Mucenski *et al.*, 1988). AKXD recombinant inbred mouse strains are derived from mating AKR/J and DBA/2J strains which have a high or low incidence of lymphoma respectively. These strains have a range of tumour incidence and have proved invaluable in identifying common sites of retroviral integrations which may represent novel proto-oncogene loci.

A particular strain, AKXD-23, was found to develop myeloid tumours at high frequency. The DNA from all the AKXD-23 myeloid tumours examined had a common ecotropic retroviral integration in the same locus, termed ecotropic viral integration site 1 (*Evi-1*). Subsequent studies identified a putative proto-oncogene, *Evi-1*, activated by provirus in IL-3 dependent myeloid leukaemia cell lines (Morishita *et al.*, 1988).

Evi-1 rearrangements have also been detected in myeloid cell lines and leukaemias in NFS/N or NFS/N hybrid mice injected as neonates with wild mouse ecotropic MCF MuLV and Cas-Br-E MuLV (Mucenski *et al.*, 1988, Bergeron *et al.*, 1992). Retroviral insertions within the *Evi-1* locus are clustered within non-coding exon I or II and are frequently orientated in the same transcriptional direction. The mechanism of *Evi-1* activation appears to that of promoter insertion since viral derived transcripts are detectable.

In a separate study, 23% of all myeloblastic leukaemias induced by infection of the friend murine leukaemia virus had proviral integrations in another common site of viral integration termed *Fim-3* (Bordereaux *et al.*, 1987). In contrast to clustered retroviral insertion observed in the *Evi-1* locus, insertions in *Fim-3* are scattered over 16kb.

A common retroviral integration site (*CB-1*) was also found in two IL-3-dependent cell lines (DA-3 and DA-34), 90kb upstream of the *Evi-1* locus. This

retroviral integration site is closely linked to the *Evi-1* locus is identical to the *Fim-3* locus (Bartholomew *et al.*, 1989). Retroviral insertions within *CB-1/Fim3* were shown to activate expression of the *Evi-1* gene from its normal promoter (Bartholomew & Ihle, 1991). Since no RNA expression was detectable from the viral LTR within the *Fim-3* region it suggests that the mechanism of *Evi-1* activation may be that of a distal enhancer type. Ectopic expression of *Evi-1* may therefore be due to retroviral insertion at either the *Fim-3* or *Evi-1* loci.

1.14 The association of the *Evi-1* proto-oncogene in human leukaemias

In humans, ectopic expression of the EVI-1 gene has been correlated with progression of some acute myelogenous leukaemias (AML) the closely related myelodysplastic syndromes (MDS) and chronic myelocytic leukaemias in blast crisis (CML-BC).

AMLs are clonal malignancies characterised by abnormal proliferation and differentiation, resulting in the production of blast and other immature granulocytes, monocytes, erythroblast, and/or megakaryocytes. As with many leukaemias, lymphomas and some sarcomas (Rowley., 1990), AML and MDS are associated with a number of recurring chromosomal abnormalities. Included in these are rearrangements at 3q26 in about 2% of patients examined. The most frequently observed chromosome translocations involving the 3q21q26 group of AMLs are inv(3)(q21q26) and t(3;3)(q21;26). The chromosomal breakpoints in 3q26 in these translocations have been located between 13-330kb 5' and upto 150kb 3' of the gene (Morishita *et al.*, 1992, Suzukama *et al.*, 1994). Characterisation of the inv(3)(q21q26) breakpoint suggests that ectopic expression of the *EVI-1* gene may be a consequence of the gene being translocated to within the control of the enhancer elements associated with the Ribophorin I gene (Suzukama *et al.*, 1994).

CML is a clonal disorder of pluripotent haematopoietic stem cells with a biphasic clinical course. The initial chronic phase is characterised by leukocytosis with maturation followed by acceleration of the disease to a acute blastic crisis (CML-BC) phase involving increased cellular proliferation, maturation arrest and karyotype clonal evolution. A characteristic of chronic phase CML is presence of the Philadelphia chromosome (Ph), t(9;22)(q34;q11) (Rowley, 1973), which results in a BCR-ABL chimeric protein with enhanced tyrosine kinase activity (Konopka *et al.*, 1994) responsible for its *in vivo* tumourigenicity (Daley *et al.*, 1990).

The chromosomal abnormalities observed in CML-BC are characteristically trisomy or isochromosome. However, the molecular events leading to blastic crises are not understood, although p53 is a relatively common target, in about a third of blastic crisis patients (Feinstein *et al.*, 1992). Interestingly, a rare reciprocal translocation, t(3;21)(q26;q22), event in CML-BC generates a fusion between the AML and EVI-1 genes which fuses the N-terminal runt homology domain of AML1 to the whole of EVI-1 (Mitani *et al.*, 1994). In addition to CML-BC, the t(3;21) translocation is also observed in MDS derived leukaemia (Rubin *et al.*, 1990) and the AML1-EVI-1 fusion transcript has also been detected in a MDS patient (Mitani *et al.*, 1994). It has been postulated that t(3;21) translocation plays a critical role in leukaemic progression as heamatopoietic stems cells transform to the acute leukaemic phase (Chen *et al.*, 1991).

In addition to the EVI-1 gene, the AML1 gene is also involved in other chromosome translocations to generate fusions with other genes, *AML1-MTG8* and *AML1-ETO* (Miyoshi *et al.*, 1991; reviewed by Nucifora & Rowley, 1995). An increasing number of chimeric fusion transcription factors resulting from reciprocal translocations, including DEK-CAN, E2A-PBX1 and PML-RARα, have also been identified showing gene fusion is an important mechanism in leukaemogenesis (reviewed by Nichols and Nimer, 1992).

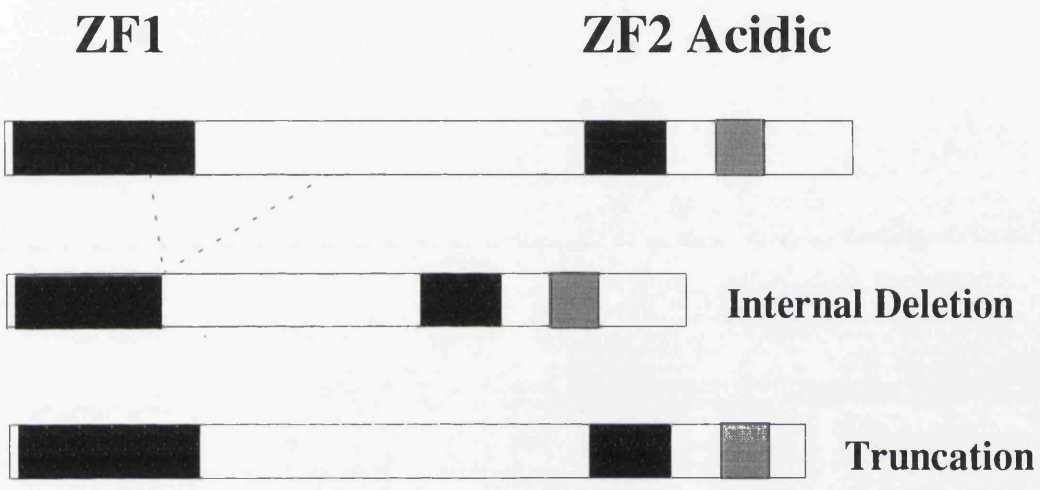
It has also been reported that ectopic expression of the *EVI-1* gene in some MDSs and other haematological malignancies appears to occur with no gross cytological abnormalities of chromosome 3q26 (Russel *et al.*, 1994), suggesting that aberrant expression of this gene may result from various mechanisms. Interestingly, EVI-1 expression in some AMLs is also accompanied by chromosome 7 deletion (Morishita *et al.*, 1992; Fichelson *et al.*, 1992). This would appear be a secondary event since, similarly to 3q26 abnormalities, chromosome 7 deletion is often found in human leukaemias following treatment with alkylating agents (Le Beau *et al.*, 1986).

Interestingly, leukaemias associated with 3q26 rearrangements involving *Evi-1* often show megakaryopoiesis and elevated platelet counts (Bitter *et al.*, 1985), suggesting that *Evi-1* might contribute to the progression of a distinct genetic disease. It should also be noted that rearrangements of 3q25 or 3q27 do not involve *Evi-1* (Morishita *et al.*, 1990, Fichelson *et al.*, 1992).

1.15 Structure of the *Evi-1* gene

The *Evi-1* gene cDNA's were originally isolated from murine IL-3 dependent myeloid leukaemia cell lines which ectopically express the gene due to retroviral insertion in the *Evi-1* locus. This identified a 5.1kb full length cDNA clone with an open reading frame of 3126 nucleotides and a predicted protein of 1042 amino acids (Morishita *et al.*, 1988). The Evi-1 protein contains ten Cys_2-Hys_2 zinc-finger repeats of 27-28 amino acids separated into two domains, an N-terminal domain of seven (designated ZF1) and a C-terminal domain of three zinc-fingers (designated ZF2) respectively. The protein also possesses an acidic rich domain proposed to be important in *trans*-activation function (Ma & Ptashne, 1987) and has been shown to be localised to the nucleus (Matsugi *et al.*, 1990) (figure 1.1). The human *EVI-1* gene has also been

Figure 1.1: Schematic representation of the Evi-1 protein



ZF1 GA(C/T)AAGATAAGATAA

ZF2 GAAGATGAG

isolated as a cDNA clone from the high *EVI-1* expressing HEC-1B cell line and shows 94% homology with the murine gene at the amino acid level.

The Evi-1 protein acts as transcription factor with both zinc-finger domains recognising distinct DNA recognition sequences. Bacterially derived fusion proteins containing ZF1, ZF2 or the full length Evi-1 were used in binding and amplification reactions to establish DNA consensus sequences for ZF1, GA(C/T)AAGA(T/C)AAGATAA (Perkins *et al.*, 1991; Delwel *et al.*, 1993), ZF2 , GAAGATGAG, (Funabiki *et al.*, 1994) and ZF1/ZF2, GACCAGATAAGATAAN$_{1\text{-}28}$ CTCATCTTC (Morishita *et al.*, 1995). The ZF1/ZF2 sequence shows the optimal recognition sequence with a normal ZF1 consensus and an inverted ZF2 consensus although there is little indication for optimal spacing requirements between the consensus sequences (Morishita *et al.*, 1995). Interestingly the ZF1 recognition sequence contains three perfect GATA repeat sequences, recognised by members of the GATA zinc-finger family of transcription factors.

In addition to the full length Evi-1 transcript, two alternatively spliced variants have also been identified. The first is an internal deletion of 324 amino acids which results in the loss of two zinc-finger motifs (6+7) of the first zinc-finger domain (Bordereaux *et al.*, 1990; Morishita *et al.*, 1990). The zinc-fingers 6 and 7 in the ZF1 domain have been shown to be important in DNA sequence recognition. Consequently, it would be predicted that this splice variant would be unable to bind to the ZF1 DNA recognition sequence and exert an effect on the transcription machinery through this site. The second is an C-terminal 105 amino acid deletion resulting in a truncated 937 amino acid protein (Bartholomew, unpublished data).

1.16 Biological effects of the *Evi-1* protein

The biological functions of the *Evi-1* gene both in normal cell proliferation/differentiation and in cellular transformation are unclear. The structural motifs identified in the Evi-1 protein is consistent with the protein being a transcription factor. Since transcription factors have been shown to act as transcription activators, repressors or both, similar functions have been assessed for the Evi-1 protein.

The EVI-1 protein acts as both a transcriptional repressor and activator

As mentioned in 1.15, the ZF1 recognition sequence is similar to GATA suggesting that Evi-1 might target a subset of genes also recognised by GATA factors. The EVI-1 protein has been shown to block GATA-1 dependent transcriptional activation from expression constructs containing the ZF1 consensus sequence in NIH3T3 cells (Kreider *et al.*, 1993) suggesting that EVI-1 acts as a transcriptional repressor via a DNA binding competition mechanism. Since GATA-1 and other members of the GATA family are expressed in haemopoietic cells it is conceivable that repression of some of the GATA target genes may contribute to leukaemogenesis. However, no known physiological GATA sites are recognised by Evi-1 (Perkins & Kim, 1996).

In a separate study, the AML-EVI1 fusion protein, which contains the AML1 *runt* homology domain and the whole of Evi-1, fails to activate a minimal tk luciferase reporter construct containing PEBP2 (PEA2) binding sites upstream (Tanaka *et al*, 1995). The AML1 protein acts as an *trans*-activator through the PEBP2 sites (Tanaka *et al.*, 1995), which were originally identified in the polyomavirus enhancer (Piettte & Yaniv, 1987). There the AML-EVI-1 fusion protein antagonises the wild type AML function, probably by a dominant negative mechanism. In addition it also indicates that

the Evi-1 protein is not a *trans*-activator protein, at least in this fusion and through this site.

In addition, data from our lab has also suggested that the EVI-1 protein can act as a active transcriptional repressor, as opposed to the probable passive repression mechanisms mentioned above. In these studies we have shown that Evi-1 is able to actively repress the function of a strong *trans*-activator, VP16, when targeted to the DNA either via the GAL4 1-147 DNA binding domain or alternatively through the wild-type ZF1 recognition site (Bartholomew *et al.*, manuscript submitted).

In contrast, the EVI-1 protein has been shown to act as a transcriptional activator via binding to a consensus sequence GACCAGATAAGATAAN$_{1\text{-}28}$ CTCATCTTC in both NIH3T3 and HEC-1B cells (Morishita *et al.*, 1995). This sequence contains the consensus binding sequences for both ZF1 and ZF2 and suggests the possibility that the combination of these two binding sites may some how alter the function of the Evi-1 protein, possibly by some conformational change. Evi-1 has also been shown to upregulate both the *c-fos* and *c-jun* promoter, although this might be an indirect effect. However, interestingly the upregulation of both promoters is dependent on the ZF2 DNA binding domain (Tanaka *et al.*, 1994; Kurokawa *et al.*, 1995).

Therefore, similarly to other transcription factors it seems that the Evi-1 protein may be bifunctional. An *in vivo* target gene, designated 150-B, has recently been identified encoding a novel protein, which is suppressed in *Evi-1* expressing myeloid cells (Matsugi *et al.*, 1995). This and other *in vivo* target genes will allow a more accurate assay for the biological function of the Evi-1 protein and the molecular mechanisms involved in transcriptional regulation.

Evi-1 is oncogenic

The oncogenic nature of the Evi-1 and AML1/EVI-1 fusion proteins, created as a result of the (3;21) translocation observed in CML-BC (see chapter 3.4), has been assessed using the Rat1 transformation assay. In these experiments EVI-1 and AML/EVI-1 have been shown to transform Rat1 cells, as measured by their ability to allow growth of the cells in soft agar. The introduction of deletion mutants of the AML1/EVI1 fusion protein indicated that the EVI-1 ZF2 domain was essential for the transforming activity (Kurokawa *et al.*, 1995). In addition, Rat1 clones were established stably expressing the BCR/ABL fusion protein ($p210^{BCR/ABL}$), a characteristic of early chronic CML. These cells formed macroscopic colonies in soft agar as described by others (Lugo & Witte., 1989). Interestingly, if $p210^{BCR/ABL}$ RAT1 cells were also transfected with constructs expressing the *AML1/EVI-1* fusion an enhanced transformation was observed with larger and more numerous colonies formed in soft agar (Kurokawa *et al.*, 1995).

Evi-1 overexpression blocks erythroid and myeloid differentiation and/or proliferation

32Dcl3 is a IL-3-dependent myeloid cell line that differentiates in response to G-CSF (Greenberger *et al.*, 1983; Migliaccio *et al.*, 1989). These cells have been used to examine the properties of proteins involved in differentiation, including c-Myb (Otten *et al.*, 1988) and Evi-1 (Morishita *et al*, 1992; Khanna-Gupta *et al.*, 1996). The introduction of retroviral constructs expressing the full length Evi-1 protein into the 32Dcl3 cells blocked the ability of the cells to differentiate in response to G-CSF (this was not due to a down regulation of the G-CSF receptor). In support of these observations the same experiments have also been performed with the AML1/EVI-1 fusion protein by a separate group, with the same conclusions (Tanaka *et al.*, 1995).

However, recent data contradicts some of these previous observations made using the 32Dcl3 cells. Firstly 32Dcl3 cells were found to constitutively express high levels of Evi-1 protein and secondly the cells were still able to respond to G-CSF (Khanna-Gupta *et al.*, 1996). Interestingly, the authors show that further overexpression of *Evi-1* in 32Dcl3 cells using retroviral *Evi-1* expression vectors prevents the cells from responding to G-CSF, which is consistent with previous observations (Morishita *et al.*, 1992). Consequently, they have proposed that it is the level of *Evi-1* expression in myeloid cells, rather than ectopic expression, which is the most important determinant in contributing to leukaemic progression.

Similarly, experiments have evaluated the response to erythropoietin (EPO) of normal mouse bone marrow progenitor cells or 32DEpol cells (a clone of 32Dcl3 cell line responsive to EPO) (Migliaccio *et al.*, 1989) following infection with retroviral *Evi-1* expression vectors (Kreider *et al.*, 1993). The results showed that ectopic *Evi-1* expression was able to block growth of EPO-responsive cells. However, it is not clear in these experiments whether the cells were unable to differentiate, proliferate or non-viable. In addition, the significance of this result to leukaemogenesis is unclear since the AML and CML-BC leukaemias, associated with ectopic *EVI-1* expression, are generally associated with increased rather than decreased proliferation.

1.17 Evi-1 shows a restricted pattern of expression in adult murine tissue

The involvement of the cellular *Evi-1* gene in leukaemic progression led workers to evaluate the involvement of the gene in normal cell proliferation/differentiation. Initially, the expression of the gene was examined in murine tissues and cell lines by Northern blot analysis. This showed that high levels of *Evi-1* expression were confined to kidney and developing oocytes, with lower levels identified in lung, uterus and heart, illustrating that the *Evi-1* gene has a restricted profile of

expression in adult murine tissues suggesting that its expression might be important in the development of these organs (Morishita *et al.*, 1990; Perkins *et al.*, 1991). *Evi-1* expression in cell lines was found to be restricted to leukaemic myeloid cells containing retroviral insertions at either the *Evi-1* locus or 90kb upstream in the *CB-1/Fim-3* locus. No expression was observed in either other myeloid, T-cell and B-cell lymphoma cell lines or carcinoma cell lines derived from kidney, lung, ovary, placenta or brain (Morishita *et al.*, 1990).

1.18 Evi-1 is essential for murine embryonic development

Evi-1 also shows both temporal and spatial expression during murine embryonic development (Perkins *et al.*, 1991). *In situ* hybridisation experiments on 9.5, 12.5 and 14.5 day mouse embryos identified high levels of *Evi-1* expression in the urinary system and the Mullerian ducts, the bronchial epithelium of the lung; the endocardial cushions and truncus swellings in the heart; focal areas within the nasal cavities; and the developing limbs. The expression patterns of *Evi-1* are consistent with the gene being important in murine embryonic development.

Recent data from *Evi-1* knockout mice has confirmed the importance of Evi-1 in murine development. Homozygous Evi-1 knock out mice die in utero at day 10.5 *post coitus* (*p.c.*). The mutants have a complex phenotype but the most obvious abnormalities include widespread haemorrhaging, the presence of large pericardial sacs, and disruption in the development of paraxial mesenchyme which is most obvious in the head. The neural ectoderm and the peripheral nervous system fail to develop. Furthermore, delayed development was observed in the kidney, heart and limb buds (Hoyt *et al.*, manuscript submitted).

Significantly, the *Evi-1 in situ* hybridisation pattern observed in normal mice showed expression in mesoderm and neural crest derived cells, associated with

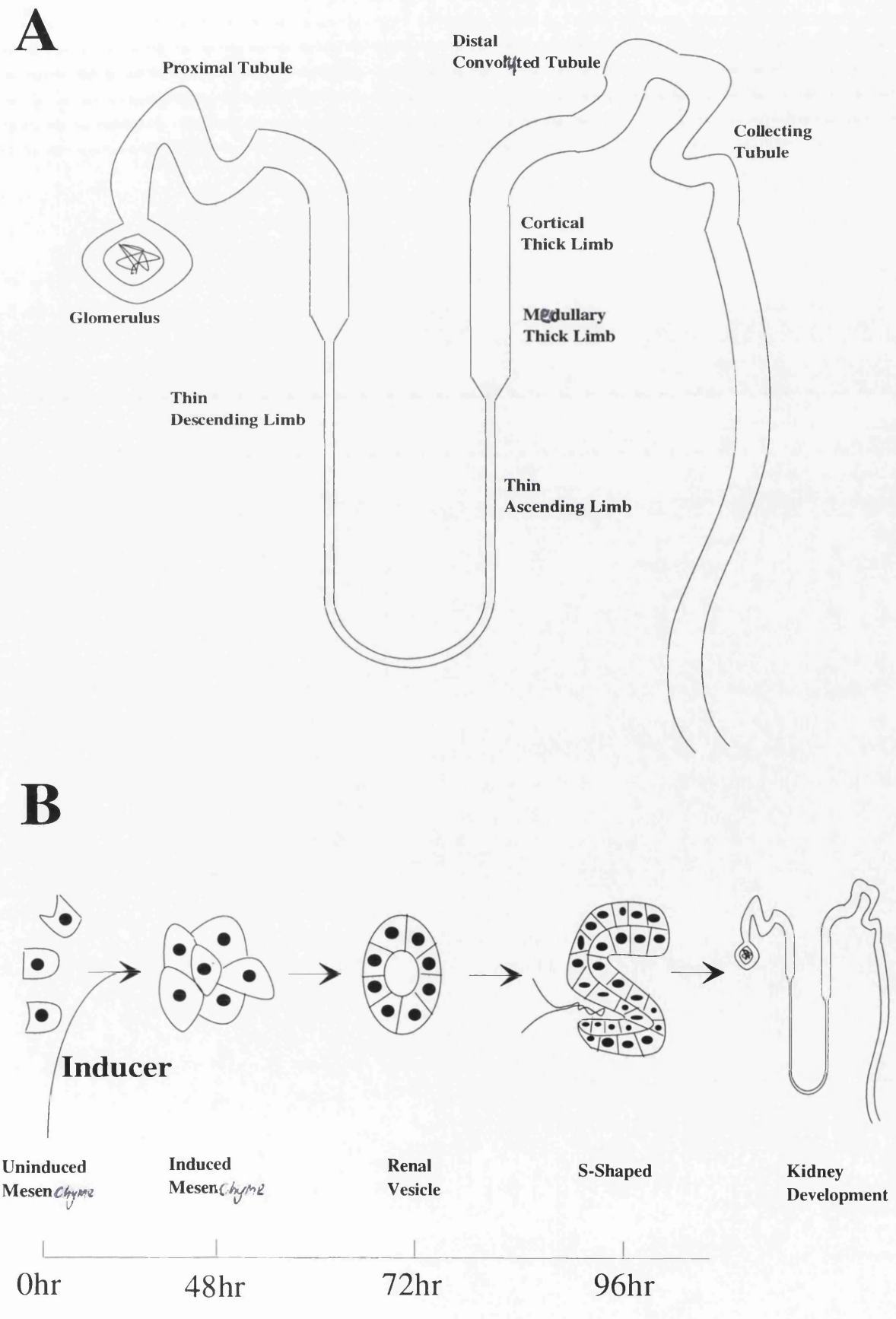
development of the peripheral nervous system and in the limb buds, kidney and heart. The observed defects are consistent with this expression (Hoyt *et al.*, manuscript submitted). However, it should be noted the knockout mice are not null mutants, because they still express the internal deletion of 324 amino acids form of the gene. It is possible a more severe phenotype might result from knocking out this transcript as well.

1.19 Murine kidney development

The high levels of *Evi-1* expression in kidney suggest that the gene plays a role in development of this organ. The kidney is a heterogeneous organ derived from mesenchymal cells which require external stimuli in order to differentiate and mature. Metanephric kidney development begins when a caudal outgrowth of the Wolffian duct, the ureteric bud, induces undifferentiated mesodermal cells of the metanephric blastema to proliferate (figure 1.2 B). Interestingly, mesodermal cell contact with the inducing tissue is required for only 24 hours for the cells to commit to advanced stages of nephrogenesis (Lehtonen *et al.*, 1983; Ekblom, 1981). This is the 'determination' point which implies that the mesodermal cells can be transplanted to another site and still differentiate to form renal tubules, so long as the appropriate extrinsic signals are supplied. In fact, inductive interaction between ureteric bud epithelium and kidney mesenchyme may be reproduced *in vitro* and has been a useful tool in the study of kidney morphogenesis (Grobstein, 1956). However, although the inducing signals are well defined the signalling molecules mediating mesenchyme induction are unclear.

Co-ordinately with 'determination', mesodermal cells begin to express characteristic epithelial cell adhesion molecules (Grobstein, 1957; Saxen *et al.*, 1983, Vestweber *et al.*, 1985; Vainio *et al.*, 1989). The inductive signals of the ureteric branches stimulate the cells to aggregate and condense to form a comma-shaped 'renal

Figure 1.2: Schematic representation of the kidney and kidney development



(adapted from Herzlinger *et al.*, 1992)

vesicle' which fuse to the ureter. This gives rise to the S-shaped body which is the first evidence of segmentation of the primitive nephron, the functional unit of the kidney (figure 1.2 B). The S-shaped body is invaded by primitive blood vessels at one pole which eventually becomes the glomerulus while the proximal tube develops from the adjoining region. The other pole, which is closer to the ureter, forms the loop of Henle and distal tubule (figure 1.2 A + B).

Intriguingly, programmed cell death appears to be widespread in kidney development and the bcl-2 protein is highly expressed in mesenchymal condensates, being down regulated during maturation of the nephron, suggesting a possible role for apoptosis in early kidney development.

The molecular events downstream from the initial inductive signalling events between the ureteric bud and the undifferentiated mesenchyme are unclear as are the signalling events. Presumably, the early proliferation and differentiation signals of the induced mesenchyme must be mediated through the activation of specific sets of transcription factors, at least some of which must be directly activated via signalling cascades from the cell surface.

A number of transcription factors have been identified which appear to play an important role in kidney development. Included in these is the *Pax-2* gene which is upregulated shortly after ureteric bud induction (Dressler *et al.*, 1992) and its suppression by the use of antisense oligonucleotides results in a severe decrease in condensing mesenchyme (Rothenpieler & Dressler, 1993). This is consistent with the gene being important specifically in the early inductive response, a point which is further supported by its downregulation later during tubular differentiation. Interestingly, the failure to suppress *Pax-2* expression using transgenic mice studies, results in immature epithelium reflective of some types of human nephric syndromes (Dressler *et al.*, 1993).

Interestingly, concomitantly with the down regulation of the Pax-2 protein an induction is observed of the Wilms' tumour suppressor gene (WT1). Wilms' tumour is an embryonal kidney tumour thought to arise through aberrant mesenchymal stem cell differentiation following the inactivation of multiple tumour suppressor genes. A candidate gene, designated *WT1,* has been shown to contribute to the development of Wilms Tumour (Haber *et al.*, 1990). The homozygous WT1 knock-out mice are embryonic lethal because of defective heart development. In these mutants the ureter bud fails to grow out of the nephric duct and the mesenchyme is not induced (Kreidberg *et al.*, 1993). Interestingly, *in-vitro* co-cultivation experiments have shown that the mesenchyme is unable to respond to wild-type inducer, suggesting a role for WT1 in both mesenchyme induction and ureter outgrowth (Kreidberg *et al.*, 1993). The WT1 zinc-finger protein has been shown to act as a transcriptional repressor (Madden *et al.*, 1991) by binding to a DNA recognition sequence, cgccccgc, similar to that of the early growth response gene (EGR-1). Interestingly, WT1 may directly repress the *Pax-2* gene through three sites present in its promoter (Ryan *et al.*, 1995).

Other transcription factors known to be involved in kidney development include the lim1 and Pax-8 proteins which are transiently and spatially induced in the renal vesicle and are down regulated as the tubular epithelium matures (Fujii *et al.*, 1994). Furthermore, N-myc is activated in the condensing mesenchyme (Magrauer & Ekblom., 1991). The importance of N-myc in mesenchymal proliferation is further supported from gene knockout studies in mice which are embryonic lethal. It was observed that the epithelial component of the embryo failed to develop (Stanton *et al.*, 1992).

1.20 The importance of Evi-1 in kidney development

In situ hybridisation has identified *Evi-1* expression throughout the mesonephric duct and mesonephric tubules in day 9.5 *p.c.* embryonic kidneys with very

high expression seen in the tubules of the met anephros at day 14.5 *post coitus* (*p.c.*). Expression in adult kidney was seen throughout the tubules of the cortex and external medulla with the expression highest between the cortical-mudullary junction (Morishita *et al.*, 1991). No expression is detected in the glomeruli (Perkins *et al.*, 1991). Consequently, *Evi-1* appears to induced early in embryonic development and expressed primarily in the epithelial ducts and tubules. *Evi-1* expression is maintained in the adult is to epithelial tubular cells.

The type of tubule cells which show high levels of *Evi-1* expression is unknown. The kidney is heterogeneous in nature containing many different tubule cells (including proximal, distal convoluted tubule, collecting tubule, and the cells of the descending and ascending limb) which possess different differentiation states, potentially different transcription factor profiles and presumably different functions (figure 1.2).

The time of *Evi-1* upregulation relative to the initial inductive signals to the mesenchyme is unclear, however, the expression pattern observed at 9.5 days *p.c.* would overlap the transcription factors mentioned previously above. In contrast to the Pax genes and WT1 gene little is known about the target genes of the Evi-1 protein in kidney cells and its interplay with other transcription factors. In addition, little is known about the mechanisms controlling tissue specific expression of the gene throughout kidney development.

1.21 Project Aims

The aim of the project was to examine the molecular mechanisms controlling tissue specifc expression in kidney cells since the *Evi-1* pattern of expression is consistent with the *Evi-1* gene having an important role in the development of this organ. As there were no known high *Evi-1* expressing kidney cell lines, we wished to screen available kidney cell lines to obtain a high expressing cell line to carry out

functional analysis of the *Evi-1* promoter. In addition, since the transcription initiation sites have been mapped in kidney and ovary (Bartholomew & Ihle, 1991) we wished to extend these studies to identify the minimal *Evi-1* promoter.

Part 2: Materials and Methods

Chapter 2: Methods

2.1 Animal Cell Culture

Cell Lines

	EACC/ATCC No.	Tisse Origin
RAG	89040605	Kidney
G402	90112715	Kidney
ACHN	88100508	Kidney
TCMK-1	90050802	Kidney
293	85120602	Kidney
A704	HTB45	Kidney
YCR	Gift from M. Oshimura	Kidney
RCC23	Gift from M. Oshimura	Kidney
STO	CRL 1503	Fibroblast
L929	CCL1	Fibroblast
NIH3T3	1658	Fibroblast
DA-3	Bartholomew *et al.*, 1989	Haemopoietic
FDCP-1	Dexter *et al.*, 1980	Haemopoietic
HEC-1A	HTB112	Endometrial
HEC-1B	HTB113	Endometrial

Cell Culture conditions

Cell lines were obtained from Beatson Institute stocks. Most cell lines were maintained in special liquid medium supplemented with 10% Faetal calf serum and 2mM glutamine, gassed to 5% CO_2 and incubated at 37^{o}C. Cells cultures were passaged every 3-4 days to maintain them in a continously growing state. Viable cell numbers were evaluated using a haemocytometer. The myeloid DA-3 cell line was grown in RMPI-1640 media supplemented with 10% foetal calf serum, glutamine, and 20 U of purified IL-3 per ml.

Storage of animal cells

Animal cells were harvested and resuspended at 10^7 cells/ml in faetal calf serum and 10% (v/v) DMSO. Cells were transfered to 1ml Nunc cryotubes and stored overnight at -70°C. The cells were then transfered to liquid nitrogen for long term storage.

2.2 Isolation of kidney cells

Adult and foetal kidney dissections

Adult kidneys were dissected from either tsT or CD1 mice and placed on ice. The kidneys were then wash with PBS several times. The capsule surrounding the kidney was removed and then washed in PBS again. Embryonic kidneys were taken at day 18.5p.c.. Each embryo was dissected and the kidneys placed in ice cold PBS. The kidneys were then cut up into small pieces using a sterile scalpel and transfered to a 50ml Falcon tube prior to collagenase/trypsin.

Isolation of kidney cells by collagenase/trypsin treatment

Kidneys (embryonic or adult) were placed into a 50ml Falcon tube containing a autoclaved magnetic flea. 7.5ml of collagenase solution (3mg/ml) and 7.5ml of trypsin (final concentration of 0.25% made up in PBS) were pipetted into the tube. The tube was then placed on a magnetic stirrer in the hot room (37°C) for exactly 8 minutes. The debris was allowed to settle and the collagense/trypsin solution was carefully removed to a universal tube. A further 7.5ml of collagenase and trypsin was

added and the procedure repeated a further 2 times. To pellet cells the universal tubes were spun at 1000rpm/5 minutes/4°C. Cell pellets were washed in 7.5ml PBS containing DNAse 1 (0.05mg/ml) and repelleted as above. The final cell pellet is resuspended in medium. All fractions are pooled and the cell number evaluated by haemocytometer. Cells were seeded out at $1x10^7$ cells in a T^{175} Falcon flask and incubated at 33°C.

Kidney cell culture media

Isolated kidney cultures were routinely passaged in SLM media. In addition, epithelial selective media was also used as peviously described:
D-Valine (Gilbert & Migeon, 1975)
Serum free medium (Taub *et al.*, 1979)

2.3 Bacterial Culture

Host Strains

E.coli strains XL1 or DH5α were obtained from Beatson Institute stocks. Bacterial strains were grown in L-broth at 37°C with good airation.

Competent cells

Competent bacterial cells (DH5α) were obtained from group stocks held by Anne-Marie Clarke.

Transformation of bacterial cells

Competent cells were thawed slowly on ice and 100µl transfered to a pre-chilled 15ml falcon tube. Plasmid or ligation mix was pipetted into the cells, mixed gently by tapping the tube, and incubated on ice for 30 minutes. The cells were heat shocked at 42°C for 2 minutes. 900µl of L-broth was then added cells incubated for at 37°C for 1 hour in a New Brunswick G25 shaker at 225rpm. 50-200 µl was then spread onto 1.5% (w/v) agar L-broth plates containing the appropriate antibiotic. The plates were left to air dry and were then inverted and incubated at 37°C overnight.

Glycerol Stocks

E.coli strains containing important plasmids were stored as glycerol stocks. Stationary cultures were mixed with an equal volume of glycerol and stored at -20°C.

2.4 Nucleic Acids: DNA Analysis

Preparation of plasmid DNA

Bacterial colonies were picked using a sterile toothpick and grown overnight in 5ml of LB medium, containing the appropriate antibiotic, in an orbital shaker. 1.5ml of this culture was spun down in a microfuge 13000rpm/2 minutes/RT. The cells were resuspended in 100µl of solution I (50mM glucose, 25mM Tris/HCl pH 8.0, 10mM EDTA) and incubated on ice for 2 minutes. 200µl of solution II (0.2N NaOH, 1% SDS) was added the solutions mixed gently and incubated for 5 minutes on ice. A 150µl of solution III (3M Potassium Acetate, 11.5% glacial acetic acid) was then added, the tubes vortexed, and placed on ice for a further 5 minutes. After

incubation the tubes were spun at 13000rpm/5 minutes/4°C. The supernatant was removed to a fresh tube and phenol/chloroform extracted. Plasmid DNA was precipitated by the addition of 800µl of cold 100% ethanol and spun 13000rpm/5 minutes/4°C. The pellet was wash with 70% ethanol, dried in a vacuum dessicater for 2 minutes and resuspended in 32µl of TE containing 10µg/ml of RNAseA.

Large scale plasmid preparations

1ml of overnight culture was inoculated into 250ml of LB medium containing the appropriate antibiotics and grown overnight in an orbital shaker at 37°C. The culture was spun down at 5000rpm/20 minutes/4°C. The supernatant is discarded, the pellet resuspended in ice cold solution I (50mM glucose, 25mM Tris/HCl pH 8.0, 10mM EDTA) and incubated for five minutes at room temperature. Add 10ml of solution II (0.2N NaOH, 1% SDS) roll tube gently to mix. Incubate on ice for 10 minutes. Add 7.5ml of ice-cold solution 3 (3M Potassium Acetate, 11.5% glacial acetic acid). Incubate on ice for 10 minutes. Spin at 5000rpm/20 minutes/4°C. Filter supernatant through several layers of miracloth into a 50ml Falcon tube. Add 0.6 volumes (12ml) isopropanol at room temperature to precipitate the DNA. Mix and incubate at room temperature for 15 minutes. Centrifuge at 5000rpm/30 minutes/room temperature. Discard the supernatant and resuspend the pellet in 8ml of TE (100mM Tris/HCl pH8.0, 1mM EDTA). Add 8.4g of Caesium Chloride and 400ul of ethidium bromide solution (10mg/ml). Centrifuge at 5000rpm/five minutes/RT to separate out ethidium bromide/protein complexes. Transfer the clear red solution to an appropriate ultracentrifuge tube using a Pasteur pipette. Spin at 55000rpm overnight at room temperature. Plasmid bands are remove with needle and syringe. The ethidium bromide is removed from the DNA solution by butan-1-ol

extraction's. The solutions are then dialysed against TE, ethanol precipitated with 1/10th volume 3M sodium acetate pH 4.6. The precipitates are resuspend in TE.

Quantification of nucleic acids

Nucleic acids were quantified by measuring absorbance (A) at 260nm and 280nm in a Beckman DU 650 spectrophotometer. 5µl of sample was added to 495µl of distilled water and added to a quartz cuvette blanked against distilled water. An A_{260} value of 1 was taken to be equivalent to 50µg/ml of plasmid or genomic DNA, 40µg/ml of RNA and 20µg/ml of oligonucleotide. The A_{260}/A_{280} ratio was taken to be a sign of purity with samples below 1.75 requiring further purification.

Restriction digests

Restriction digests of plasmid DNA were carried out using the buffered solutions sent with the enzymes from the supplier. Typically, plasmid DNA of concentration <3ug were digested in a total of 20µl with 5-10units of enzyme per µg of DNA in 1x buffer for 1-2 hours at the appropriate temperature (usually 37°C). The total volume of the reaction was proportionally increased for larger concentrations of DNA. Reactions were terminated by the addition of 1/10th volume of agarose gel loading buffer (0.25% (w/v) xylene cyanol, 0.25% (w/v) bromophenol blue, 25% (v/v) Ficoll in water).

Gel electrophoresis

Typically 1% agarose gels were made to resolve restriction enzyme digested DNA. 2g of agarose (low melting point agarose was used for preparative gels) was

added to 1x TAE (50x TAE is 2M Tris, 50mM EDTA, 57.1ml/l glacial acetic acid) upto a volume of 200ml. The solution was then heated in a microwave and placed on a stirring block to cool. When hand hot ethidium bromide was added to 0.5µg/ml and the solution poured into an appropriate gel former. Gels were allowed to set at room temperature and transferred to an electrophoresis tank containing 1x TAE. Samples, containing 1/10th volume agarose gel loading buffer (section) were loaded into separate wells and run at 40-120 volts depending on the size of the fragments and the time available. The molecular weight standard used was 1kb DNA ladder. The DNA bands were visualised using a U.V. source and photographed.

Purification of DNA fragments from agarose gels

DNA fragments were excised from the agarose gels and placed into an eppendorf tube. DNA was extracted using a Geneclean kit according to the manufacturer's protocol. Prior to further manipulation, the quality and quantity of the purified fragment(s) were assessed by electrophoresis of a small aliquot of the DNA on the agarose gel.

2.5 Subcloning of DNA

Dephosphorylation of vector DNA

Plasmid vector were digested with the appropriate enzyme (section). At the end of the reaction 1ul of calf intestinal alkaline phosphatase (CIAP) was added and the reaction incubated for a further 30 minutes. For the generation of blunt-ended vectors, the reaction was transferred to 56^{o}C for 30 minutes with the addition of a further 1ml of CIAP. The vector was then agarose gel purified.

Generation of inserts

Inserts were generated by restriction enzyme digestion of either plasmid or PCR DNA and were agarose gel purified.

Ligation of DNA fragments into plasmids

The vector and insert DNA was then included in the following reaction mix,

20ng of dephosphorylated plasmid

5 molar excess of DNA insert

2ul of 5x Ligation buffer (0.25M Tris.HCl (pH7.6), 50mM $MgCl_2$, 5mM dATP 5mM dithiothreitol, 25% (w/v) polyethylene glycol-8000)

1µl of T4 DNA ligase (1U/µl)

H_2O upto 10µl

The reaction mix was incubated at 16°C overnight.

2.6 Polymerase Chain Reaction (PCR)

PCR primers

Oligonucleotide primers were designed such that they were complementary to opposite strands and opposite ends of the DNA of interest. In some cases extra non-specific bases were synthesised at the 5' end to facilitate subsequent cloning into the appropriate restriction enzyme site in the plasmid DNA.

Synthesis of DNA oligonucleotides

Oligonucleotides were synthesised at the Beatson Institute on an Applied Biosystems model 381A DNA synthesiser according to the manufacturers instructions. 5'trityl groups were removed by the machine, and the DNA immobilised on a column. The DNA was eluted in 29% (v/v) ammonia by passing the solution through a column approximately 20 times every 10 minutes for 1.5 hours. This solution was sealed in a glass vial an incubated overnight at 55°C. The DNA was precipitated by the addition of ammonium acetate to 0.3M and 3 volumes of 100% ethanol. The DNA was pelleted at 7000rpm for 15 minutes in a Sorvall SS-34 rotor, resupended in 0.2M sodium acetate, and reprecipitated in 3 volumes of ethanol. The pellet was then washed in 70% (v/v) ethanol and dissolved in 1ml of distilled water. The oligonucleotides were quantified by spectrophometry (section).

Oligonucleotide PCR Primers

5' and 3' Functional Constructs

1	ccagatcttttaagtagaaacttaa
198	ccagatctcctgaaagcaattctcc
242	ccagatctctctaacacagactctc
347	cctgtgctcaggacatc
401	cctggagaaatttcccc
431	cggcagaaatctacatg
488	ggacttgggtcttggcg

553 tttactgaaagaggagc

630 ccagatctctttctggatggccgag

670 ccagatctgaaaccctgcggctggg

R955 ccaagcttgagctcgcagctattcc

Footprinting -474 to -135 construct

SFP ctccaccctatctttcc

SFP3' ttaaagtgacagcagcc

PCR reactions

The appropriate plasmid template of interest was added at a concentration of 100ng/reaction together with 20ng of each PCR primer. The reaction also contained 10ml Taq DNA polymerase buffer, 2µl from 5mM stock solutions of dATP, dTTP, dGTP and dCTP, 6µl of 25mM $MgCl_2$, 0.5µl of Taq polymerase and distilled water to 50µl. The reaction is then overlaid with 50µl of paraffin oil and incubated in a PCR machine (Perkin-Elmer) for 30 cycles 1) 94°C for 1 minute; 2) 50°C for 1 minute; 3) 72°C for 90 seconds. Reactions were incubated for 15 minutes at 72°C after the final cycle, then cooled to 4°C. Each reaction was purified from agarose gel (section).

2.7 Sequencing Plasmid DNA

Sequencing of plasmid DNA was performed using the Sequenase Version 2.0 kit supplied by USB according to the manufacturers instructions. The plasmid DNA was denatured as followed, 4µg of plasmid DNA was added to a total volume of 20µl 0.2M NaOH and the solution incubated at room temperature for 5 minutes. The

DNA was precipitated by the addition of 100µl of ethanol and 8µl of 5M ammonium acetate pH 5.2 and the solution incubated on ice for 10 minutes. The DNA is then pelleted in a microfuge at 13,000rpm for 5 minutes at 4°C. The pellet is washed in 70% ethanol, allowed to air dry and resuspended in 7 µl of distilled water.

Sequencing Primers

Footprinting -474 to -135 construct

SFP	ctccaccctatctttcc
SFP3'	ttaaagtgacagcagcc

DHS I region

T3	attaaccctcactaaag
T7	aatacgactcactatag
PJ1	gaggtaggagcatcct
PJ2	gtggtgggttcacctagg
PJ3	gtcgctccgcctccc
PJ4	ggggtctaatgaacat
PJ5	ccgtcccgagataccc
RPJ1	ggtatttaaccttacg
RPJ2	gggtttggtaaggttc
RPJ3	taaattcacatttgtc
RPJ4	cctaggtgaacccaccac

2.8 Southern blot analysis

Following DNA separation by electrophoresis the gel was soaked for 15 minutes in HCl then washed with distilled water and soaked in alkaline buffer (1.5M NaCl, 0.5M NaOH) for 30 minutes. The DNA was transferred to Hybond N+ nylon membrane in alkaline buffer for 2 hours using a vacuum blotter. The filter was washed in 2xSSC and the DNA fixed using a UV Stratagene 1800 crosslinker.

Random-primed radiolabelling of DNA probes

DNA probes used in Northern to Southern blotting were prepared as directed in the manufacturers instructions. Following labelling unincorporated radionucleotides were removed using Sephadex G50 columns. The probes were denatured by boiling for 5 minutes immediately prior to addition to the filter.

Hybridisation conditions

All hybridisations were carried out at 42^{o}C in 50% formamide in accordance with the manufacturers instructions. Pre-hybridisations were carried out for a minimum of 4 hours and then replaced with denatured labelled probe at 10^{6}CPM/ml in fresh hybridisation solution. The filters were incubated overnight and washed in accordance with the manufacturers instructions

2.9 Nucleic Acids: RNA Analysis

Isolation of RNA

Total cytoplasmic RNA was prepared using the RNAzol B method (Biogenesis Ltd) according to the manufacturers instructions.

Agarose electrophoretic separation of RNA

To 87ml of DEPC treateed H_2O, 1g of agarose and 10ml 10xMOPS (3-(N-morpholino)propanesulphonic acid, 50mM sodium acetate, 10mM EDTA pH7.0) was added. This was disolved by heating in a microwave, cooled to hand hot and then 5.1ml formaldehdye solution was added. The gel was then cast and run in 1xMOPS buffer. Samples of RNA were dissolved in DEPC treated water and 10µl sample buffer (which was made up fresh before use: 0.75ml deionised formamide, 0.15 10xMOPS, 0.24ml formaldehyde, 0.1ml water, 0.1ml glycerol, 0.08% bromophenol blue) was added per 2µl of sample prior to electrophoresis. An RNA ladder (BRL) was run to provide markers on each gel.

Northern Transfer of RNA

Following RNA separation by electrophoresis the gel was soaked in distilled water for 10 minutes and then 1hour in 20x SSC. The RNA was transferred onto Hybond N+ nylon membrane (Amersham International) in 20x SSC using a vacuum blotter. The RNA was fixed onto the membrane using a UV Stratagene 1800 crosslinker.

2.10 DNAse 1 hypersensitivity analysis of chromatin structure

Isolation of nuclei from tissue

Frozen tissue samples are ground in the presence of liquid nitrogen and homogenised in a dounce homgeniser in A/NT/L (20:30:50) (Solution A: 0.6M sucrose, 120mM KCl, 15mM NaCl, 0.3mM spermine, 2mM spermidine, 28mM β-mercaptoethanol, 4mM EDTA, 2mM EGTA, 0.2mM PMSF, 2mM DTT, 0.2% Triton X-100, 10mM Tris.HCl pH7.9) (Solution NT: 15mM NaCl, 10mM Tris.HCl pH7.9) (Solution L: 10mM NaCl, 0.1% NP-40 (v/v), 10mM Tris.HCl pH7.9). The homogenate is filtered through several layers of mira cloth. Nuclei are precipitated by centrifugation at 1000rpm for 10 minutes at 4^{o}c. The nuclei are stained using trypan blue and counted for viability using a haemocytometer. The nuclei are resuspended in A/NT (50:50) and then pelleted as above. The nuclei are resuspended at 10^{8}/ml in storage buffer and stored at -70^{o}c.

Isolation of Nuclei from Cell Lines

1 x10^{8} cells are pelleted at 1000rpm at 4^{o}C. Cells are washed in Hanks balanced salt solution and pelleted as above. Cells were resuspended in 17.5ml of ice cold nuclear isolation buffer 1 (NIB 1) (60mM KCl, 15mM NaCl, 5mM $MgCl_2$, 0.1mM EGTA, 15mM Tris.HCl pH 7.4, 0.5mM DTT, 0.1mM PMSF, 0.3M sucrose). Cells were disrupted by addition of NIB 2 (60mM KCl, 15mM NaCl, 5mM $MgCl_2$, 0.1mM EGTA, 15mM Tris.HCl pH 7.4, 0.5mM DTT, 0.1mM PMSF, 0.3M sucrose, 10% NP-40). Pipette 15ml of NIB3 (60mM KCl, 15mM NaCl, 5mM $MgCl_2$, 0.1mM EGTA, 15mM Tris.HCl pH 7.4, 0.5mM DTT, 0.1mM PMSF, 1.7M sucrose) into a 36ml Sorvall polypropylene tube. Layer NIB1/2 lysate onto NIB 3 solution. Spin

13000rpm for 15 minutes in a pre-chilled centrifuge and rotor at 4°C. Resuspend nuclei in 5ml NIB 4 (60mM KCl, 15mM NaCl, 5mM $MgCl_2$, 0.1mM EGTA, 15mM Tris.HCl pH 7.4, 0.5mM DTT, 0.1mM PMSF, 0.3M sucrose, 5% glycerol). Nuclei are either DNase1 digested immediately or stored at -70°c.

Digestion of nuclei with DNase1 enzyme

Immediately prior to digestion the nuclei are thawed (if from -70°C) and spun at 12,000rpm in a microfuge. The nuclei were resuspended in digestion buffer (60mM KCl, 15mM NaCl, 5mM $MgCl_2$, 0.1mM EGTA, 15mM Tris.HCl pH 7.4, 0.5mM DTT, 0.1mM PMSF, 0.3M sucrose, 5% glycerol) and split into 500µl aliquots. The DNAse1 enzyme was serially diluted in a range from 20µg/ml to 0.1 µg/ml. The diluted enzyme was then added to the appropriate tube and incubated at 37°C for 3 minutes. The reaction was stopped by the addition of 50µl of stop buffer (5% SDS, 100mM EDTA). 50µl of proteinase K was then added and the samples incubated at 37°C overnight. DNA was isolated by sequential phenol and then chloroform extractions and ethanol precipitated.

2.11 Protein Analysis: Isolation of Nuclear extracts

Nuclear protein preparation

All manipulation were carried out at 4°C, and all solutions contained 0.5mM PMSF, 0.5mM benzamidine, 10mM sodium butyrate, 10mM β-glycerolphosphate, 2mM levamisole, 50mM sodium orthovandate pH8.0, and 1µg (each) of leupeptin, aprotinin, bestatin and pepstatin per ml.

Minipreparation of nuclear extracts

Nuclear extracts were prepared using $1x10^7$ cells grown in the appropriate media. Suspension cells were spun down and washed twice with 1.5ml ice cold PBS, adherent cells are scraped of into 1.5ml of ice cold PBS. Cells were spun in a microfuge for 10 seconds and resuspended in 400ul of buffer A (10mM HEPES-KOH pH7.9, 1.5mM $MgCl_2$, 10mM KCl, 0.5mM DTT, 0.2mM PMSF) by flicking the tube. The cells were allowed to swell on ice for 10 minutes, and then vortexed for 10 seconds. The samples were then span for 10 seconds in a microfuge, resupended in 100µl of buffer B (20mM HEPES-KOH pH7.9, 25% glycerol, 420mM NaCl., 1.5mM $MgCl_2$, 0.2mM EDTA, 0.5mM DTT, 0.2mM PMSF) and incubated on ice for 20 minutes. Cellular debris was removed by spinning the cells at 12000rpm for 2 minutes at 4°C. The supernatant was removed, aliquoted and stored at -70°C. The yield was typically between 50-75µg protein /106 cells.

Large scale preparation of protein nuclear extracts

10^{10} cells were washed in phosphate-buffered saline and then twice in TMS (5mM Tris-HCl pH 7.5, 2.5mM $MgCl_2$, 125mM sucrose). Cells were lysed in 200ml of TMS plus 0.25% Trion-X-100, and the nuclei were harvested by centrifugation at 2000rpm for 20 minutes at 4°C. Nuclei were washed three times in 200ml of TMS and resuspended in approximately 5ml of TMS (5-10mg of DNA per ml), and 0.1 volumes of 4M NaCl was added dropwise while stirring. The solution was centrifuged at 17,000 rpm for 30 minutes, and the supernatant spun at 35,000 rpm for 60 minutes. Solid ammonium sulphate was added to 0.35 g/ml and left on ice for 30 minutes. The precipitate was pelleted at 17,000rpm for 30 minutes and redisolved in 5ml of E50 buffer (50mM ammonium sulphate, 20mM HEPES pH7.9, 5mM $MgCl_2$,

0.1mM EDTA, 0.1% (v/v) glycerol, 1mM DTT). The crude protein extract was cleared by centrifugation at 35,000rpm for 60 minutes, and aliquots were stored at -70°C.

2.12 Protein Analysis: Electrophoretic mobility shift assays

Oligonucleotides

The following oligonucleotides were used in EMSA analysis;

-341	caccettgtgctcaggacatcaaaaccagagagattt
-314	agagagatttccctggggaaacaaatcctgcc
uv	agagaga555ccc5ggggaaacaaa5cc5gcc
-303	cctggggaaacaaatcctgcctggagaaatttccccatt
FpIII	agagagatttccctggggaaacaaatcctgcctggagaaatttccccatt
-277	aaatttccccattggttgtttatcggcagaaatctacat
-237	tttctggggatggtgcatctataatcagtctgtccctatagg
-171	tgacctctcccgccagaggaggctgctgtcactttaaa
-110	cgtctggcttccgaccactctggag
+83	ccgtcctgcctggcaggcccctaccttcg
+105	taccttcgcacactttcctcctgcgggtct
-294	gatcaaacaaatcctgcc
-305	gatctccctggggaaaca
-296	gatcagagagatttccct
fb	gatctttccctggggaaa

5= bromodeoxyuridine

Competitor Oligonucleotides

AP2	ccggccccaggct
ZF1	ggatccctggacaagataagataaaggcagatct
ZF2	gatccctagctatgtgcctcatcttcgacaggctgtct
C/EBP	gatcggatcccaattgggcaatcaggggatcc
NF-1	cttttggattgaagccaatatgaga
SP1	tgggaatcctaactgggcggagttatgctggtggtg
NF-κB	agttgaggggactttcccagg
GATA-1	ccgggcaactgataaggattccctg
PU1	gatccataacctctgaaagaggaacttggttaggt
PEA3	gatcctcgagcaggaagttcga
E74	gatctctagctgaataaccggaagtaactcatccta
E Box	cccccaccacgtggtgcctga
H4TF-1	tttagatttcccctcccccaccggggcggga
Eα	ctacacctataaaccaatcac

Labelling of double stranded probe

5 μg of each complimentary oligonucleotides were annealed in TE to give a final concentration of 100ng/μl. 200ng of the annealed oligonucleotide was radiolabelled with [γ-^{32}P]ATP using T4 polynucleotide kinase according to the suppliers instructions. The labelled oligonucleotide was purified after separation by PAGE. The appropriate band was visualised by autoradiography, cut out and eluted in 1 ml of TE overnight at 37^{o}C.

Binding reaction

500pg of radiolabelled oligonucleotide probe was incubated with 5µl of 4x reaction buffer (100mM KCl, 10mM Tris/HCl pH7.6, 5mM $MgCl_2$,1mM EDTA, 1mM DTT, 12.5% glycerol), 6µg poly (dI.dC), 5µl nuclear extract (typically approximately 10ug) and water to 25µl. The reaction was incubated at room temperature for 20 minutes. Where appropriate, cold competitor was added prior to the addition of labelled oliognucleotide and the reaction pre-incubated for 10 minutes on ice. Samples were run on a 6% acrylamide gel in 0.5xTBE running buffer for 2 hours/120 volts/4^{o}C. The gel was then dried prior to autoradiography.

2.13 DNAse 1 footprinting

The appropriate footprinted region was generated by PCR and subcloned into a bluescript KS+ vector. Footprinting probes were prepared by 5' labelling of DNA restriction fragments with T4 polynucleotide kinase and [γ-^{32}P]ATP and isolated after secondary restriction (Plumb and Goodwin, 1988). DNAse 1 footprint protection assays were performed in a volume of 100µl of storage buffer in the presence of 6ug of poly(dI-dC), 2ng of end-labelled restriction fragment, and upto 80µl of nuclear protein extract. Following partial DNAse digestion the nucleic acid was purified and resolved by denaturing PAGE and autoradiography (Plumb and Goodwin, 1988). Marker sequencing reactions were performed on the plasmid DNA using the Sequenase version II kit.

2.14 Western bloting

Isolation and quantification of protein samples

Monolayer cultures were washed twice with PBS. The cells were then lysed in RIPA buffer (150mM NaCl, 1% NP-40, 0.5% Deoxycholate, 0.1% SDS, 50mM Tris-HCl pH7.6) and scraped off the culture flask to a 1.5ml conical tube. The cells were spun at 4000rpm for 10 minutes at 4^{o}C. The lysate is removed to a fresh tube and stored at -70^{o}C. Protein concentrations were quantified using a BioRad protein assay kit according to the manufactureres instructions.

SDS-PAGE electrophoretic separation

SDS-PAGE gels were cast and protein samples electrophorised in vertical polyacrylamide gels as previously described (Sambrook et al, 1989). Samples were run with pre-stained size markers.

Westen blotting

Protein samples were tranfered to ECL membrane (Amersham) using a Sartorius semi-dry blotter according to the manufacturers instructions.

Immunoprecipitation

Protein samples were pre-cleared with normal rabbit serum before addition of the appropriate primary antibody. Protein/antibody complexes were collected by the addition of protein A sepharose as described previously (Lane and Harlow, 1988).

2.15 UV Crosslinking

FpIII oliogonucleotides were synthesized with bromodeoxyuridine substituted for thymidine bases, annealed, rediolabelled and used as probes in gel retardation reactions as described previously. Protein/DNA coplexes were resolved on a 6% nondenaturing PAGE gel. The relevant band was excised from the gel, illuminated with U.V. light (312 nm) for 30 minutes and then soaked in SDS loading buffer (62.5 mM Tris-HCl, pH6.8, 5% SDS, 5% β-mercaptoethanol, 10% glycerol, 2% bromophenol blue) for 1 hour. The gel peices containing either U.V. treated or untreated cross-linked complexes were placed directly into the wells of a 7.5% SDS-polyacrylamide gel and electrophorised. The gel was then dried, autoradiographed and typically exposed for several days.

2.16 Transfection into mammalian cell lines

DOTAP

60mm cuture dishes were seeded with $1x10^6$ cells the day prior to transfection so that the cells were approximately 80% confluent. DOTAP reaction were set up as described in the maufacturers protocol (Boehringer Mannheim). Typically 5ug of reporter plasmid and 5ug of the internal control containing a β-galactosidase gene (pHSV β-gal) plamid were used in each reaction. In reactions which were corrected for molar ratio, the total amount of DNA was kept constant by the addition of pUC19 to the appropiate samples.

$CaPO_4$

60mm cuture dishes were seeded with $1x10^6$ cells the day prior to transfection so that the cells were approximately 80% confluent. Reaction were set up as described in the manufacturers protocol.

2.17 Reporter Gene Assays

Luciferase Assays

60mm cuture dishes were seeded with $1x10^6$ cells the day prior to transfection so that the cells were approximately 80% confluent. Reaction were set up as described in the manufacturers protocol (Promega).

β-*galactosidase assays*

Cell extracts were isolated using the lysis buffer supplied in the luciferase assay kit. Typically, 50µl of extract was incubated with 500µl of Solution I (60mM Na_2HPO_4, 40mM NaH_2PO_4, 10mM KCl, 1mM $MgCl_2$, 50mM β-mercaptoethanol) and 100ml of Solution II (60mM Na_2HPO_4, 40mM NaH_2PO_4,, 2mg/ml ONPG). The reactions were then incubated at 37°C for 10-90 minutes or until the appearance of a yellow colouration. The reactions were then stopped by the addition of 250µl of 1M Na_2CO_3. The samples were read at an absorbance of 420nm in plastic 2ml cuvettes.

Chapter 3: Materials

3.1 Tissue culture media and supplies

Supplier: Beatson Institute Central Services

Sterile PE

Sterile PBS

Sterile glassware and pipettes

Sterile water

Sterile CT buffer

Penicillin (7.5mg/ml)

Streptomycin (10mg/ml)

Supplier: Gibco Europe Life Technologies Ltd., Paisley, Scotland.

Special Liquid Medium

Foetal Calf Serum

200mM Glutamine

2.5% (w/v) Trypsin

Sodium bicarbonate

Non-essential amino acids (100x)

Gentomycin (10mg/ml)

Supplier: Fisons Scientific Equipment, Loughborough, Leics., England.

DMSO

Supplier: A/S Nunc, Roskilde, Denmark.

Tissue culture flasks

Nunc tubes

Supplier: Becton Dickenson Labware, Plymouth, England.

100mm/60mm diameter tissue culture dish

Supplier: Northumbria Biologicals Ltd., Cramlington, England.

RPMI-1640

3.2 Bacterial media

Unless otherwise stated bacterial media components were purchased from BDH Chemicals Ltd., Poole, Dorset, England.

Supplier: Beatson Institute central services

L-Broth (prepared as according to Sambrook et al., 1989)

Sterile glassware

Supplier: Difco, Detroit, Michigan, USA.

Agar

Supplier: Sigma Chemical Co., Poole, Dorset, England.

X-gal

Ampicillin

3.3 Plasticware

Supplier: Bibby-Sterilin Ltd., Stone, Staffs, England.

30ml universal tubes

Bacteriological dishes

Supplier: Becton Dickenson Labware., Plymouth, England.

Falcon Tubes

3.4 Plasmids

Supplier: Promega., Madison, Wisconsin, USA

pGL2 luciferase reporter vectors

3.5 Kits

Supplier: Promega., Madison, Wisconsin, USA

Sequenase version 2.0 kit

Luciferase Reaction Mix, Lysis Buffer Kit

Supplier: Boehringer Mannheim UK., Lewes, East Sussex, England.

Random-primed DNA labelling kit

Supplier: Invitrogen

TA cloning Kit

3.6 Membranes, paper and X-ray film

Supplier: Amersham International plc., Amersham, Bucks, England.

Hybond N^+ nylon membrane

Supplier: Bethesda Research Laboratories, Gibco Ltd., Paisley, Scotland.

Dialysis tubing

Supplier:

Dialysis tubes

Supplier: Whatman International Ltd., Maidstone, Kent, England.

3MM filter paper

Supplier: Eastman Kodak Co., Rochester, New York, USA.

X-ray film (X-OMAT AR5)

Supplier: Presentation Technology Ltd., Clydebank, Scotland.

3.7 Nucleotides, polynucleotides, RNA and DNA

Supplier: Amersham International plc., Amersham, Bucks, England.

Redivue [α-^{32}P] dCTP 3000 Ci/mmol

Redivue [γ-^{32}P] dATP 5000 Ci/mmol

Redivue [α-^{35}S] dATP 1000 Ci/mmol

Unlabelled dATP, dGTP, dCTP, dTTP

Supplier: Parmacia Ltd., Milton Keynes, Bucks, England.

Poly [dI-dC]

Supplier: Sigma Chemical Co. Ltd., Poole, Dorset, England.

Salmon sperm DNA

Supplier: Bethesda Research Laboratories, Gibco Ltd., Paisley, Scotland.

RNA ladder 0.24-9.5kb

Bacteriophage 1 DNA markers (HindIII cut)

3.8 Enzymes

Supplier: Northumbria Biologicals Ltd., Cramlington, Northumberland, England.

Klenow DNA polymerase (1U/µl)

T4 polynucleotide kinase (10U/µl)

DNA Ligase (10U/µl)

Supplier: Lorne Laboratories Ltd., Twyford, Reading, Barks, England

DNAse 1 (Molecular biology grade, RNase and protease free)

Supplier: Boehringer Mannheim UK., Lewes, East Sussex, England.

RNAse A

Calf intestinal alkaline phosphatase (1U/µl)

Supplier: Bethesda Research Laboratories, Gibco Ltd., Paisley, Scotland.]

All enzymes not listed above.

3.9 Chemicals

All chemicals not listed both here and above were obtained (AnalaR grade) from BDH Chemicals Ltd., Poole, Dorset, England.

Supplier: Sigma Chemical Co. Ltd., Poole, Dorset, England.

Spermidine

Spermine

Dithiothreitol

Triton-X-100

Diethylpyrocarbonate (DEPC)

Bromophenol blue

Xylene cyanol

MOPS

Supplier: Pharmacia Ltd., Milton Keynes, Bucks, England.

Ficoll 400

Supplier: Rathburn Chemicals Ltd., Walkerburn, Scotland.

Water-saturated phenol

Supplier: Northumbria Biologicals Ltd., Cramlington, Northumberland, England.

Bovine Serum Albumin (20% w/v)

Supplier: James Burrough Ltd., Witham, Essex, England.

Ethanol

Supplier: Fluka Chemika-Biochemika AG, Buchs, Switzerland.

Formamide

Supplier: Fisons Scientific Equipment, Loughborough, England.

Formaldehyde

Supplier: Cinna/Biotecx Laboratories Inc., Houston, Texas, USA.

RNAzol B

3.9 Water

Distilled water for buffers and general solutions was obtained from Milipore MilliRO 15 system. Water for protein/enzyme work or recombinant DNA protocols was further purified on a Millipore MilliQ system to 18M cm.

Part 3: Results

Chapter 4: DNAse1 Hypersensitivity Site Analysis of the *Evi-1* Locus

DNA is generally tightly associated in condensed higher order structures with chromatin. In this state the DNA is less accessible to transcription factor protein interaction and, as a consequence, unable to initiate transcription. In contrast to the bulk of chromatin, transcriptionally active regions of DNA acquire a more relaxed, or unwound, structure with the chromatin framework. It has been shown that structural relaxation surrounding transcriptionally active genes also correlates with a general increase in the sensitivity to nuclease digestion and that small 15-200bp regions exhibit extreme DNAse1 sensitivity termed DNAse1 Hypersensitivity Sites (DHS) (Wu *et al.*, 1979 a & b). These DHS regions are generally associated with active promoter and enhancer elements (reviewed by Gross and Garrard, 1988). Consequently, mapping DNAse1 sensitivity is a useful method of identifying gaps in the nucleosomal array and therefore the location of potentially transcriptionally active genes.

4.1 Chromatin structure of the 5' *Evi-1* locus in murine kidney tissue

The *Evi-1* gene is expressed in a tissue specific manner, the kidney having one of the highest levels of expression (Morishita *et al*, 1990, Perkins *et al*, 1991). To identify the regions of DNA important in controlling this tissue specific expression DHS analysis was performed with nuclei isolated from murine adult kidney. The nuclei were digested with DNAse1 and subsequently the genomic DNA was isolated, digested with appropriate restriction enzymes and examined by Southern blot analysis with the appropriate ^{32}-P labelled probe. Two DNA probes were used in all the DHS analysis which covered an 18kb region of the 5' *Evi-1* locus: either a 2kb EcoRI/BamHI probe (probe b) or a 500bp EcoRI/SstI probe (probe a); both originally derived from a λ

bacteriophage clone, designated λCB9 (figure 4.1, adapted from Bartholomew & Ihle, 1991).

Initially, DHS analysis was performed on a 5kb EcoRI fragment using probe b (figure 4.2B). Southern blot analysis with probe b detects a unique 5kb EcoRI fragment (figure 4.2 A, lane 1). However, increasing DNAse1 digestion reveals the presence of two hypersensitivity sites which are clearly visible (figure 4.2A, lanes 7 and 8). An intense 2.3kb band, designated hypersite I, maps approximately 2kb upstream of the major transcription initiation sites while the less intense hypersites, designated hypersite II, are located within exon I. The differences in intensity between the hypersites may be due to: 1) a diluting effect of DNAse1 digestion at hypersite I since this hypersite is closer to the probe or: 2) the relative sensitivity of the region to DNAse1 which can be affected by the proteins binding to the DHS regions.

As mentioned previously DHS sites are generally associated with active promoter and enhancer elements. The location of the DHS II hypersites over the multiple transcription initiation sites, previously identified in *Evi-1* expressing kidney and ovary tissues (Bartholomew & Ihle, 1991), is consistent with this association.

In contrast, DHS I was mapped to approximately 2kb upstream of the major initiation sites. To confirm this location the filter was striped and reprobed with probe a, located at the opposite end of the 5kb EcoRI fragment. (figure 4.3 B). Increasing DNAse1 digestion reveals a 2.7kb band (in contrast to the 2.3kb band with probe b) representing hypersite I (figure 4.3 A, lanes 7 and 8). This confirms the location of hypersite I as 2.7kb upstream from the 3' EcoRI site and approximately 2kb upstream of the major initiation sites (figure 4.3 B). DHS I may represent an enhancer element which regulates the transcriptional activity of the putative promoter region.

DHS regions located considerable distances from the promoter may regulate transcriptional activity. To identify other hypersites outside of the 5kb EcoRI region we extended our analysis both 5' and 3'. The same DNAse1 treated kidney genomic DNA

Figure 4.1 Schematic representation of the probes used for DHS analysis

The diagram shows the *Evi-1* locus and the genetically linked *fim-3* locus.

DHS analysis of the *Evi-1* locus
Two probes, a and b, were used in DHS analysis of the *Evi-1* locus both of which were isolated from the same λCB9 clone. Probe a is 0.5kb SstI/EcoRI fragment with probe b a 2kb EcoRI/BamHI fragment.

DHS analysis of the *Fim-3* locus
A 1.2kb EcoRI/HindIII fragment, probe c, isolated from a λCB24 clone, was used in DHS analysis of the *Fim-3* locus.

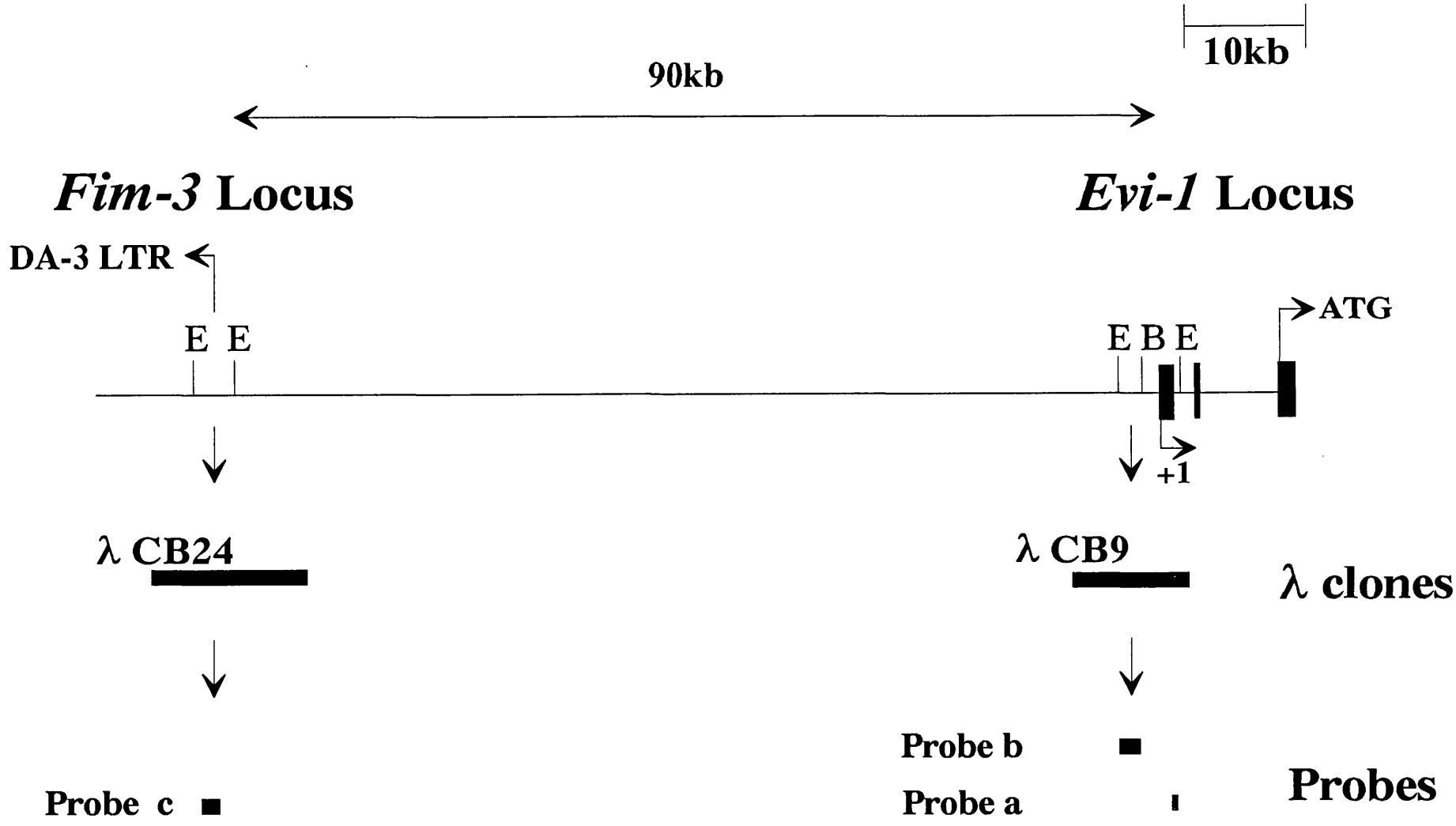

10kb
90kb
Fim-3 Locus
Evi-1 Locus
DA-3 LTR
E E
E B E
ATG
+1
λ CB24
λ CB9
λ clones
Probe b
Probe a
Probe c
Probes

Figure 4.2: DNase 1 hypersite analysis of the 5' *Evi-1* Locus in kidney tissue with a 2kb EcoRI/BamHI probe following EcoRI digestion.

A

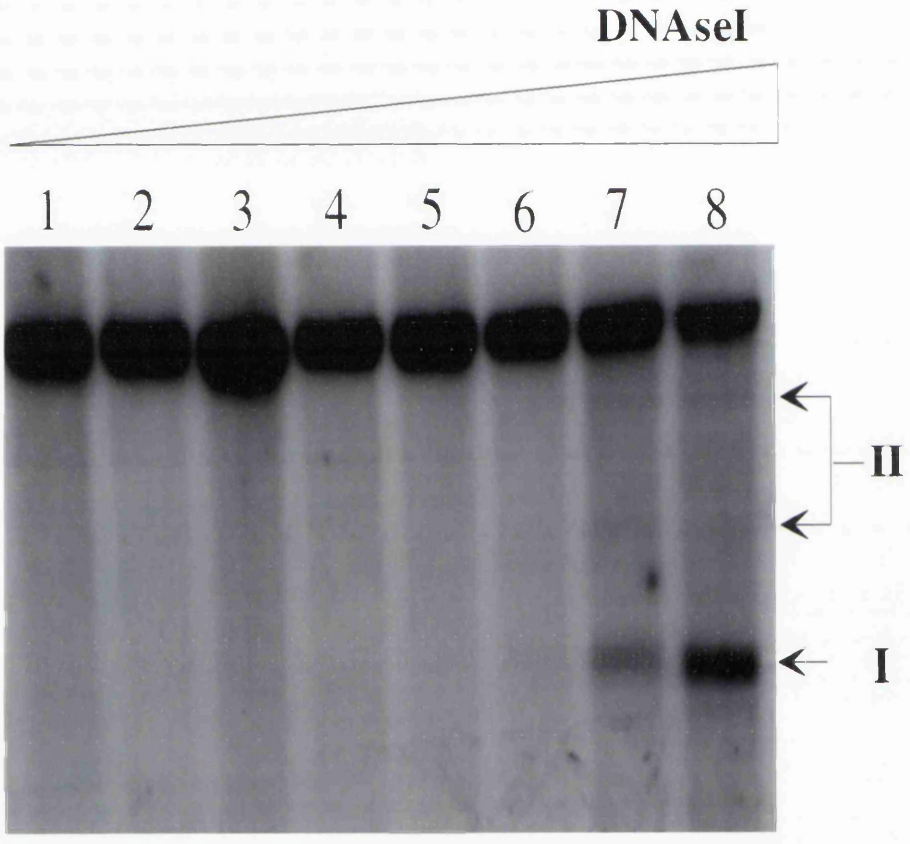
The DNA was digested with EcoRI restriction enzyme. The 2kb EcoRI/BamHI labelled probe shown in figure 1.2, B (**Probe b**) hybridises to a 5kb EcoRI fragment present at the top of all the lanes 1-8. DNAse 1 hypersensitive sites within this 5kb region become apparent with increasing DNAse 1 enzyme concentration. A strong hypersite (**I**) and two less intense hypersites (**II**) are seen in lanes 7 + 8 and are discussed in the text.

DNAse1 Concentrations
Lane 1: Untreated, Lane 2: 0.312 µg/ml, Lane 3: 0.625 µg/ml, Lane 4: 1.25 µg/ml, Lane 5: 2.5 µg/ml, Lane 6: 5 µg/ml, Lane 7: 10 µg/ml, Lane 8: 20 µg/ml.

B

Schematic representation of the 5' Evi-1 locus. BamHI (B) and EcoRI (E) restriction enzyme sites are shown. The approximate location within the *Evi-1* locus of hypersites **I** and **II** are also indicated. Hypersite **I** maps to approximately 2kb upstream of exon I while hypersite **II** is located within exon I.

A


DNAseI
1 2 3 4 5 6 7 8
II
I

B

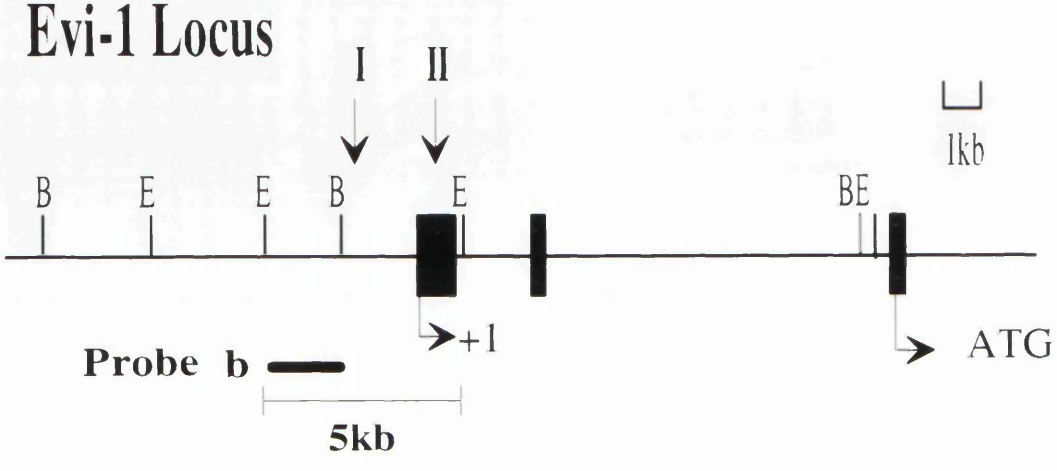

Evi-1 Locus
I II
1kb
B E E B E BE
+1
ATG
Probe b
5kb

Figure 4.3: Confirmation of the location of DHS I with a 500bp EcoRI/SstI probe following EcoRI digestion.

A

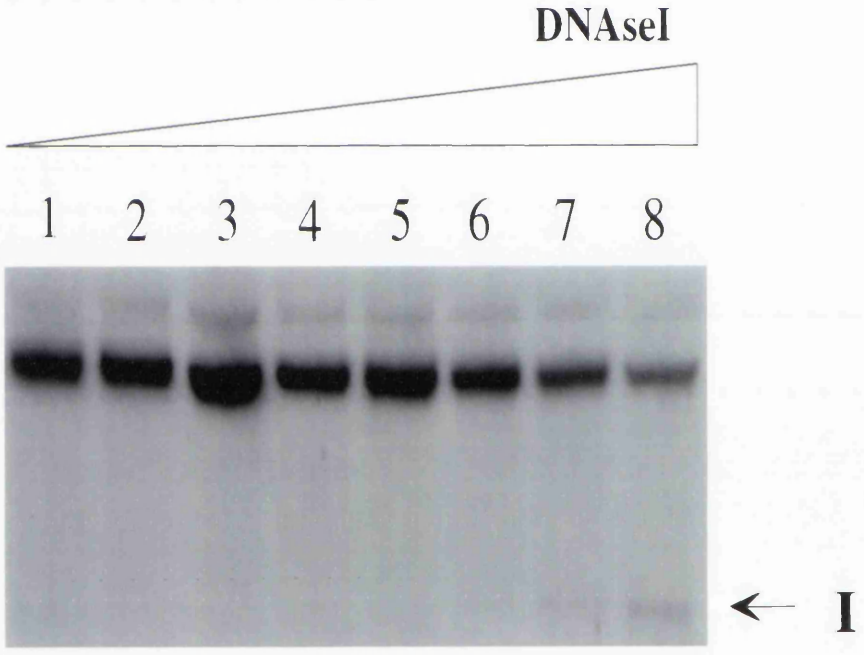
The DNA was digested with EcoRI restriction enzyme. The EcoRI/SstI labelled probe shown in figure 1.3, B (**Probe a**) hybridises to a 5kb EcoRI fragment present at the top of all the lanes 1-8. DNAse 1 hypersensitive sites within this 5kb region become apparent with increasing DNAse 1 enzyme concentration. A 2kb frament corresponding to hypersite (**I**) is observed in lanes 7 + 8 and is discussed in the text.

DNAse1 Concentrations
Lane 1: Untreated, Lane 2: 0.312 μg/ml, Lane 3: 0.625 μg/ml, Lane 4: 1.25 μg/ml, Lane 5: 2.5 μg/ml, Lane 6: 5 μg/ml, Lane 7: 10 μg/ml, Lane 8: 20 μg/ml.

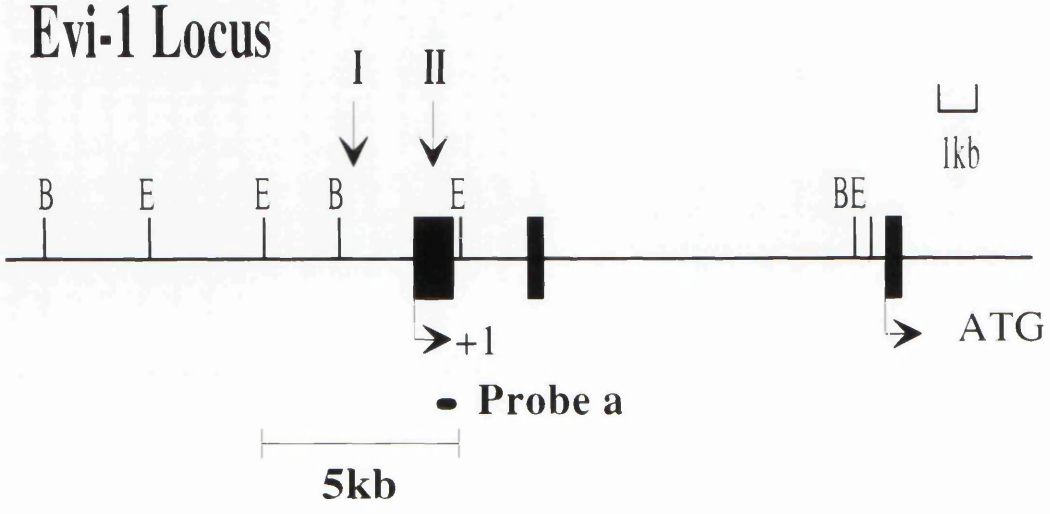
B

Schematic representation of the 5' Evi-1 locus. BamHI (B) and EcoRI (E) restriction enzyme sites are shown. The approximate location within the *Evi-1* locus of hypersites **I** and **II** are also indicated. Hypersite **I** maps to approximately 2kb upstream of exon I while hypersite **II** is located within exon I.

A


DNAseI
1 2 3 4 5 6 7 8
I

B


Evi-1 Locus
I
II
1kb
B
E
E
B
E
BE
+1
ATG
Probe a
5kb

isolated previously was digested with BamHI and Southern analysis performed using probe b, which hybridises to a unique 7kb fragment (figure 4.4 A, lane1). The DHS analysis failed to identify any hypersites further upstream of the EcoRI fragment. Consequently, only two DHS regions, DHS I and II, are present in an approximately 9.3kb of DNA upstream of the major sites of transcription initiation.

Similarly, to identify any DHS sites 3' of the EcoRI fragment DNAse1 treated kidney genomic DNA isolated previously was digested with SstI. Southern analysis using probe a detects a unique 8.2kb fragment (figure 4.5 A, lane1). Increasing DNAse1 digestion reveals the presence of a faint hypersite (figure 4.5, lane 6 + 7). This hypersite probably maps to the 3' of exon I and overlaps the border with intron I and reflecting one of the sites in DHS II previously identified (figure 4.2). However, it is conceivable that the hypersite may be located at the other end of the SstI fragment, within intron II.

4.2 Chromatin structure of the 5' *Evi-1* locus in DA-3 and L929 cells

To see if there is a correlation between the pattern of the DHS I and II and the level of *Evi-1* expression, DHS analysis was performed with nuclei isolated from DA-3 cells (high *Evi-1* expressing cells) and L929 cells (relatively low *Evi-1* expressing cells). DA-3 cells ectopically express high levels of *Evi-1* due a retroviral insertion 90kb upstream in the *Fim-3* locus (Bartholomew *et al.*, 1989). Since no RNA transcripts were identified from the LTR the mechanism of ectopic expression has been suggested to occur by a distal enhancer model. Importantly, the induced expression occurs by activating the normal *Evi-1* promoter (Bartholomew & Ihle, 1991).

To compare the DHS I and II pattern observed in kidney tissue DNAse1 treated genomic DNA from either DA-3 or L929 cells were digested with EcoRI and the blots hybridised with probe b. Again Southern analysis with probe b detects a 5kb EcoRI

Figure 4.4: DNase 1 hypersite analysis of the 5' *Evi-1* Locus in Kidney with a 2kb EcoRI/BamHI probe following BamHI digestion.

A

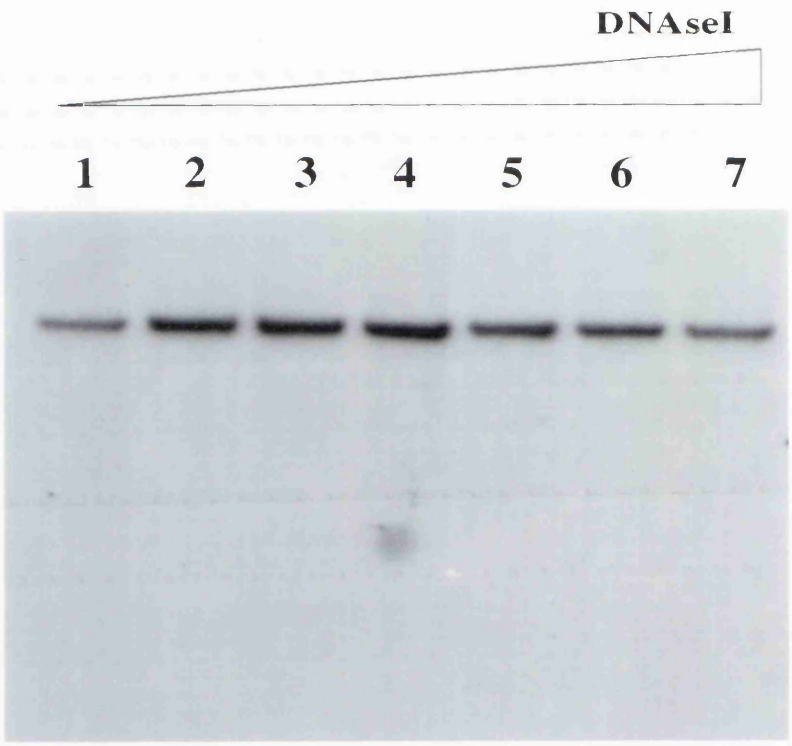
The DNA was digested with BamHI restriction enzyme. The EcoRI/BamHI labelled probe shown in figure 1.4, B (**Probe b**) hybridises to a 7kb BamHI fragment present at the top of all the lanes 1-7. There were no DNAse 1 hypersensitive sites seen within this 7kb region.

DNAse 1 Concentrations
Lane 1: Untreated, Lane 2: 0.625 μg/ml, Lane 3: 1.25 μg/ml, Lane 4: 2.5 μg/ml, Lane 5: 5 μg/ml, Lane 6: 10 μg/ml, Lane 7: 20 μg/ml.

B

Schematic representation of the 5' Evi-1 locus. BamHI (B) and EcoRI (E) restriction enzyme sites are shown. The approximate location within the *Evi-1* locus of the previously identified hypersites **I** and **II** are also indicated.

A



B

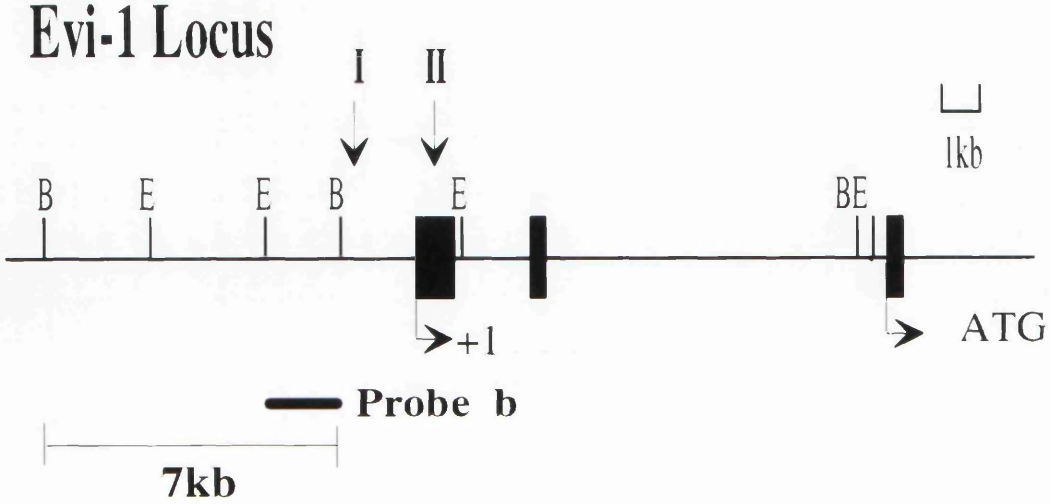


Figure 4.5: DNase 1 hypersite analysis of the 5' *Evi-1* Locus in Kidney with a 500bp EcoRI/SstI probe following SstI digestion.

A

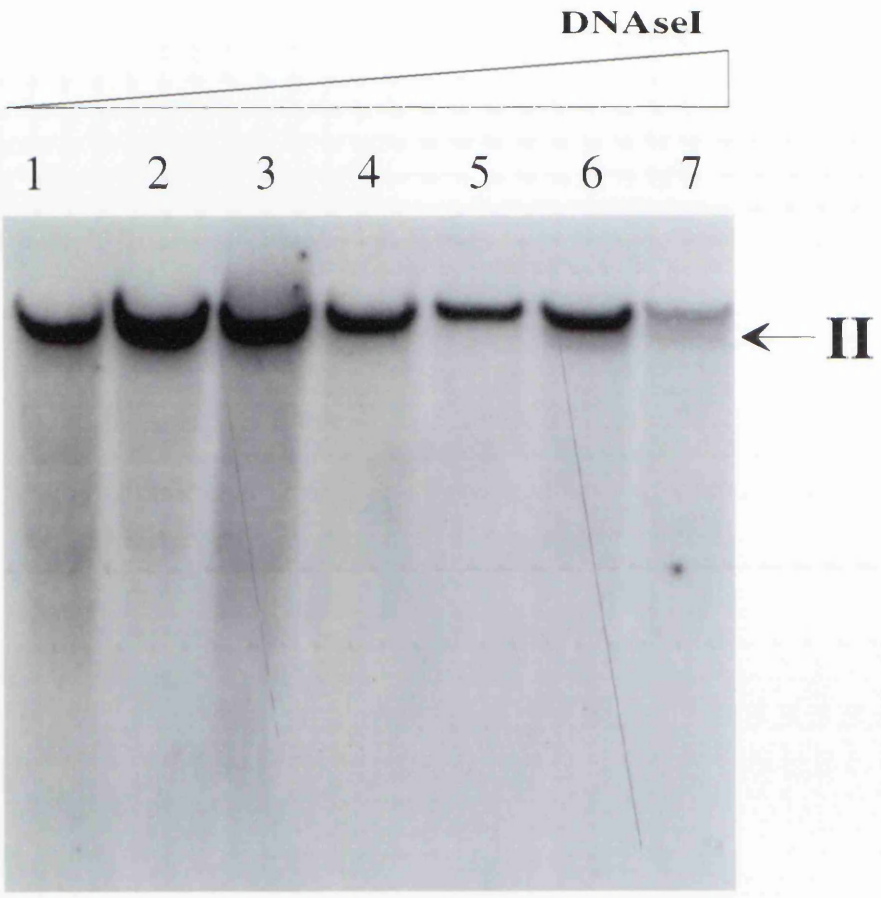
The DNA was digested with SstI restriction enzyme. The EcoRI/SstI labelled probe shown in figure 1.5, B (**Probe a**) hybridises to a 8.2kb SstI fragment present at the top of all the lanes 1-7. There were no DNAse 1 hypersensitive sites seen within this 8.2kb region

DNAse 1 Concentrations
Lane 1: Untreated, Lane 2: 0.625 μg/ml, Lane 3: 1.25 μg/ml, Lane 4: 2.5 μg/ml, Lane 5: 5 μg/ml, Lane 6: 10 μg/ml, Lane 7: 20 μg/ml.

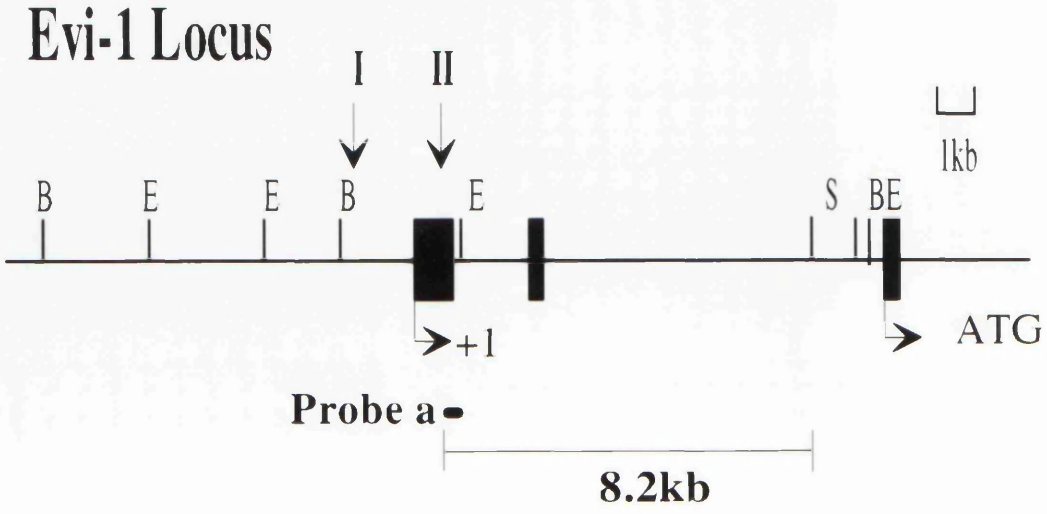
B

Schematic representation of the 5' Evi-1 locus. BamHI (B), EcoRI (E) and SstI (S) restriction enzyme sites are shown. The approximate location within the *Evi-1* locus of the previously identified hypersites **I** and **II** are also indicated.

A


DNAseI
1
2
3
4
5
6
7
II

B


Evi-1 Locus
I
II
1kb
B
E
E
B
E
S
BE
+1
ATG
Probe a
8.2kb

fragment (figure 4.6 A & B, lane 1). DHS sites are observed with increasing DNAse1 concentration (figure 4.6 A & B) which map to the same location as DHS I and II previously identified in kidney tissue. The results show that DHS I and II are present in both DA-3 and L929 cells. In conclusion, DHS I and II correlate with expression of the *Evi-1* gene, being present in low and high expressing cells, but not with the level of expression.

4.3 Chromatin structure of the 5' *Evi-1* locus in Spleen Tissue and FDCP-1 cells

Our analysis has identified DHS I and II in kidney, DA-3 and L929 cells which all express *Evi-1*. The tissue specific nature of *Evi-1* expression might lead us to predict the presence of tissue specific hypersites present only in cells that express *Evi-1*, as is the case for other genes such as the β-globin genes (reviewed by Evans *et al*, 1990). To examine if DHS I and II correlates with *Evi-1* expression we examined hypersites in non-expressing cell lines and tissue from FDCP-1 and spleen respectively.

As previously observed in kidney, DA-3 and L929 cells, EcoRI digestion of spleen or FDCP-1 genomic DNA results in a 5kb EcoRI fragment which hybridises to probe b (figure 4.7 A & B, lane 1). DHS I is visible in both spleen, (figure 4.7 A, lane 4) and FDCP-1 cells (figure 4.7 B, lane 4) whereas DHS II is not observed to be present in either spleen or FDCP-1 cells. The two bands marked N present in figure 4.7 A, lane 4 are non-specific bands which also appear in the DNAse1 untreated lane 1. In conclusion, DHS I is a ubiquitous hypersite being present in all the cell lines and tissues tested and, by contrast, DHS II is present only in *Evi-1* expressing cells.

Figure 4.6: DNase 1 hypersite analysis of the 5' *Evi-1* Locus in DA-3 and L929 cell lines with a 2kb EcoRI/BamHI probe following EcoRI digestion.

A

DNA was digested with EcoRI restriction enzyme. The 2kb EcoRI/BamHI labelled probe shown in figure 1.6, B (**Probe b**) hybridises to a 5kb EcoRI fragment present at the top of all the lanes 1-4. DNAse 1 hypersensitive sites within this 5kb region become apparent with increasing DNAse 1 enzyme concentration. A strong hypersite (**I**) and two less intense hypersites (**II**) are seen in lanes 7 + 8 and are discussed in the text.

DNAse Concentrations
Lane 1: Untreated, Lane 2: 5 μg/ml, Lane 3: 10 μg/ml, Lane 4: 20 μg/ml.

B

DNA was digested with EcoRI restriction enzyme. The 2kb EcoRI/BamHI labelled probe shown in figure 1.6, B (**Probe b**) hybridises to a 5kb EcoRI fragment present at the top of all the lanes 1-4. DNAse 1 hypersensitive sites within this 5kb region become apparent with increasing DNAse 1 enzyme concentration. A strong hypersite (**I**) and two less intense hypersites (**II**) are seen in lanes 3 + 4 and are discussed in the text.

DNAse Concentrations
Lane 1: Untreated, Lane 2: 5 μg/ml, Lane 3: 10 μg/ml, Lane 4: 20 μg/ml.

C

Schematic representation of the 5' Evi-1 locus. BamHI (B) and EcoRI (E) restriction enzyme sites are shown. The approximate location within the *Evi-1* locus of the previously identifed hypersites **I** and **II** in kidney cells are also indicated.

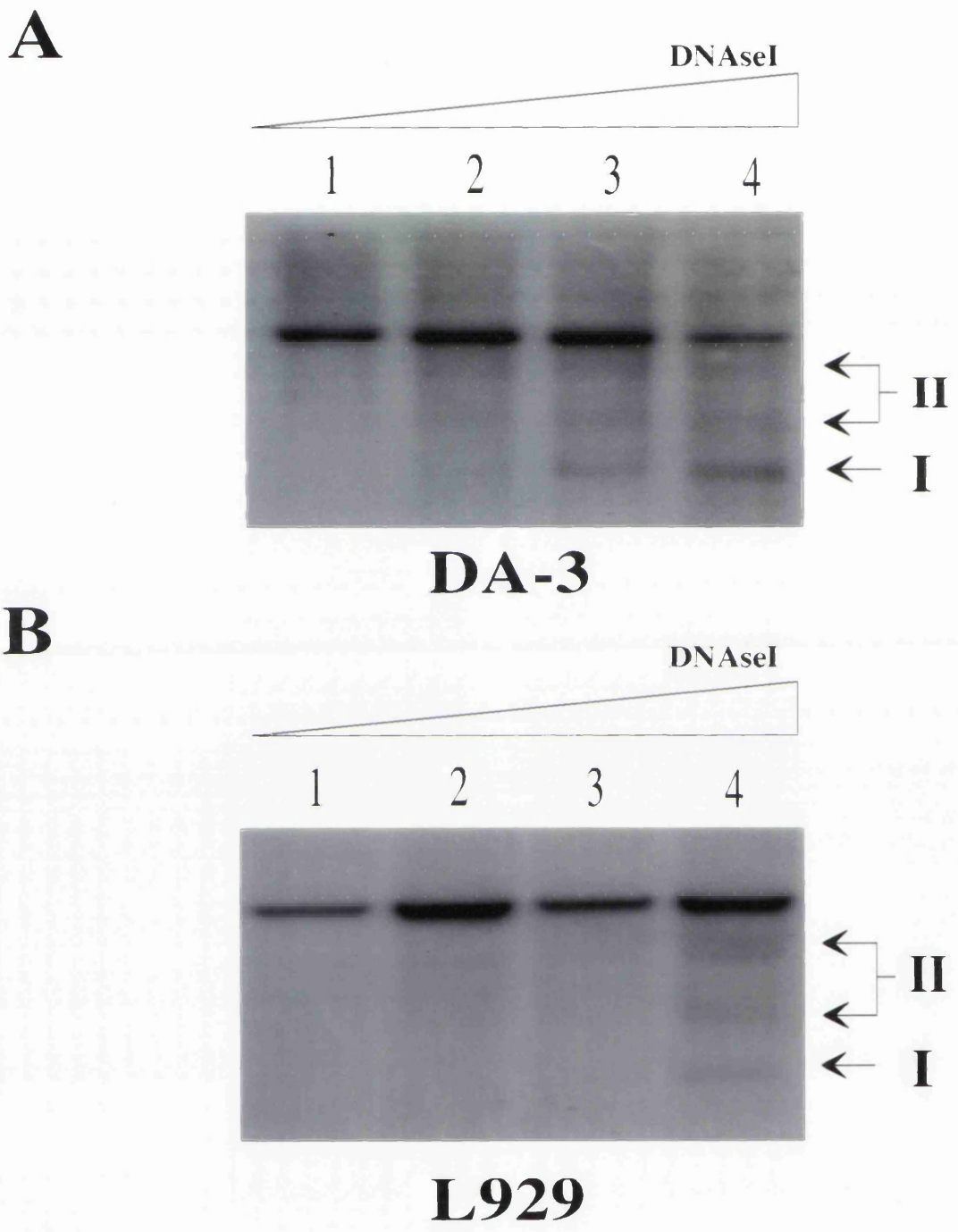

A
DNAseI
1
2
3
4
II
I
DA-3
B
DNAseI
1
2
3
4
II
I
L929
C
Evi-1 Locus
I
II
1kb
B
E
E
B
E
BE
+1
ATG
Probe b
5kb

Figure 4.7: DNase 1 hypersite analysis of the 5' *Evi-1* Locus in Spleen tissue and the FDCP-1 cell line with a 2kb EcoRI/BamHI probe following EcoRI digestion.

A

DNA was digested with EcoRI restriction enzyme. The 2kb EcoRI/BamHI labelled probe shown in figure 1.7 B (**Probe b**) hybridises to a 5kb EcoRI fragment present at the top of all the lanes 1-4. DNAse 1 hypersensitive sites within this 5kb region become apparent with increasing DNAse 1 enzyme concentration. A strong hypersite (**I**) was seen in lanes 3 + 4 and is discussed in the text.

DNAse Concentrations
Lane 1: Untreated, Lane 2: 5 μg/ml, Lane 3: 10 μg/ml, Lane 4: 20 μg/ml.

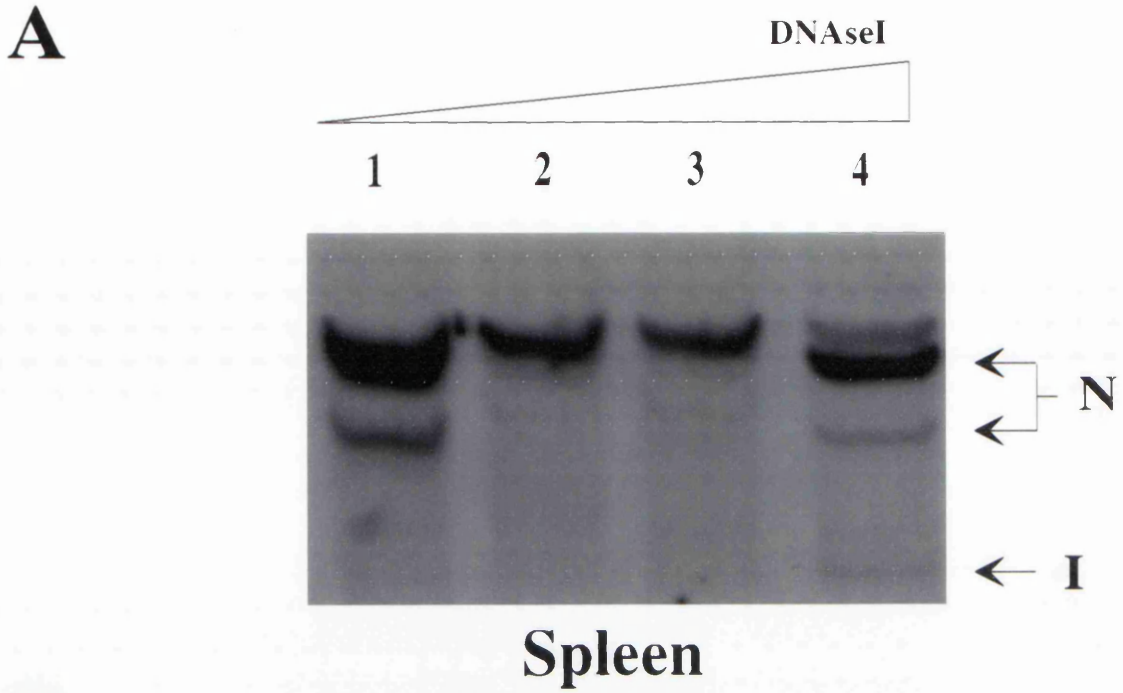
B

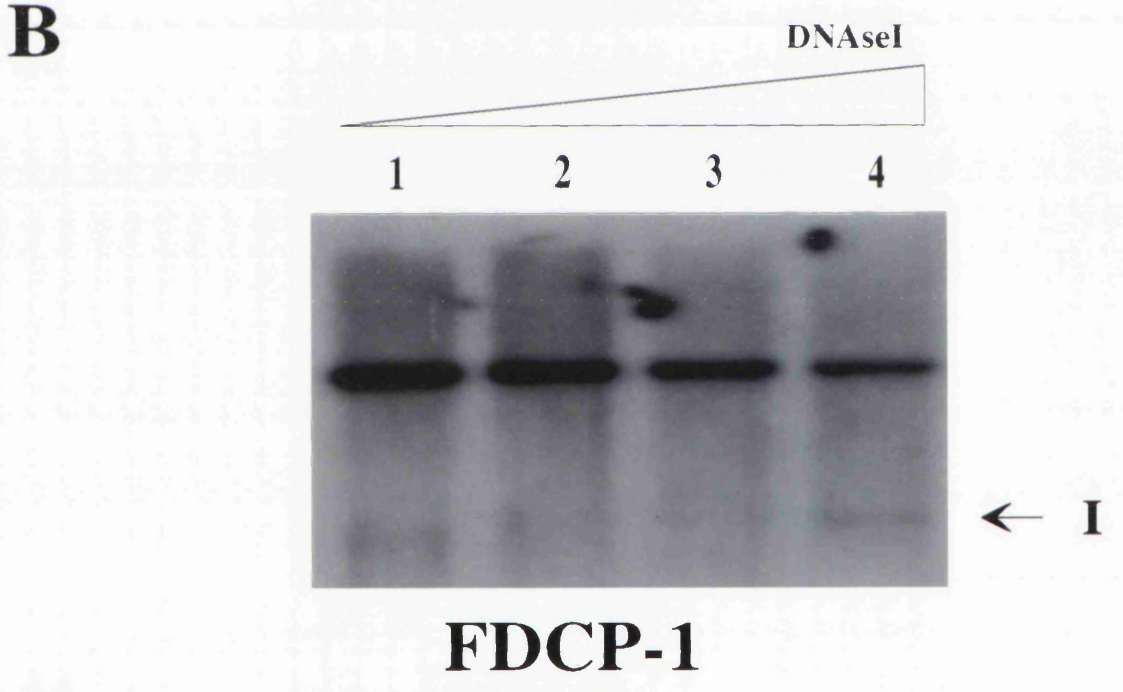
DNA was digested with EcoRI restriction enzyme. The 2kb EcoRI/BamHI labelled probe shown in figure 1.7, B (**Probe b**) hybridises to a 5kb EcoRI fragment present at the top of all the lanes 1-4. DNAse 1 hypersensitive sites within this 5kb region become apparent with increasing DNAse 1 enzyme concentration. A strong hypersite (**I**) was seen in lanes 3 + 4 and is discussed in the text.

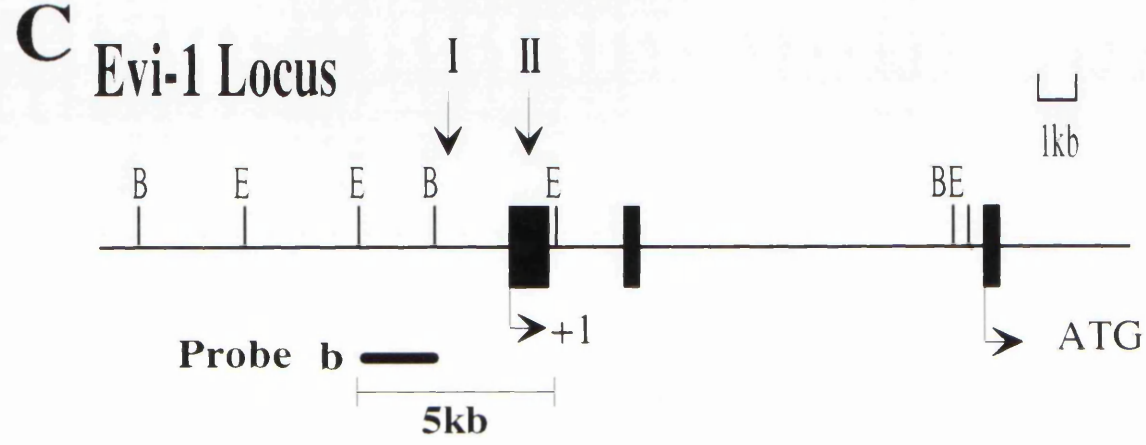
DNAse Concentrations
Lane 1: Untreated, Lane 2: 5 μg/ml, Lane 3: 10 μg/ml, Lane 4: 20 μg/ml.

C

Schematic representation of the 5' Evi-1 locus. BamHI (B) and EcoRI (E) restriction enzyme sites are shown. The approximate location within the *Evi-1* locus of the previously identifed hypersites **I** and **II** in kidney cells are also indicated.


A
DNAseI
1
2
3
4
N
I
Spleen


B
DNAseI
1
2
3
4
I
FDCP-1


C
Evi-1 Locus
I
II
1kb
B
E
E
B
E
BE
+1
ATG
Probe b
5kb

4.4 Summary of DHS analysis of the 5' *Evi-1* locus

Figure 4.8 shows a schematic representation of the *Evi-1* locus and a summary of the hypersites found in the different tissues and cell lines examined in this study. An approximately 18kb region of the *Evi-1* locus has been examined for the DHS sites which might be involved in controlling the expression of the *Evi-1* gene. DHS analysis in *Evi-1* expressing and non-expressing cells has identified an ubiquitous hypersite, I, located approximately 2kb upstream of the major transcription initiation site. In contrast a second region has DHS sites located over exon I which are present only in *Evi-1* expressing cells.

The location of the ubiquitous DHS I might suggest that this region may act as an enhancer element. However, the appearance of DHS enhancer sites are generally correlated with increased expression of the gene. Alternatively, it is conceivable that DHS I may be important *in vivo* in maintaining a more relaxed chromatin framework around the *Evi-1* gene.

The presence of DHS II over exon I presumably reflects that this is the promoter region of the gene since transcriptional initiation by RNA polymerase II is generally mediated by a TATA box and transcription factors immediately upstream. DHS II may also contain regulatory *cis*-elements important in tissue specific expression of the gene. The functional importance of both these regions in transcriptional regulation of the *Evi-1* promoter are discussed in later chapters.

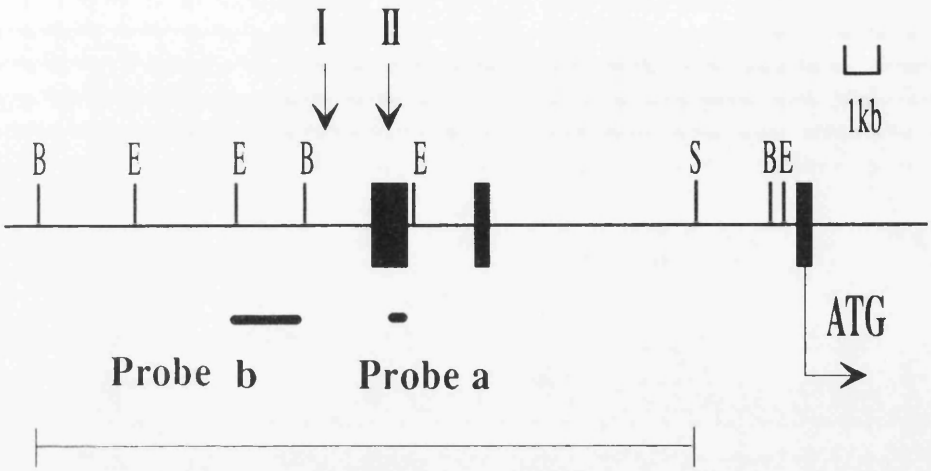
4.5 Chromatin structure in the *Fim-3* Locus

Ectopic expression of the *Evi-1* gene in DA-3 cells is due to a retroviral insertion in the *Fim-3* locus 90kb further upstream (Bordereaux *et al*, 1987, Bartholomew *et al.*, 1989). Expression occurs by activating the normal *Evi-1* promoter and the mechanism

Figure 4.8: DHS anlysis of the *Evi-1* locus in different cell types

A schematic representation of the *Evi-1* locus showing the location of DHS I and II, and the 18kb probed region covered using the probes a and b. The presence (+) or absence (-) of the DHS I and II in the tissues or cell lines examined is also shown.

Evi-1 Locus



	Kidney		Spleen		L929		DA-3		FDCP-1	
Hypersite	I	II	I	II	I	II	I	II	I	II
	+	+	+	-	+	+	+	+	+	-

has been postulated to occur by a distal enhancer model. The *Fim-3* region was examined surrounding the retroviral insertion site to see if there were any DHS sites which might contribute to the increase in *Evi-1* expression levels. The DHS pattern of this region in DA-3 cells was also compared with kidney tissue to evaluate if the *Fim-3* region contributes to normal tissue specific expression of the gene.

Therefore, DNAse1 treated DA-3 and Kidney genomic DNA were digested with EcoRI and examined by Southern blot analysis with the 1.2kb EcoRI/HindIII ^{32}-P labelled probe c, from the *Fim-3* locus (figure 4.9, B). In DA-3 probe c detects two EcoRI fragments of 3.5kb, (normal allele), and 4kb, (rearranged allele containing a retroviral insertion) (figure 4.9 A, lane 1). Increasing DNAse1 digestion identified a DHS site, designated DHS III, of 2.2kb which maps to within the retroviral LTR. The location of DHS III is confirmed by the observation that the intensity of the 3.5kb relative to the 4kb band at lower DNAse1 digestion is the same (figure 4.9 A, lanes 1 to 5) whereas on appearance of the hypersite the upper 4kb fragment containing the LTR becomes less intense than the lower 3.5kb fragment, due to increasing DNAse1 digestion within the LTR.

In contrast, Southern analysis of EcoRI digested kidney genomic DNA with probe c detects a single 3.5kb EcoRI fragment (figure 4.10, A, lane 1). However, no DHS sites were seen with increasing DNAse1 digestion. In conclusion, no DHS sites are located within the 3.5kb fragment of the normal allele in kidney and presumably this region of *Fim-3* is not important for high levels of tissue specific expression of the *Evi-1* promoter observed in kidney tissue.

4.6 Summary of DHS Analysis in the *Fim-3* locus

A summary of the analysis to investigate DHS sites in a 3.5kb EcoRI region of the *Fim-3* locus in either retroviral LTR containing DA-3 cells or kidney tissue is

Figure 4.9: DNase 1 hypersite analysis of the *Fim-3* Locus in DA-3 cells with a 1.2kb EcoRI/HindIII probe following EcoRI digestion.

A

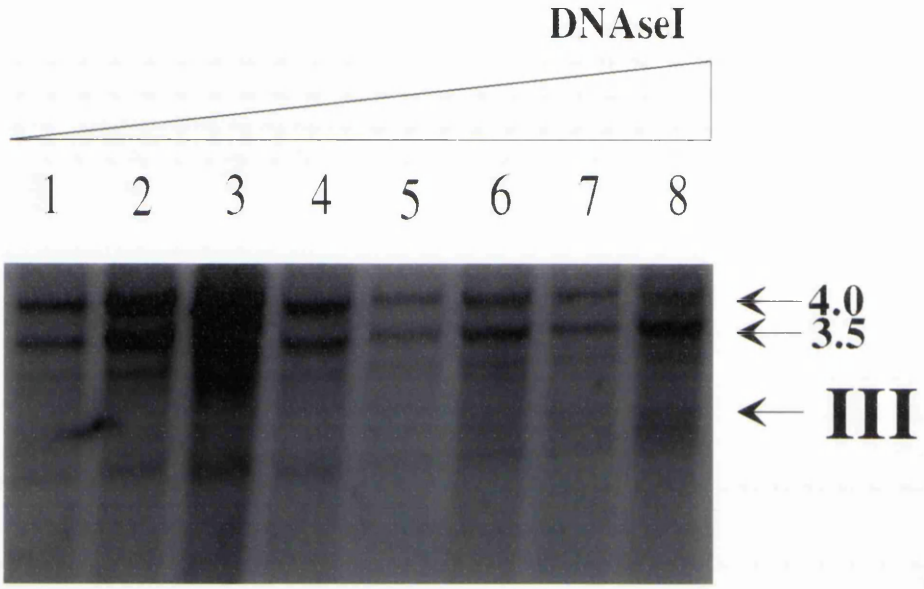
DNA was digested with EcoRI restriction enzyme. The 1.2kb EcoRI/HindIII labelled probe shown in figure 1.9, B (**Probe a**) hybridises to a 3kb and 3.5kb EcoRI fragment present at the top of all the lanes 1-8. The doublet is seen as a result of a retroviral insertion at one allele and the subsequent presence of a Long Terminal Repeat (**LTR**). DNAse 1 hypersensitive sites within this 5kb region become apparent with increasing DNAse 1 enzyme concentration. A hypersite (**I**) is seen in lanes 7 + 8 and are discussed in the text.

DNAse Concentrations
Lane 1: Untreated, Lane 2: 0.312 µg/ml, Lane 3: 0.625 µg/ml, Lane 4: 1.25 µg/ml, Lane 5: 2.5 µg/ml, Lane 6: 5 µg/ml, Lane 7: 10 µg/ml, Lane 8: 20 µg/ml.

B

Schematic representation of the *Fim-3* locus. BamHI (B) and EcoRI (E) restriction enzyme sites are shown. The approximate location within the *Fim-3* locus of the identified hypersite (**I**) is also indicated.

A


DNAseI
1 2 3 4 5 6 7 8
4.0
3.5
III

DA-3

B

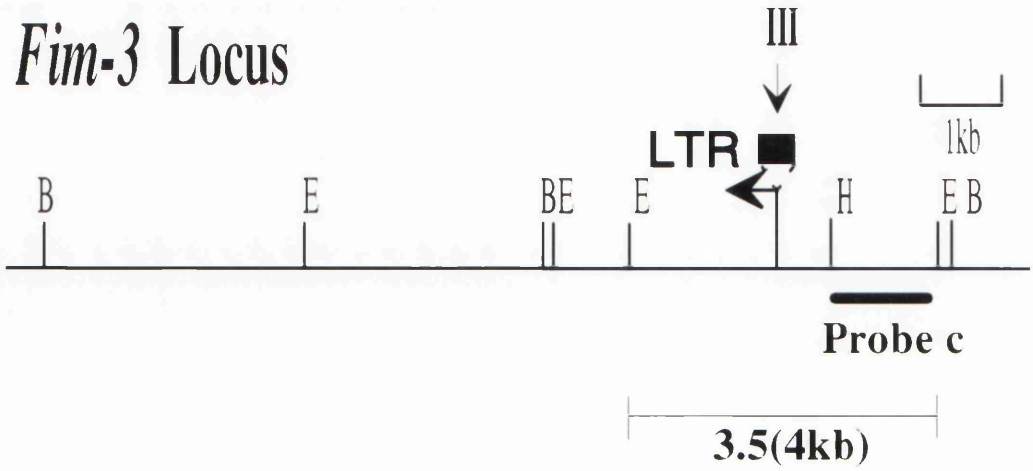

Fim-3 Locus
III
LTR
1kb
B E BE E H E B
Probe c
3.5(4kb)

Figure 4.10: DNase 1 hypersite analysis of the *Fim-3* Locus in Kidney with a 1.2kb EcoRI/HindIII probe following EcoRI digestion.

A

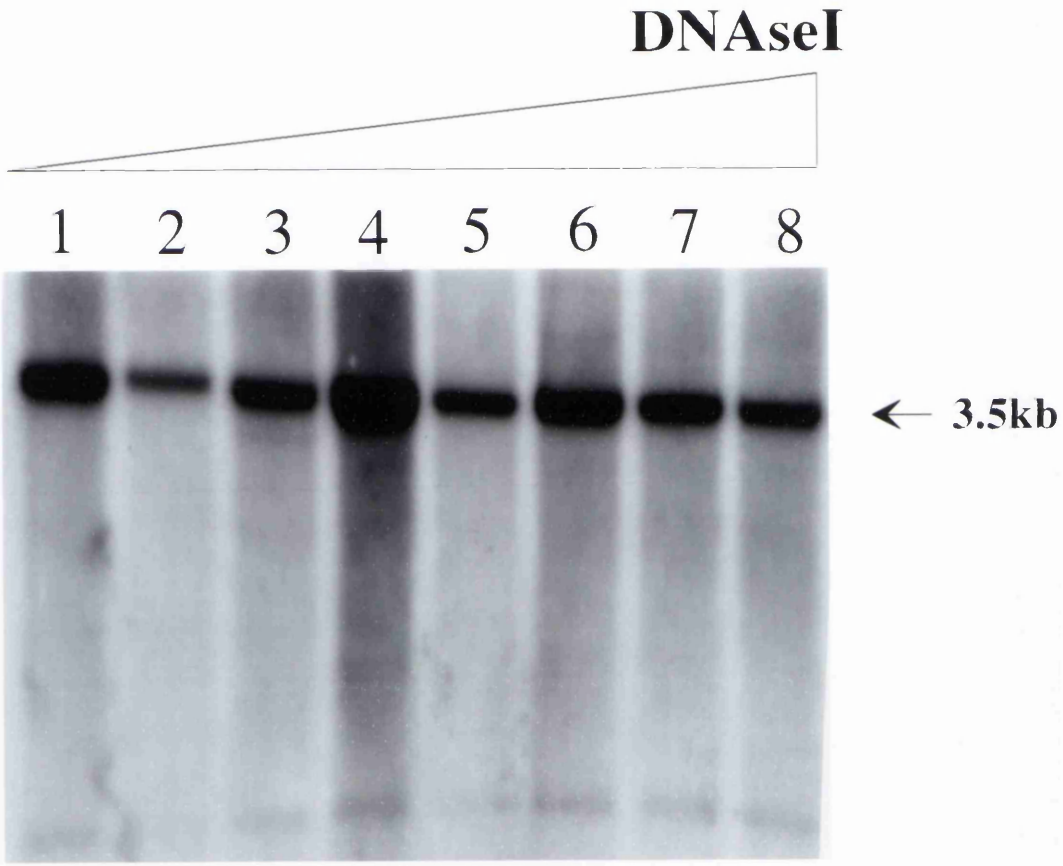
DNA was digested with EcoRI restriction enzyme. The 1.2kb EcoRI/HindIII labelled probe shown in figure 1.10, B (**Probe a**) hybridises to a 3kb EcoRI fragment present at the top of all the lanes 1-8. There were no DNAse 1 hypersites identified in this region and this is discussed further in the text.

DNAse Concentrations
Lane 1: Untreated, Lane 2: 0.312 μg/ml, Lane 3: 0.625 μg/ml, Lane 4: 1.25 μg/ml, Lane 5: 2.5 μg/ml, Lane 6: 5 μg/ml, Lane 7: 10 μg/ml, Lane 8: 20 μg/ml.

B

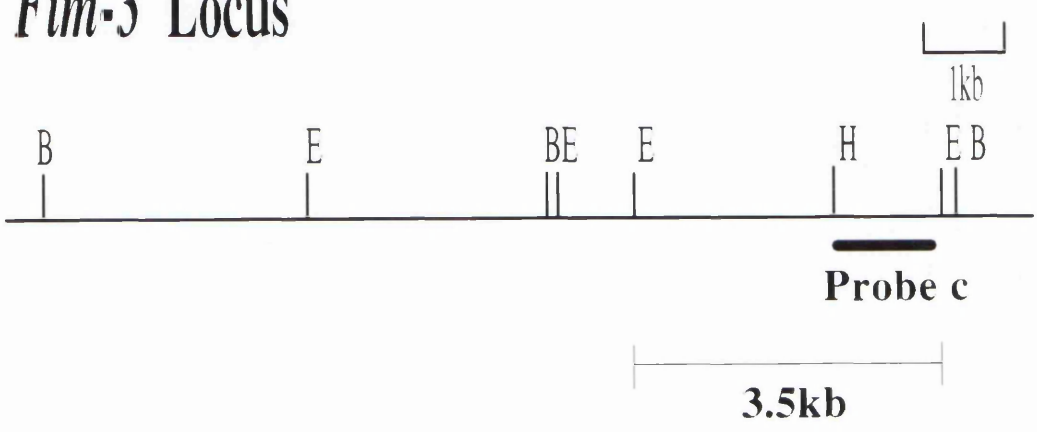
Schematic representation of the *Fim-3* locus. BamHI (B) and EcoRI (E) restriction enzyme sites are shown. The location of **probe a** is also shown.

A



Kidney

B

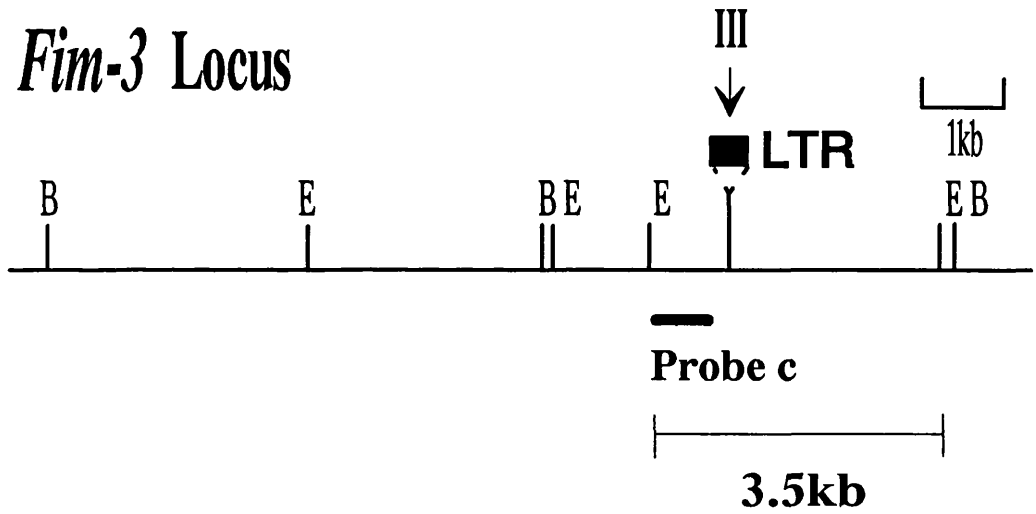


illustrated in a schematic representation of the *Fim-3* locus (figure 4.11). The induction of a DHS site, designated DHS III, within the LTR of the virus in DA-3 cells correlates with ectopic expression of the *Evi-1* gene from its normal promoter. This is consistent with a putative long range enhancer mechanism suggested previously to be a possible method for inducing *Evi-1* promoter expression in DA-3 cells (Bartholomew & Ihle., 1991). However, no DHS sites were identified within this region in kidney confirming that DHS III is due to retroviral insertion at this site in DA-3 cells. Therefore, the *Fim-3* region presumably plays no part in tissue specific expression of the *Evi-1* promoter in kidney tissue although it is still plausible that there are DHS sites present outside the region examined in this study.

Figure 4.11: DHS anlysis of the *Fim-3* locus in different cell types

A schematic representation of the *Fim-3* locus showing the location of DHS III, and the 3.5kb probed region covered using probe c. The presence (+) or absence (-) of the DHS I and II in DA-3 cells or kidney tissue is also shown.

Fim-3 Locus



	DA-3	Kidney
Hypersite	III	III
	+	-

Chapter 5: Functional analysis of the *Evi-1* promoter

Previous studies have shown that the *Evi-1* promoter has multiple transcriptional initiation sites with two major sites occurring at two adjacent nucleotides in kidney and ovary tissue (Bartholomew & Ihle, 1991). However, the functional elements controlling *Evi-1* promoter activity are unknown. DHS analysis around the promoter region has identified two DHS regions I and II (chapter 4.4). DHS II is located over exon I and is found exclusively in *Evi-1* expressing cell lines and tissues examined (chapter 4.4). In contrast, DHS I is found in all cells tested. DHS sites are generally associated with active promoter and enhancer elements (reviewed by Gross and Garrard, 1988), and therefore the role of these regions in *Evi-1* transcriptional regulation was investigated.

5.1 The generation of *Evi-1* promoter luciferase reporter constructs

To identify the regions important in controlling *Evi-1* promoter activity relatively large restriction fragments from the *Evi-1* locus were inserted upstream of a luciferase reporter gene. The largest reporter construct examined, designated pGL2E8, is an 8kb EcoRI/SmaI restriction fragment which extends approximately 4.2kb upstream of the major sites of transcription initiation and 3.8kb downstream to the end of exon II (figure 5.3). This construct was generated as shown in figure 5.2 and contains both DHS I and II regions. A second 3' truncated construct, designated pGL2E5, contains a 4.5kb EcoRI/SstI fragment which extends 4.2kb upstream of exon I and approximately 300bp downstream to an SstI within exon I. The pGL2E5 construct also has DHS I but may partially delete the DHS II region (figure 5.3). Two additional constructs, designated pGL2E2.5 and pGL2E2, possess the same 3' SstI site as pGL2E5 but extend only 1.75kb and 1.1kb respectively, upstream of the major sites of transcription initiation. The pGL2E2.5 construct has partial loss of DHS I whereas pGL2E2 deletes the whole

Figure 5.1: Schematic representation of the steps involved in cloning portions of the *Evi-1* locus upstream of a luciferase reporter.

A 5kb portion of the *Evi-1* locus containing exon I and approximately 4kb of upstream sequence had previously been cloned into a pKSII+ bluescript vector (Chris Bartholomew).

A schematic representation of the *Evi-1* locus is shown with exon I, II and III represented as shaded boxes.

Reporter Constructs

pGL2E5 construct- an approximately 4.7kb region of the *Evi-1* locus was cut out of the pKSII+E5 bluescript vector on a SstI restriction fragment (one SstI site is within exon I and another further upstream in the vector polylinker). This SstI fragment was cloned into a pGL2basic luciferase reporter vector (Promega) which had previously been cut with SstI and the ends phosphorylated to prevent re-ligation. The orientation of the insert was then evaluated.

pGL2E2.5 construct- an approximately 2.5kb HindIII restriction fragment was cut out of the 5kb pGL2 reporter vector (one HindIII site cuts within the previously identified DHS I of the *Evi-1* locus and the other further downstream in the vector polylinker). The HindIII fragment was cloned into a pGL2basic luciferase reporter vector (Promega) which had previously been cut with HindIII and the ends phosphorylated to prevent re-ligation. The orientation of the insert was then evaluated.

pGL2E2 construct- an approximately 2.kb BglII restriction fragment was cut out of the 5kb pGL2 reporter vector (one BglII site cuts downstream of the previously identified DHS I of the *Evi-1* locus and the other further 3' within vector polylinker). The BglII fragment was cloned into a pGL2basic luciferase reporter vector (Promega) which had previously been cut with BglII and the ends phosphorylated to prevent re-ligation. The orientation of the insert was then evaluated.

Restriction enzyme sites:

S = SstI
B =BamHI
E =EcoRI

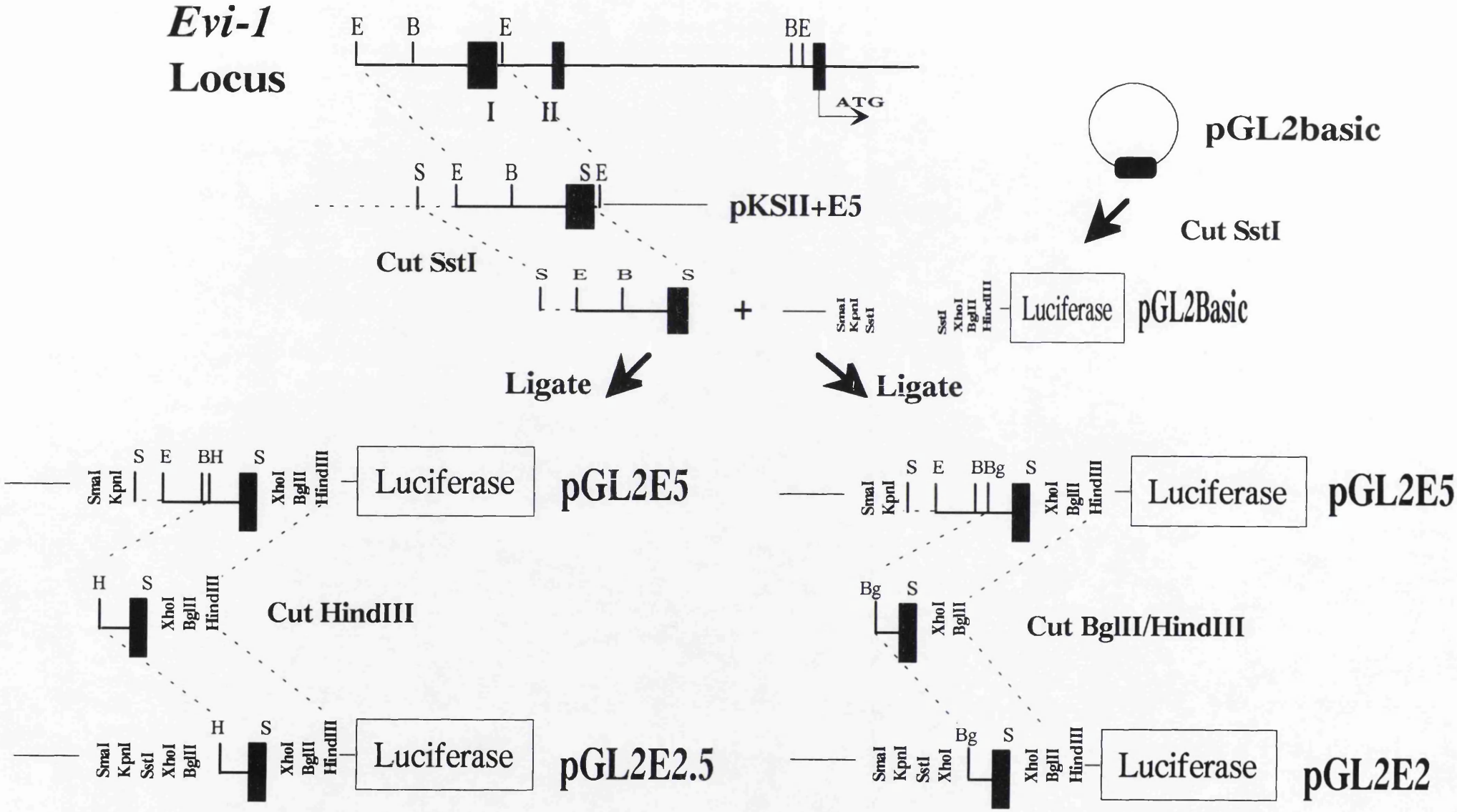

Evi-1
Locus
E B E BE
I II
ATG
S E B SE
pKSII+E5
Cut SstI
S E B S
+
SmaI KpnI SstI
pGL2basic
Cut SstI
SstI XhoI BglII HindIII
Luciferase
pGL2Basic
Ligate
Ligate
SmaI KpnI S E BH S XhoI BglII HindIII
Luciferase
pGL2E5
H S XhoI BglII HindIII
Cut HindIII
SmaI KpnI SstI XhoI BglII H S XhoI BglII HindIII
Luciferase
pGL2E2.5
SmaI KpnI S E BBg S XhoI BglII HindIII
Luciferase
pGL2E5
Bg S XhoI BglII
Cut BglII/HindIII
SmaI KpnI SstI XhoI Bg S XhoI BglII HindIII
Luciferase
pGL2E2

Figure 5.2: Schematic representation of the steps involved in cloning an 8.0kb region of the *Evi-1* locus, containing both exon I and II, upstream of a luciferase reporter.

A 5kb portion of the *Evi-1* locus containing exon I and approximately 4kb of upstream sequence had previously been cloned into a pKSII+E5 bluescript vector (Dr Chris Bartholomew).

A schematic representation of the *Evi-1* locus is shown with exon I, II and III represented as shaded boxes.

Reporter Construct

pGL2E8 construct- A 3kb region of the *Evi-1* locus from an SstI site within exon I to the end of exon II was generated by PCR as described in Materials and methods. The 3' PCR primer (exon II) contained extra nucleotides to incorporate a SmaI restriction enzyme site. The 3.2kb PCR fragment was cut with SmaI enzyme. This produces a 3kb fragment which was sub-cloned into a SmaI cut pKSII+E5 bluescript vector which already contains 5kb of upstream *Evi-1* sequence.

The resulting contstruct, designated pKSII+E8, is a 8.0kb stretch of the *Evi-1* locus which contains exon I, intron I and exon II. Subsequenlty, a 6kb fragment containing exon I, intron I and exon II was excised as a BamHI/NotI restriction fragment from pKSII+E8, ready for cloning into pGL2E5.

However, before this could be accomplished it was necessary to convert a HindIII site within the polylinker of the pGL2E5 luciferase reporter vector to a NotI restriction site in order to create an appropriate 3' restriction enzyme site for cloning a 6.0kb fragment. The subsequent vector, designated pGL2E5Not, was then cut with BamHI and NotI which removes a 3kb region of *Evi-1* sequence, including exon I. The 6.0kb BamHI/NotI fragment was then cloned into pGL2E5not to generate an luciferase reporter construct containing 8kb of the 5' *Evi-1* locus, designated pGL2E8.

Restriction Enzyme Sites:

S = SstI
Sm = SmaI
B =BamHI
E =EcoRI
N = NotI

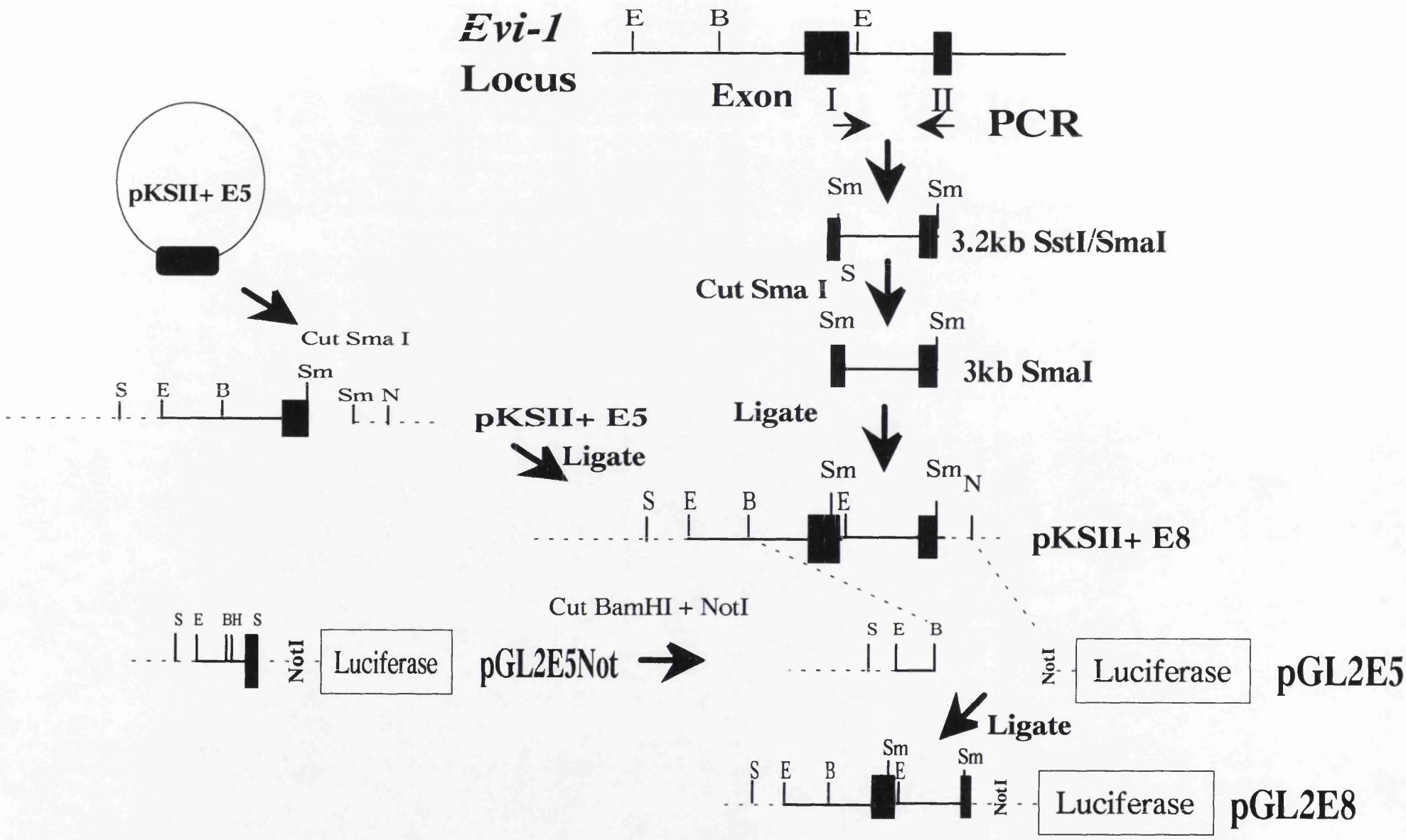
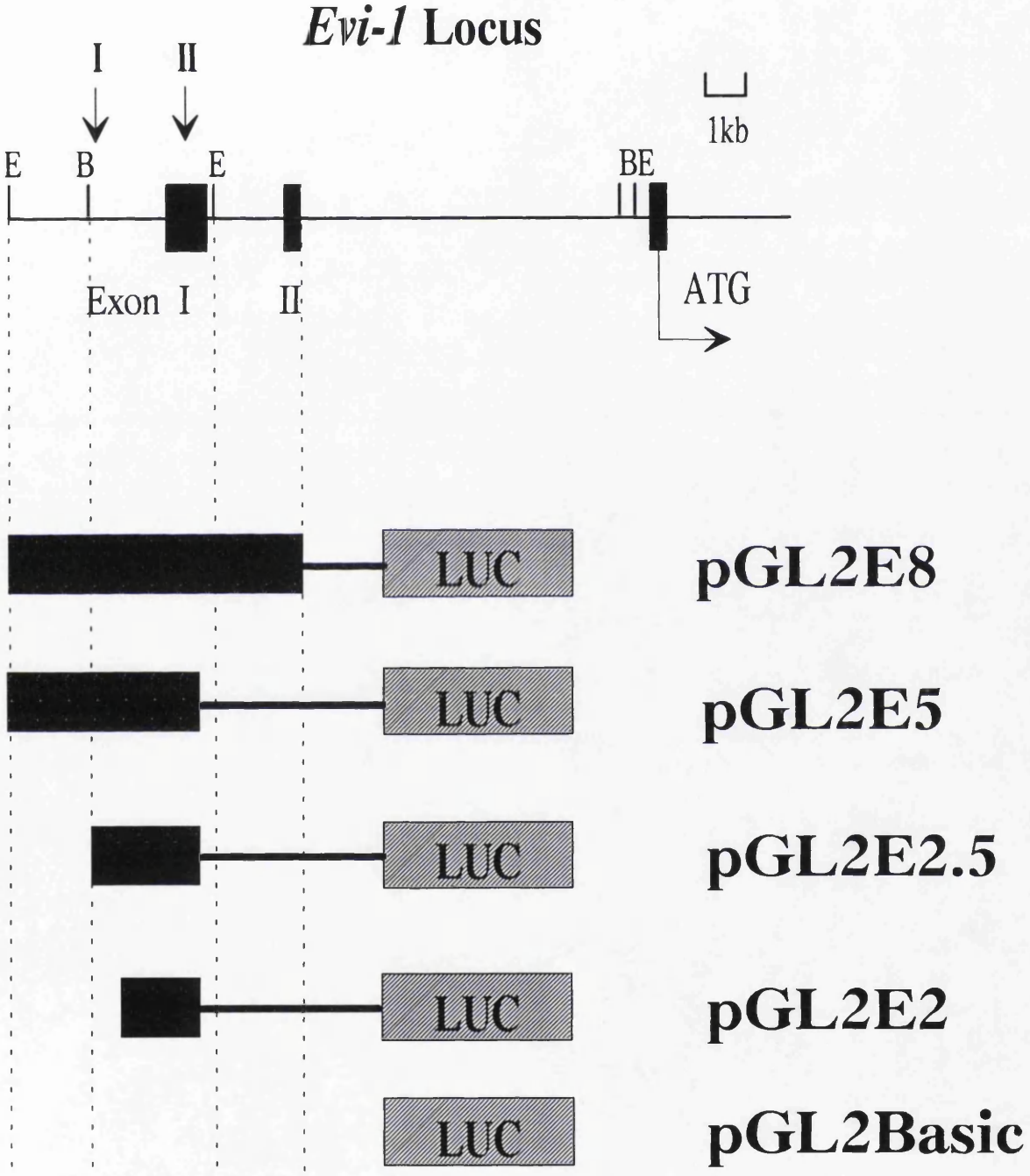

Evi-1
Locus
E
B
E
Exon I
II
PCR
Sm
Sm
3.2kb SstI/SmaI
Cut Sma I
S
Sm
Sm
3kb SmaI
Ligate
pKSII+ E5
Cut Sma I
Sm
S E B
Sm N
pKSII+ E5
Ligate
Sm
Sm
N
S E B
E
pKSII+ E8
Cut BamHI + NotI
S E BH S
NotI
Luciferase
pGL2E5Not
S E B
NotI
Luciferase
pGL2E5
Ligate
Sm
Sm
S E B
E
NotI
Luciferase
pGL2E8

Figure 5.3: Preliminary reporter constructs containing 8.0kb to 2kb of the *Evi-1* locus.

Schematic Representation of the *Evi-1* locus and the cloned reporter constructs. Exon I and II are labelled and represented as shaded boxes. The approximate location of DHS sites I and II are shown with arrows.

Preliminary Constructs



region (figure 5.3). The generation of the pGL2E5, pGL2E2.5 and pGL2E2 reporter constructs are described in figure 5.1.

5.3 Activity of the *Evi-1* reporter constructs in 293, L929 and HEC-1B cells

To identify the region required for basal transcription of the *Evi-1* gene the reporter constructs described above were transfected into 293 cells using DOTAP (Material & methods). These cells were chosen because they could be transfected relatively efficiently and were of kidney origin although they express only low levels of *Evi-1*. The results are shown in figure 5.4, where the promoter activity is calculated relative to the largest 8kb fragment, pGL2E8, which represents 100% activity. Deletion of the 8kb fragment, removing approximately 3kb of the 3' region to an SstI site at +272 within exon I (pGL2E5), does not effect the level of expression from the *Evi-1* promoter. This deletion removes a portion of DHS II suggesting the deleted region is not important for basal expression of the *Evi-1* promoter in these cells. Subsequent 5' deletions to just before exon I (pGL2E2.5 and pGL2E2) also showed no drop in the promoter activity. Since these deletions result in the loss of the constitutive DHS I region it is assumed that DHS I does not contribute to basal transcription levels in these cells.

Similar results were also obtained if the same constructs were examined in L929 and HEC-1B cells (figure 5.5 & 5.6). L929 cells were chosen since they represent a low Evi-1 expressing fibroblastic cell line and HEC-1B cells since they express high-levels of Evi-1 protein although the reason for this level is unclear since HEC-1B cells contains 96 chromosomes and have rearrangements within chromosome 3 (Morishita *et al.*, 1990). In addition to the conclusions reached in 293 cells the *Evi-1* promoter is also not responsive to auto-regulation, at least with the constructs used here, since the basal expression pattern is the same in high *Evi-1* expressing HEC-1B cells.

Figure 5.4: A histogram showing relative luciferase activity of preliminary *Evi-1* promoter constructs in 293 cells.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

293
% Relative Luciferase Activity
120
100
80
60
40
20
0
LUC
LUC
LUC
LUC
LUC
pGL2E8
pGL2E5
pGL2E2.5
pGL2E2
pGL2Basic
Luciferase Constructs

Figure 5.5: A histogram showing relative luciferase activity of preliminary *Evi-1* promoter constructs in HEC-1B cells.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

HEC-1B
% Relative Luciferase Activity
120
100
80
60
40
20
0
LUC
LUC
LUC
LUC
pGL2E8
pGL2E5
pGL2E2
pGL2Basic
Luciferase Constructs

Figure 5.6: A histogram showing relative luciferase activity of preliminary *Evi-1* promoter constructs in L929 cells.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

L929
% Relative Luciferase Activity
120
100
80
60
40
20
0
LUC
LUC
LUC
LUC
pGL2E8
pGL2E5
pGL2E2
pGL2Basic
Luciferase Constructs

The results show that DHS I plays no part in modulating basal promoter activity. It is plausible that DHS I may be important in inducible expression of the *Evi-1* gene or alternatively that the region has some intrinsic enhancer function which is being masked, either by other factors or by lack of appropriate tissue specific factors. Importantly, the results indicate that promoter activity resides around exon I of the gene and correlates with the presence of part of the DHS II region.

5.4 Defining the 5' border of the *Evi-1* minimal promoter

To more accurately define the location of the *Evi-1* promoter, 5' deletions of pGL2E2 were created by PCR. The sequence of exon I and immediately upstream have previously been published (Bartholomew & Ihle, 1991). The sequence from, -685 to +565, relative to the major transcription initiation site was used to design primers for the PCR reactions (figure 5.7). The 5' constructs generated by PCR were initially sub-cloned into a TA cloning vector (Invitrogen. TM) and then subsequently into the pGL2basic vector (Promega) used previously to generate the pGL2E vectors (Materials and methods). All the 5' deletion constructs contain the same 3' end as pGL2E and are summarised in figure 5.8.

5.5 Functional analysis of the PCR generated 5' deletion constructs in 293, L929, and HEC-1B cells

The relative luciferase activities of the 5' deletion constructs in 293 cells are shown (figure 5.9). The pGL2E2 construct, which represents 100% activity, was used to allow direct comparison of the 5' deletion PCR fragments with the preliminary constructs. No significant difference in activity is observed with deletions to: -684 (pGL2-684), -487 (pGL2-487) which removes a putative AP2 site; deletion to -443

Figure 5.7: PCR primers designed for generating 5' deletion reporter constructs

The 5' sequence of exon I with respect to the major initiation site +1 is shown. The 5' PCR primers are underlined with arrows. Putative DNA transcription factor binding sites are boxed and labelled appropriately. All constructs have the same 3' end at +272.

PCR primers

All constructs used the same 3' reverse primer

RP ccaagcttgagctcgcagctattcc

The 5' primers are also shown below;

-684 tttaagtagaaacttaa
-487 cctgaaagcaattctcc
-443 ctctaacacagactctc
-338 cttgtgctcaggacatc
-284 cctggagaaatttcccc
-254 cggcagaaatctacatg
-197 ggacttgggtcttggcg
-132 atttactgaaagaggag
-15 catgaaatagaaaccctgcgg

Upstream *Evi-1* Basal Promoter -

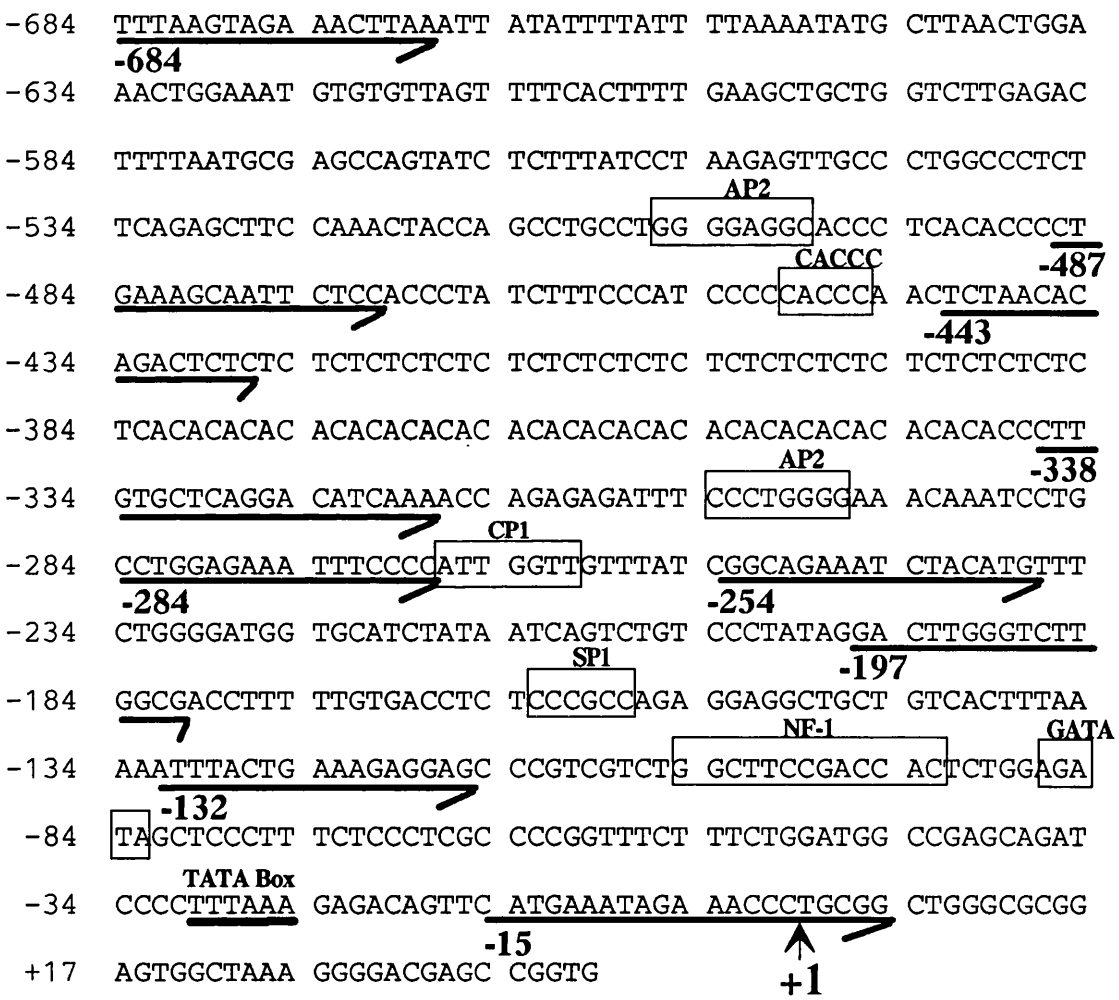


Figure 5.8: The 5' deletion reporter constructs generated by PCR.

5' deletions of the *Evi-1* locus with respect to the major site of transcriptional initiation were generated by PCR. A schematic representation of exon I is shown with the major initiation sites +1 shown. All constructs have the same 3' end at +272.

PCR generated 5' deletions

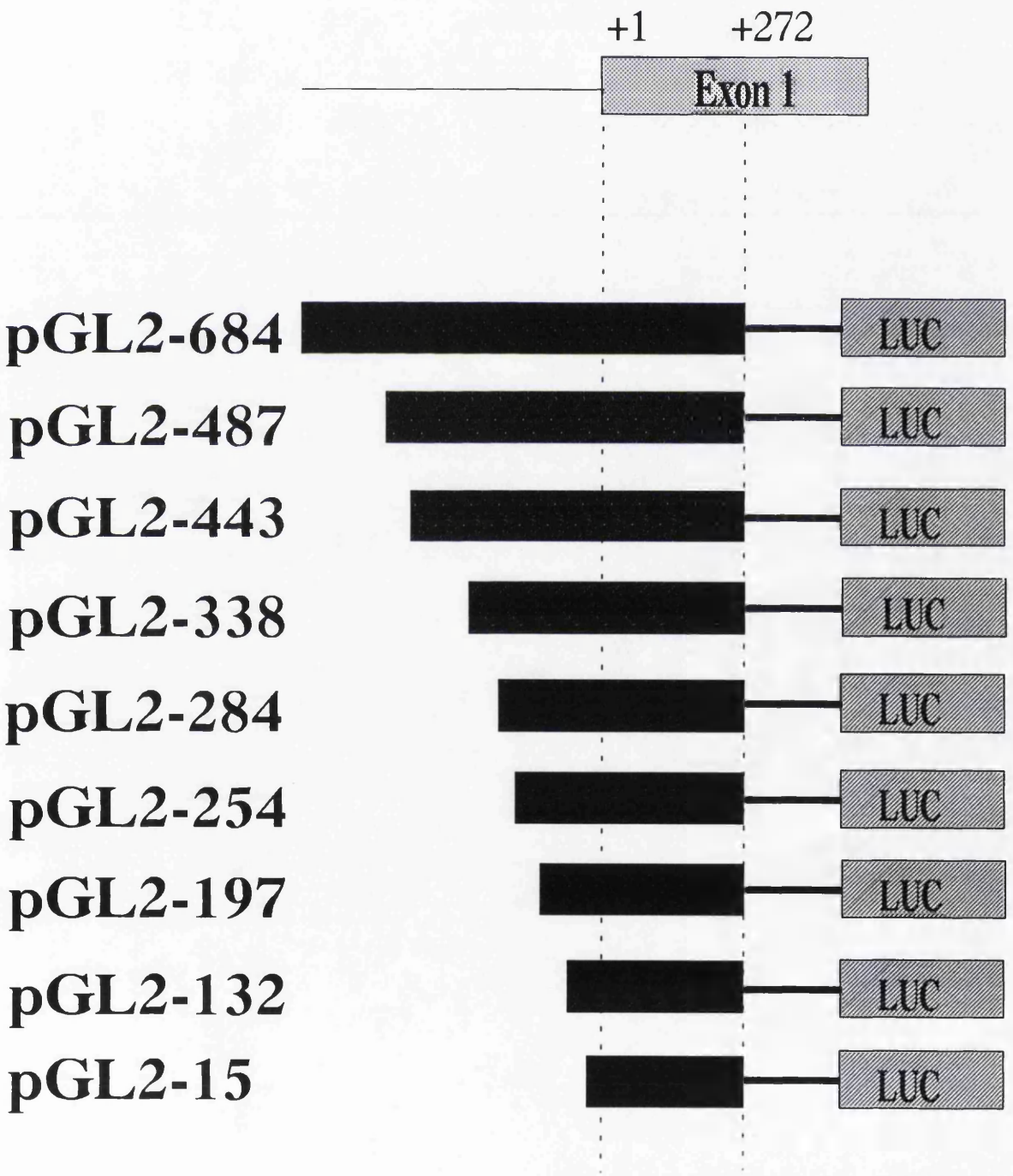


Figure 5.9: A histogram showing relative luciferase activity of 5' deletion *Evi-1* promoter constructs in 293 cells.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

293
% Relative Luciferase Activity
120
100
80
60
40
20
0
LUC
pGL2E2
pGL2-685
pGL2-487
pGL2-443
pGL2-338
pGL2-284
pGL2-254
pGL2-197
pGL2-132
pGL2-15
pGL2
Luciferase Constructs

(pGL2-443) which removes a putative CACCC box and deletion to -338 (pGL2-338) which removes a series of CTCT (GAGA) and CACA repeats. Presumably, the putative sites indicated in figure 5.7 are not essential for basal transcription. However, removal of another 52 base pairs to -284 (pGL2-284), results in an approximately 50% drop in activity. This correlates with the loss of a region with sequence homology to AP2, AP3, C/EBP and ets binding sites. Deletion of another 30 base pairs to -254 (pGL2-254), which removes a putative CP1 site, results in a further 20% drop in activity. The deletion from -338 to -254 represents the largest drop in functional activity and presumably reflects the loss of binding of *trans*-acting factors which are required for *Evi-1* basal promoter activity.

Further deletions result in a steady drop in activity down to -15 which removes putative, Sp1, NF-1, GATA and TATA box sequences, but maintains the initiation sites. These observations suggests that other elements within this region are important in basal promoter activity. The contribution of any given site in isolation to overall basal activity will require site directed mutagenesis of appropriate sites. The -15 construct is still approximately 4 to 5 fold higher than the pGL2basic reporter construct (which contains just the luciferase gene) and this may reflect the presence of *trans*-activating proteins between +1 and the SstI site at +272.

Similar results were also obtained in L929 and HEC-1B cells (figure 5.10 & 5.11). suggesting that the 5' border for optimal basal *Evi-1* promoter activity is at -338 and is the same in different cell types. In addition the majority of the basal activity is lost with deletion from -338 to -254 in all cell types. Therefore, the upstream activating sequences important in basal transcription within this region presumably bind similar ubiquitous factors or alternatively there is some redundancy and different proteins have the same effects in different cell lines or tissues. In addition, the results also suggest that other sequences between -254 and +272 may also contribute to basal activity.

Figure 5.10: A histogram showing relative luciferase activity of 5' deletion *Evi-1* promoter constructs in HEC-1B cells.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

HEC-1B
% Relative Luciferase Activity
120
100
80
60
40
20
0
LUC
pGL2E2
pGL2-443
pGL2-338
pGL2-284
pGL2-254
pGL2-197
pGL2-132
pGL2-15
pGL2
Luciferase Constructs

Figure 5.11: A histogram showing relative luciferase activity of 5' deletion *Evi-1* promoter constructs in L929 cells.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

L929
% Relative Luciferase Activity
140
120
100
80
60
40
20
0
LUC
pGL2E2
pGL2-338
pGL2-284
pGL2-15
pGL2
Luciferase Constructs

5.6 Defining the 3' border of the Evi-1 promoter

The majority of the constructs generated terminate at a 3' SstI site (+272) within exon I. To further define the 3' border of the basal *Evi-1* promoter two 3' deletions were generated. The first construct, -443 to +163, was made by PCR using primers designed from the published sequence (figure 5.12) and results in the loss of two putative GATA motifs. The construct was sub-cloned into a TA cloning vector and then into a pGL2basic vector to generate pGL2+163 (Material and methods). The second 3' deletion construct, -443 to +48 was generated by deleting an AvaI restriction enzyme fragment (there is an AvaI site at +48) from the -443 to +163 fragment located in the TA cloning vector, and subsequently cloning this fragment vector into pGL2basic to generate pGL2+48. This results in the loss of putative AP2 and PEA3 sites (a member of the ets family), in addition to GATA motifs. Both of the 3' deletion constructs contain the same 5' end at -443 and are summarised in figure 5.13.

5.7 Functional analysis of the 3' Evi-1 promoter deletion constructs in 293 and HEC-1B cells

The relative luciferase activity of the 3' deletion constructs in 293 cells are shown in figure 5.14 where they are compared with the optimal pGL2-443 vector The pGL2+163 vector shows similar levels of luciferase activity to the pGL2-443 construct. This shows that the two putative 3' GATA motif are not involved in basal transcription. However, the GATA sequences may be important in mediating signals from more distal elements in tissue specific regulation. However, pGL2+48 results in a dramatic 75-80% drop in activity. This deletion results in the loss of putative binding sites for AP2 and PEA3 although the region is generally GC rich and may therefore also be targets for Sp1 Similar results were also obtained with HEC-1B cells (figure 5.15).

Figure 5.12: The creation of the 3' deletion reporter constructs.

The 3' sequence of exon I with respect to the major initiation site +1 is shown.
The 3' +168 PCR primer is underlined with an arrow.
The +48 reproter construct was generated by digesting the -434 PCR product in the TA cloning vector with AvaI enzyme to remove 104bp, the location of the AvaI site within exon I is shown.
The vector was then re-ligated and the remaining PCR fragment cloned into in a pGL2basic luciferase reporter vector.
Putative DNA transcription factor binding sites are boxed and labelled appropriately.
All constructs have the same 5' end at -443.

+168 reverse primer `tcccactcctctgtcgc`

Downstream *Evi-1* Basal Promoter

Figure 5.13: The 3' deletion reporter constructs.

3' deletions of the *Evi-1* locus with respect to the major site of transcriptional initiation were generated by PCR. A schematic representation of exon I and the major initiation sites, **+1**, are shown. Both constructs have the same 5' end at -434.

PCR generated 3' deletions

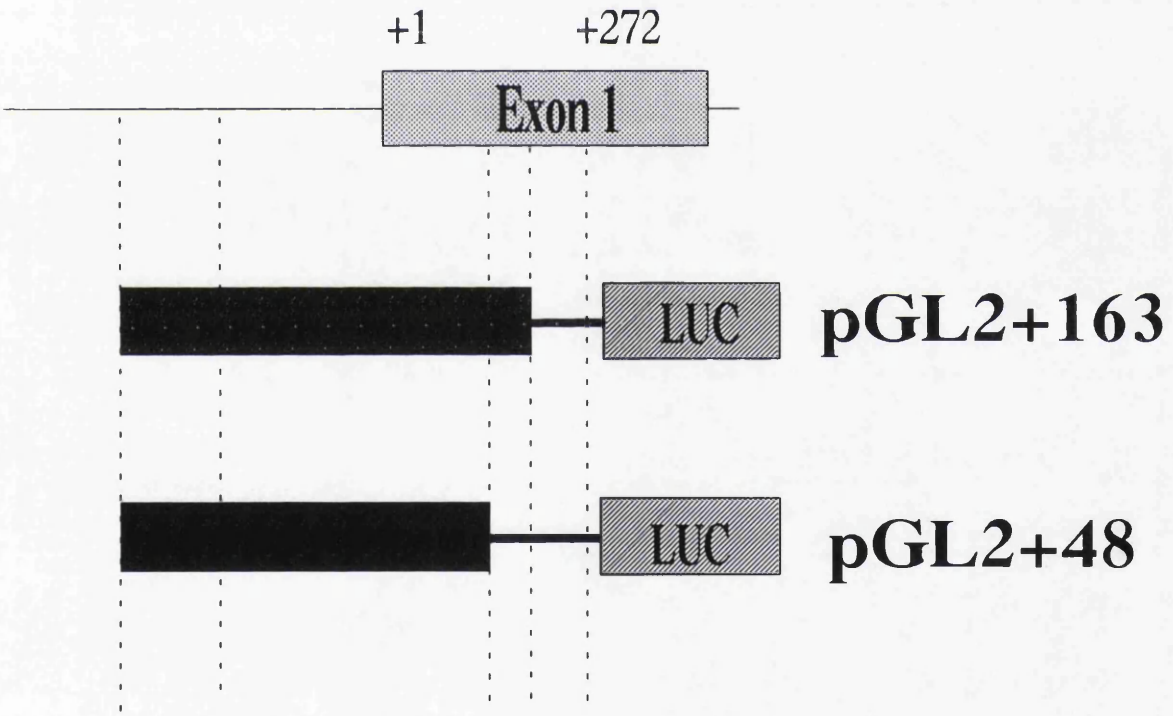


Figure 5.14: A histogram showing relative luciferase activity of 3' deletion *Evi-1* promoter constructs in 293 cells.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

293
% Relative Luciferase Activity
120
100
80
60
40
20
0
LUC
LUC
LUC
LUC
pGL2-443
pGL2
pGL2+48
pGL2+163
Luciferase Constructs

Figure 5.15: A histogram showing relative luciferase activity of 3' deletion *Evi-1* promoter constructs in HEC-1B cells.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

HEC-1B
Relative Luciferase Units
120
100
80
60
40
20
0
LUC
LUC
LUC
LUC
pGL2-443
pGL2
pGL2+48
pGL2+163
Luciferase Constructs

The results presented here conclude that the minimal basal *Evi-1* promoter is located between -338 to +168 and that two regions appear to be particularly significant in regulating transcriptional activity, a region 5' between -338 to -254 and a region 3' between +163 and +48.

5.8 Analysis of the DHS I region

DHS analysis has identified two DHS regions, I and II (chapter 4.4). DHS II, present exclusively in *Evi-1* expressing cells, represents the promoter region of the gene. By contrast, the DHS I region is constitutively present in all cells examined, but does not contribute to basal activity of the *Evi-1* promoter. As previously mentioned, DHS sites are generally associated with active promoter and enhancer elements (reviewed by Gross and Garrard, 1988). Therefore the function of the DHS I site was examined.

To characterise the DHS I region an 880bp region containing DHS I was sequenced to identify putative *cis*-acting sequences. Two 440bp restriction fragments (BamHI/HindIII and HindIII/BglII) each containing part of the DHS I region were ligated into an appropriately cut pKSII+bluescript vector and designated pKS-DIBH and pKS-DIHB respectively. T3 and T7 primers were used to sequence each end of the cloned fragments using the Sequenase version II kit (Materials and methods) (figure 5.16). The sequence information obtained allowed design of oligonucleotide primers for further sequencing (figure 5.16). To confirm the sequence data each of the two 440bp fragments were sequenced from both the 5' and 3' end.

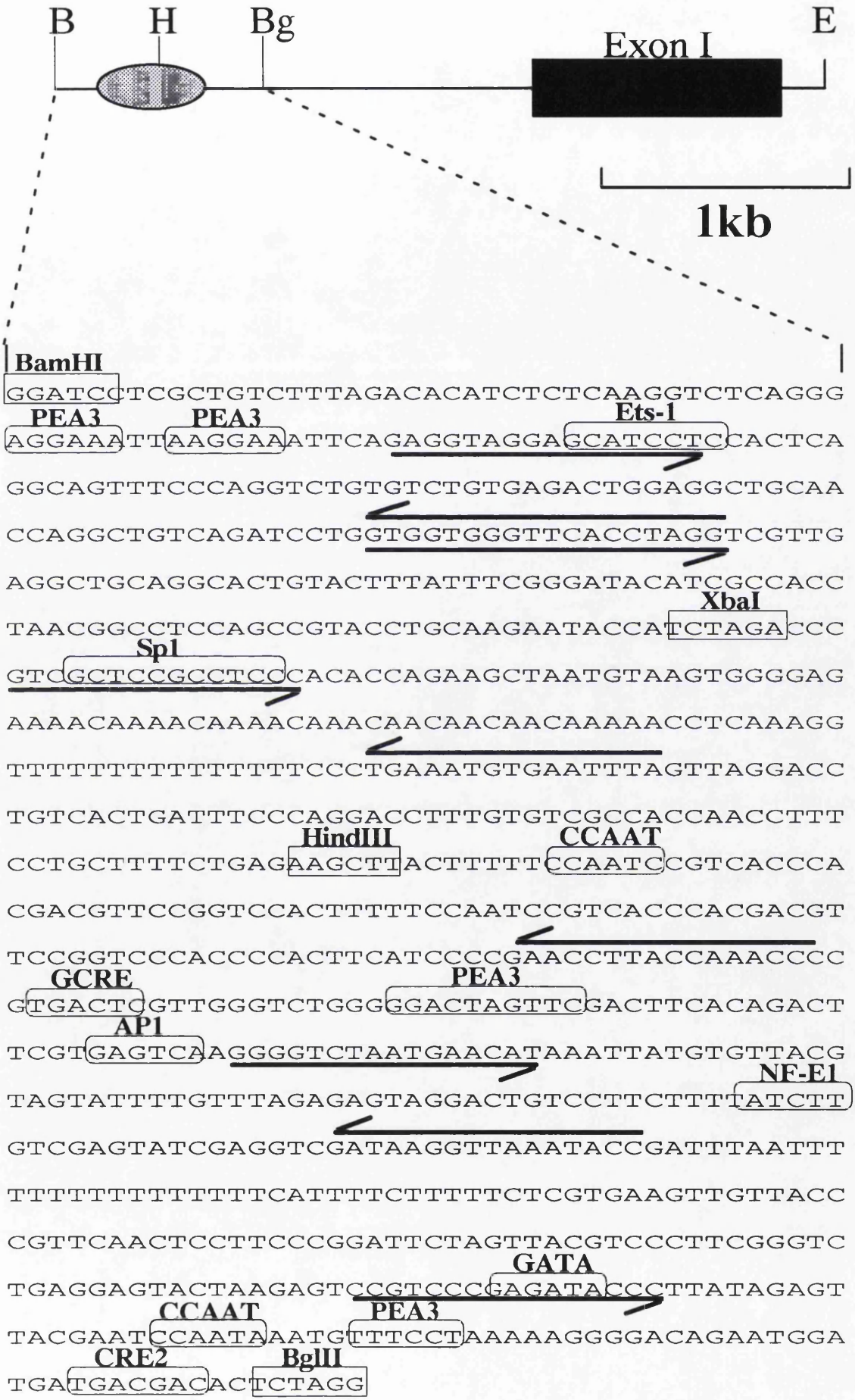
Figure 5.16: Sequence of the 880bp DHS I

The sequence of the 880bp BamHI/BglII fragment is shown. Pertinent restriction enzyme sites are boxed and labelled appropriately. In addition, putative protein recognition sequences identified within this region are boxed and labelled.

A schematic representation of the approximate location of DHS I relative to exon I is shown.

⬭ Denotes the approximate location of the DHS I region.

Evi-1 Locus


B
H
Bg
E
Exon I
1kb
BamHI
GGATCCTCGCTGTCTTTAGACACATCTCTCAAGGTCTCAGGG
PEA3
PEA3
Ets-1
AGGAAATTAAGGAAATTCAGAGGTAGGAGCATCCTCCACTCA
GGCAGTTTCCCAGGTCTGTGTCTGTGAGACTGGAGGCTGCAA
CCAGGCTGTCAGATCCTGGTGGTGGGTTCACCTAGGTCGTTG
AGGCTGCAGGCACTGTACTTTATTTCGGGATACATCGCCACC
XbaI
TAACGGCCTCCAGCCGTACCTGCAAGAATACCATCTAGACCC
Sp1
GTCGCTCCGCCTCCCACACCAGAAGCTAATGTAAGTGGGGAG
AAAACAAAACAAAACAAACAACAACAACAAAAACCTCAAAGG
TTTTTTTTTTTTTTTTCCCTGAAATGTGAATTTAGTTAGGACC
TGTCACTGATTTCCCAGGACCTTTGTGTCGCCACCAACCTTT
HindIII
CCAAT
CCTGCTTTTCTGAGAAGCTTACTTTTTCCAATCCGTCACCCA
CGACGTTCCGGTCCACTTTTTCCAATCCGTCACCCACGACGT
TCCGGTCCCACCCCACTTCATCCCCGAACCTTACCAAACCCC
GCRE
PEA3
GTGACTCGTTGGGTCTGGGGGACTAGTTCGACTTCACAGACT
AP1
TCGTGAGTCAAGGGGTCTAATGAACATAAATTATGTGTTACG
NF-E1
TAGTATTTTGTTTAGAGAGTAGGACTGTCCTTCTTTTATCTT
GTCGAGTATCGAGGTCGATAAGGTTAAATACCGATTTAATTT
TTTTTTTTTTTTTCATTTTCTTTTTCTCGTGAAGTTGTTACC
CGTTCAACTCCTTCCCGGATTCTAGTTACGTCCCTTCGGGTC
GATA
TGAGGAGTACTAAGAGTCCGTCCCGAGATACCCTTATAGAGT
CCAAT
PEA3
TACGAATCCAATAAATGTTTCCTAAAAAGGGGACAGAATGGA
CRE2
BglII
TGATGACGACACTCTAGG

5.9 Creation of DHS I luciferase reporter constructs to assess the transcriptional activity of DHS I

Although deletion of DHS I had no impact on basal transcription activity (chapter 5, section 5.3) it is plausible that the factors present in the *Evi-1* promoter repress activation by DHS I. To investigate if DHS I had enhancer activity, an 880bp BamHI/BglII fragment containing the DHS I region was cloned into a luciferase reporter gene immediately upstream of either a minimal -32 tk promoter, designated pGL2tkBB, or an SV40 promoter, designated pGL2svBB (figure 5.17). In addition, the 880bp fragment was sub-cloned as two 440bp BamHI/HindIII and HindIII/BglII fragments into a pKSII+bluescript vector (figure 5.18, A) and subsequently into a pGL2promoter vector, designated pGL2svB/H and pGL2svH/Bg (figure 5.17, B), or a minimal -32tk vector. The DHS I reporter constructs are summarised in figure 5.18.

5.10 Functional activity of the DHS I region in 293 and HEC-1B cells

DHS I reporter constructs were transfected into 293 cells using DOTAP (Materials and methods). The relative luciferase activity of the DHS I containing constructs were compared with the parental minimal tk, SV40 promoter and pGL2basic constructs (figure 5.19 A & B). All the DHS I constructs in both pGL2tk and pGL2sv vectors have the same relative luciferase activities as the parental vectors. Consequently, the DHS I region is unable to enhance the activity of a heterologous promoter in 293 cells. Similar observations were also seen in HEC-1B cells (figure 5.20 A & B).

The relative luciferase activity of the -32 tk or SV40 promoters in both 293 and HEC-1B cells was not significantly altered, either positively or negatively, by the presence of DHS I sequences located immediately upstream. This result is consistent

Figure 5.17: Schematic representation of the steps involved in cloning the DHS I region into a pKSII+ bluescript vector and upstream of a luciferase reporter vector.

An 880bp BamHI/BglII fragment containing the DHS I region was cloned directly into a pGL2basic vector at a BglII site immediately upstream of either a minimal tk or SV40 promoter.

A

The 880bp fragment was then sub-cloned as two 440bp BamHI/HindIII and HindIII/BglII fragments into a pKS+bluescript, designated pKS+B/H and pKS+H/Bg respectively, and used to sequence the DHS I region.

B

In addition, The 440bp BamHI/HindIII and HindIII/BglII fragments were clone upstream of either a minimal tk or SV40 promoter in a pGL2 luciferase reporter vector.

Evi-1 Locus

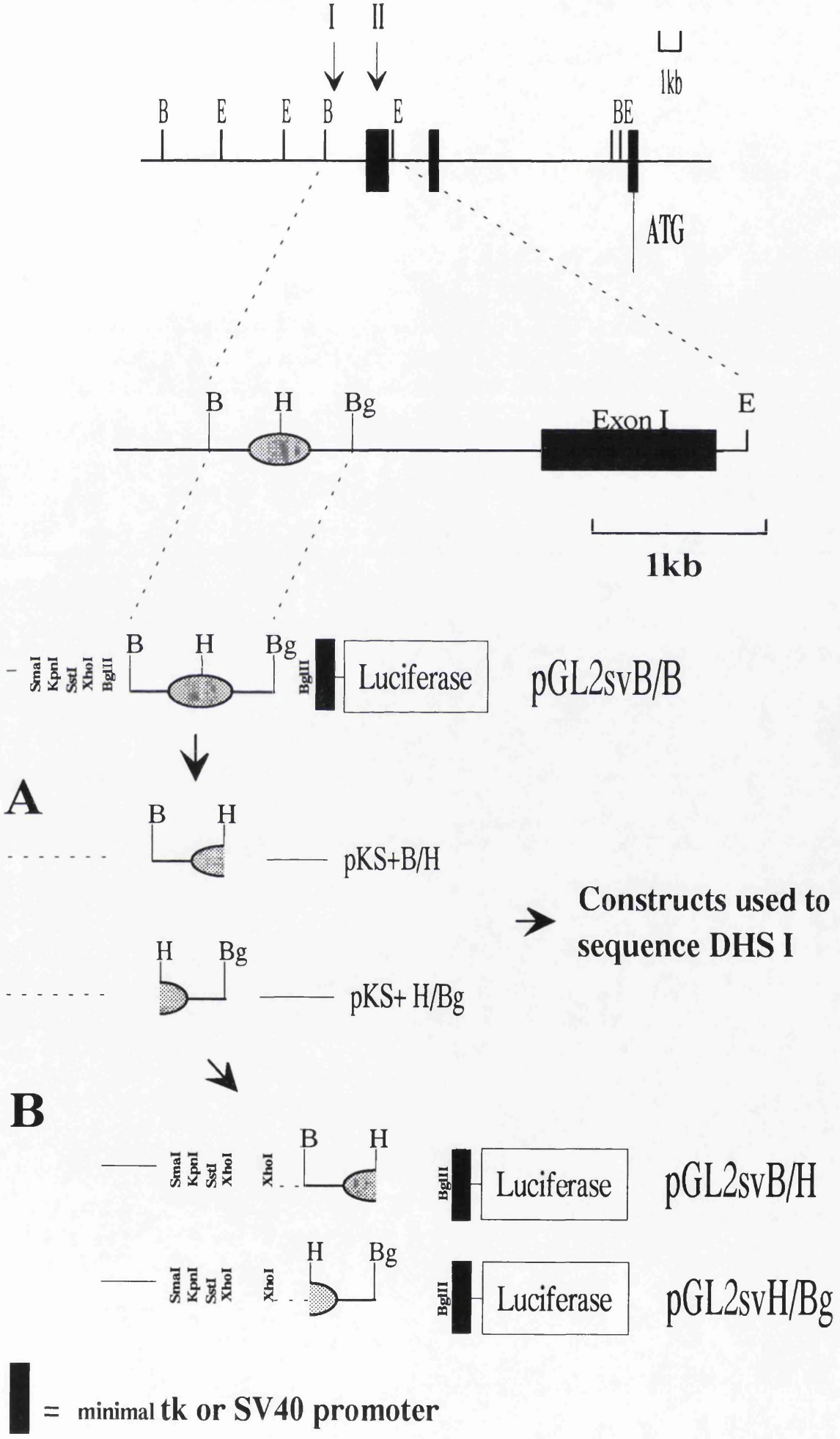
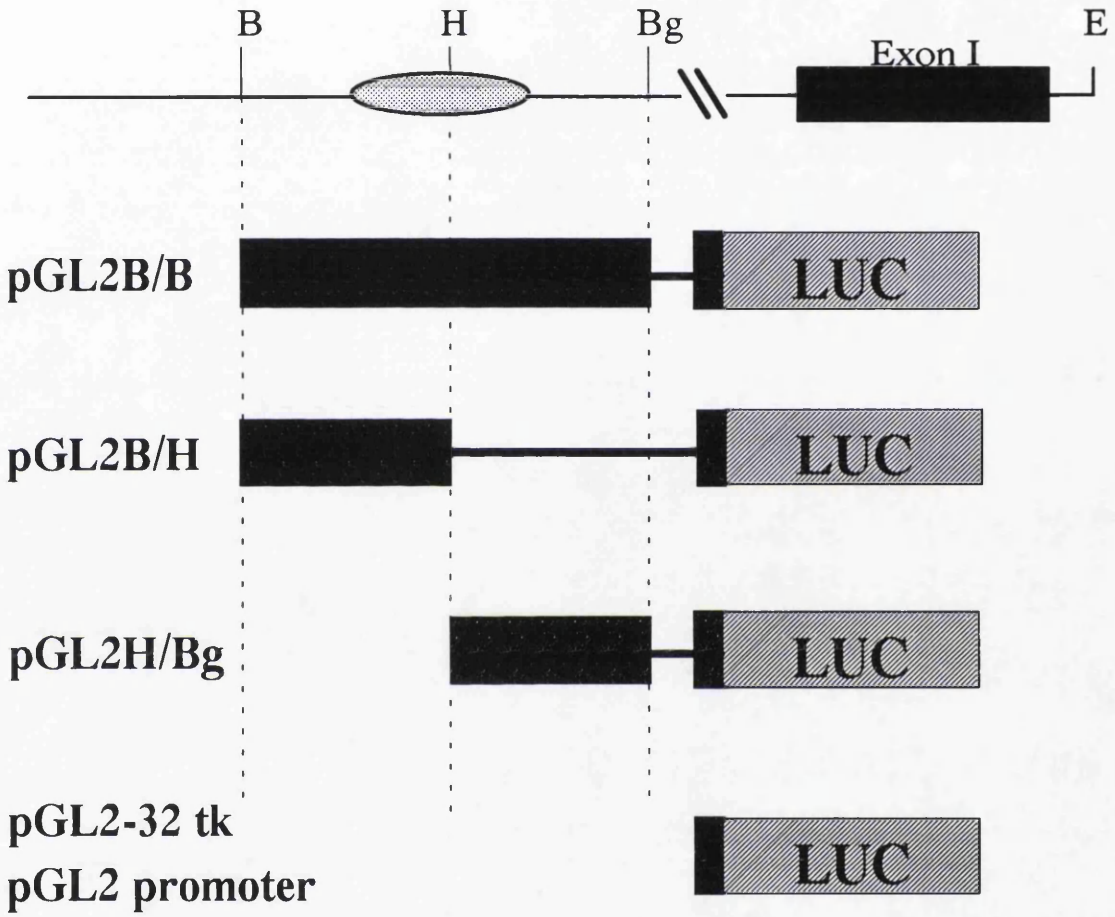

I
II
1kb
B
E
E
B
E
BE
ATG
B
H
Bg
Exon I
E
1kb
B
H
Bg
SmaI
KpnI
SstI
XhoI
BglII
BglII
Luciferase
pGL2svB/B
A
B
H
pKS+B/H
Constructs used to sequence DHS I
H
Bg
pKS+ H/Bg
B
SmaI
KpnI
SstI
XhoI
XhoI
B
H
BglII
Luciferase
pGL2svB/H
H
Bg
BglII
Luciferase
pGL2svH/Bg
= minimal tk or SV40 promoter

Figure 5.18: Summary of DHS I luciferase reporter constructs.

Schematic representation of the DHS I reporter constructs created. All the constructs were cloned upstream of an SV40 and minimal -32 tk promoter.

Summary of DHS I constructs



= minimal tk or SV40 promoter

= DHS I Region

B = BamHI H = HindIII Bg = BglII E = EcoRI

Figure 5.19: A histogram showing the promoter activity of the minimal -32 tk and SV40 promoters in the presence of the DHS I region in 293 cells.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

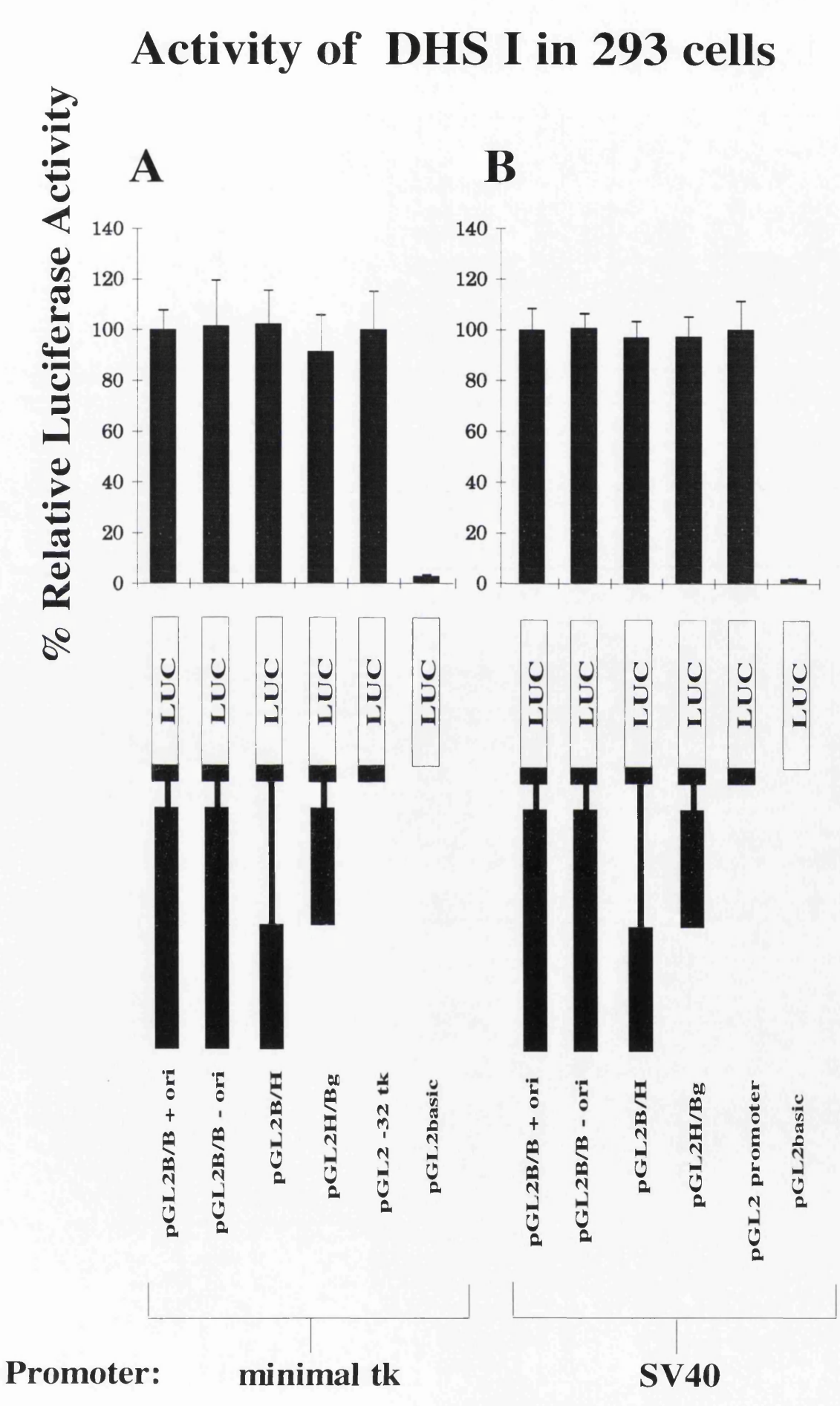

Activity of DHS I in 293 cells
% Relative Luciferase Activity
A
B
140
120
100
80
60
40
20
0
LUC
pGL2B/B + ori
pGL2B/B - ori
pGL2B/H
pGL2H/Bg
pGL2 -32 tk
pGL2basic
pGL2 promoter
Promoter:
minimal tk
SV40
Luciferase Constructs

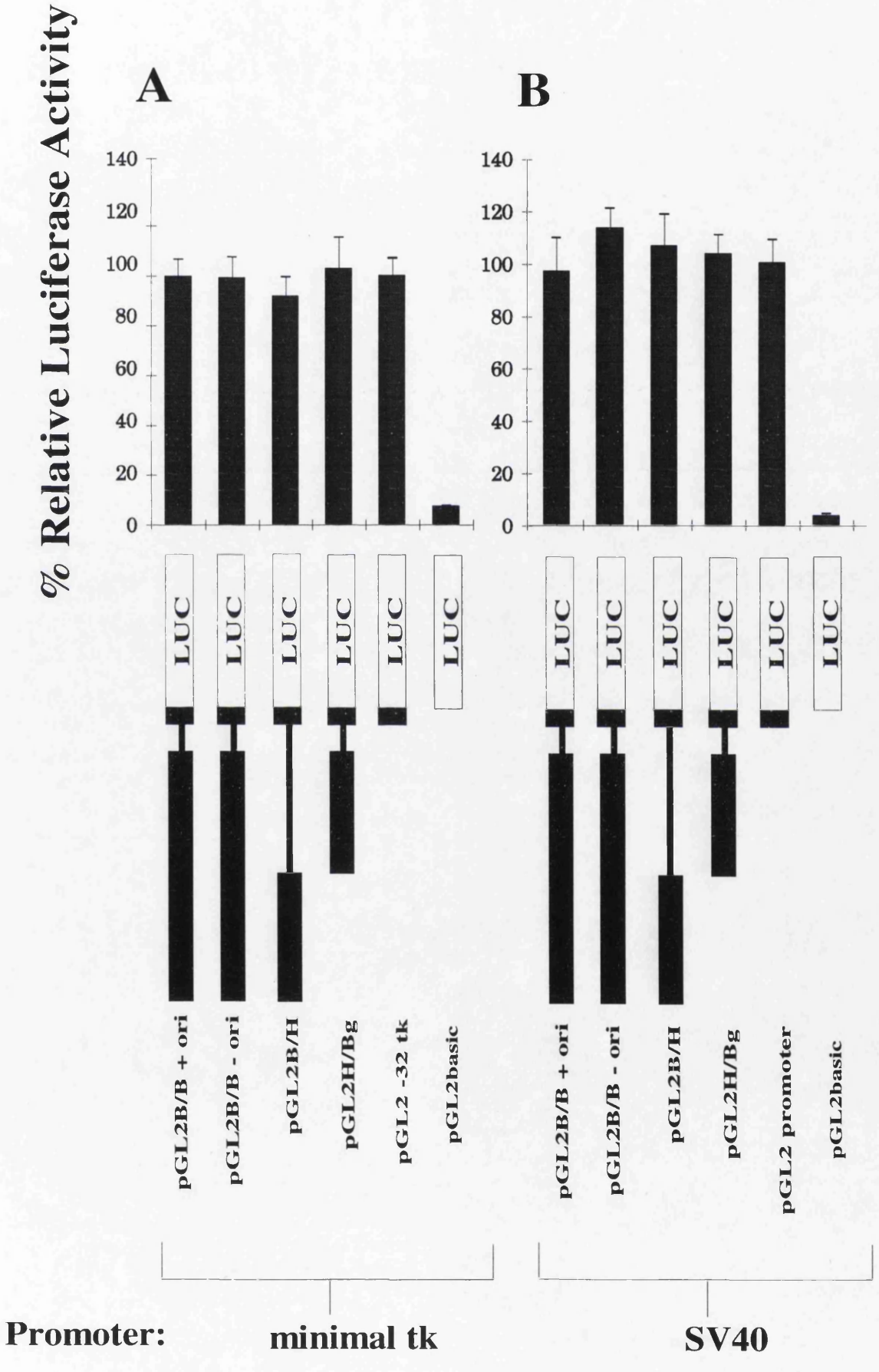
Figure 5.20: A histogram showing the promoter activity of the minimal -32 tk and SV40 promoters in the presence of the DHS I region in HEC-1B cells.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

Activity of DHS I in HEC-1B cells



with the transfection studies performed where the removal of DHS I had no detectable effect on the authentic *Evi-1* promoter (chapter 5.6). In all the assays performed no function for DHS I in controlling transcriptional levels has been identified. As mentioned previously, DHS I may be important in inducible or tissue specific expression in the appropriate cell type. Alternatively, DHS I may be important in the maintenance of nucleosomal positioning within the chromatin framework, which may not be apparent in transient assays.

Chapter 6: *In vitro* analysis of transcription factor/DNA interactions within the minimal *Evi-1* promoter

DNAse1 hypersensitivity site analysis implicated DHS II as being important in controlling transcriptional regulation of the *Evi-1* promoter since the hypersite was found exclusively in *Evi-1* expressing cells. In addition, transfection studies have identified the minimal promoter located between -338 to +163 (figure 6.1). To examine this region in more detail DNA/protein interactions have been investigated and their importance in basal transcription assessed.

6.1 DNAse1 footprinting of the -474 to -135 region

DNAse 1 footprinting was performed over the -474 to -135 region to investigate if there are DNA/protein complexes which correlate with changes in basal promoter activity observed in functional studies (Chapter 5.5). Since transfection studies had already shown functional activity of the *Evi-1* promoter in HEC-1B and 293 cells, these cells were chosen to isolate nuclear extracts for footprinting analysis. A probe for the -474 to -135 region was generated by PCR and sub-cloned into a pKSII+ vector. The 339bp EcoRI/HindIII probe was ^{32}P-end labelled at the 3' (HindIII) end and footprint reactions performed as described in Materials and methods.

Two footprinted regions have been identified with nuclear extracts derived from HEC-1B or 293 cells which are labelled FpIII and FpIV (figure 6.2). Two lanes are shown for each extract which represents increasing concentrations of DNAse 1. FpIII protects a region of approximately 37bp from -321 to -284 while FpIV protects approximately 69bp from -277 to -208. To confirm the location of FpIII and FpIV the same -474 to -135 region was also ^{32}P-end labelled from the -474 (5') end and

Figure 6.1 Sequence of -338 to +163 minimal *Evi-1* basal promoter

The major sites of transcriptional initiation (nucleotides CT) are indicated by arrows underneath. The sequence is numbered relative the first, C nucleotide, site of transcriptional initiation.
The putative TATA box is underlined.

The symbol ▢ represents the 5' or 3' deletion reporter constructs assayed in transfection studies (Chapter 5.5 & 5.7).

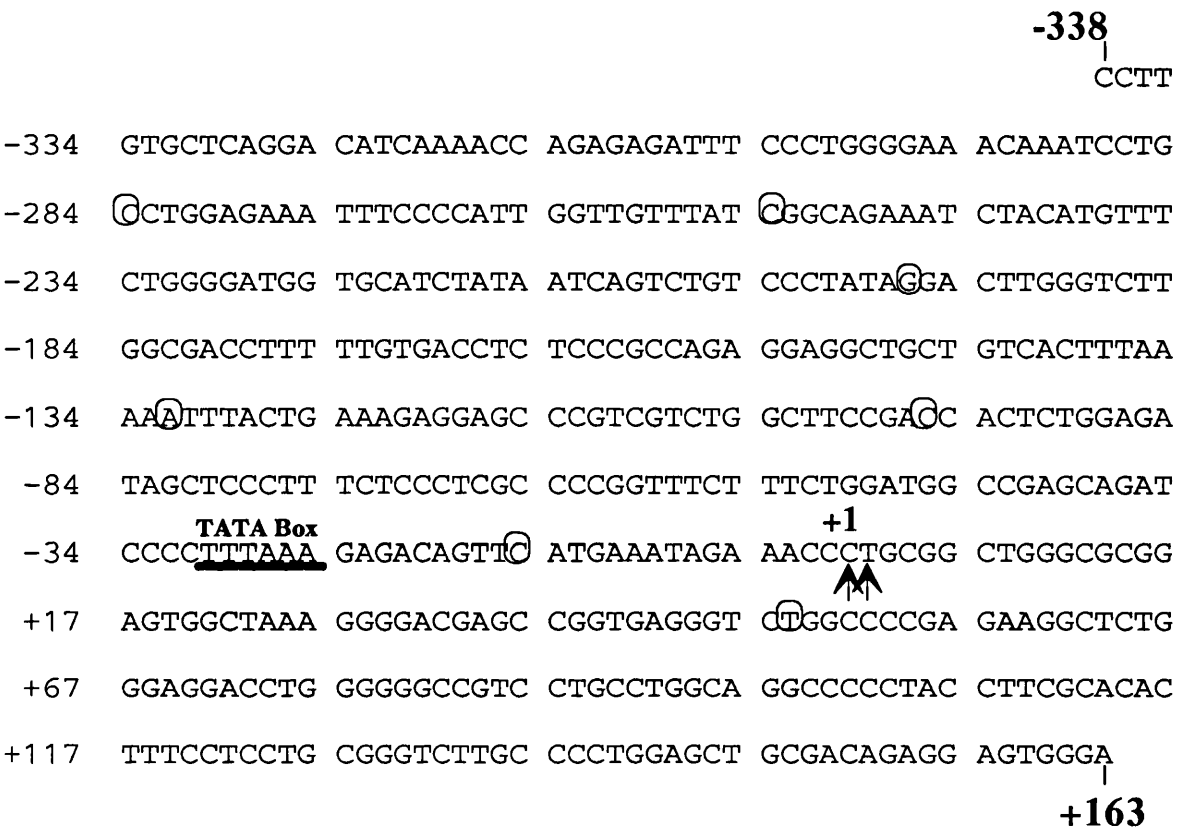

-338
CCTT
-334 GTGCTCAGGA CATCAAAACC AGAGAGATTT CCCTGGGGAA ACAAATCCTG
-284 CCTGGAGAAA TTTCCCCATT GGTTGTTTAT CGGCAGAAAT CTACATGTTT
-234 CTGGGGATGG TGCATCTATA ATCAGTCTGT CCCTATAGGA CTTGGGTCTT
-184 GGCGACCTTT TTGTGACCTC TCCCGCCAGA GGAGGCTGCT GTCACTTTAA
-134 AAATTTACTG AAAGAGGAGC CCGTCGTCTG GCTTCCGACC ACTCTGGAGA
-84 TAGCTCCCTT TCTCCCTCGC CCCGGTTTCT TTCTGGATGG CCGAGCAGAT
TATA Box
+1
-34 CCCCTTTAAA GAGACAGTTC ATGAAATAGA AACCCTGCGG CTGGGCGCGG
+17 AGTGGCTAAA GGGGACGAGC CGGTGAGGGT CTGGCCCCGA GAAGGCTCTG
+67 GGAGGACCTG GGGGGCCGTC CTGCCTGGCA GGCCCCCTAC CTTCGCACAC
+117 TTTCCTCCTG CGGGTCTTGC CCCTGGAGCT GCGACAGAGG AGTGGGA
+163

Figure 6.2: DNAse 1 Footprinting of -343 to -138 of the *Evi*-1 promoter. Labelling from 3', -138 end.

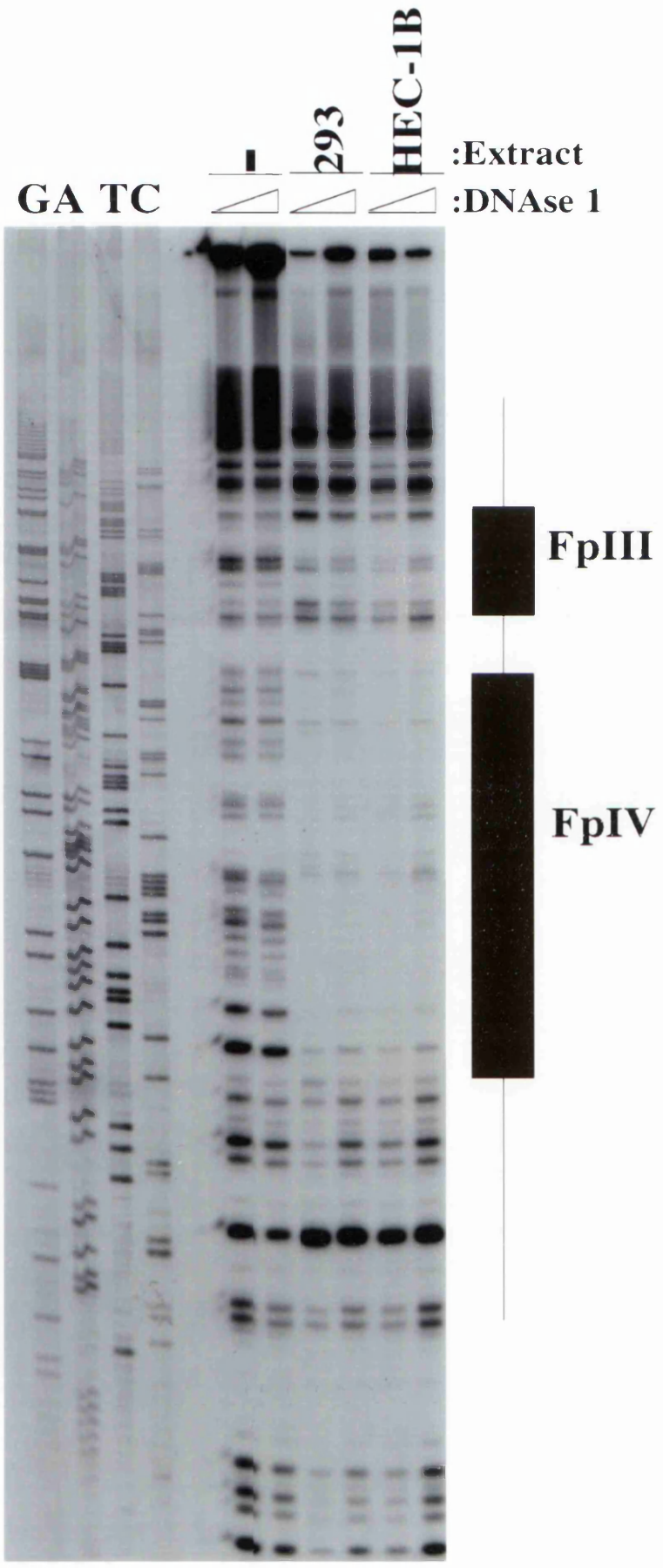
The -343 to -138 region of the *Evi-1* promoter was generated by PCR and cloned into a bluescript KS+ vector. The PCR primers were constructed so that the 5' and 3' ends could be subsequently cut with EcoRI and HindIII restriction enzymes respectively. DNA footprinting probes were ^{32}P end-labelled and purified as previously described. Labelled probe was incubated in the presence or absence of nuclear extract, digested with DNAse 1 enzyme and separated by electrophoresis on a denaturing PAGE gel as described by Plumb and Goodwin, 1988. Gels were dried and exposed to autoradiograph film.

Labelled probe was incubated either in the absence of nuclear extract, ▮ , or with either nuclear extract from HEC-1B or 293 cells. The nuclear extract used in the reaction is shown above the appropriate lane. Approximately 180µg of nuclear extract was used in each reaction.

◿ illustrates increasing DNAse 1 concentration.

GATC sequencing tracks are labelled and were performed using the Sequenase version II kit (USB).

Footprinted regions are represented by shaded boxes.


GA TC
-
293
HEC-1B
:Extract
:DNAse 1
FpIII
FpIV

footprinting reactions performed again as above. The FpIII and FpIV regions are shown in figure 6.3.

A summary of the DNAse1 protected regions are shown in figure 6.4. Computer analysis of this region identified putative binding sites for AP2 and CP1 transcription factors which correlate with the footprinted region, although other putative binding sites were also identified in this region (data not shown).

Other footprints have previously been identified outside this region and are labelled within this chapter as FpI, II and V (Bartholomew, unpublished data). FpI and FpII are further upstream than FpIII and IV identified here. FpI and FpII are lost in the deletions between -487 to -338. However, no change in luciferase activity (Chapter 5.5) is observed and therefore is assumed to play no part in maintaining basal promoter expression. FpV maps over the transcriptional initiation sites and presumably represents the initiation complex.

6.2 Electrophoretic Mobility Shift Assay (EMSA) analysis of protein binding activity

In addition to DNAse1 footprinting to analyse *in vitro* protein binding to specific DNA sequences another technique commonly used is electrophoretic mobility shift assay analysis (EMSA). This technique is advantageous because it, can distinguish different DNA/protein complexes binding to the same sequence by virtue of their relative mobility, is a simpler and more rapid technique, and requires lower concentrations of nuclear extract. As a consequence EMSA was employed to examine regions of the minimal promoter for protein binding activity. EMSA analysis was undertaken with sequence derived from both FpIII and FpIV to identify the proteins binding to these regions.

Figure 6.3: DNAse 1 Footprinting of -343 to -138 of the *Evi*-1 promoter. Labelling from 5', -343 end.

The -343 to -138 region of the *Evi-1* promoter was generated by PCR and cloned into a bluescript KS+ vector. The PCR primers were constructed so that the 5' and 3' ends could be subsequently cut with EcoRI and HindIII restriction enzymes respectively. DNA footprinting probes were ^{32}P end-labelled and purified as previously described. Labelled probe was incubated in the presence or absence of nuclear extract, digested with DNAse 1 enzyme and separated by electrophoresis on a denaturing PAGE gel as described by Plumb and Goodwin, 1988. Gels were dried and exposed to autoradiograph film.

Labelled probe was incubated either in the absence of nuclear extract, [redacted] , or with either nuclear extract from HEC-1B or 293 cells. The nuclear extract used in the reaction is shown above the appropriate lane. Approximately 180µg of nuclear extract was used in each reaction.
GATC sequencing tracks are labelled and were performed using the Sequenase version II kit (USB).
Footprinted regions are represented by shaded boxes.

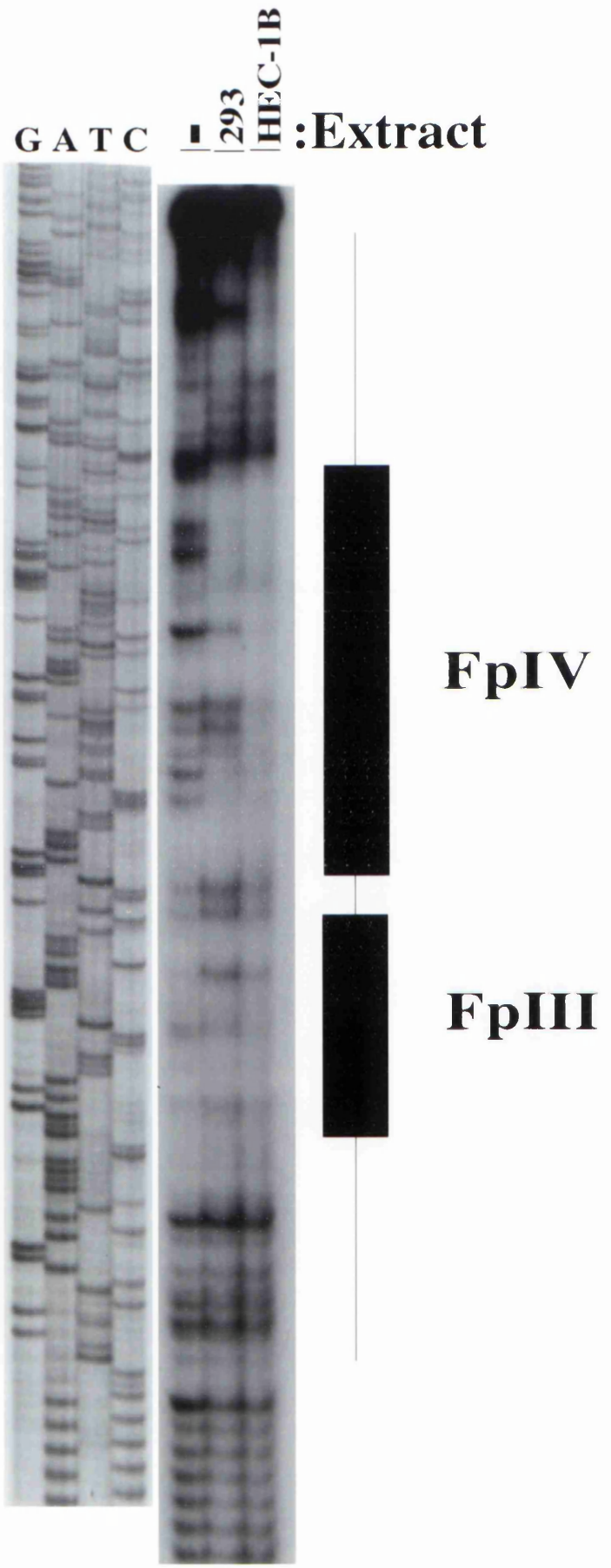
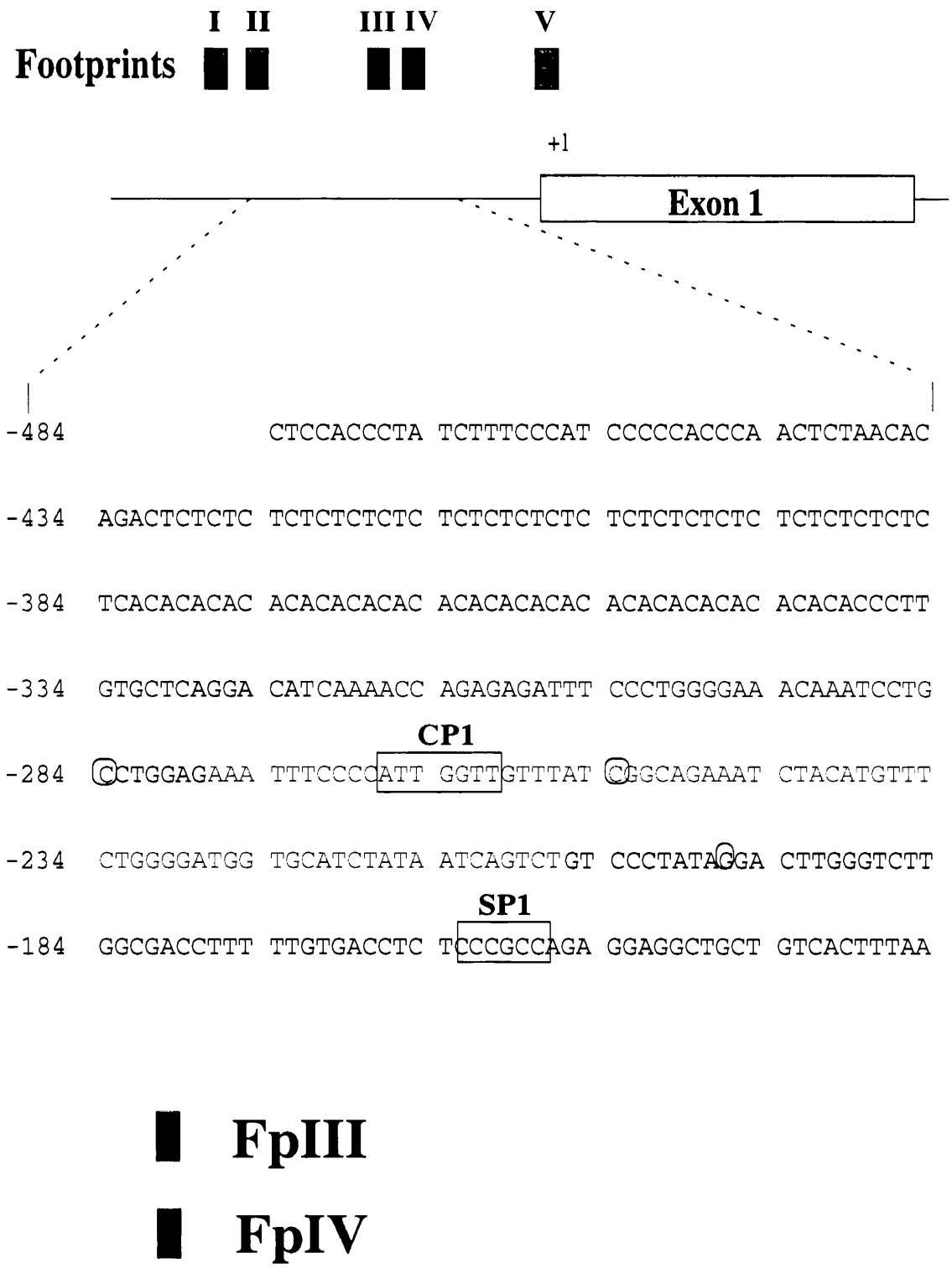

G A T C
-
293
HEC-1B
:Extract
FpIV
FpIII

Figure 6.4: Summary of DNAse 1 footprinting

A shematic representation of the location of the DNAse1 footprints, FpIII and FpIV relative to exon 1. FpIII and IV are highlighted in either red or blue respectively. The sequence is numbered relative to the first major site of transcriptional initiation, +1. The putative binding sites for either CP1 or SP1 are boxed.

-474 to -135 Footprinted Region



6.3 Protein interactions with the FpIII region

Figure 6.5 is a schematic representation of the FpIII region highlighting the -321 to -284 protected sequence and EMSA probes used in the analysis of this region. Figure 6.6 shows EMSA analysis using the FpIII probe, which covers most of FpIII and a small region 3' down to the start of FpIV. Five complexes are observed in EMSA using nuclear extracts from 293 cells. A major specific slow mobility complex labelled EvBP1 (Evi-1 binding protein 1) was identified which competed with both cold FpIII oligonucleotide as well as cold 5' FpIII and 3' FpIII oligonucleotides. Faint high mobility complexes A and B are not competed with cold 3' FpIII oligonucleotides but show slight competition with cold FpIII and 5' FpIII oligonucleotides. Two non-specific complexes labelled N appeared in all the lanes. It should be noted that the relative intensities of some of the minor complexes show some minor variation between experiments, the significance of which is unclear.

Functional analysis showed that the first major drop in basal transcriptional activity occured between -338 and -284. FpIII covers most of this region, however to determine if there are other proteins upstream EMSA analysis was performed between -341 to -305. However, this analysis failed to detect any DNA/protein complexes in this region (figure 6.7). EMSA analysis with FpIII, was also performed in parallel with the same extracts showing the expected EvBP1 complex (figure 6.7). EMSA analysis of the FpIII region has, therefore, identified a major EvBP1 complex. Functional studies show that loss of this binding activity correlates with a significant (50%) drop in basal activity.

Figure 6.5: Schematic representation of the FpIII region

The sequence -321 to -284 represents the FpIII region. The probes used in the following EMSA experiments are indicated by shaded boxes.
The arrows show an inverted repeat sequence within the FpIII region.

Footprints I, II and V have previously been identified by (Bartholomew, unpublished data).

EMSA Probes

-341	caccttgtgctcaggacatcaaaaccagagagattt
FpIII	agagagatttccctggggaaacaaatcctgcctggagaaatttccccatt
5' FpIII	agagagatttccctggggaaacaaatcctgcc
3'FpIII	cctggggaaacaaatcctgcctggagaaatttccccatt

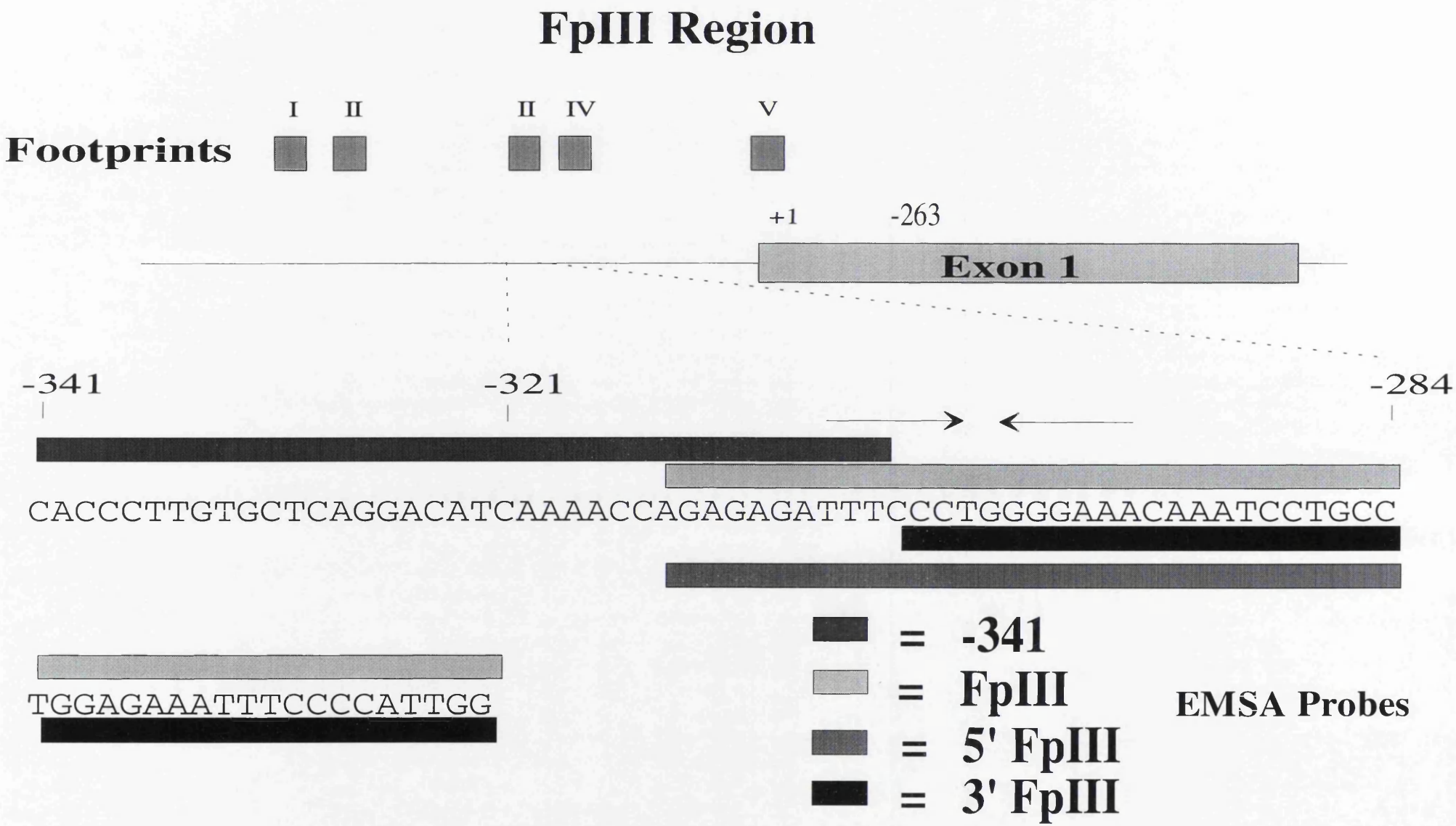

FpIII Region
Footprints
I
II
II
IV
V
+1
-263
Exon 1
-341
-321
-284
CACCCTTGTGCTCAGGACATCAAAACCAGAGAGATTTCCCTGGGGAAACAAATCCTGCC
TGGAGAAATTTCCCCATTGG
= -341
= FpIII
= 5' FpIII
= 3' FpIII
EMSA Probes

Figure 6.6: EMSA showing protein binding to FpIII region.

Nuclear extracts were prepared from 293 cells as described in Materials and Methods. A double stranded oligonucleotide corresponding to FpIII region was ^{32}P end-labelled. Binding reactions were set up using 500pg of labelled oligonucleotide and equal amounts of protein nuclear extract (approximately 10μg per reaction). Reactions were run on non-denaturing PAGE gels, dried down and exposed to X-ray film overnight (Materials and methods).

The labelled oligonucleotide used in all reactions was FpIII. Unlabelled competitor oligonucleotides are shown above each lane: ▮ - indicates no competitor, **Cold** indicates unlabelled FpIII region, **5'**- indicates 5' FpIII region, **3'**- indicates 3' region of FpIII region. All competitors are added at 200 fold excess (100ng). 293 nuclear extract was used in all reactions. Unbound labelled FpIII oligonucleotide is observed at the bottom of the gel.

Specific complexes are labelled EvBP1, A and B.

Non-specific complexes which are not competed with cold competitor are labelled N.

Oligonucleotide sequences:

FpIII	agagagatttccctggggaaacaaatcctgcctggagaaatttccccatt
5' FpIII	agagagatttccctggggaaacaaatcctgcc
3' FpIII	cctggggaaacaaatcctgcctggagaaatttccccatt

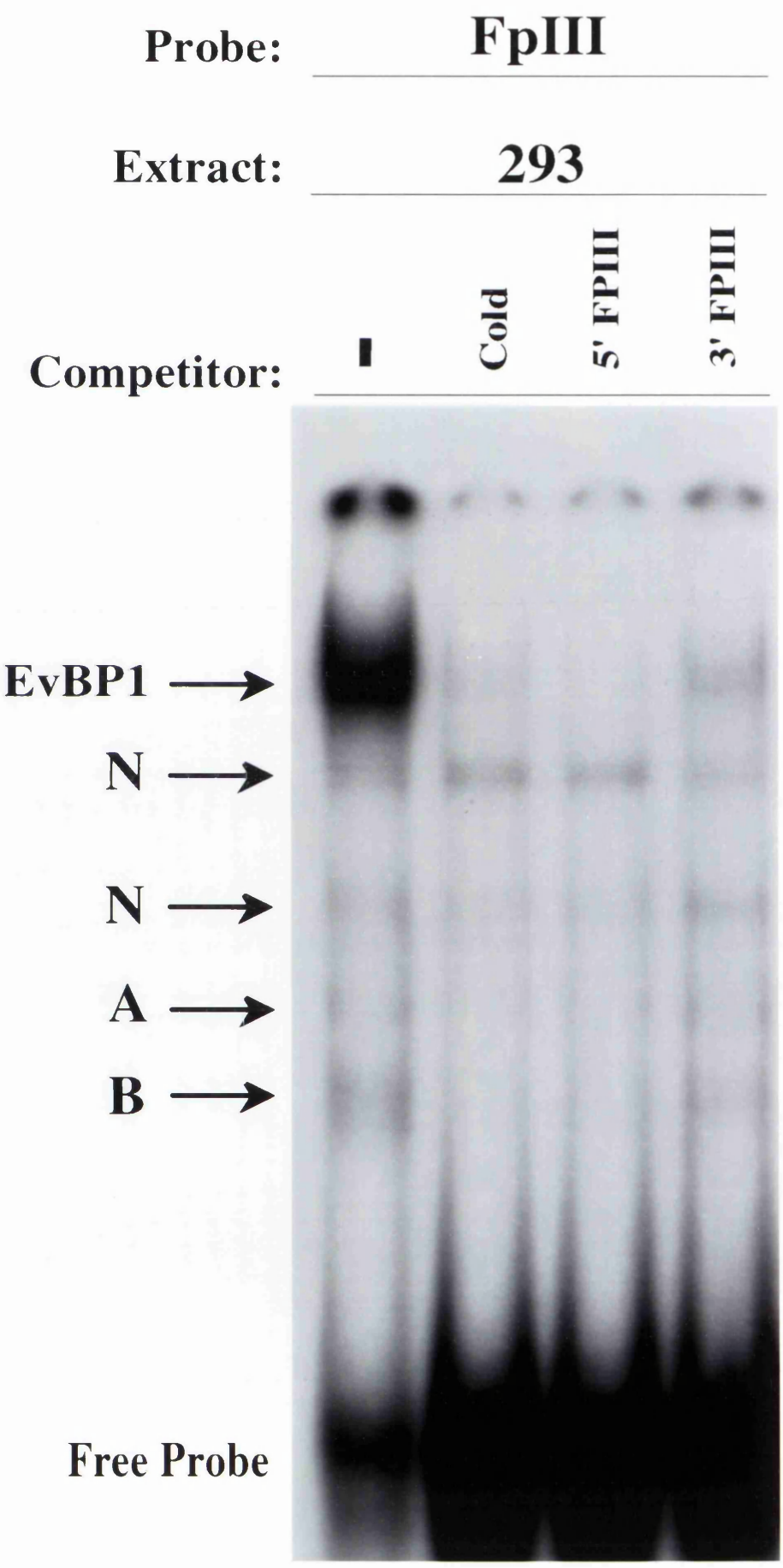

Probe:
FpIII
Extract:
293
Competitor:
-
Cold
5' FPIII
3' FPIII
EvBP1
N
N
A
B
Free Probe

Figure 6.7: EMSA showing protein binding from -341 to -305.

Nuclear extracts were prepared from 293 cells as described in Materials and Methods. A double stranded oligonucleotide corresponding to the FpIII and -341 region was ^{32}P end-labelled. Binding reactions were set up using 500pg of labelled oligonucleotide and equal amounts of protein extract (approximately 10μg per reaction). Reactions were run on non-denaturing PAGE gels, dried down and exposed to X-ray film overnight (Materials and methods).

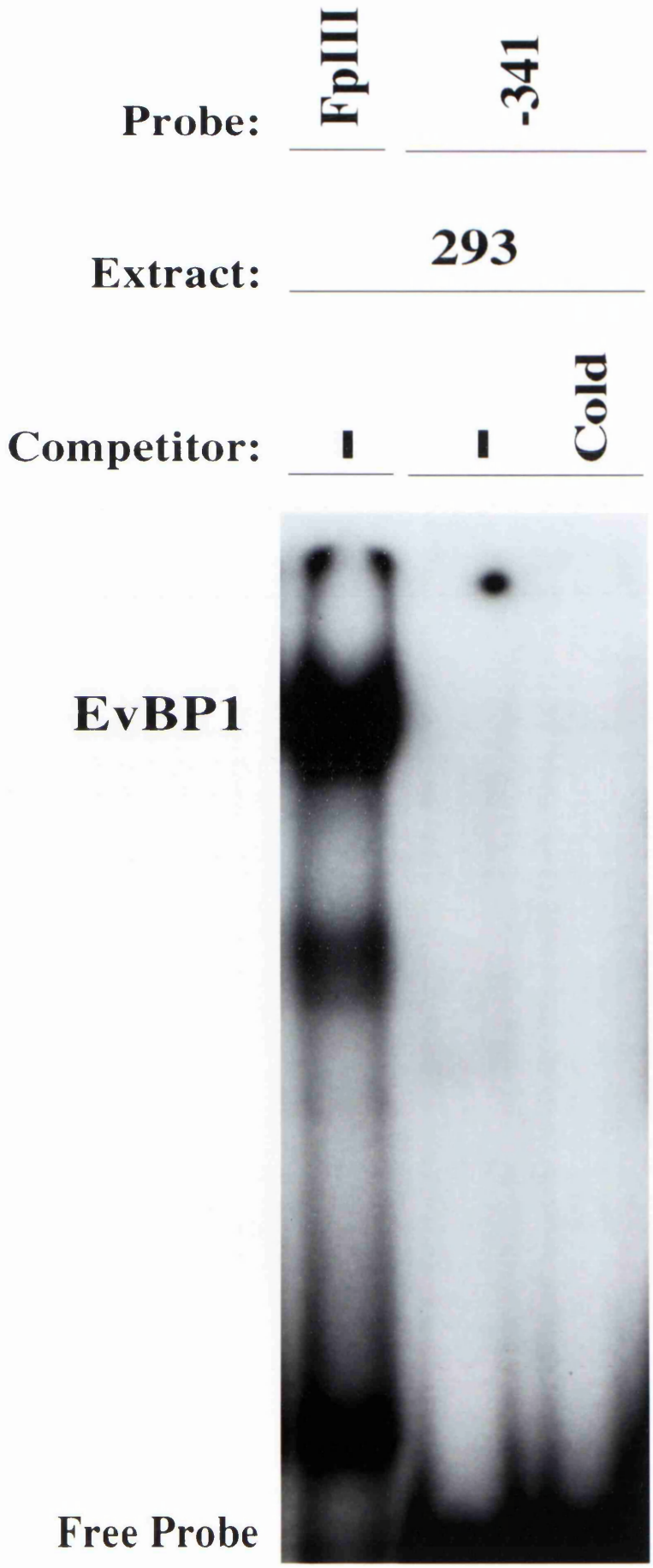
The labelled oligonucleotides used were either FpIII and -341 as shown above the lane. Unlabelled competitor oligonucleotides are shown above each lane: **▮** - indicates no competitor, **Cold** - indicates unlabelled -341 region. All competitors are added at 200 fold excess (100ng). 293 nuclear extract was used in all reactions. Unbound labelled -341 and FpIII oligonucleotide is observed at the bottom of the gel.

FpIII was added as a positive control for the reactions. The major EvBP1 complex is clearly retarded with a higher-mobility non-specific complex running below.

Oligonucleotide sequences:

-341 caccttgtgctcaggacatcaaaaccagagagattt

FpIII agagagatttccctggggaaacaaatcctgcctggagaaatttccccatt


Probe:
FpIII
-341
Extract:
293
Competitor:
-
-
Cold
EvBP1
Free Probe

6.4 Characterisation of FpIII binding activity

To identify the proteins binding to FpIII, particularly the EvBP1 complex, a competition experiment was performed. This involves incubating excess cold candidate consensus binding site oligonucleotides for known transcription factors at 200 fold excess with labelled FpIII oligonucleotides and nuclear extract. If EvBP1 contains a component which binds to the consensus binding site oligonucleotide it will compete out the EvBP1 complex observed with labelled FpIII. Figure 6.8 shows the results of this analysis using 293 cell extracts. The particular consensus binding site oligonucleotide cold competitor used are indicated.

The EvBP1 complex is competed with cold FpIII, as shown previously in figure 6.8. Cold consensus binding site oligonucleotides for E74, PU1, PEA3, SP1, NF-1, E box, H4TF-1, NF-κB, ZFI, ZF2, GATA-1 and AP2 all failed to compete for the EvBP1 complex. However, the C/EBP consensus binding site oligonucleotide showed effective competition for EvBP1 binding activity, suggesting that a member of this family of transcription factors might be a component of the complex. Interestingly, this portion of the 5' FpIII region identified a sequence tggggaaac which shows some homology to the C/EBP consensus t(t/g)nng(c/t)aa(t/g).

6.5 EvBP1 is not a C/EBP homo/hetero dimer

To further investigate the EvBP1 complex a C/EBP consensus binding site was ^{32}P-end labelled and EMSA performed to see if the FpIII cold oligonucleotide could compete for C/EBP binding (figure 6.9). The C/EBP ^{32}P-labelled probe forms two complexes when incubated with 293 extract. The slower mobility complex is competed with cold C/EBP. The higher mobility complex only partially competes suggesting that this complex may be non-specific.

Figure 6.8: EMSA showing competition FpIII oligonucleotide protein binding activity

Nuclear extracts were prepared from 293 cells as described in Materials and Methods. A double stranded oligonucleotide corresponding to FpIII region was ^{32}P end-labelled.

Binding reactions were set up using 500pg of labelled oligonucleotide and equal amounts of protein nuclear extract (approximately 10μg per reaction). Reactions were run on non-denaturing PAGE gels, dried down and exposed to X-ray film overnight (Materials and methods).

The labelled oligonucleotide used in all reactions is FpIII. Unlabelled competitor oligonucleotides are shown above each lane: ▮ - indicates no competitor, **Cold** indicates unlabelled FpIII region. All competitors are added at 200 fold excess (100ng). 293 nuclear extract was used in all reactions. Unbound labelled FpIII oligonucleotide is observed at the bottom of the gel.

Oligonucleotide sequences:

FpIII	agagagatttccctggggaaacaaatcctgcctggagaaatttccccatt
5' FpIII	agagagatttccctggggaaacaaatcctgcc
3' FpIII	cctggggaaacaaatcctgcctggagaaatttccccatt
AP2	ccggccccaggct
ZF1	ggatccctggacaagataagataaaggcagatct
ZF2	gatccctagctatgtgcctcatcttcgacaggctgtct
C/EBP	gatcggatcccaattgggcaatcaggggatcc
NF-1	cttttggattgaagccaatatgaga
SP1	tgggaatcctaactgggcggagttatgctggtggtg
NF-κB	agttgaggggactttcccagg
GATA-1	ccgggcaactgataaggattccctg
PU1	gatccataacctctgaaagaggaacttggttaggt
PEA3	gatcctcgagcaggaagttcga
E74	gatctctagctgaataaccggaagtaactcatccta
E Box	cccccaccacgtggtgcctga
H4TF-1	tttagatttcccctcccccaccggggcggga

ZF1 and ZF2 represent the two DNA recognition sequences for zinc-finger domains I and II of the EVI-1 protein.

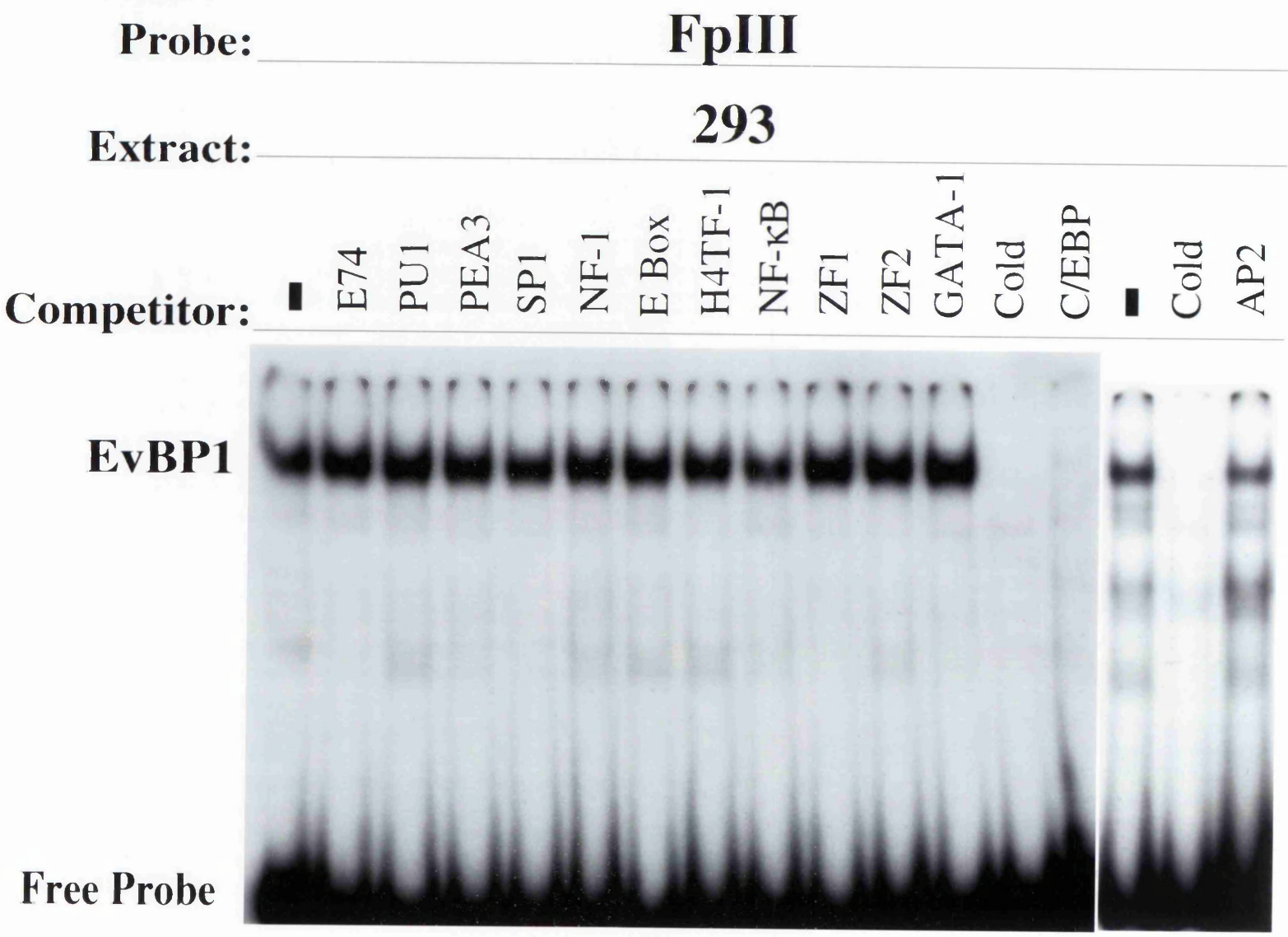

Probe:
FpIII
Extract:
293
Competitor:
-
E74
PU1
PEA3
SP1
NF-1
E Box
H4TF-1
NF-κB
ZF1
ZF2
GATA-1
Cold
C/EBP
-
Cold
AP2
EvBP1
Free Probe

Figure 6.9: EMSA showing competition of a labelled C/EBP consensus oligonucleotide with the FpIII region.

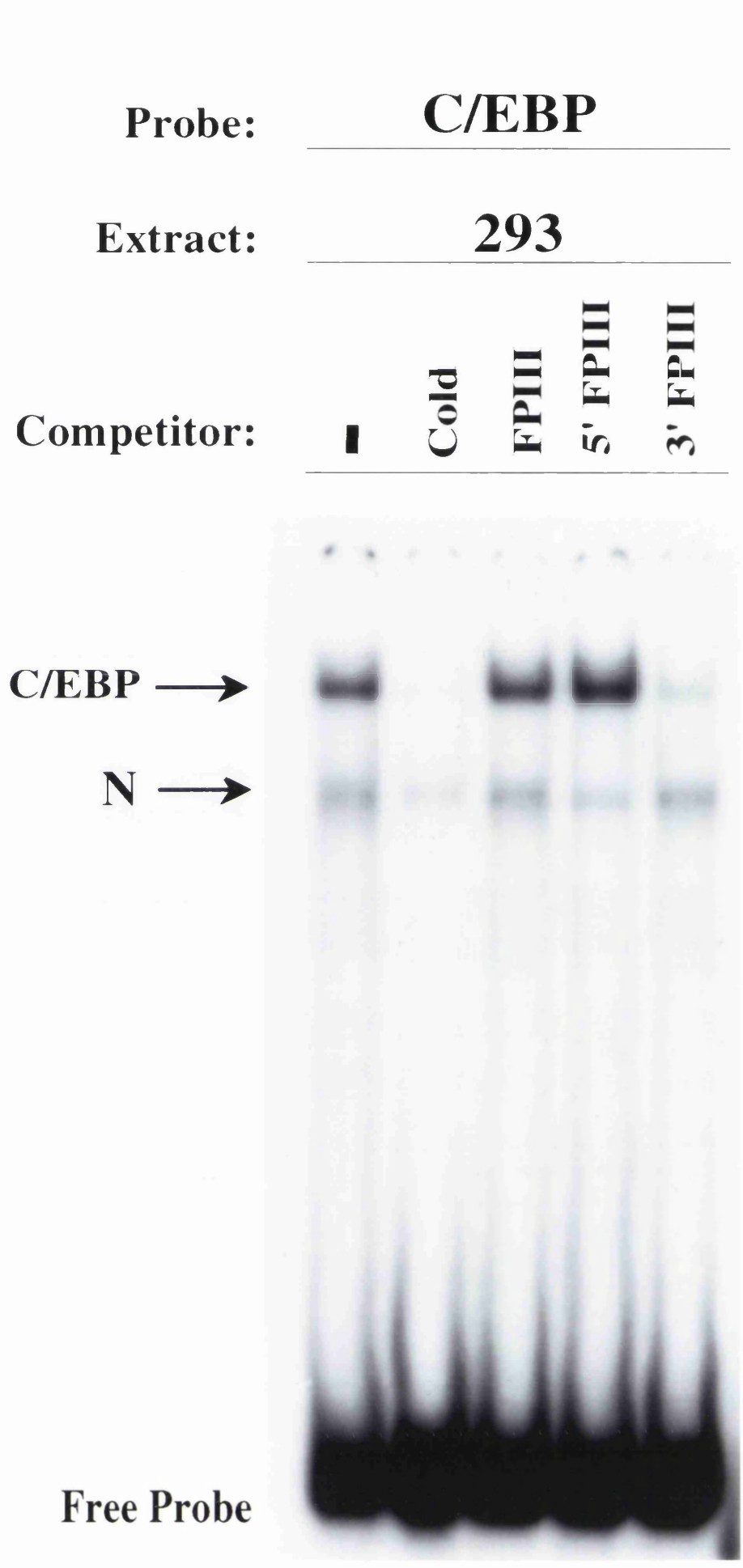
Nuclear extracts were prepared from 293 cells as described in Materials and Methods. A double stranded oligonucleotide corresponding to an C/EBP consensus binding site was ^{32}P end-labelled. Binding reactions were set up using 500pg of labelled oligonucleotide and equal amounts of protein nuclear extract (approximately 10μg per reaction). Reactions were run on non-denaturing PAGE gels, dried down and exposed to X-ray film overnight (Materials and methods).

The labelled oligonucleotide used in all reactions was C/EBP. Unlabelled competitor oligonucleotides are shown above each lane: ▮- indicates no competitor, **Cold** indicates unlabelled C/EBP region, **5'**- indicates 5' FpIII region, **3'**- indicates 3' region of FpIII region. All competitors are added at 200 fold excess (100ng). Unbound labelled C/EBP oligonucleotide is observed at the bottom of the gel.

A specific, C/EBP related complex, and a non-specific (N) complex are labelled appropriately.

Oligonucleotide sequences:

C/EBP	gatcggatcccaattgggcaatcaggggatc
FpIII	agagagatttccctggggaaacaaatcctgcctggagaaatttccccatt
5' FpIII	agagagatttccctggggaaacaaatcctgcc
3' FpIII	cctggggaaacaaatcctgcctggagaaatttccccatt


Probe:
C/EBP
Extract:
293
Competitor:
-
Cold
FPIII
5' FPIII
3' FPIII
C/EBP
N
Free Probe

The C/EBP complex was examined by competition with either cold FpIII, 5' FpIII or 3' FpIII. Interestingly, only the 3' FpIII region was able to compete for the C/EBP binding complex. The 3' FpIII region disrupts a potential inverted repeat sequence which binds the EvBP1 complex (chapter 6.7). An explanation for the observed difference in specificity for the C/EBP complex may therefore be that the DNA forms a secondary structure which usually masks any affinity of C/EBP for this region. Alternatively, it might be that C/EBP competes for EvBP1 because of similarity of sequence, although the affinity is low for the C/EBP site. This may also explain the observation that the 3' FpIII competes for the EvBP1 complex with decreased affinity relative to cold FpIII or 5' FpIII (figure 6.6) and the 3' FpIII ^{32}P-end-labelled probe also has substantial decreased ability to form the EvBP1 complex, figure 6.10. Both the identity and function of complex C in figure 6.10 is unknown.

A characteristic of C/EBP family members is that they are heat stable. To further prove that the EvBP1 complex is not a C/EBP family homo- or heterodimer a heat stability experiment was performed to see if the EvBP1 complex is still present following heat treatment, figure 6.11. Non-treated 293 extracts form either EvBP1 or C/EBP complexes in the presence of the appropriate labelled probe which are also competed with cold competitor. If the labelled probes are incubated with extracts heat treated at 50^{o}C for 5 minutes then only the C/EBP complex is formed. This suggests that the EvBP1 complex is not heat stable and consequently is not a C/EBP complex. However, the experiment cannot rule out the possibility that C/EBP is a component of the EvBP1 complex and that the dimerisation partner is heat sensitive.

6.6 The EvBP1 complex is ubiquitous

EMSA analysis was performed with ^{32}P end-labelled FpIII and different cell extracts, to evaluate if the EvBP1 complex components are ubiquitously expressed, as

Figure 6.10: EMSA showing binding to labelled 3' FpIII region

Nuclear extracts were prepared from 293 cells as described in Materials and Methods. A double stranded oligonucleotide corresponding to FpIII region was ^{32}P end-labelled. Binding reactions were set up using 500pg of labelled oligonucleotide and equal amounts of protein nuclear extract (approximately 10μg per reaction). Reactions were run on non-denaturing PAGE gels, dried down and exposed to X-ray film overnight (Materials and methods).

The labelled oligonucleotide used in all reactions is 3' FpIII. Unlabelled competitor oligonucleotides are shown above each lane: ▮ - indicates no competitor, **Cold** indicates unlabelled 3' FpIII region. All competitors are added at 200 fold excess (100ng). 293 nuclear extract was used in all reactions. Unbound labelled FpIII oligonucleotide is observed at the bottom of the gel.

Specific complexes are labelled A, B, C, and D are discussed in the text.

Oligonucleotide sequences:

FpIII	agagagatttccctggggaaacaaatcctgcctggagaaatttccccatt
5' FpIII	agagagatttccctggggaaacaaatcctgcc
3' FpIII	cctggggaaacaaatcctgcctggagaaatttccccatt
C/EBP	gatcggatcccaattgggcaatcaggggatc

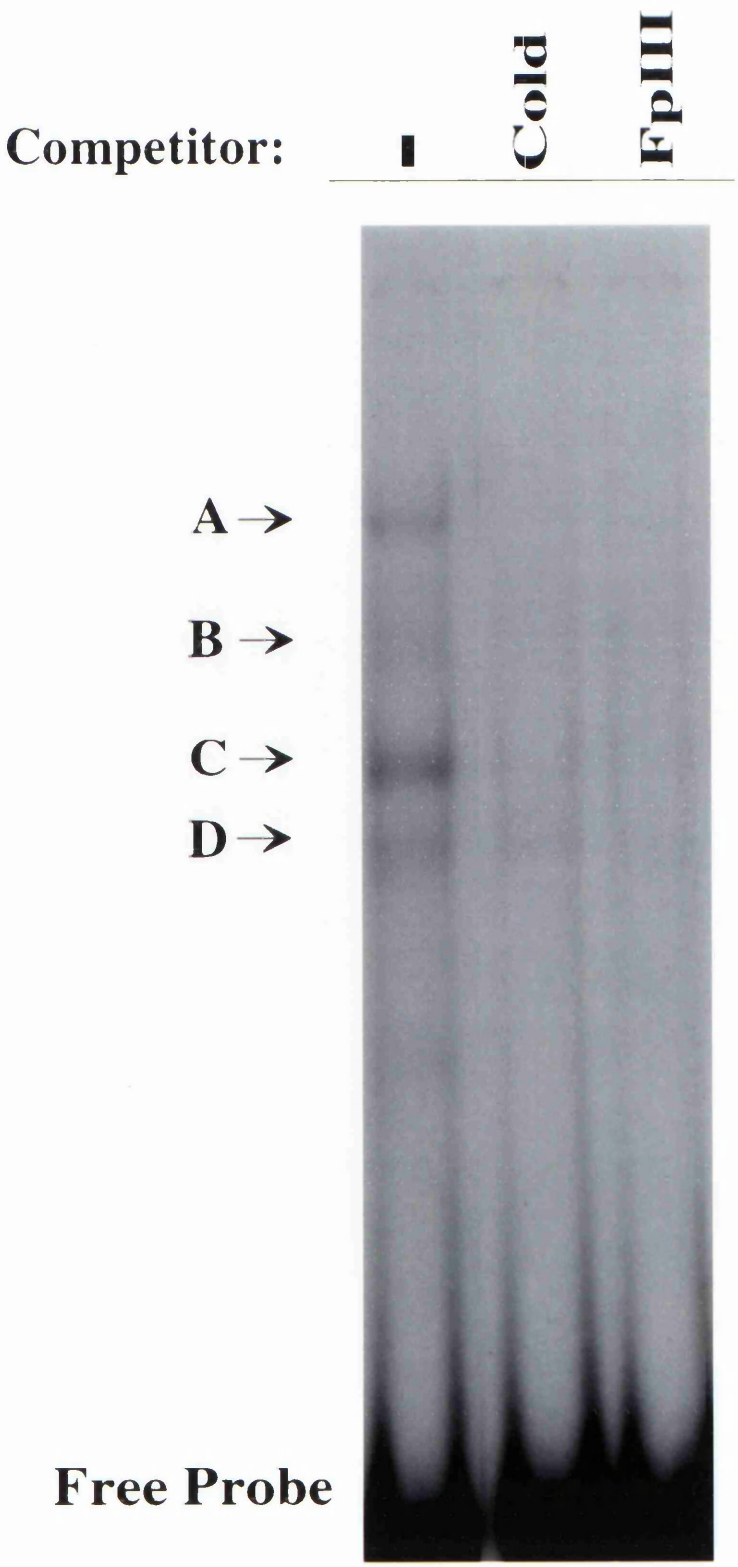

Probe: 3' FpIII
Extract: 293
Competitor: - Cold FpIII
A →
B →
C →
D →
Free Probe

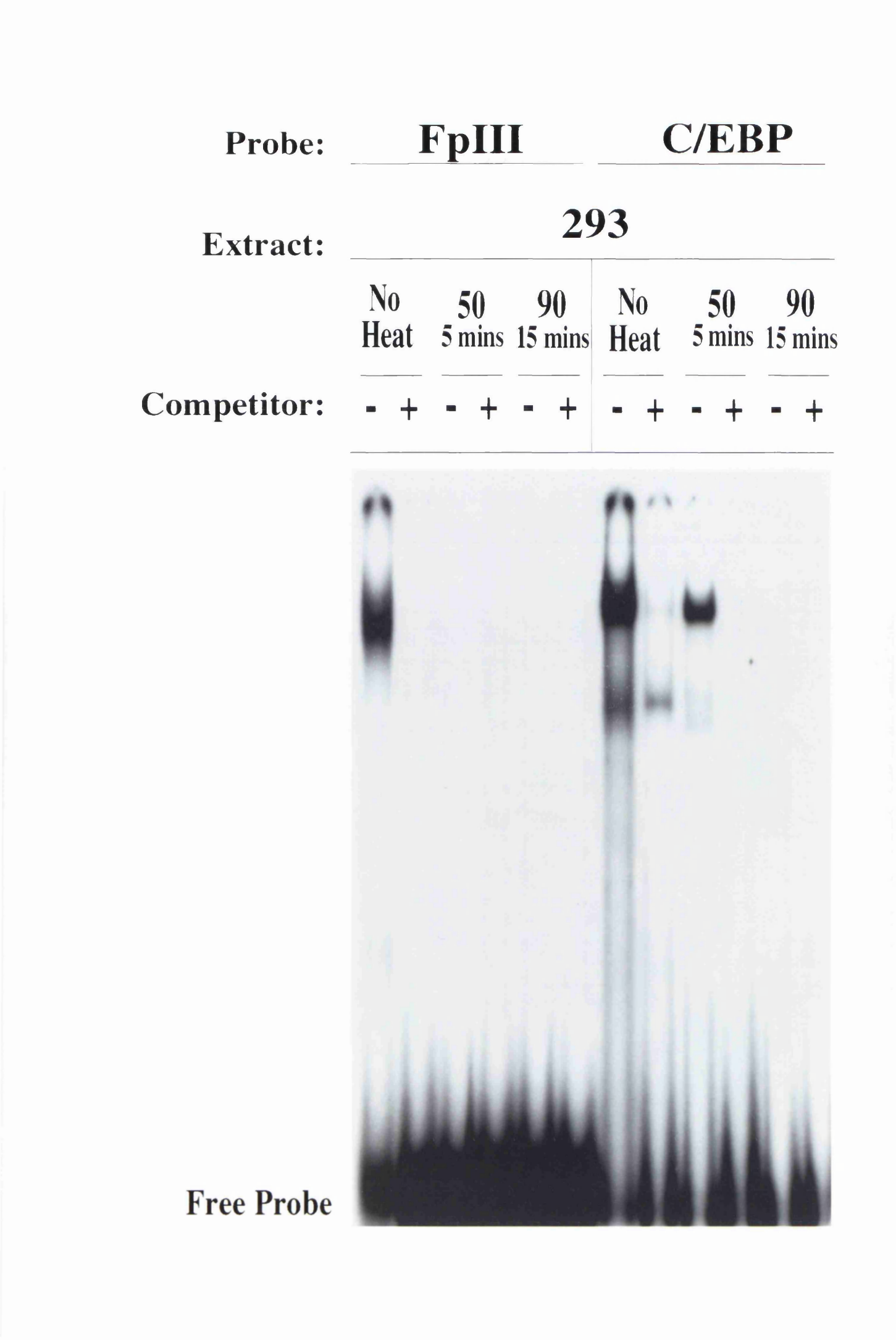
Figure 6.11: EMSA showing the heat stability of the EvBP1 and C/EBP protein complexes.

Nuclear extracts were prepared from 293 cells as described in Materials and Methods. A double stranded oligonucleotide corresponding to a C/EBP consensus binding site and the FpIII region were ^{32}P end-labelled. Two aliquots of 293 nuclear extracts were heated at 50°Cc for five minutes or 90°C for 15 minutes, prior to setting up the binding reactions. Binding reactions were set up using 500pg of labelled oligonucleotide and equal amounts of the appropriate protein nuclear extract (approximately 10μg per reaction). Reactions were run on non-denaturing PAGE gels, dried down and exposed to X-ray film overnight (Materials and methods).

The labelled oligonucleotide used was either FpIII or C/EBP and is shown above the lanes. Unlabelled competitor oligonucleotides are shown above each lane: ▮ - indicates no competitor, **Cold** indicates unlabelled C/EBP region. All competitors are added at 200 fold excess (100ng). Unbound labelled C/EBP oligonucleotide is observed at the bottom of the gel.

Oligonucleotide sequences:

C/EBP	gatcggatcccaattgggcaatcaggggatcc
FpIII	agagagatttccctggggaaacaaatcctgcctggagaaatttccccatt
5' FpIII	agagagatttccctggggaaacaaatcctgcc
3' FpIII	cctggggaaacaaatcctgcctggagaaatttccccatt


Probe:
FpIII
C/EBP
Extract:
293
No Heat
50 5 mins
90 15 mins
No Heat
50 5 mins
90 15 mins
Competitor:
- + - + - + - + - + - +
Free Probe

would be predicted if the protein complex was important in basal promoter activity. Figure 6.12 shows the presence of the EvBP1 complex in a variety of cell types including human kidney; 293 and A704 cells, murine fibroblasts; L929 cells, bovine palate keratinocytes, PalK cells, human endometrial carcinoma cells; HEC-1B and murine myeloid; DA-3 cells. In addition, murine kidney cultures (isolated from transgenic mice harbouring a temperature sensitive T-antigen, Chapter 7) also contain the EvBP1 complex (figure 6.13).

6.7 EvBP1 complex binds a 14bp inverted repeat sequence

To identify the DNA recognition sequence for the EvBP1 protein, 14bp stretches of the 5' FpIII region were synthesised and tested in EMSA to evaluate their ability to compete for the EvBP1 complex. The results of this competition are shown in figure 6.14. The -314 and -296 oligonucleotides did not compete for EvBP1 whereas -305 and IR (inverted repeat region) oligonucleotides competed for the EvBP1 complex with the same efficiency as cold FpIII oligonucleotide. Partial competition is also observed with the 3' FpIII region.

The inverted repeat region, tttccctggggaaa, is also able to bind the EvBP1 complex if it is ^{32}P-end labelled (data not shown). This 14bp sequence contains a putative ets core, ggaa, a rel half site, gggaaa, homology to an inverted AP2 site, and tggggaaac shows some homology to the C/EBP consensus t(t/g)nng(c/t)aa(t/g). However, as already described consensus NF-κB, PEA3, E74, PU1, and AP2 binding site oligonucleotides failed to compete for the EvBP1 complex, figure 6.8. This analysis has eliminated a number of candidate factors and therefore the component(s) of the EvBP1 complex are unknown.

Consistent with the importance of the 14bp EvBP1 binding sequence in basal promoter expression, it is also interesting to note that the inverted repeat region is

Figure 6.12: EMSA showing EvBP1 binding to the FpIII region with nuclear extracts from differing cell lines.

Nuclear extracts were prepared from 293, L929, HEC-1B, PalK, A704 and DA-3 cells as described in Materials and Methods. A double stranded oligonucleotide corresponding to FpIII region was ^{32}P end-labelled. Binding reactions were set up using 500pg of labelled oligonucleotide and equal amounts of protein nuclear extract (approximately 10μg per reaction). Reactions were run on non-denaturing PAGE gels, dried down and exposed to X-ray film overnight (Materials and methods).

The labelled oligonucleotide used in all reactions is FpIII. Unlabelled competitor oligonucleotides are shown above each lane: ▮ - indicates no competitor, **Cold** indicates unlabelled FpIII region. All competitors are added at 200 fold excess (100ng). The appropriate nuclear extract used in each reaction is indicated above the lane. Unbound labelled FpIII oligonucleotide is observed at the bottom of the gel.

Oligonucleotide sequences:

FpIII	agagagatttccctggggaaacaaatcctgcctggagaaatttccccatt
5' FpIII	agagagatttccctggggaaacaaatcctgcc
3' FpIII	cctggggaaacaaatcctgcctggagaaatttccccatt

Cell Lines:

293	Human kidney epithelial cell line
A704	Human renal cell carcinoma cell line
HEC-1B	Human endometrial carcinoma cell line
PalK	Bovine epithelial keratinocytes
DA-3	Murine myeloid cell line
L929	Murine fibroblasts

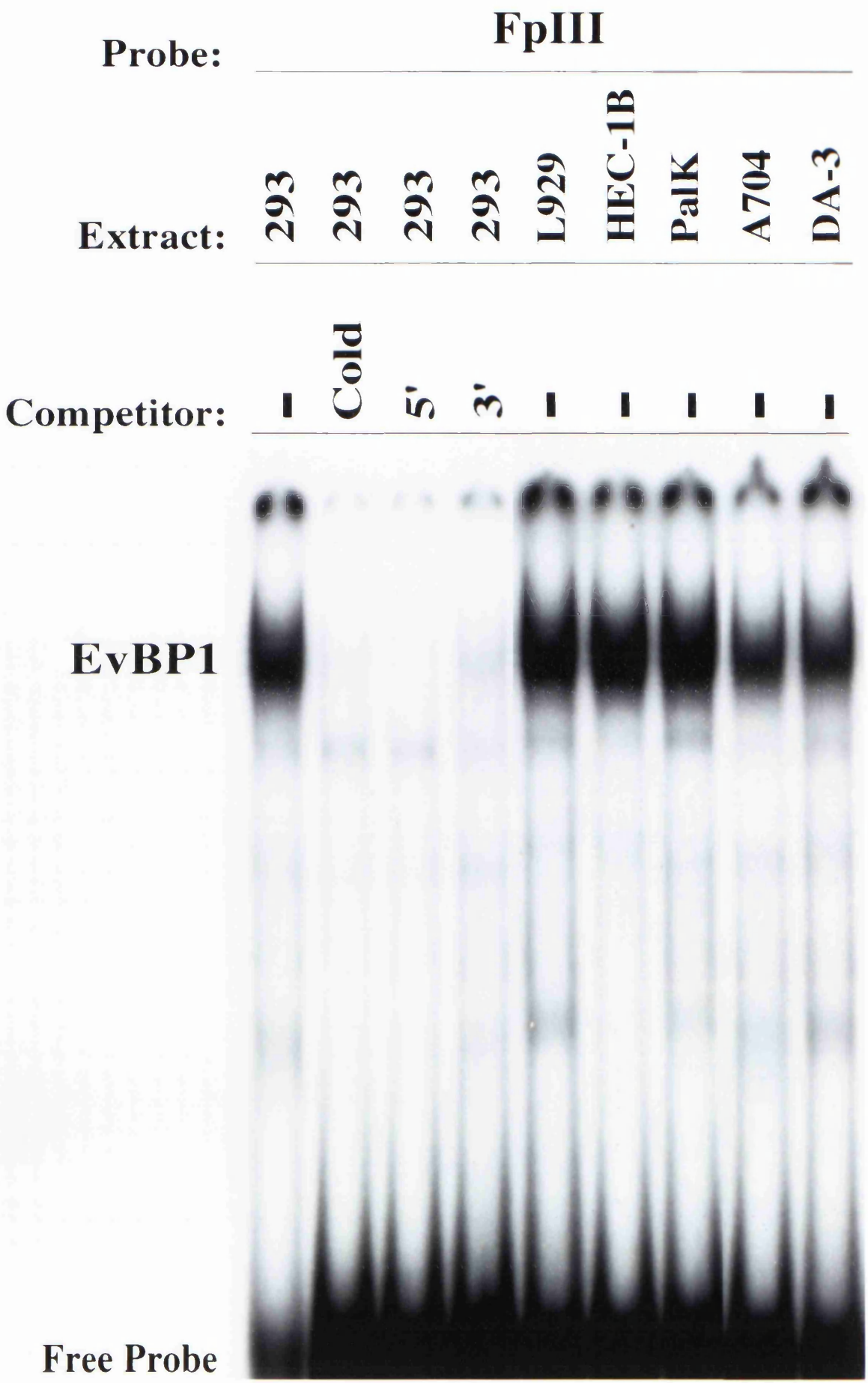

Probe:
FpIII
Extract:
293
293
293
293
L929
HEC-1B
PalK
A704
DA-3
Competitor:
-
Cold
5'
3'
-
-
-
-
-
EvBP1
Free Probe

Figure 6.13: EMSA showing binding of EvBP1 in kidney extracts

Nuclear extracts were prepared from 293, primary adult kidney (PAK), primary embryonic kidney (PEK), adult kidney cultures (AK), embryonic kidney cultures (EK) cells as described in Materials and Methods.

PAK and PEK were kidney cultures harvested after four days in culture and were unpassaged. AK and EK were kidney cultures harvested after two weeks in culture and after several passages. All kidney cultures were from transgenic mice harbouring a temperature sensitive T antigen, previously described in Chapter 2.

A double stranded oligonucleotide corresponding to FpIII region was ^{32}P end-labelled. Binding reactions were set up using 500pg of labelled oligonucleotide and equal amounts of the appropriate protein nuclear extract (approximately 10μg per reaction). Reactions were run on non-denaturing PAGE gels, dried down and exposed to X-ray film overnight (Materials and methods).

The labelled oligonucleotide used in all reactions is FpIII. Unlabelled competitor oligonucleotides are shown above each lane: ▮ - indicates no competitor, **Cold** indicates unlabelled FpIII region. All competitors are added at 200 fold excess (100ng). The nuclear extract used in each reaction is shown above each lane. Unbound labelled FpIII oligonucleotide is observed at the bottom of the gel.

Oligonucleotide sequences:

FpIII	agagagatttccctggggaaacaaatcctgcctggagaaatttccccatt
5' FpIII	agagagatttccctggggaaacaaatcctgcc
3' FpIII	cctggggaaacaaatcctgcctggagaaatttccccatt

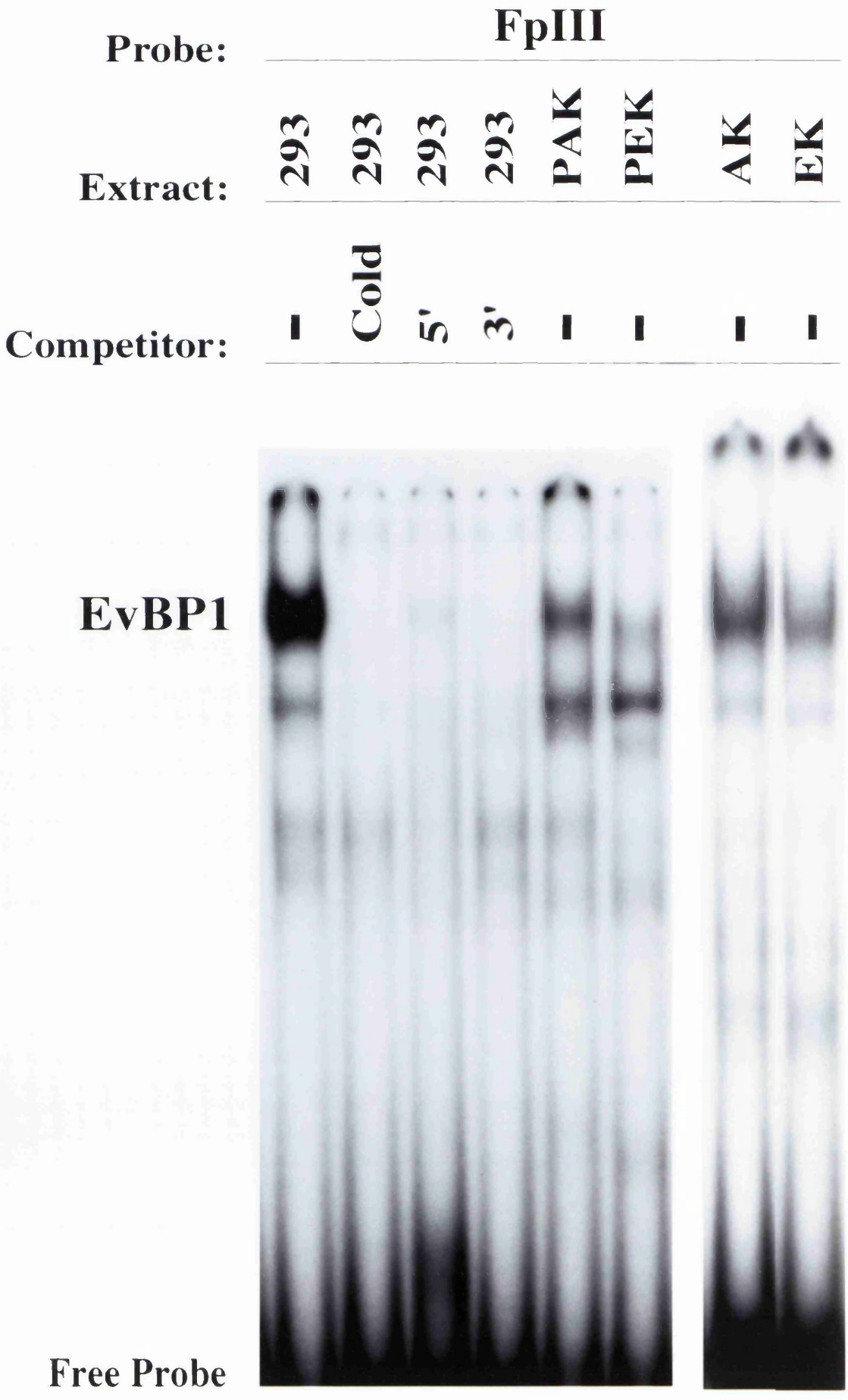

Probe:
FpIII
Extract:
293
293
293
293
PAK
PEK
AK
EK
Competitor:
-
Cold
5'
3'
-
-
-
-
EvBP1
Free Probe

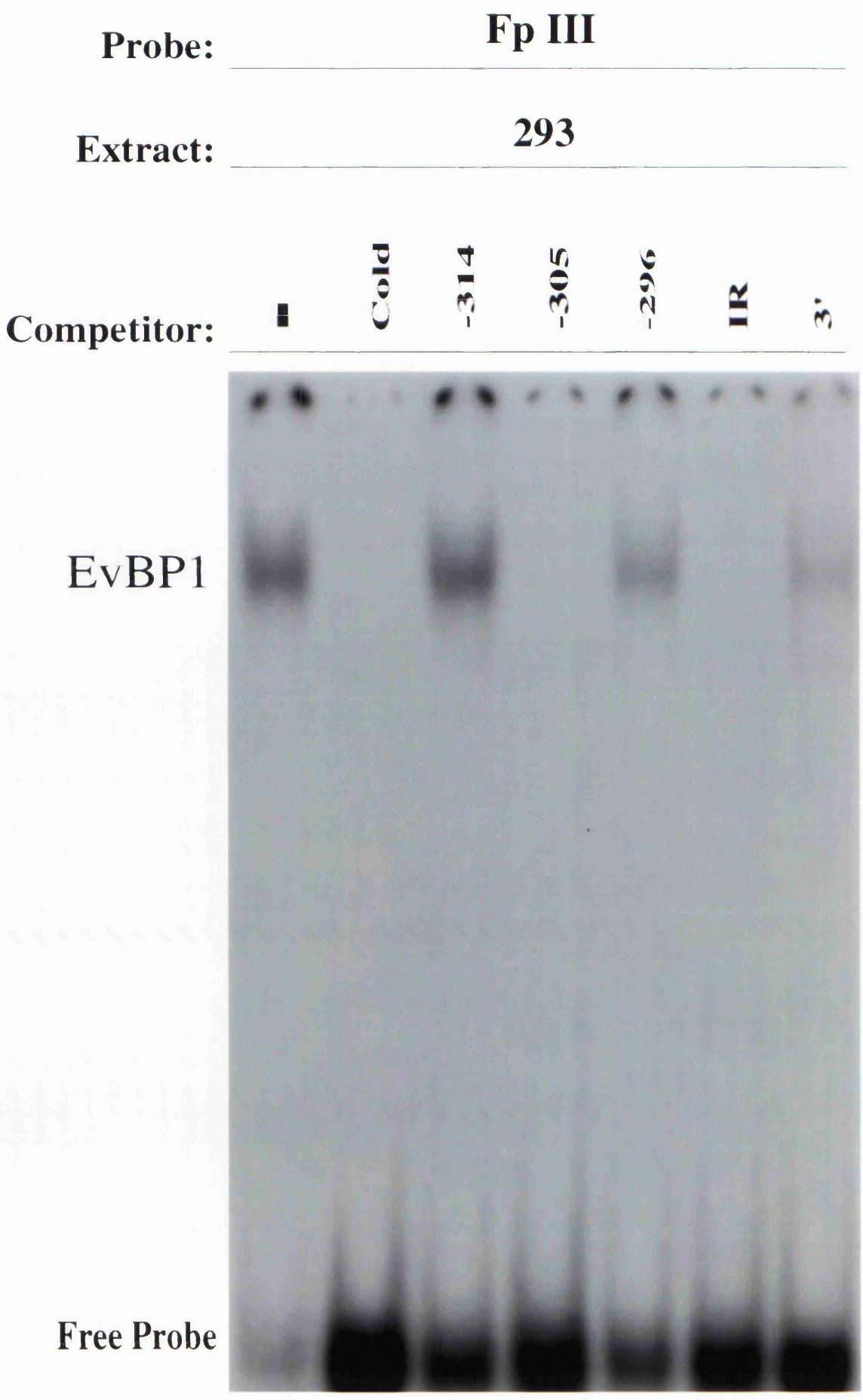
Figure 6.14: EMSA showing competition of EvBP1 with small 14bp stretches of the 5' FpIII region

Nuclear extracts were prepared from 293 cells as described in Materials and Methods. A double stranded oligonucleotide corresponding to the FpIII region was ^{32}P end-labelled. Binding reactions were set up using 500pg of labelled oligonucleotide and equal amounts of protein nuclear extract (approximately 10μg per reaction). Reactions were run on non-denaturing PAGE gels, dried down and exposed to X-ray film overnight (Materials and methods).

The labelled oligonucleotide used in all reactions was FpIII as shown. Unlabelled competitor oligonucleotides are shown above each lane. **-** indicates no competitor, **Cold** - indicates unlabelled FpIII region. All competitors are added at 200 fold excess (100ng). 293 nuclear extract was used in all reactions Unbound labelled FpIII oligonucleotide is observed at the bottom of the gel.

Oligonucleotide sequences:

FpIII	agagagatttccctggggaaacaaatcctgcctggagaaatttccccatt
3' FpIII	cctggggaaacaaatcctgcctggagaaatttccccatt
-296	gatcaaacaaatcctgcc
-305	gatctccctggggaaaca
-314	gatcagagagatttccct
IR	gatctttccctggggaaa


Probe:
Fp III
Extract:
293
Competitor:
-
Cold
-314
-305
-296
IR
3'
EvBP1
Free Probe

absolutely conserved in the human sequence (a gift from J.N. Ihle, unpublished data) (figure 6.15).

6.8 U.V. Cross-linking of the EvBP1 Complex

To identify the components of EvBP1 the retarded complex was U.V. cross-linked (Material and methods). The FpIII region oligonucleotide was re-synthesised replacing the adenine bases with BrdU, the oligonucleotide ^{32}P-end-labelled and EMSA performed confirming that the oligonucleotide is still able to bind EvBP1 (data not shown). This analysis reveals a U.V. light dependent specific band of approximately 97 kDa in size (figure 6.16). A very faint band is also observed with an approximate size of 200 kD which probably reflects a dimer of the lower band. The probability of cross-linking the dimerised form is much reduced and hence its fainter intensity. A protein/oligonucleotide complex of approximately 97kD in size is too large to be C/EBP (42 kD), AP2 (50 kD), AP3 (48/57 kD) or ets (54/68 kD) transcription factors and is further evidence that these factors are unlikely to be part of the EvBP1 complex.

6.9 Binding activity of the FpIV region

The second substantial drop in promoter activity corresponds with FpIV. Figure 6.17 is a schematic representation of the FpIV region showing the -277 to -208 protected region and two EMSA probes used in this analysis. Figure 6.18 shows EMSA analysis using the -277 probe, which covers most of FpIV. Six complexes, A to F, are observed in EMSA using nuclear extracts from 293 cells. All six complexes competed with cold -277 probe. Data base analysis of the sequence in this region identified a putative CP1 (CCAAT box protein 1) site which is boxed (figure 6.17),

Figure 6.15: Comparison of human and murine sequence over the FpIII and FpIV region

Murine DNA sequence is shown on top of the human.
The DNAse1 footprints identified, FpIII and FpIV, are highlighted in either red or blue respectively.
The sequence is numbered relative to the first major site of transcriptional initiation, +1.

The symbol | represents conserved nucleotides between murine and human.

A putative CCAAT box within this region is also highlighted in green.

Comparison of Murine and Human Sequence

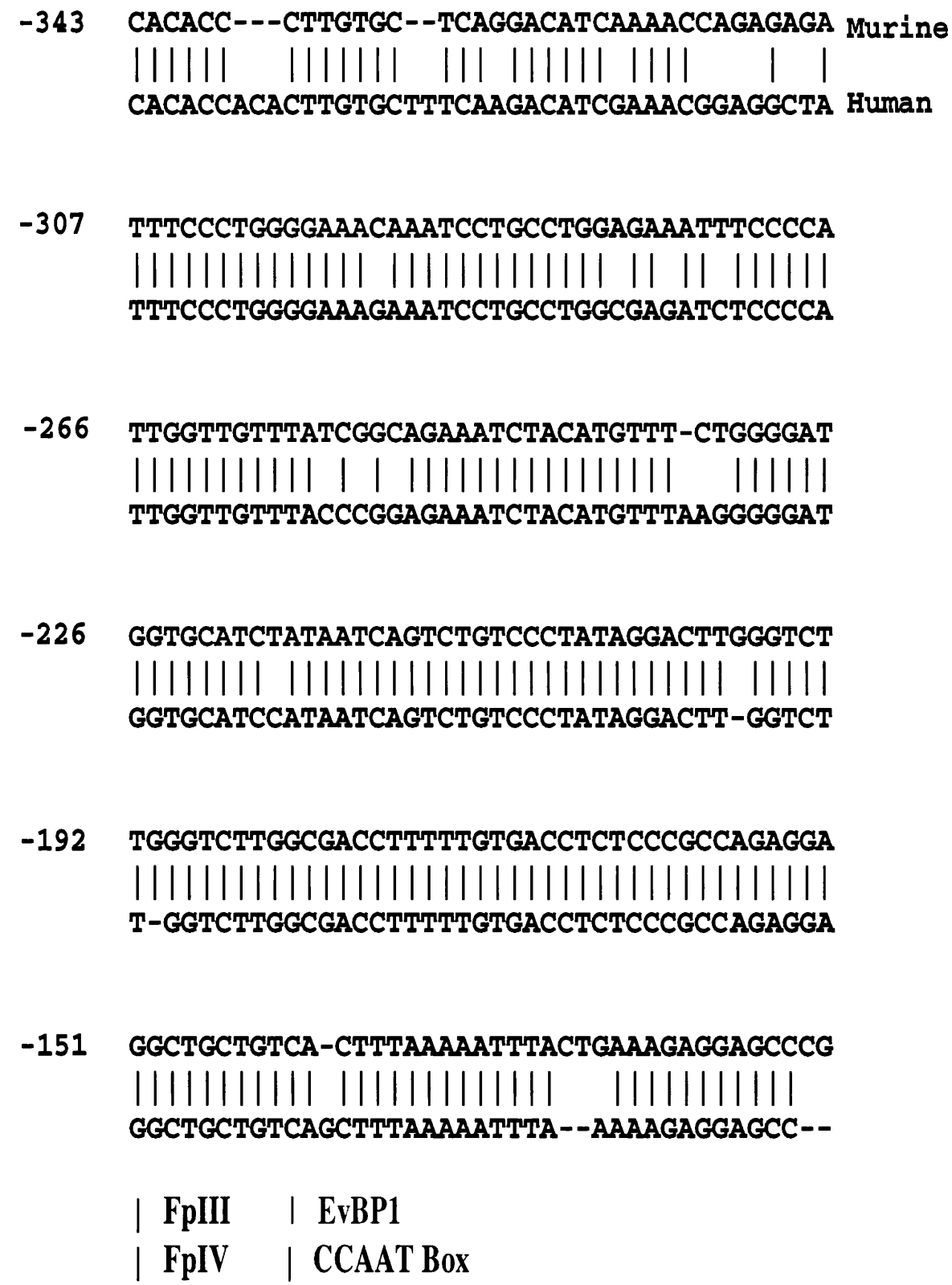


Figure 6.16: U.V. crosslinking of EvBP1

Gel slices containing DNA/protein complexes either treated or untreated with U.V. light. The gel pieces were placed directly into the wells of a 7.5% SDS-polyacrylamide gel and electrophorised. The gel was then dried, autoradiographed and typically exposed for several days.

An approximately 97kD protein is observed specifically in the U.V. treated lane. A fainter approximately 200kD band is also observed at the top of the gel probably representing the dimerised form of the protein.

The 200, 97.4 and 69 kD bands from ^{14}C-labelled molecular weight rainbow markers are also shown.

U.V. Crosslinking of EvBP1

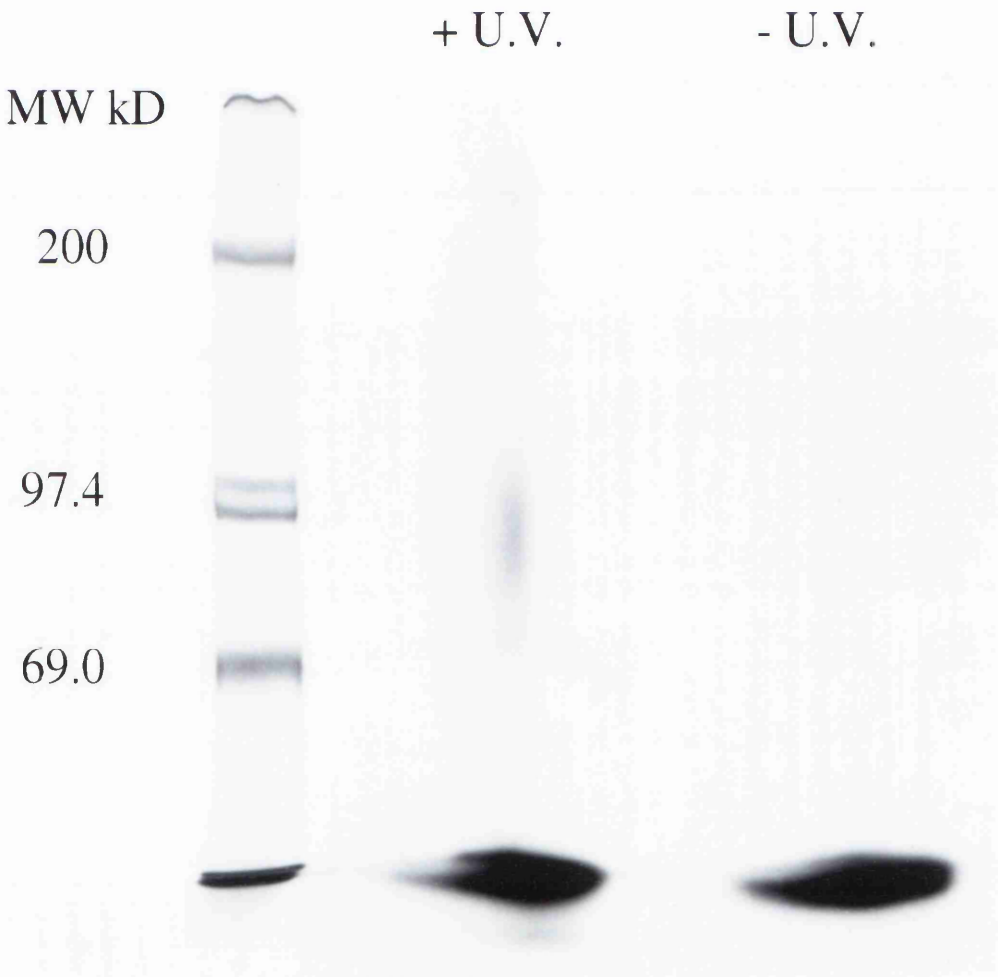


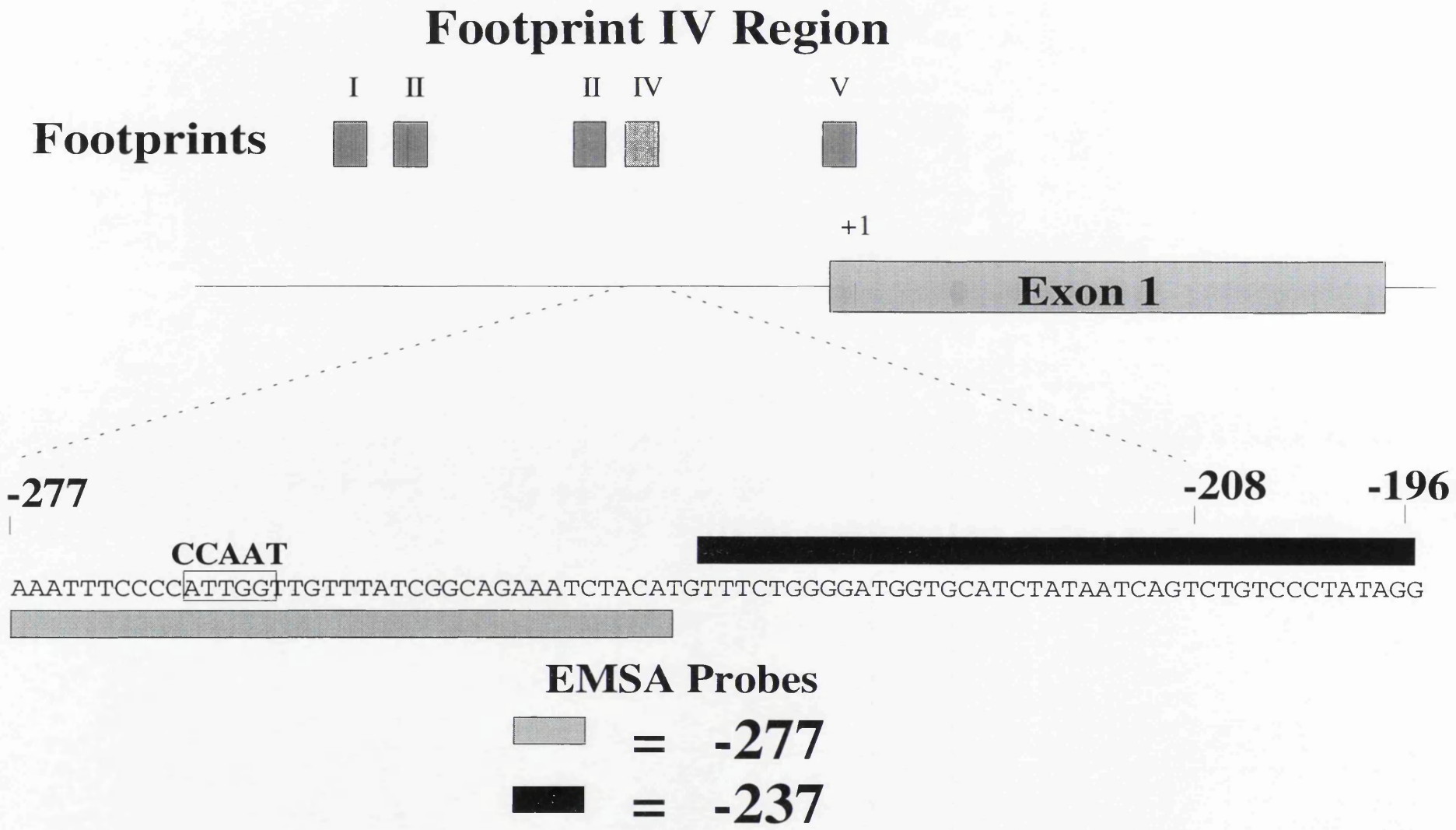
Figure 6.17: Schematic representation of the FpIV region

The sequence -277 to -206 represents the FpIV region. The probes used in the following EMSA experiments are indicated by shaded boxes.

Footprints I, II and V have previously been identified by (Bartholomew, unpublished data).

EMSA Probes

```
-277        aaatttccccattggttgtttatcggcagaaatctacat
-237        tttctggggatggtgcatctataatcagtctgtccctatagg
```


Footprint IV Region
I
II
II
IV
V
Footprints
+1
Exon 1
-277
-208
-196
CCAAT
AAATTTCCCCATTGGTTGTTTATCGGCAGAAATCTACATGTTTCTGGGGATGGTGCATCTATAATCAGTCTGTCCCTATAGG
EMSA Probes
= -277
= -237

the CCAAT site is in the reverse orientation. Again competition experiments were carried out with candidate oligonucleotides to identify the proteins, the particular consensus binding site oligonucleotide cold competitor used being shown above the appropriate lane (figure 6.18). No competition was observed with consensus binding site oligonucleotides from E74, PEA3 or SP1. However, oligonucleotides NF-1, C/EBP and Eα all compete for complex A and C, with NF-1 also competing complex B and E. All three consensus binding site oligonucleotides may bind CCAAT binding proteins. The CP1 protein shows varied specificity for different CCAAT boxes suggesting that the region immediately surrounding the CCAAT box is also important in CP1 specificity. The Eα oligonucleotide has been shown to bind the CP1 protein. This correlation is consistent with the presence of a putative CP1 site within this region but further work is required to show that it is definitely able to bind. Interestingly, deletion between -284 to -254 transfection studies which removes this putative CP1 site results in a significant drop in basal promoter activity in 293 cells (Chapter 5.3). The identity and function of complexes B, D, E, and F are unknown.

Figure 6.19 shows EMSA analysis using the -237 probe, which covers most of the remainder of FpIV (figure 6.17). Three complexes, A, B and C, are observed in EMSA using nuclear extract from 293 cells. All three complexes are competed with cold -237 probe. No competition is observed with consensus binding site oligonucleotides from E74, PEA3, NF-1 or NF-κB, SP1, C/EBP and Eα only partially compete for complexes A, B and C. There is no obvious known transcription factor binding sites within this region except for a putative TATA box. Since the promoter has multiple initiation sites it is plausible that this TATA box might also be used for TFIID binding and subsequent pre-initiation complex formation.

Figure 6.18: EMSA showing protein binding to -277 to -239 region

Nuclear extracts were prepared from 293 cells as described in Materials and Methods. A double stranded oligonucleotide corresponding to the -277 region was ^{32}P end-labelled. Binding reactions were set up using 500pg of labelled oligonucleotide and equal amounts of protein extract (approximately 10μg per reaction). Reactions were run on non-denaturing PAGE gels, dried down and exposed to X-ray film overnight (Materials and methods).

The labelled oligonucleotides used were FpIII and -277 as shown. Unlabelled competitor oligonucleotides are shown above each lane: ▮ - indicates no competitor, **Cold** - indicates unlabelled -277 region. All competitors are added at 200 fold excess (100ng). 293 nuclear extract was used in all reactions. Unbound labelled -277 oligonucleotide is observed at the bottom of the gel.

Specific complexes are labelled A, B, C, D, E and F are discussed in the text.

Oligonucleotide sequences:

-277	aaatttccccattggttgtttatcggcagaaatctacat
C/EBP	gatcggatcccaattgggcaatcaggggatcc
NF-1	cttttggattgaagccaatatgaga
SP1	tgggaatcctaactgggcggagttatgctggtggtg
GATA-1	ccgggcaactgataaggattccctg
PU1	gatccataacctctgaaagaggaacttggttaggt
PEA3	gatcctcgagcaggaagttcga
E74	gatctctagctgaataaccggaagtaactcatccta
H4TF-1	tttagatttcccctcccccaccggggcggga
Eα	gtgtgaaacatttttctgattgtttaaaagttgagtgct

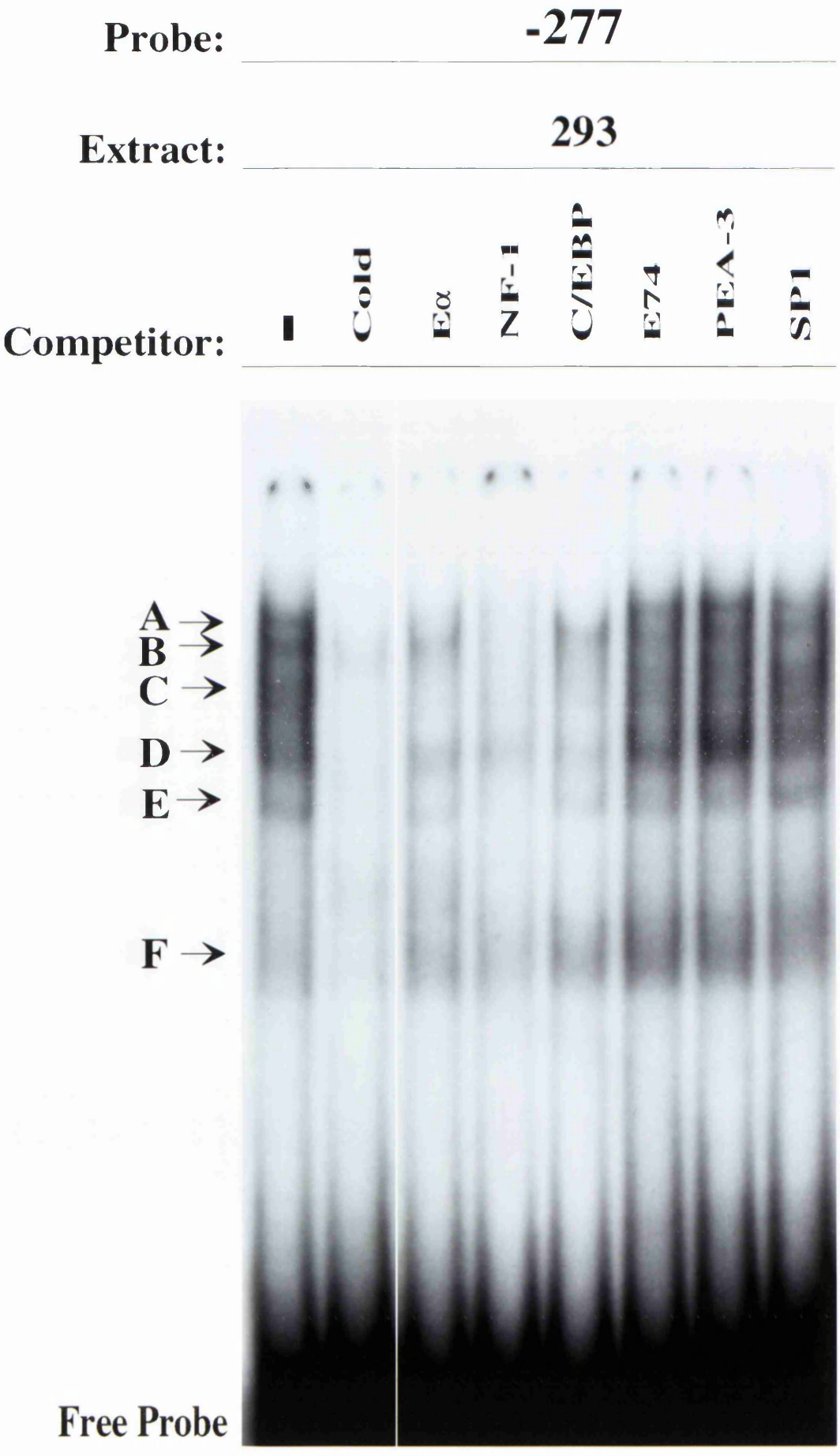

Probe:
-277
Extract:
293
Competitor:
-
Cold
Eα
NF-1
C/EBP
E74
PEA-3
SP1
A
B
C
D
E
F
Free Probe

Figure 6.19: EMSA showing protein binding to -237 to -196.

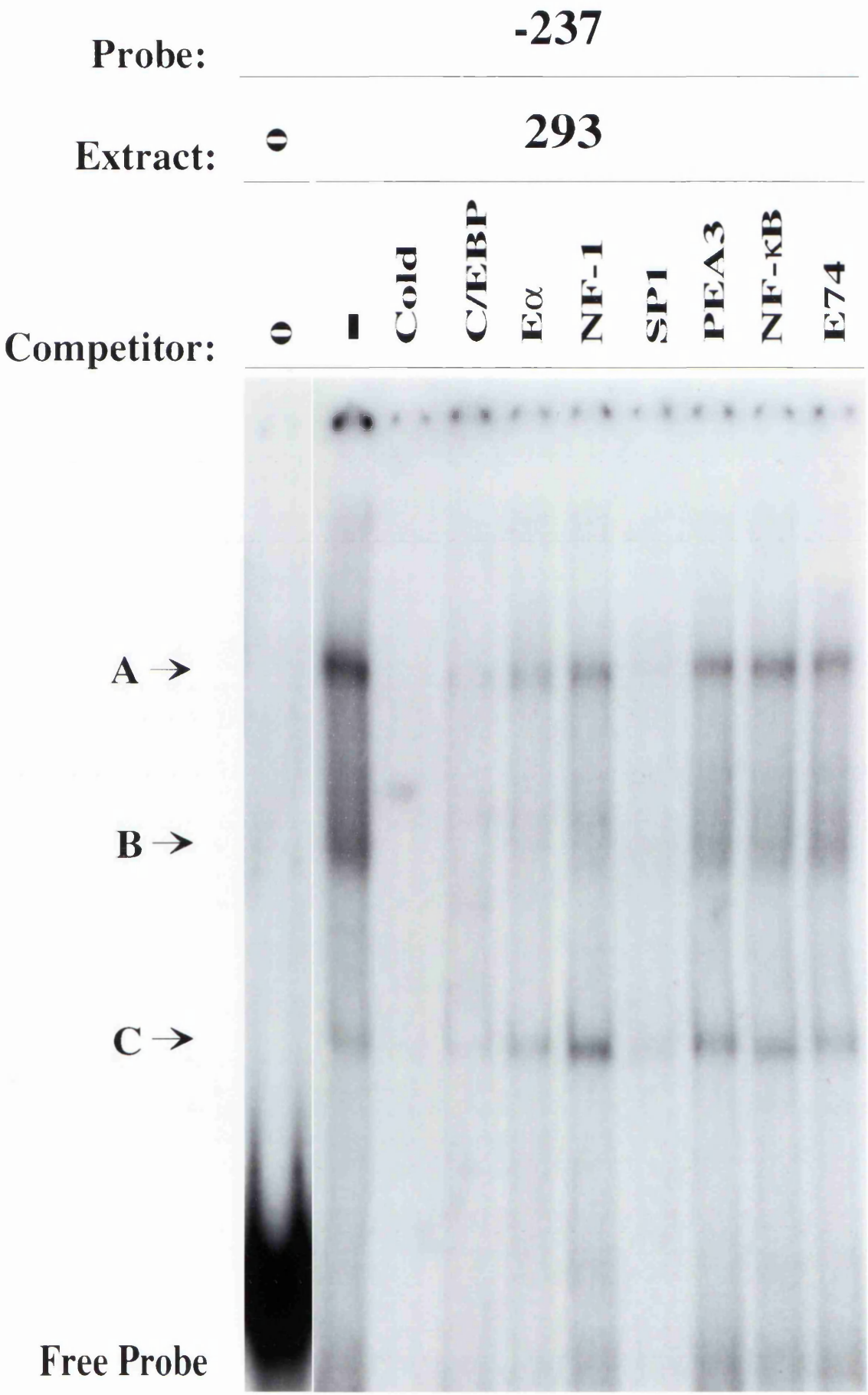
Nuclear extracts were prepared from 293 cells as described in Materials and Methods. A double stranded oligonucleotide corresponding to the -237 region was ^{32}P end-labelled. Binding reactions were set up using 500pg of labelled oligonucleotide and equal amounts of protein nuclear extract (approximately 10μg per reaction). Reactions were run on non-denaturing PAGE gels, dried down and exposed to X-ray film overnight (Materials and methods).

The labelled oligonucleotides used were FpIII and -237 as shown. Unlabelled competitor oligonucleotides are shown above each lane: ▮ - indicates no competitor, **Cold** - indicates unlabelled -237 region. All competitors are added at 200 fold excess (100ng). 293 nuclear extract was used in all reactions. Unbound labelled -237 oligonucleotide is observed at the bottom of the gel.

Specific compexes labelled A, B and C are discussed in the text.

Oligonucleotide sequences:

-237	tttctggggatggtgcatctataatcagtctgtccctatagg
C/EBP	gatcggatcccaattgggcaatcaggggatcc
NF-1	cttttggattgaagccaatatgaga
SP1	tgggaatcctaactgggcggagttatgctggtggtg
NF-κB	agttgaggggactttcccagg
PEA3	gatcctcgagcaggaagttcga
E74	gatctctagctgaataaccggaagtaactcatccta
Eα	gtgtgaaacatttttctgattgtttaaaagttgagtgct


Probe:
-237
Extract:
0
293
Competitor:
0
-
Cold
C/EBP
Eα
NF-1
SP1
PEA3
NF-κB
E74
A →
B →
C →
Free Probe

6.10 Conclusions

The results have identified two footprinted regions, FpIII and FpIV, 5' of the major sites of transcriptional initiation. Sequential deletion of these footprinted regions results in a substantial drop in relative promoter activity suggesting that the proteins binding to these regions are important for basal promoter activity. EMSA analysis of FpIII has identified binding activity over a 14bp inverted repeat sequence, designated EvBP1. Interestingly, this recognition sequence is absolutely conserved in the human sequence. The components of the EvBP1 complex are still unknown but U.V. crosslinking has identified a monomeric protein/oligonucleotide complex of approximately 97kD.

EMSA analysis of the FpIV region identified multiple protein binding activities some of which are competed with C/EBP, Eα and NF-1. Interestingly, a putative CP1 binding site (CCAAT) is present in this region and the loss of this CCAAT site also correlates with a significant drop in basal promoter activity.

Chapter 7: Tissue Specific Expression of the *Evi-1* Gene

7.1 *Evi-1* expression in murine tissues

Previous work has shown that the *Evi-1* gene is expressed in a tissue specific manner (Chapter 1.17). To study tissue specificity of the *Evi-1* promoter we required a suitable cell line expressing high levels of the endogenous gene. The pattern of *Evi-1* expression was examined in murine tissues by Northern blot analysis of different adult tissues from CD1 mice using an 2kb SstI fragment Evi-1 cDNA ^{32}P-labelled probe (figure 7.1). The full length 5kb transcript and alternatively spliced 4.5kb transcript (Bordereaux *et al.*, 1990; Morishita *et al.*, 1990) are indicated with arrows. Highest levels of expression were observed in adult lung, ovaries and kidney with high levels also detected in embryonic kidney (17.5 and 15.5 days *p.c.*) as seen by others (Morishita *et al.*,1990; Perkins *et al.*, 1991). Therefore we examined murine kidney derived cell lines to identify cells suitable for studying the high level tissue specific expression of the murine *Evi-1* gene.

7.2 Evi-1 expression in kidney cell lines

In view of the high *Evi-1* expression levels seen in adult kidney we assessed the levels of *Evi-1* expression in kidney cell lines. The abundance of *Evi-1* mRNA in a panel of kidney cell lines was monitored by Northern blot analysis using a 2kb SstI fragment *Evi-1* cDNA probe (figure 7.2).

High levels of expression were observed in DA-3 cells (which ectopically express the gene due to a retroviral insertion) kidney tissue and HEC-1B. The human endometrial carcinoma cell line HEC-1B expresses high levels of Evi-1 as compared with parental HEC-1A cells (figure 7.2 B). However, the reason for the expression is

Figure 7.1: Northern blot showing *Evi-1* expression in different murine tissues.

Northern blot analysis using a 2kb SstI fragment Evi-1 cDNA ^{32}P-labelled probe.

The appropriate tissue from which RNA was isolated is given above each lane. The RNA from embryonic kidneys were isolated at 15.5 and 17.5 days *post coitus*.

5kb and 4.5kb *Evi-1* transcripts are labelled with arrows.

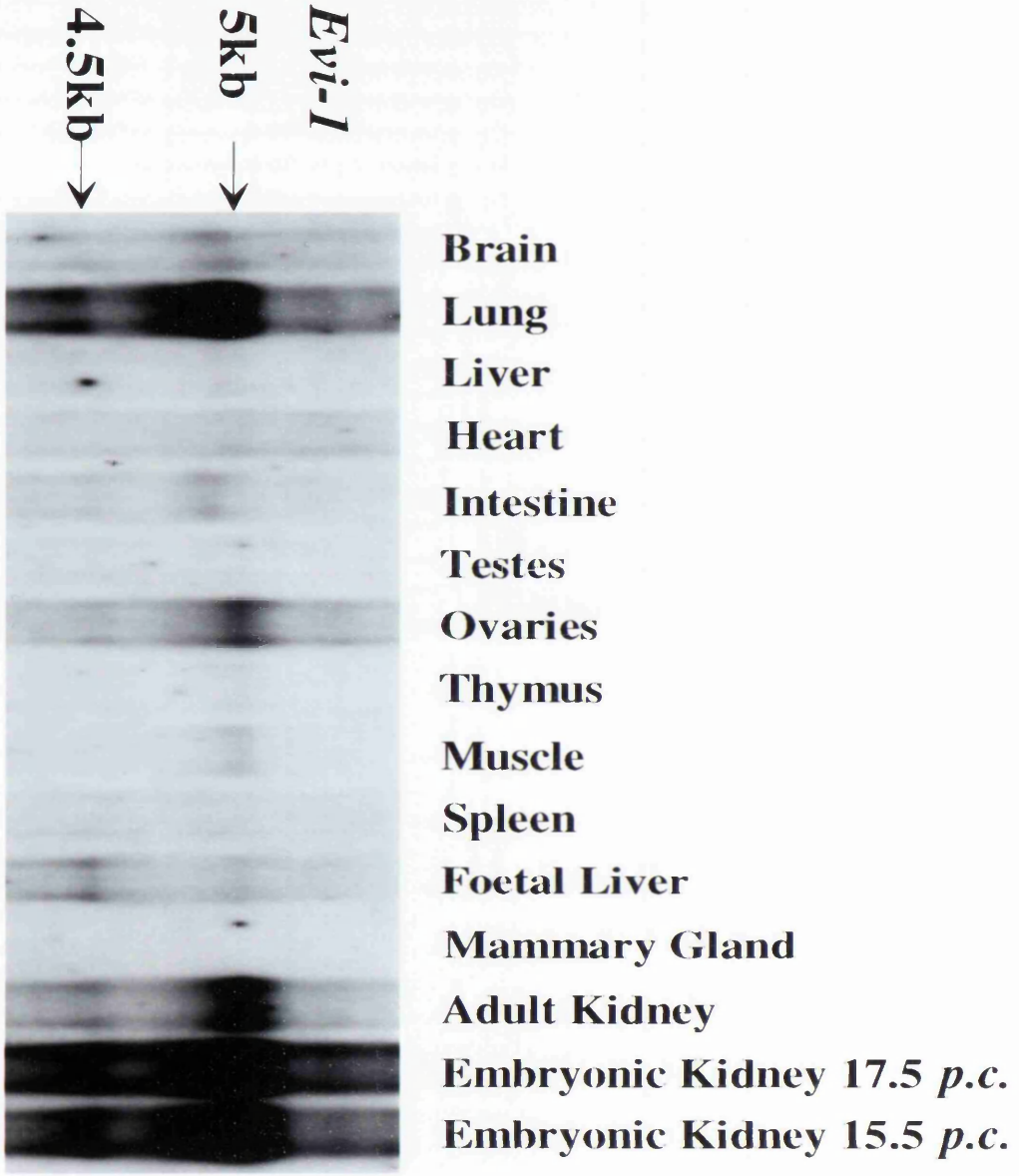

Evi-1
5kb
4.5kb
Brain
Lung
Liver
Heart
Intestine
Testes
Ovaries
Thymus
Muscle
Spleen
Foetal Liver
Mammary Gland
Adult Kidney
Embryonic Kidney 17.5 p.c.
Embryonic Kidney 15.5 p.c.

Figure 7.2: Northern blot showing *Evi-1* expression in different kidney and other cell lines.

Northern blot analysis using a 2kb SstI fragment Evi-1 cDNA ^{32}P-labelled probe.

The appropriate tissue or cell line from which RNA was isolated is given above each lane.

The cell lines used are listed below:
Kidney:
- RAG
- G402
- ACHN
- TCMK-1
- 293
- A704
- YCR
- RCC23

Fibroblast:
- STO
- L929
- NIH3T3

Haemopoietic:
- DA-3
- FDCP-1

Endometrial:
- HEC-1A
- HEC-1B

RNA loading was controlled by stripping the blots and re-probing with a cDNA β-actin 32P-labelled probe.

A
Kidney
Spleen
DA-3
NIH3T3
STO
L929
RAG
G402
ACHN
TCMK-1
293
Evi-1→
β-actin
B
Kidney
DA-3
FDCP-1
NIH3T3
HEC-1A
HEC-1B
A704
YCR
RCC23
Evi-1→
β-actin

unclear, but HEC-1B cells contains 96 chromosomes and have rearrangements within chromosome 3 (Morishita *et al.*, 1990). Unfortunately, all the kidney cell lines examined (RAG, G402, ACHN, TCMK-1, 293, A704, YCR and RCC23) showed no or relatively low levels of *Evi-1* expression similar to that observed in the fibroblastic cell lines (L929, STO and NIH3T3). Although expression levels in STO cells appear higher no consistent difference is observed between the fibroblastic cell lines. The relatively low levels of *Evi-1* expression in all the kidney cell lines examined suggests that none of the cell lines are appropriate for investigating *Evi-1* promoter activity.

7.3 Isolation of kidney cultures from transgenic mice harbouring a temperature sensitive Simian virus 40 large T-antigen gene.

The absence of a suitable high expressing cell line meant we had no cell line with optimal activity of the *Evi-1* promoter. Although no *Evi-1* expression is observed in kidney cell lines, expression is retained in 5 day primary kidney cultures established from 18.5*p.c.* foetal kidneys (Bartholomew & Clarke., 1994). However, the cultures are difficult to maintain and although adult kidney cultures may survive for several weeks at confluence if the medium is routinely changed the cells do not passage and attempts to do so results in massive cell death and tiny cell colonies which fail to continue to proliferate (data not shown). To circumvent this problem a new approach was undertaken in an attempt to generate *Evi-1* expressing primary cell cultures which would be more suitable for examining the *Evi-1* promoter activity.

In previous studies primary immortalised cell lines have been established by co-transfecting the cultures with immortalising viral proteins such as the SV40 T-antigen and a antibiotic resistance marker. A number of studies have suggested that immortalisation of various cell types with the SV40 large T-antigen can maintain cell type specific functions (Paul *et al.*, 1988., Williams *et al.*, 1988.). A temperature

sensitive mutant of the large T-antigen from the SV40 virus, designated tsA58, has also been used successfully to isolate stable cell lines which are temperature sensitive for expression of tsA58 T-antigen (Jat, P.S. & Sharp, P.A., 1989, Robinson *et al*, 1994, Piedagnel, *et al*, 1994). The mutant is stable at the permissive temperature of 33°C but becomes unstable and is degraded at the non-permissive temperature of 39°C.

Interestingly, a transgenic mouse harbouring the tsA58 SV40 large T-antigen (tsT) gene under the control of an interferon-inducible Class I antigen promoter has been developed. Recently the H-2KbtsA58 transgenic mouse has been proposed to be a new tool in rapid generation of novel cell lines (Noble *et al*, 1995). Conditionally immortal cell lines have been successfully established from these transgenic mice which maintain some cell-type specific responses. (Yanai *et al.*, 1991a, Yanai *et al.*, 1991b, Jat *et al.*, 1991).

In an attempt to establish our own kidney cell lines and to maintain primary cultures which express the *Evi-1* gene for extended periods of time the H-2KbtsA58 mouse was used to establish kidney cultures from both adult and foetal kidneys (day 18.5*p.c.*) (Taub *et al.*, 1979). Adult and foetal kidneys were dissected, incubated with collagenase/trypsin and put into culture, (see Materials and methods). Adult and foetal kidney cultures were initially established in either a general media, SLM, or epithelial selective, serum free or D-valine, media (Material and methods), which has been used successfully by other workers to establish primary cultures (Chung *et al.*, 1982). The cultures were observed to grow best in SLM media which was therefore routinely used in all the experiments (data not shown). Fibroblastic cells made up a component of the kidney cultures but never proved a problem with overgrowing the culture, which is similar to observations by other workers when establishing primary kidney cultures (Ebert *et al.*, 1990).

A range of Interferon-γ concentrations to induce SV40 T-antigen expression comparable to those used by others in similar studies (Jat *et al.*, 1991) was also added

to the culture media with both adult and foetal kidney cultures. However, the interferon-γ inducible promoter constitutively expressed the tsA58 T-antigen in our kidney cultures, negating the necessity for interferon γ supplementation.

Interestingly, both adult and foetal kidney cultures were observed to be particularly sensitive to trypsin treatment on the first couple of passages. However, collagenase treatment alone was found to be less harsh on the cells and therefore the protocol was modified accordingly for the first few passages.

Figure 7.3 shows the morphology of the isolated adult kidney cultures at 33°C at sub-confluence or 33°C and 39°C at either confluence or 7 days post-confluence. Adult kidney cultures at 33°C grow as typical epithelial islands at sub-confluence (figure 7.3 A), forming a monolayer of cobblestone appearance when confluent (figure 7.3 B), similar to observations by others (Ebert *et al.*, 1990). Even after 7 days post-confluence at 33°C the morphology of the cells remained unchanged. Upon switching the confluent cultures to the non-permissive temperature (39°C) the cells were still similar in appearance. However, after 7 days post-confluence at 39°C the cells were generally larger and elongated representing a more spindle appearance, the cells were also more difficult to focus. Similar observations were made with foetal kidneys (figure 7.4 A, B and C).

Interestingly both adult and foetal kidney cultures grown at 33°C appeared to survive for several weeks at confluence, so long as the medium was changed, which is similar to our observations with adult kidneys from 'normal' mice. However, unlike the normal primary kidney the tsT cells also survived being passaged. Adult and foetal kidney cultures grown for longer than seven days at 39°C were found to die out slowly. Interestingly, if cells grown for seven days at 39°C and then split and put back at 33°C or 39°C the cells at 33°C recovered, while the cells at 39°C died out, showing that the cells require the SV40 T-antigen to survive.

Figure 7.3: Morphology of adult kidney cultures at 33°C and 39°C.

A

Pasaged kidney cells were grown at 33°C and used to seed two T^{75} flasks. The flasks were returned to 33°C until the cultures were just sub-confluent and the cells were photograph. At this point one of the two flasks were switched to 39°C.

B

The flasks were then incubated at the permissive or non-permissive temperature until confluent, at which point they were photographed.

C

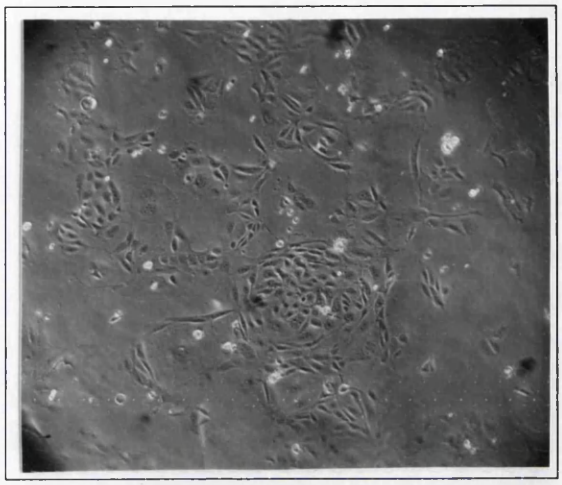
The flasks were then incubated a further seven days and photographed again.

Adult Kidney

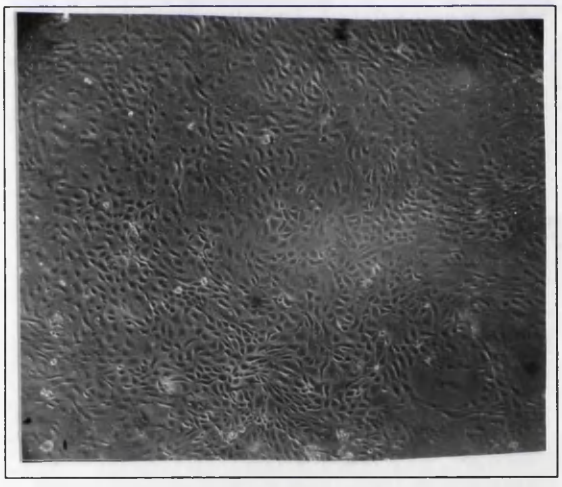
33°C

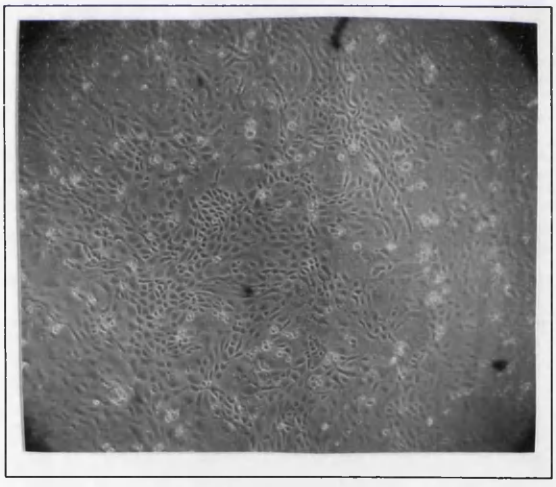
39°C

A

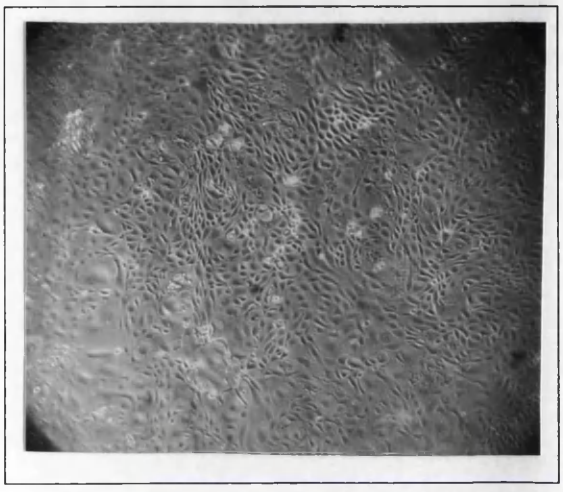


B





C



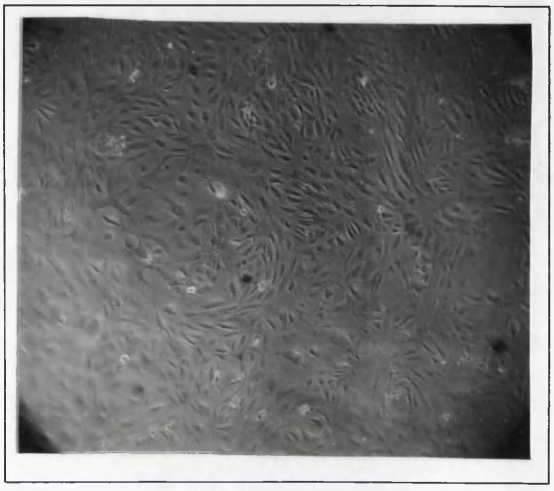


Figure 7.4: Morphology of embryonic kidney cultures at 33°C and 39°C.

A

Pasaged embryonic kidney cells were grown at 33°C and used to seed two T^{75} flasks. The flasks were returned to 33°C until the cultures were just sub-confluent and the cells were photograph. At this point one of the two flasks were switched to 39°C.

B

The flasks were then incubated at the permissive or non-permissive temperature until confluent, at which point they were photographed.

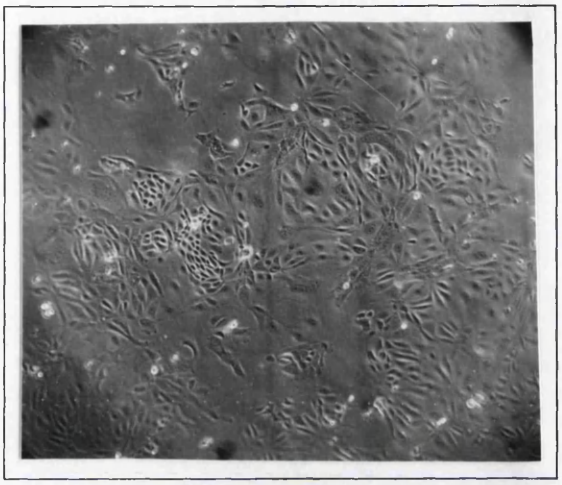
C

The flasks were then incubated a further seven days and photographed again.

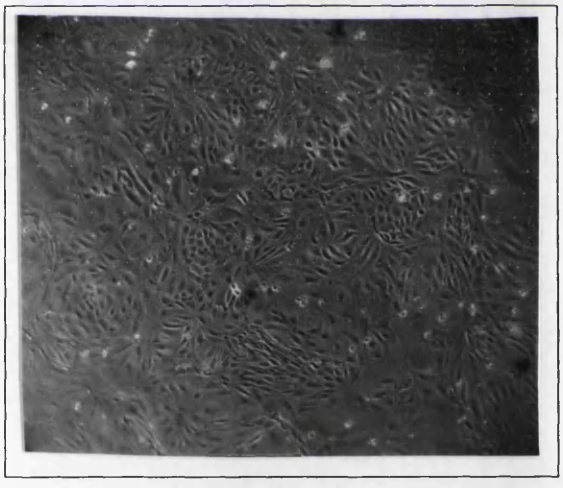
Embryonic Kidney

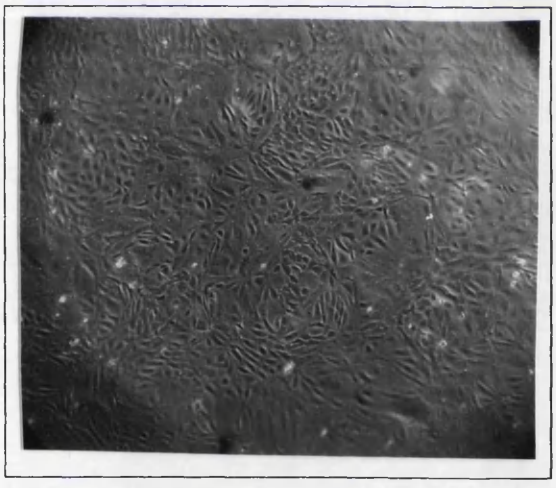
33°C 39°C

A

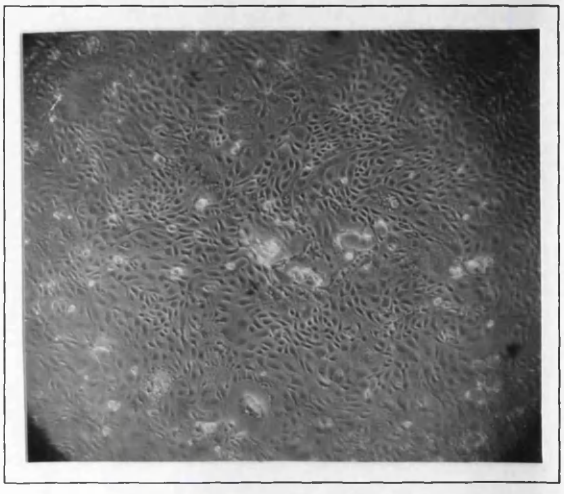


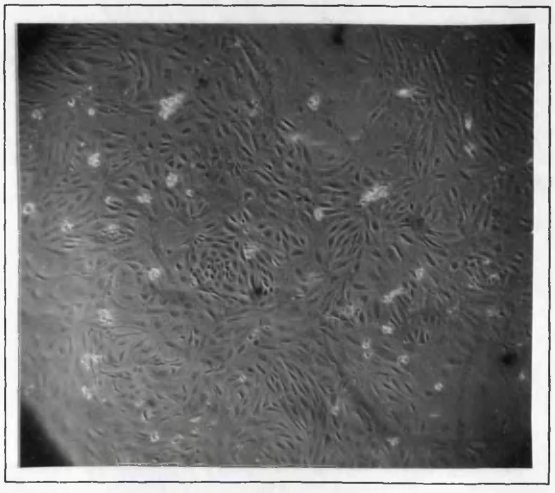
B





C





It has been observed by others that some conditionally immortal hepatocyte and kidney tubule cell lines are growth arrested at the non-permissive temperature of 39°C and undergo apoptosis. The authors suggest that an accumulation of p53 which binds to T-antigen at 33°C is released at 39°C resulting in apoptosis (Yanai & Obinata, 1994). We find that kidney cultures grown at 33°C which are then split and switched to 39°C grow to reach confluence, will survive for about a week and then begin to die out. However, the mechanism of cell death in these cultures has not been examined.

7.4 Heat Stability of the tsT at 33°C and 39°C

To confirm that tsT has the expected properties in our culture system the levels of protein were assessed in both foetal and kidney cultures. Adult or foetal kidney cultures were grown to semi-confluence and then incubated at either 33°C or 39°C for seven days. Protein extracts were prepared with RIPA buffer (Materials and methods) and equal protein concentrations used in immunoprecipitation reactions with an SV40 T-antigen monoclonal antibody (Sigma). Following SDS-PAGE electrophoresis the proteins were examined by Western blot analysis using the same SV40 T-antigen monoclonal antibody.

The results of this analysis are shown in figure 7.5. The 97.5kD tsT protein is clearly visible in embryonic kidney incubated at 33°C but not at 39°C (figure 7.5 A). Similar results are also observed with adult kidney cultures (figure 7.5 B). These results conclude that tsT is abundant in cells grown at the permissive temperature (33°C) but is degraded at the non-permissive temperature (39°C). This supports the observation that continual survival of tsT adult and foetal kidney cultures is due to the expression of the 97.5kD tsT-antigen present in these cells.

Figure 7.5: Heat stability of the 97.5kD T antigen in conditionally immortalised embryonic and adult kidney cultures.

Kidney cultures isolated from transgenic mice harbouring a temperature sensitve T antigen (Jat *et al*, 1989) were grown to just below confluence and then incubated at either 33°C or 39°C for seven days. Protein extracts were made using RIPA buffer and equal protein concentrations used in immuneprecipitation reactions. Sepharose A/Antibody/protein complexes were boiled and electrophorised on an SDS-PAGE gel (Material and methods). The proteins were then blotted onto ECL nitrocelluose membrane and the levels of T antigen determined by ECL detection according to the manufacturers protocol.

A

Expression of 97.5kD T antigen in embryonic cultures at either the permissive temperature of 33°C or non-permissive temperature of 39°C.

B

Expression of 97.5 kD T antigen in adult kidney cultures at either the permissive temperature of 33°C or non-permissive temperature of 39°C.

A

Embryonic Kidney

33 39

T antigen

B

Adult Kidney

33 39

T antigen

7.5 *Evi-1* expression in adult and day 18.5p.c. foetal kidney cultures

To evaluate the *Evi-1* expression levels in the isolated tsT adult kidney cultures a Northern blot was performed using an 2kb SstI fragment Evi-1 cDNA ^{32}P-labelled probe (figure 7.6). Established kidney cultures were incubated at 33^{o}C, the culture was split and then allowed to reach sub-confluence or confluence for 3 or 7 days, at both 33^{o}C and 39^{o}C, before RNA was isolated.

Relatively high levels of the full length 5kb *Evi-1* transcript were observed in 'normal' adult kidney, tsT kidney and tsT primary adult kidney cultures (a less intense 4.5kb transcript is typically seen in the cultures on longer exposure but in the following experiments only the full length 5kb transcript is shown). Passaged kidney cultures incubated at 33^{o}C showed relatively low levels of expression, *Evi-1* transcripts being detectable on longer exposure. The same expression pattern was also seen with passaged cultures incubated at 39^{o}C. Similar observations were made with day 18.5p.c. foetal kidney cultures (figure 7.7).

The results show that *Evi-1* expression is high when cells are first put into culture but is dramatically down regulated over a period of time. Therefore, we wished to establish when the down regulation occured. RNA was isolated from tsT primary kidney cultures, after one, two, three or four passages. To assess the abundance of the full length 5kb *Evi-1* transcript at each split, Northern blot analysis was performed using a 2kb SstI fragment cDNA ^{32}P-labelled probe (figure 2.8). The moderate level of full length *Evi-1* expression observed in primary kidney is maintained until the fourth passage at which point the level drops to undetectable. As indicated by the 7S loading control Passage 3 RNA was significantly underloaded, a faint *Evi-1* band could be seen on the original autoradiograph within this lane suggesting that there may be some *Evi-1* expression still present at this passage. In conclusion, the primary kidney cultures

Figure 7.6: Northern blot showing *Evi-1* expression in adult kidney cultures isolated from transgenic mice.

Northern blot analysis using a 2kb SstI fragment Evi-1 cDNA ^{32}P-labelled probe.

The appropriate tissue or kidney culture from which RNA was isolated is given above each lane and is explained below,

Adult kidney- Kidney RNA was isolated from adult kidneys.

All the kidneys and cultures below were isolated from transgenic mice harbouring a temperature sensitive T antigen.

'T' adult kidney- RNA was isolated from adult kidneys
Primary Kidney-kidney cultures established and RNA isolated at confluence

The following procedure was carried out at both 33^{o}C and 39^{o}C.

Sub-confluence-kidney cultures established over several passages were grown to just below confluence and then harvested for RNA.
3 days-kidney cultures established over several passages were grown to confluence, left for three days, and then harvested for RNA.
7 days-kidney cultures established over several passages were grown to confluence, left for seven days, and then harvested for RNA.

RNA loading was controlled by stripping the blots and re-probing with a 7S ^{32}P-labelled probe.

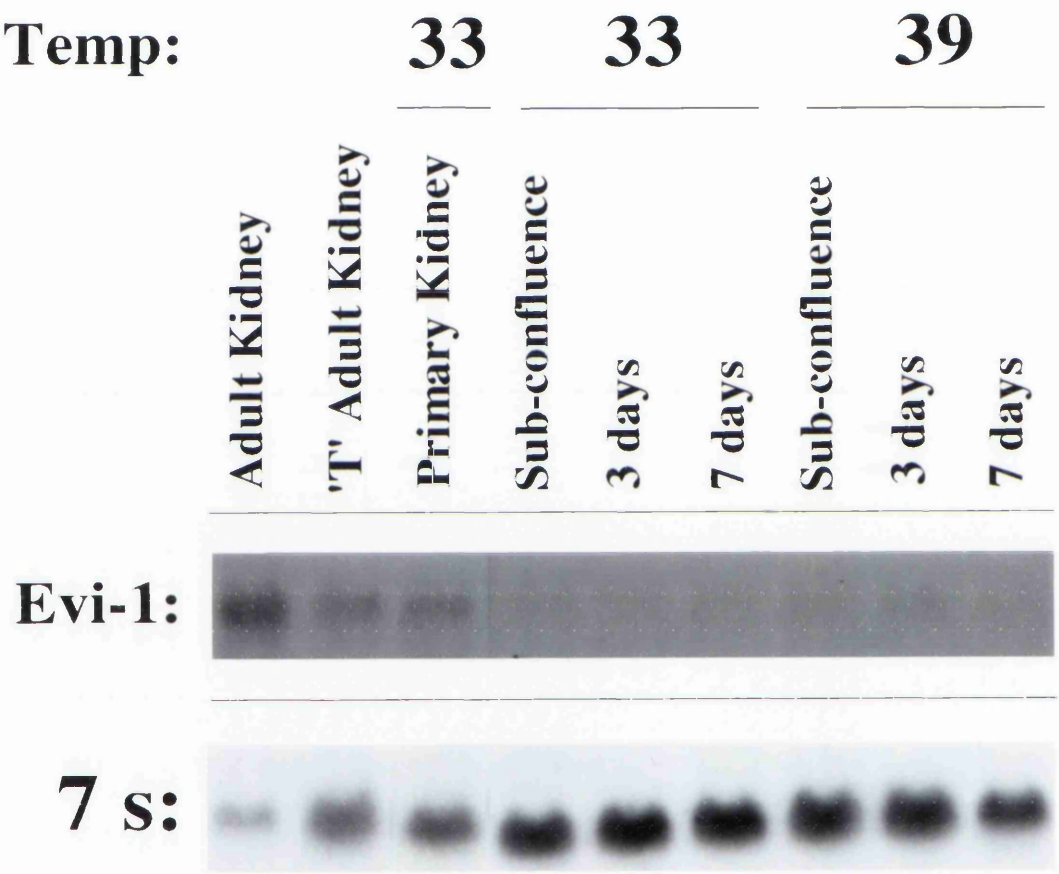

Temp:
33
33
39
Adult Kidney
'T' Adult Kidney
Primary Kidney
Sub-confluence
3 days
7 days
Sub-confluence
3 days
7 days
Evi-1:
7 s:

Figure 7.7: Northern blot showing *Evi-1* expression in embryonic kidney cultures isolated from transgenic mice.

Northern blot analysis using a 2kb SstI fragment Evi-1 cDNA ^{32}P-labelled probe.

The appropriate tissue or kidney culture from which RNA was isolated is given above each lane and is explained below,

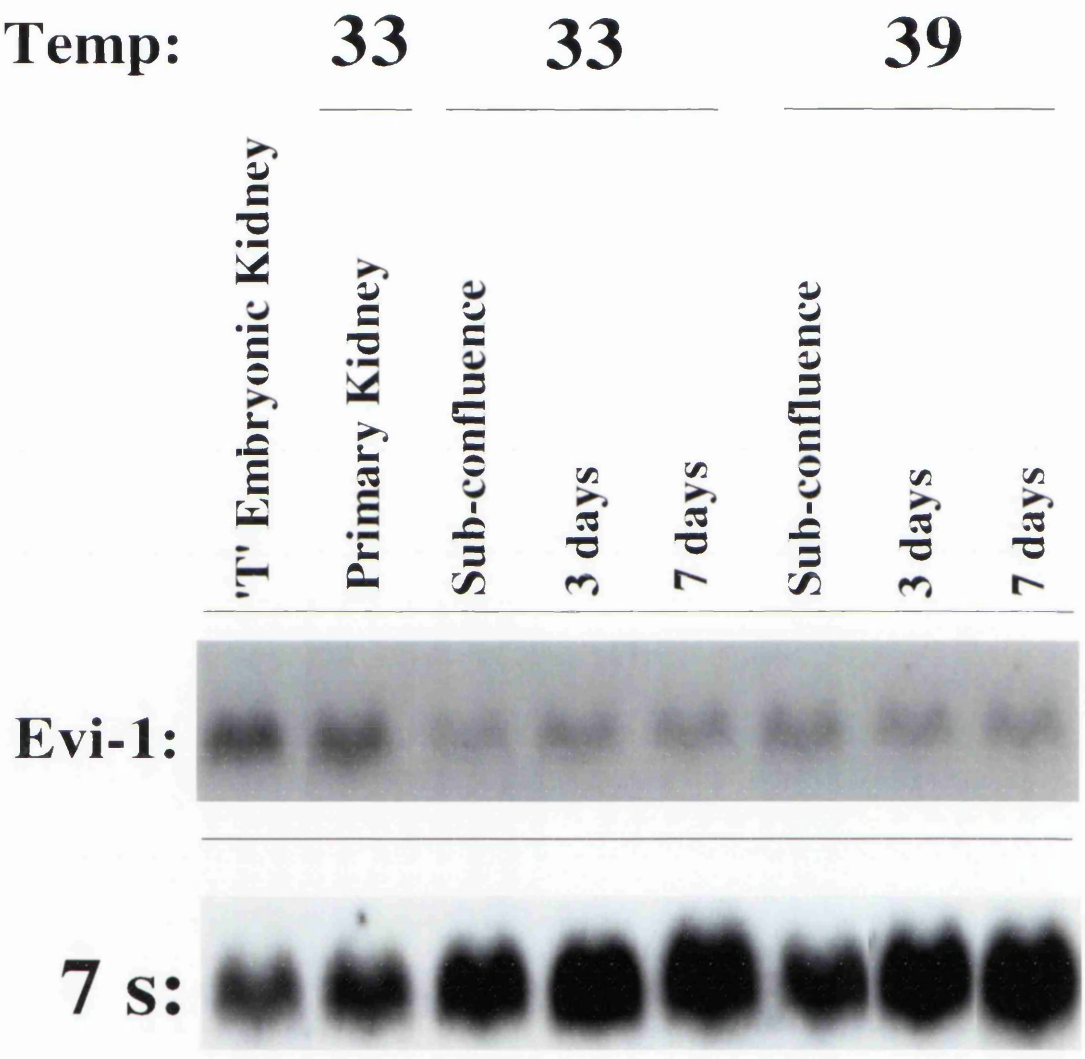
All the kidneys and cultures below were isolated from transgenic mice harbouring a temperature sensitive T antigen.

'T' adult kidney-RNA was isolated from embryonic kidneys taken at 18.5 days *post-coitus.*
Primary Kidney-kidney cultures established and RNA isolated at confluence

The following procedure was carried out at both 33^{o}C and 39^{o}C.

Sub-confluence-kidney cultures established over several passages were grown to just below confluence and then harvested for RNA.
3 days-kidney cultures established over several passages were grown to confluence, left for three days, and then harvested for RNA.
7 days-kidney cultures established over several passages were grown to confluence, left for seven days, and then harvested for RNA.

RNA loading was controlled by stripping the blots and re-probing with a 7S ^{32}P-labelled probe.


Temp:
33
33
39
'T' Embryonic Kidney
Primary Kidney
Sub-confluence
3 days
7 days
Sub-confluence
3 days
7 days
Evi-1:
7 s:

express moderate levels of *Evi-1* which is dramatically down regulated in established kidney cultures.

The results show that *Evi-1* expression levels drop between the initial primary kidney cultures and the fourth passage (figure 7.8). The changes in expression may reflect the increasing predominance of low expressing cell types as the culture is passaged. Alternatively it may reflect an alteration in the high expressing cells as it is passaged, although cell morphology looks very similar, even after several passages. Although the diversity of cell types present in our cultures are unknown previous studies by other groups have identified similar cultures to contain proximal and distal tubular cells and endothelial and fibroblasts cells (Yanai *et al* 1991a; Ebert *et al.*, 1990).

7.6 Functional analysis of *Evi-1* reporter constructs in primary kidney cultures

The primary kidney cultures described above were used for the functional analysis of *Evi-1* tissue specific gene expression. *Evi-1* reporter constructs were transfected into primary kidney cultures which have been shown to express relatively high levels of *Evi-1* (Chapter 7.5).

A selection of reporter constructs used in the basal promoter expression studies were chosen to analyse tissue specific expression of the *Evi-1* promoter. These included the largest two constructs, pGL2E8 and pGL2E5, which contain DHS I & DHS II, pGL2E2 which has lost DHS I, the pGL2-338 construct which represents the minimum optimal basal promoter, pGL2-284, which results in a 50% reduction in basal promoter activity, pGL2-15 which has no TATA box but has the major initiation sites and the pGL2basic vector (Chapter 5.5).

Evi-1 reporter constructs were transfected into primary adult kidney cells from 'normal' mice using DOTAP (Materials and methods). The primary kidney cultures are difficult to transfect and the luciferase activities are considerably lower than those seen

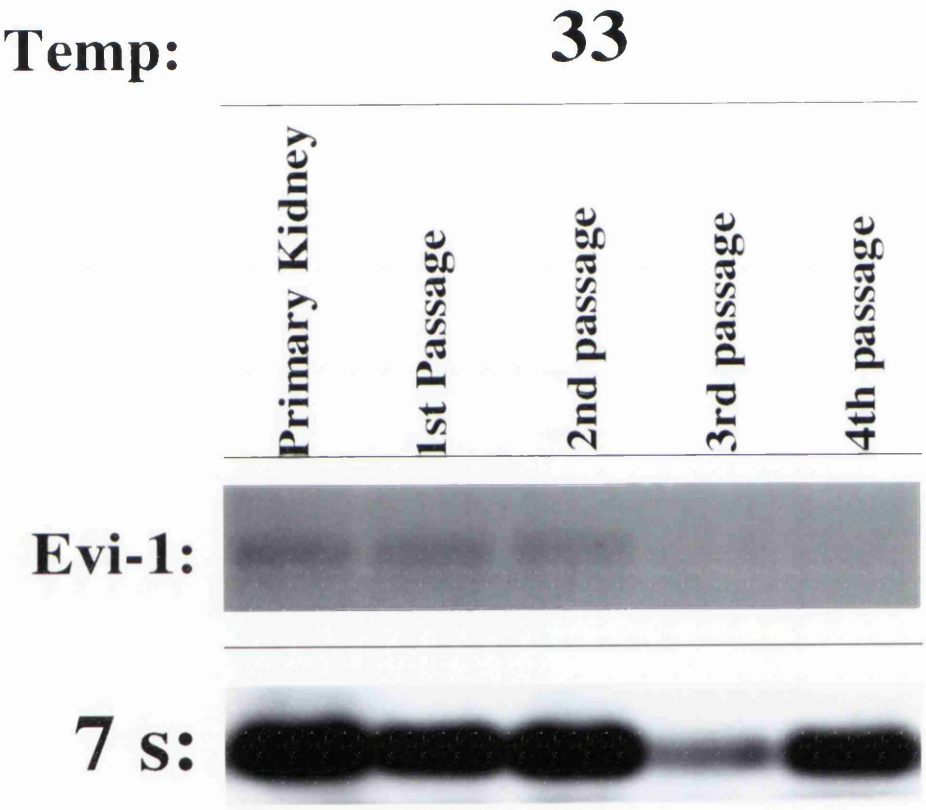
Figure 7.8: Northern blot showing *Evi-1* expression in embryonic kidney cultures over time in culture.

Northern blot analysis using a 2kb SstI fragment Evi-1 cDNA ^{32}P-labelled probe.

All the kidney cultures were isolated from transgenic mice harbouring a temperature sensitive T antigen.

RNA was isolated from primary embryonic kidney cultures or kidney cultures after the 1st, 2nd, 3rd or fourth passage. The appropriate RNA is indicated above each lane.

RNA loading was controlled by stripping the blots and re-probing with a 7S ^{32}P-labelled probe.


Temp:
33
Primary Kidney
1st Passage
2nd passage
3rd passage
4th passage
Evi-1:
7 s:

in the cell lines 293, HEC-1B and L929. The results of promoter activity are shown in figure 7.9. The histogram shows promoter activity relative to the largest 8kb fragment which represents 100% activity. The results suggest that the pGL2E8 construct is approximately 50% more active than the other reporter constructs tested, which is in contrast to the pattern of expression observed in the basal promoter studies. Furthermore, the drop in activity between -338 to -284 observed in basal expression studies is not seen. These data suggest that there are functional *cis*-acting sequences, located within an approximately 3kb region from within exon I, into intron I and down to the end of exon II, which are responsible for the increased promoter activity observed. These preliminary experiments are also supported by similar results observed in primary tsT adult and embryonic kidney cultures (figure 7.10 & 7.11).

7.7 Activity of reporter constructs in passaged tsT kidney cultures

Results from Northern blots had suggested that the *Evi-1* expression levels in tsT kidney cultures is significantly decreased over time in culture such that by the fourth passage *Evi-1* expression is undetectable (Chapter 7.8). To see if loss of expression also correlated with a change in the activity of the *Evi-1* promoter constructs the same reporter constructs were transfected into tsT adult kidney cultures which had been passaged for several weeks (Material & methods). The results are shown in figure 7.12. In contrast to primary kidney cultures, there was no significant change in the relative luciferase activity between the pGL2E8 and pGL2E5 constructs suggesting that the 3' 3kb region was not contributing to the transcriptional activity in these cells. There was also no difference in activity down to pGL2-338 and then an approximately 50 % drop in promoter activity is seen between pGL2-338 and pGL2-284. This trend in promoter activity is the same as that in 293, HEC-1B and L929 cell lines, described previously (Chapter 5.5). A similar trend is also seen in passaged tsT embryonic kidney cultures

Figure 7.9: A histogram showing relative luciferase activity from *Evi-1* promoter constructs in primary adult kidney cultures.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

Primary Adult Kidney
% Relative Luciferase Activity
140
120
100
80
60
40
20
0
LUC
LUC
LUC
LUC
LUC
LUC
LUC
pGL2E8
pGL2E5
pGL2E2
pGL2-338
pGL2-284
pGL2-15
pGL2
Luciferase Constructs

Figure 7.10: A histogram showing relative luciferase activity from *Evi-1* promoter constructs in primary adult tsT kidney cultures.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

Primary tsT Adult Kidney
% Relative Luciferase Activity
140
120
100
80
60
40
20
0
LUC
pGL2E8
pGL2E5
pGL2E2
pGL2-338
pGL2-284
pGL2-15
pGL2
Luciferase Constructs

Figure 7.11: A histogram showing relative luciferase activity from *Evi-1* promoter constructs in primary embryonic tsT kidney cultures.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

Primary tsT Embryonic Kidney
% Relative Luciferase Activity
140
120
100
80
60
40
20
0
LUC
LUC
LUC
LUC
LUC
LUC
LUC
pGL2E8
pGL2E5
pGL2E2
pGL2-338
pGL2-284
pGL2-15
pGL2
Luciferase Constructs

Figure 7.12: A histogram showing relative luciferase activity from *Evi-1* promoter constructs in passage adult tsT kidney cultures.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

Passaged tsT Adult Kidney
% Relative Luciferase Activity
140
120
100
80
60
40
20
0
LUC
pGL2E8
pGL2E5
pGL2E2
pGL2-338
pGL2-284
pGL2-15
pGL2
Luciferase Constructs

transfected with the same constructs (figure 7.13). Consequently, there is a correlation between expression of the endogenous Evi-1 and activity of 3kb region.

7.8 Conclusions of transfections into primary and passaged kidney cultures

DHS analysis using an SstI digest revealed a faint DHS site which may map over the border between exon I and intron I. Deletion of this region in previous basal transcription studies and now also in passaged tsT kidney cultures has no effect on transcriptional activity. In contrast, preliminary results in primary kidney cultures has shown that an 8kb EcoRI/SmaI fragment within the pGL2E8 reporter construct has significantly higher activity (50-80%) in primary kidney cultures than the pGL2E5 construct which has a 3' 3kb deletion including the exon I/intron I boundary. This suggests that tissue specific *cis*-acting sequences are present within a 3kb region from an SstI site at +272 to the end of exon II and that previous DHS analysis might suggest that this activity is located over the border of exon I/intron I.

In support of these observations are the results in passaged tsT kidney cultures, which have lost endogenous *Evi-1* expression over time in culture, and the pattern of exogenous gene expression reverts to basal transcriptional regulation. Consequently, there is a correlation between *Evi-1* expressing cultures and enhanced activity of pGL2E8 suggesting tissue specific expression of this construct in a proportion of primary kidney cells.

Figure 7.13: A histogram showing relative luciferase activity from *Evi-1* promoter constructs in passaged embryonic tsT kidney cultures.

5μg of pGL2basic or equimolar amounts were transfected with 5μg of pHSV-βgal, which contains a β-galactosidase gene under the HSV-2 IE5 gene promoter. The total amount of DNA transfected was then made upto 20μg with pKSII+ bluescript vector.
Transfections were carried out using DOTAP as described by the manufacturer.

The histogram illustrates the results obtained from a typical experiment carried out at least three times. Shaded bars on the histogram represent the average value from at least four independent transfections. Error bars show the maximum error from the mean.

The particular luciferase reporter construct used in each reaction is shown schematically below each bar.

Passaged tsT Embryonic Kidney
% Relative Luciferase Activity
140
120
100
80
60
40
20
0
LUC
pGL2E8
pGL2E5
pGL2E2
pGL2-338
pGL2-284
pGL2-15
pGL2
Luciferase Constructs

Part 4: Discussion

Chapter 8: Discussion

Previous studies have implicated *Evi-1* as being an important transcription factor in the development of specific murine tissues (kidney, lung and developing oocytes) and in the progression of some murine and human leukaemias, following ectopic expression of the gene. The aim of this project was to assess the factors controlling transcriptional regulation of the *Evi-1* gene. The primary observations from this work are summarised below;

We have defined a minimal promoter activity between nucleotides -341 and +163 relative to the major initiation site, which coincides with the presence of a DHS, designated DHS II. Sequential 5' deletion of the optimal basal promoter delineates two regions -341 to -284 and -284 to -254 which are critical for maximal promoter activity in 293, HEC-1B and L929 cells. A DNAse 1 footprint was identified in each of these regions, and designated FpIII and FpIV respectively. EMSA analysis identified a potentially novel factor (complex), designated EvBP1 which is required for optimal activity and coincides with FpIII. The EvBP1 monomer/oligo is approximately 97kD, recognises the sequence tttccctggggaaa and is ubiquitous. EMSA analysis of FpIV identified a CCAAT binding protein, which might be CP1, binding to this region. In addition we have identified a 3' region between +48 to +163 important to basal activity.

We have established *Evi-1* expressing kidney cultures as a system to examine tissue specific expression. This allowed the identification of a 3kb region, located between exon I and II, necessary for higher activity in the cultures. This activity might correlate with a DHS site which is part of the DHS II complex. Furthermore, a second ubiquitous DHS I site was identified, although a function has not been defined for this region yet.

8.1 Minimal basal promoter activity

Functional analysis of the *Evi-1* promoter identified a minimal promoter -341 to +163 which correlates with the presence of DHS II. In addition two footprinted regions, FpIII and FpIV, have been identified whose combined loss correlated with a drop in 70% of the total activity of the promoter. The protein interactions which occur over FpIII and FpIV are discussed below;

The FpIII region binds a potentially novel transcription factor

EMSA of the FpIII region identified a major binding complex, designated EvBP1 in all cell types examined. Deletion of this site, from -338 to -284, results in 50% drop in activity suggesting that the EvBP1 complex is also a positive regulator of basal promoter activity. The majority of transcription factors bind to palindromic sequences which influences the way proteins line up with respect to the major/minor grooves of DNA (Pabo & Sauer, 1992). Consistent with this the EvBP1 complex binds an 14bp imperfect palindromic sequence (chapter 6.7), tttccctggggaaa, which is absolutely conserved in the human *Evi-1* promoter sequence (Ihle, unpublished data) suggesting that this recognition site may play an important part in regulation the *Evi-1* gene. There were no other proteins identified upstream of EvBP1 and the minor complexes downstream were observed to alter in intensity between experiments and cell lines (chapter 6.3). Consequently, we therefore propose that EvBP1 is the major complex binding over this region contributing to basal promoter activity. To confirm this we intend to specifically delete the EvBP1 binding site and examine the affect on optimal promoter activity. Alternatively, multiple copies of EvBP1 binding sequence may be cloned upstream of an minimal promoter to see if it activates a heterologous promoter.

Attempts to identify the components of the EvBP1 complex by competition with consensus binding site oligonucleotides for other transcription factors revealed a potential link with C/EBP. However, the EvBP1 complex is composed of a 97kD monomer/oligo, which may form dimers, and is heat labile which contrast with the properties exhibited by members of the C/EBP family. Competition of the EvBP1 complex with excess C/EBP consensus site oligonucleotides may be due to fortuitous similarities in the sequence and represents a non-specific interaction. It is also plausible that other factors are present in the complex but perhaps only EvBP1 binds DNA and therefore is cross-linked. The identity of the competitor of the EvBP1 complex is unknown. Therefore, EvBP1 might be a novel transcription factor which is involved in the transcriptional regulation of the *Evi-1* gene.

In addition to EvBP1 this region contains a number of putative recognition sequences for other proteins, including AP3, C/EBP, *ets* sequence and an inverted AP2 sequence. AP2 (α and β), members of the *ets* and C/EBP family have distinct expression patterns and have been implicated in tissue specific regulation (Mitchell *et al.*, 1991; Moser *et al.*, 1995; Seth *et al.*, 1992; Birkenmeier *et al.*, 1989). It is conceivable that the proteins binding over this region may differ depending on the cellular environment, and thereby influence tissue specific, inducible or basal expression of the EVI-1 protein. In support of this hypothesis is an observation from the competition experiments which suggest that the complex was composed of members of the C/EBP family. EMSA analysis using a C/EBP ^{32}P-labelled probe, showed that the 3' FpIII region was able to compete for a C/EBP specific complex, which contrasted with the results using the whole FpIII or 5'FpIII region (chapter 6.5). The 3'FpIII probe contains only half of the EvBP1 palindrome region and is unable to bind C/EBP. It is possible that in the absence of EvBP1 that this unmasks a C/EBP site. However, if so, the excess FpIII and 5'FpIII should bind all the EvBP1 but still have oligonucleotides free to compete for C/EBP. An explanation for this might be that EvBP1 is particularly abundant. In support of this, the

majority of labelled FpIII is shifted (figure 6.6) into the EvBP1 complex, suggesting the levels of EvBP1 may be particularly high in 293 cells. Alternatively, the intact inverted repeat may form secondary structure which prevents interaction of a C/EBP-like factor. Nonetheless, the data suggest that there are circumstances when this region can bind C/EBP.

DNA/Protein interaction over the FpIV region

As mentioned previously there are a selection of common motifs often found within the promoter of class II genes. Consistent with this, there are putative CCAAT and GC box elements within the *Evi-1* promoter. Deletion of FpIV (-284 to -254) results in substantial reductions in basal promoter activity. Significantly a CCAAT box located upstream of the major initiation sites of the *Evi-1* gene, was located in the middle of the FpIV region. A computer-based search of the Site-data data base of the FpIV region identified the CCAAT box as a putative CP1 factor binding site. The CP1 complex is composed of two sub-units CP1a and CP1b which are both required for appropriate DNA binding (Chodosh *et al.*, 1988). In addition to CP1 a number of other proteins may also recognise the CCAAT site, including members of the C/EBP family, CP2 and NF1/CTF. However, the CCAAT proteins show selectivity for different sequences with the region immediately surrounding the CCAAT box important in specificity (Chodosh *et al.*, 1988). CP1 or related factors have been shown to bind to other promoter regions including the mouse α globin, the MHC class II Eα and human hsp70 promoter (Cohen *et al.*, 1986; Wu *et al.*, 1987) and activate transcription in a positive manner. Further experiments are required to confirm the binding of CP1 to this region which would initially include heat stability experiments to eliminate C/EBP factors binding to the FpIV region.

FpIV is an extensive footprint from -277 to -208 suggesting that other proteins are binding over this region and this has been confirmed by EMSA analysis. Consequently, we cannot yet attribute the reduction in optimal promoter activity with the loss of only the CCAAT site. Once again it will be necessary to specifically delete the CCAAT site to conclusively determine its role in optimal basal promoter activity.

Protein interaction with other regions of the minimal promoter

Functional analysis also identified other regions important in optimal basal promoter activity. For example, a steady reduction in activity was observed in 5' deletions between -254 to -15. It should be noted that deletions from -154 also result in the loss of minor initiation sites which may contribute to the reduction in activity. Furthermore, a substantial drop in basal promoter activity was also observed in the 3' deletion between +163 to +48. Preliminary EMSA analysis has suggested that there are protein interactions within these regions of the *Evi-1* promoter (data not shown). For example, the Sp1 transcription factor may appear to target a putative Sp1 site between -163 to -158. Although there is no consensus Sp1 site between +163 to +48 the region is quite GC-rich and preliminary evidence that Sp1 may also bind in this region. The next logical step would be to use a combination of footprinting, EMSA and antibody supershift analysis of these regions to identify where the protein interactions occur and the factors binding.

8.2 Tissue specific expression of the *Evi-1* promoter

Since DHS sites are generally associated with active promoter and enhancer elements (Gross and Garrard, 1988) we looked for a correlation between expression of the gene in a basal-, tissue specific- , and ectopic- manner and the chromatin structure

surrounding the 5' *Evi-1* gene. DHS analysis was performed in a number of different cell lines or tissues, i.e. DA-3, L929, FDCP-1 cells and kidney and spleen tissue and identified two DHS sites I and II over an 18kb region in the 5' *Evi1* locus.

Functional analysis of the 5' *Evi-1* locus identified a minimal basal promoter which correlated with the DHS II region. As already mentioned a number of protein interactions have been observed over this region *in vitro* which presumably contribute to basal transcriptional activity. However, the functional significance of the DHS II region in tissue specific expression of the *Evi-1* gene is less clear. Classically, tissue specific or lineage expression of a gene often correlates with the presence of a specific DHS site Although DHS II is present exclusively in *Evi-1* expressing cells, its presence did not correlate with high level tissue specific expression of the gene, since it was also identified in low expressing *Evi-1* cells (L929) (figure 4.7). This might suggest that tissue specific expression of the *Evi-1* promoter may be dependent purely on the interaction of tissue specific transcription factors with DHS I or II rather, than being reliant on the relief of repressive effects mediated by the chromatin framework, following induction of tissue specific DHS sites. In support of this, a region of DHS II may confer tissue specific expression in primary murine kidney cultures.

It is possible that tissue specific expression of the *Evi-1* gene may be dependent on the appropriate interaction of more ubiquitous factors. This form of regulation is also observed in other promoters. For example, the *Pim-1* gene is selectively transcribed in haemato-lymphoid cell lines despite G/C-rich housekeeping promoter (Meeker *et al.*, 1990). Ectopic expression of the *Evi-1* promoter in DA-3 cells illustrates that high levels of *Evi-1* promoter expression is presumably not absolutely dependant on tissue specific factors, but may be mimicked by the integrated LTR. Consistent with this hypothesis is the presence of a DHS site at the *Fim-3* locus mapping within the integrated LTR, found in DA-3 cells but not kidney tissue (Chapter 4.5). Similarly, ectopic expression of *EVI-1*

in the translocation, inv (3)(q21q26), is believed to bring the gene under the regulation of the highly expressed Ribophorin I gene (Suzukama *et al.*, 1994). Since the Ribophorin I gene is a constitutively active gene presumably ubiquitous factors regulate the promoter and subsequently the *Evi-1* promoter in this translocation.

Presumably, if the *Evi-1* promoter was regulated by constitutive factors the gene would have to be repressed in non-expressing cells, possibly by chromatin structure or methylation. Since there is no evidence for tissue specific DHS sites this may suggest that methylation is important in *Evi-1* repression. Intriguing, one allele of the *Evi-1* gene is methylated in DA-3 cells, the high expressing murine myeloid cell line (Bartholomew & Ihle, 1991). This question remains unsolved although some preliminary data suggests that the *Evi-1* gene is not upregulated in 293 and RAG kidney cells following treatment with 5-azacytidine (data not shown).

Functional analysis of tissue specific expression of the Evi-1 gene

The *Evi-1* gene is normally expressed in a tissue specific manner, and therefore, one of the objectives of this study was to identify *cis*- and *trans*-acting factors involved. However, the absence of appropriate cell lines which express the endogenous *Evi-1* gene has hindered this analysis. Therefore, a kidney cell culture system has been established to address this issue. Although normal kidney cultures express high levels of *Evi-1* they have limited growth. By contrast, the tsT kidney cultures have extensive growth although, unfortunately, high levels of *Evi-1* expression are maintained only for a limited period of time (3/4 passages). Attempts to isolate cell lines from this kidney culture have proved unsuccessful, and in any case since *Evi-1* expression is lost after 3/4 passages it is likely that any cell lines established will have relatively low levels of *Evi-1* expression.

Since these cultures initially consist of multiple cell types continued culture may favour establishment of cells which don't express *Evi-1*. Furthermore, it is interesting that

only low *Evi-1* expression is observed in the kidney cell lines examined suggesting that these cell lines might be established from cells not expressing *Evi-1* or which have lost *Evi-1* expression. Although *In situ* hybridisation experiments have suggested that high levels of *Evi-1* are restricted mainly to kidney tubules the kidney is a heterogeneous organ and the precise cell type which expresses are unknown. Consequently, it is not immediately clear whether high levels of expression of *Evi-1* in kidney is due to a general upregulation in a number of cell types or alternatively a small number of cells with high levels of expression.

A number of genes are commonly downregulated over time in culture by a mechanism that is thought to be due to the inhibitory effects of methylation over the appropriate promoter region. Examples include the thymidine kinase (Wise & Harris)[1988] and metallothionein (Compere & Palmittier, 1981) genes. Consequently, it is possible that *Evi-1* may be repressed in a similar manner. Since it is not clear precisely what biological role *Evi-1* plays in both adult and embryonic kidney development it is difficult to assess whether this may conflict with the establishment of high expressing cell lines. Interestingly, a separate study observed that *Evi-1* mRNA levels are consistently reduced in a selection of kidney tumours (Bartholomew & Oshimura, unpublished data) suggesting that this event may be important in kidney tumour progression. Presumably, this effect must be tissue type dependant since, by contrast, ectopic expression of *Evi-1* is believed to be an important determinant in the progression of some leukaemias and high *Evi-1* expression is observed in established myeloid cell lines. Any of these possibilities, either singularly or in combination, may contribute to the technical problems of establishing and maintaining high *Evi-1* expressing cell lines.

Functional analysis of the *Evi-1* promoter in the primary kidney cultures has implicated a 3kb region 3' to the major transcriptional initiation sites as being important in tissue specific expression of the gene. The location of tissue specific elements within this 3kb region which includes exon I and intron I and exon II is unknown. However, our

analysis suggests a DHS site maps to this region at the exon I/intron I boundary. To identify putative protein binding sites within the exon I/intron I boundary an approximately 200bp of DNA from the end of exon I to an EcoRI site within intron I has been sequenced. The sequence is particularly G/C-rich and contains a putative site for Sp1. Interestingly, these factors have been implicated in tissue specific expression of other genes and Sp1 has been proposed to be able to stabilise DNA looping, and enhancer mediated activation (Su *et al.*, 1991). It is therefore conceivable that this region may direct tissue specific expression independently or alternatively by interaction with enhancer elements, possibly located within DHS I or II. However, further constructs will need to be made to assess: 1) the contribution of DHS I to tissue specific expression; 2) the region within the 3' 3kb region responsible for tissue specific expression. These studies could be performed in the primary kidney cultures, particularly if the transfection conditions are improved.

8.3 cAMP regulation of the *Evi-1* promoter

Previous studies have shown that the *Evi-1* promoter has been shown to be responsive to the levels of cAMP within cells and may be induced upto 16-fold by the addition of forskolin in the epithelial renal carcinoma cell line, A704 (Bartholomew & Clarke., 1993). However, the sequences reponsible for this regulation have not been discribed. Response to intracellular cAMP levels can be mediated by two cis-acting elements, cyclic AMP response elements (CREs) and TPA response elements (TREs) which bind CREB (reviewed by Sassone-Corsi, 1995) and the AP2 proteins respectively, although both sequences may also bind other factors (reviewed by Karin, 1988). This is particularly interesting with regard to *Evi-1* since there are three putative AP2 sites within the *Evi-1* promoter and also putative CRE sites further upstream within DHS I (chapter 5.16). The putative AP2 sites are located, (1) over the EvBP1 recognition site (2) further

upstream partially overlapping a previously identified footprinted region, FpI (Bartholomew, unpublished data) (3) further downstream within the region lost in the deletion +163 to +48. A related factor to AP2, designated AP2β has recently been cloned, suggesting that there may be a family of related proteins (Moser *et al.*, 1995). Interestingly, expression of AP2β is particularly high in kidney cells.

It would be of interest to identify which *cis*-acting elements are responsible for cAMP induction of *Evi-1* but further investigation have been hampered by the low transfection efficiency of the A704 cell line (data not shown). Preliminary studies in 293 cells, using forskolin to induce cAMP levels, did not have an effect on a 5kb *Evi-1* pGL2E5 reporter construct (data not shown). However, since forskolin does not induce the endogenous gene in 293 cells it is possible that there are other tissue specific factors present in A704 cells which contribute to the induction response. Consistent with this, a number of tissue specific regulated promoters require cAMP inducible elements in addition to a tissue specific factors to achieve optimal expression. For example, the tissue specific enhancer of the human glycoprotein hormone α-subunit gene requires cAMP-inducible elements to confer tissue specificity (Delgeane *et al.*, 1987).

It is not clear whether the cAMP responsive pathway is important *in vivo* for the tissue specific expression pattern of the *Evi-1* gene. However, it may be significant that the stimulation of adenylate cyclase activity by putative stimulating agents appears to be different for specific cell types in the kidney (Wilson *et al.*, 1987). It would be interesting to see whether the cAMP response observed in A704 cells (16-fold) and normal primary kidney (2-fold) (Bartholomew & Clarke, 1993), is observed in the tsT kidney cultures.

8.4 The function of DHS I

The location of the DHS I approximately 2 kb upstream of exon I suggested that this region may act as an enhancer of the *Evi-1* gene, however, we have found no activity, either positive or negative, for this region on the basal expression of the *Evi-1* promoter or on the heterologous SV40 or minimal tk promoters in transient transfection assays. Transient transfection experiments were used to identify transcriptional control elements and allow relatively rapid and comparative analysis of a range of constructs. However, a disadvantage of this technique is that the transient constructs are not organised into appropriate nucleosomal units, which may be essential for the function of some *cis*-elements on transcriptional expression. It is therefore possible that the function of DHS I region is not observed in transient assays. Examples of this type of regulation include the SAR (Klehr *et al.*, 1991) or the functionally related *Drosophila* Scs regions, whose presence may activate transcription if stably integrated into the genome, but have no effect in transient transfection assays (chapter 1.8). Furthermore, there is evidence suggesting that some enhancers *in vivo* may function by suppressing PEV (chapter 1.7), altering the chromatin framework surrounding a gene resulting in a more accessible promoter region (Walters *et al.*, 1996).

Although, the function of DHS I is as yet unknown, it is conceivable that it may function in a manner similar to that mentioned above. As a consequence, future experiments should involve the stable transfection of *Evi-1* reporter constructs into cells to assess if DHS I has an effect on basal transcription levels. There are also a number of disadvantages about the use of stabile transfections. It is time consuming, particularly with a range of constructs, constructs may show positional dependent expression depending on the integration site, and the integrated constructs may form concatamers, altering the copy number. However, this can be controlled for by analysing a pool of colonies.

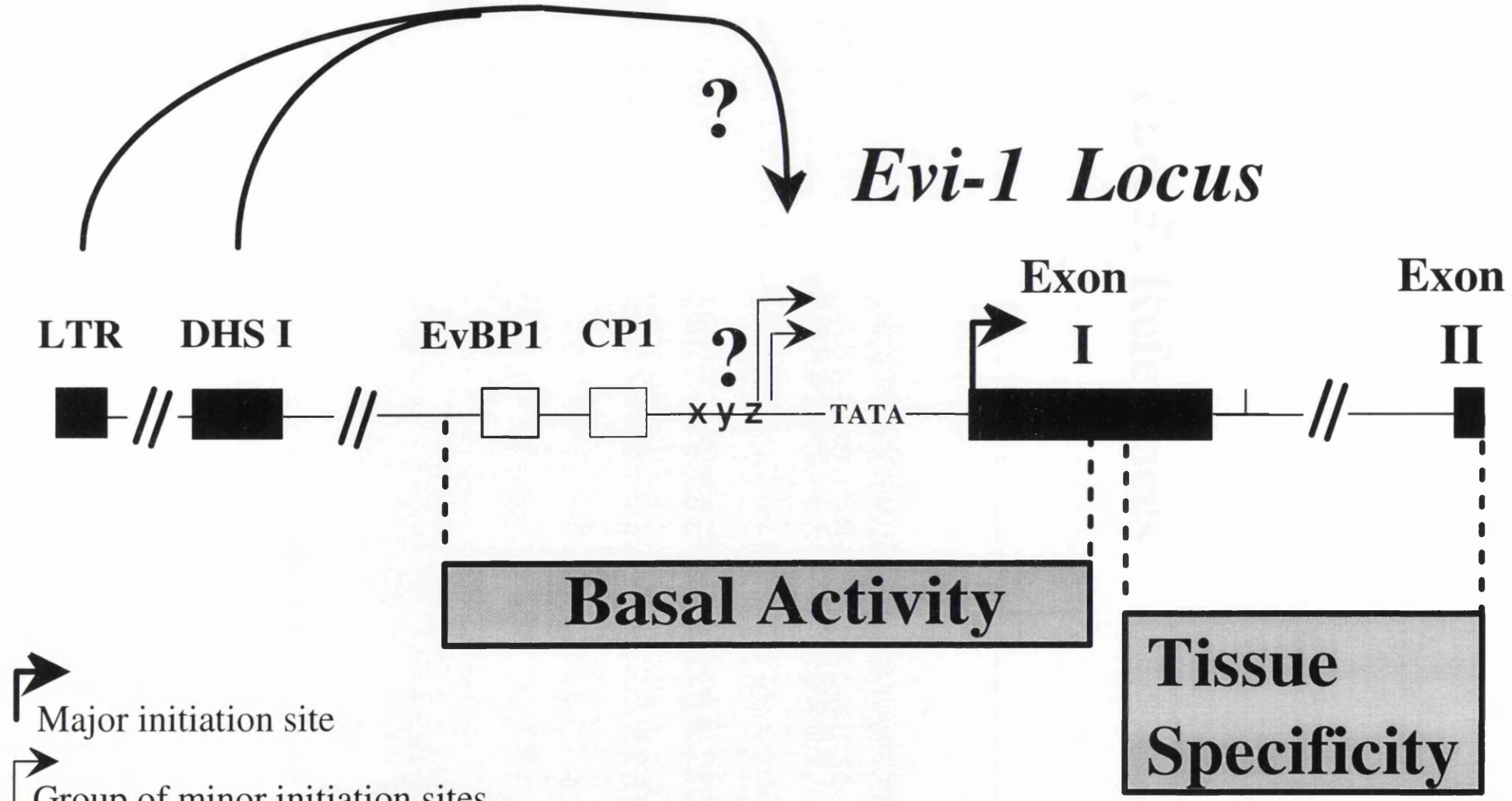
Alternatively, the DHS I region may be important in the cAMP inducible nature of the *Evi-1* gene. This is particularly significant because as mentioned previously the DHS I region contains putative AP1, CRE and TRE elements. However, it is also possible that DHS I has no enhancer capacity and that DHS I may reflect the presence of an alternative promoter, possibly even a tissue specific promoter. However, loss of DHS I in transient transfection assays had no effect on the overall promoter activity, suggesting that this probably isn't the case in basal expressing cells. It should also be noted, that the primary function of DHS I may be independent of transcription altogether, since DHS sites are also associated with sites for origins of DNA replication and chromosome segregation (Wolffe & Pruss, 1996).

Do other tissue specific DHS sites exist ?

It is conceivable that there are unidentified tissue specific DHS sites within the 18kb region, but due to the heterogeneous nature of the kidney the proportion of high expressing cells within the population may be too small to detect them. Alternatively, since distal enhancers/silencers may function at considerable distances up- and downstream of the promoter there may be other DHS sites outside the 18kb region examined in this study. It is therefore still possible that the *Evi-1* promoter region may be a target for other tissue specific enhancers.

We can conclude that the transcriptional regulation of the *Evi-1* gene is complicated. The temporal and spatial expression pattern of *Evi-1* in embryonic development suggests that the regulation of *Evi-1* in development is tightly controlled. De-regulation of expression has profound consequences, this is dramatically demonstrated when *Evi-1* is aberrantly expressed. For example, ablation of the *Evi-1* gene in knock-out studies leads to embryonic while ectopic expression of the gene is associated the progression of some leukaemias. As a consequence it is perhaps not

Figure 8.1: Transcriptional Control of the *Evi-1* gene



Chapter 9: References

Adams, J., Gerondakis, S., Webb, E., Corcoran, L.M., Cory, S. (1983). Cellular *myc* oncogene is altered by chromosome translocation to the immunoglobulin locus in murine plasmacytomas and is rearranged similarly in human Burkitt lymphomas. Proc. Natl. Acad. Sci. USA. 80, 1982-1986.

Adams, J.M., Cory, S. (1992). Oncogene co-operation in leukaemogenesis. Cancer. Surveys: oncogenes in the development of leukaemia. 15, 119-141.

Agre, P., Johnson, P.F., McKnight, S.L. (1989). Cognate DNA binding specificity retained after leucine zipper exchange between GCN4 and C/EBP. 246, 922-926.

Alitalo, K., Winqvist, R., Lin, C.C., de la Chapelle, A., Schwab, M., Bishop, J.M. (1984). Aberrant expression of an amplified c-myb oncogene in two cell lines from a colon carcinoma. Proc. Natl. Acad. Sci. 81, 4534-4538.

Almer, A., Horz, W. (1986). Nuclease hypersensitive regions with adjacent positioned nucleosomes mark the gene boundaries of the PHO5/PHO3 locus in yeast. EMBO. J. 5, 2681-2687.

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

Amati, B., Dalton, S., Brooks, M.W., Littlewood, T.D., Evan, G.I., Land, H. (1992). Transcriptional activation of the human c-Myc oncoprotein in yeast requires interaction with Max. Nature. 359, 423-426.

Amati, B., Land, H. (1994). Myc-Max-Mad: a transcription factor network controlling cell cycle progression and death. Curr. Opin. Genet. Dev. 4, 102-108.

Anderson, K.P., Kern, C.B., Crable, S.C., Lingrel, J.B. (1995). Isolation of a gene encoding a functional zinc-finger protein homologous to erythroid kruppel-like factor: identification of a new multigene family. Mol. Cell. Biol. 15, 5957-5965.

Anderson, W.F., Ohlendorf, D.H., Takeda, Y., Mattthews, B.W. (1981). Structure of the cro repressor from the bacteriophage λ and its interaction with DNA.

Annunziato, A.T., Frado, L-L. Y., Seale, R.L., Woodcock, C.L.F. (1988). Treatment with sodium butyrate inhibits the complete condensation of interphase chromatin. Chromosoma, 96, 132-138.

Austin, R.J., Biggin, M.D. (1995). A domain of the *even skipped* protein represses transcription by preventing TFIID binding to a promoter; repression by cooperative blocking. Mol. Cell. Biol. 15, 4683-4693.

Ausio, J., Dong, F., van Holde, K.E. (1986). Use of selectively trypsinised nucleosome core particles to analyse the role of the histone tails in the stabilisation of the nucleosome. J. Mol. Biol. 206, 451-463.

Ayer, D.E., Kretzner, L., Eisenman, R.N. (1993). Mad: a heterodimeric partner for Max that antagonises myc transcriptional activity. Cell. 72, 211-222.

Ball, D.J., Gross, D.S., Garraed, W.T. (1983). 5-methylcytosine is localised in nucleosomes that contain histone H1. Proc. Ntl. Acad. Sci. USA. 80, 5490-5494.

Baniahmad, A., Muller, M., Steiner, C., Renkawitz, R. (1987). Activity of two different silencer elements of the chicken lysozyme gene can be compensated by enhancer elemenrs. EMBO. J. 6, 2297-2303.

Barberis, A., Super-Furga, G., Busslinger, M. (1987). Mutually exclusive interaction of the CCAAT-binding factor and a displacement protein with overlapping sequences of a histone gene promoter. Cell. 50, 347-359.

Barnhart, K.M., Kim, C.G., Banerji, S.S., Sheffery, M. (1988). Identification and characterisation of multiple erythroid cell proteins that interact with the promoter of murine alpha-globin gene. Mol. Cell. Biol. 8, 3215-3226.

Bartholomew, C., Morishita, K., Askew, D., Buchberg, A., Jenkins, N.A., Copeland, N.G., Ihle, J.N. (1989). Retroviral intsertions in the *CB-1/Fim-3* common site of integration activate expression of the *Evi-1* gene. Oncogene 4, 529-534.

Bartholomew, C., Ihle, J.N. (1991) Retroviral insertions 90 kilobases to the *Evi-1* myeloid transforming gene activate transcription from the normal promoter. Mol. Cell. Biol. 11, 1820-1828.

Bartholomew, C., Clark, A.M. (1994) Induction of two alternatively spliced *evi-1* proto-oncogene transcripts by cAMP in kidney cells. Oncogene. 9, 939-942.

Ben-David, Y., Giddens, E.B., Letwin, K., Bernstein, A. (1991). Erythroleukaemia induction by Friend murine leukaemia virus: insertional activation of a new member of the ets gene family, Fli-1, closely linked to ets-1. Genes Dev, 5, 908-919.

Berg., J.M., Shi, Y. (1996). The galvanisation of biology: a growing appreciation for the role of zinc. Science. 271, 1081-1085.

Bergeron, D., Poliquin, L., Houde, J., Barbeu, B., Rassart, E. (1992) Analysis of proviruses integrated in the *Fli-1* and *Evi-1* regions in Cas-Br-E MuLV-induced non-T, non-B-cell leukaemias. Virology. 191, 661-669.

Bieker, J.J., Southwood, C.M. (1995). The erythroid Kruppel-like factor transactivation domaain is a critical component for cell-specific inducibility of a b-globin promoter. Mol. Cell. Biol. 15, 852-860.

Binetruy, B., Smeal, T., Karin, M. (1991). Ha-Ras aguments cJun activity and stimulates phosphorylation of its activation domain. Nature. 351, 122-127.

Bird, A.P. (1986). CpG-rich islands and the function of DNA methylation. Nature (London) 321, 209-213.

Bird, A.P. (1987). CpG islands as gene markers in the vertebrate nucleus. Trends Genet. 3, 342-347.

Birkenmeier, E.H., Gwynn, B., Howard, S., Jerry, J., Gordon, J.I., Landscultz, W.H., McKinght, S. (1989). Tissue specific expression, developmental regulation, and genetic mapping of the gene encoding CCAAT/enhancer binding protein. Genes. Dev. 3, 1146.

Bitter, M., Neilly, M., LeBeau, M., Pearson, M., Rowley, J. (1985). Rearrangements of chromsome 3 involving bands 3q21 and 3q26 are associated with normal and elevated platelet counts in acute nonlymphocytic luekaemia. Blood. 66, 1362-1370.

Blackwood, E.M., Eisenmann, R.N. (1991). Max: a helix-loop-helix zipper protein that forms a squence-specific DNA-binding complex with Myc. Science 251, 1211-1217.

Blair, W.S., Bogerd, H.P, Madore, S.J., Cullen, B.R. (1994). Mutational analysis of the transcriptional activation domain of RelA: identification of a highly synergistic minimal acidic activation domain. Mol. Cell. Biol. 14, 7226-7234.

Blank, V., Kourilsky, P., Israel, A. (1992). NF-kB and related proteins: Rel/dorsal homologies meet ankyrin-like repeats. Trends. Biochem. Sci. 135-140.

Bohmann, D., Bos, T.J., Admon, A., Nishimura, T., Vogt, P.K., Tjian, R. (1987). Human proto-oncogene c-jun encodes a DNA binding protein with structural and functional properties of trnascription factor AP-1. Science. 238, 1386.

Bopp, D., Burri, M., Baumgartner, G., Frigerio, G., Noll, M. (1986). Conservation of a large protein domain in the segmentation gene *paired* and in functionally related genes of *Drosophila.* Cell. 47, 1033-1040.

Bordereaux, D., Fichelson, S., Sola, B., Tambourin, P.E., Gisselbrecht, S. (1987). Frequent involvement of the fim-3 region in friend murine leukaemia virus-induced mouse myeloblastic leukaemias. J. Virol. 61, 4043-4045.

Bordereaux, D., Fichelson, S.,Tambourin, P.E., Gisselbrecht, S. (1990). Alternative splicing of the Evi-1 zinc-finger gene generates mRNAs which differ by the number of zinc-finger motifs. oncogene. 5, 925-927.

Boyes, J., Bird, A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell. 1991. 64, 1123-1134.

Boyle, W.J., Smeal, T., Denlize, L.H.K., Angel, P., Woodgett, J.R., Karin, M., Hunter, T. (1991). Activation of protein kinase C decreases phosphorylation of c-jun at sites that negatively regulate its DNA-binding activty. Cell. 64, 573-584.

Breathnach, R., Chambon, P. (1981). Organisation and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. 50, 349-383.

Bresnick, E.H., Bustin, M., Marsaud, H., Richard-Foy, H., Hager, G.L. (1992). The transcriptionally active MMTV promoter is depleted of histone H1. Nucleic. Acids. Res. 20, 273-278.

Brodeur, G.M., Seeger, R.C., Schwab, M., Varmus, H.E., Bishop, J.M. (1984). Amplifcation of N-myc in untreated human neuroblastomas correlates with advanced disease stage. Science. 224, 1121-1124.

Bugge, T.H., Pohl, J., Lonnoy, O., Stunnenberg, H.G. (1992). RXRα a promiscuous partner of retinoic acid and thyroid hormone receptors. EMBO. J. 11, 1409-1418.

Buratowski, S. (1994). The basics of basal transcription by RNA polymerase II. Cell. 77, 1-3.

Burk, O., Mink, S., Ringwald, M., Klempnauer, K-H. (1993). Synergistic activation of the chicken *mim-1* gene by *v-myb* and C/EBP transcription factors. EMBO. J. 12, 2027-2038.

Carcamo, J., Maldonado, E., Cortes, P., Ahn, M-H., Ha, I., Kasai, Y., Flint, J., Reinberg, D. (1990). A TATA-like sequence located downstream of the transcription initiation site is required for expression of an RNA plymerase II transcribed gene. Genes. Dev. 4, 1611-1622.

Chahal, S.S., Matthews, H.R., Bradbury, E.M. (1980). Acetylation of histone H4 and its role in chromatin structure and function. Nature. 287, 76-79.

Chambers, S.A.M., Shaw, B.R. (1984). Levels of histone H4 deactylation decrease dramatically during sea urchin embryonic development and correlate with cell doubling in rate. J. Biol. Chem. 259, 13458-13463.

Chan, S., Attisano, L., Lewis, P.N. (1988). Histone H3 thiol reactivity and acetyltransferases in chicken erythrocyte nuclei. J. Biol. Chem. 263, 15643-15651.

Chen, Z. , Morgan, R., Baer, M.R., Ligorsky, R., Sandberg, A.A. (1991). Translocation (3;21) characterises crises in myeloid stem cell disorders. Cancer. Gener. Cytogenet. 57, 153-159.

Chicoine, L.G., Richman, R., Cook, R.G., Gorovsky, M.A., Allis, C.D. (1987). A single histone acetyltransferase from Tetrahymena macro nuclei catalyses deposition-related acetylation of free histones and transcription -related acetylation of nucleosomal histones. J. Cell. Biol. 105, 127-135.

Chodosh, L.A., Baldwin, A.S., Carthew, R., Sharp, P.A. (1988). Human CCAAT binding proteins have heterolgous subunits. Cell. 55, 11-24.

Choy, B., Green, M.R. (1993). Eukaryotic activators function during multiple steps of preinitiation complex assembly. Nature. 366, 531-536.

Christy, B.A., & Nathans, D. (1989). DNA binding site of the growth factor-inducible protein Zif268. Proc. Natl. Acad. Sci. USA. 85, 8737-8741.

Chung, S.D., Alavi, N., Livingston, D., Hiller, S., Taub, M. (1982). Characterization of primary rabbit kidney cultures that express proximal tubule functions in a hormonally defined medium. J. Cell. Biol. 95, 118-126.

Clark, K.L., Halay, E.D., Lai, E., Burley, S.K. (1993). Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. Nature. 364, 412-420.

Cohen, R.B., Sheffery, M., Kim, C.G. (1986). Partial purification of a nuclear protein that binds to the CCAAT box of the mouse alpha 1 globin gene. Mol. Cell. Biol. 6, 821-832.

Compere, S.J., Palmiter, R.D. (1981). DNA methylation controls the inducibility of the mouse metallothionein gene in lymphoid cells. Cell. 25, 233-240.

Cordingley, M.G., Riegel, A.T., Hager, G.L. (1987). Steroid-dependent interaction of transcription factors with the inducible promoter of mouse mammary tumour virus *in vivo*. Cell. 48, 261-270.

Courey, A.J., Tjian, R. (1988). Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell. 55, 887-898.

Cousens, D.J., Greaves, R., Golding, C.R., O'Hare, P. (1989). The C-terminal 79 amino acids of the herpes simplex virus regulatory protein, Vmw65, efficiently activate transcription in yeast and mammalian cells in chimeric DNA-binding proteins. EMBO.J. 8, 2337-2342.

Cross, M., Dexter, T.M. (1991). Growth factors in development, transformation, and tumourigenesis. Cell. 271-280.

Crossley, M.A., Tsang, A.P., Bieker, J.J., Orkin, S. (1994). Regulation of the erythroid Kruppel-like factor (EKLF) gene promoter by the erythroid transcription factor GATA-1. J. Biol. Chem. 269, 15440-15444.

Daley, G.Q.A. Van Etten, R., Baltimore, D. (1990). Induction of chronic myelogeneous leukaemia in mice by the P210bcr/abl gene of the Philedelphia chromosome. Science. 247, 824-830.

Dalla-Favera, R., Bregni, M., Erickson, J., Patterson, D., Gallo, R.C., Croce, C.M. (1982). Human *c-myc* oncogene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc. Natl. Acad. Sci. USA. 79, 7824-7827.

Dalla-Favera, R., Martinotti, S., Gallo, R., Erickson, J., Croce, C. (1983). Translocations and rearrangements of the *c-myc* oncogene locus in human undifferentiated B-cell lymphomas. Science. 219, 963-997.

Davis, R.J., D'Cruz, C.M., Lovell, M.A., Biegel, J.A., Barr, F.G. (1994). Fusion of PAX7 to FKHR by the variant t(1;13)(p36;q14) translocation in alveolar rhabdomyosarcoma. Cancer Res, 54, 2869-2872.

de Boer, E., Antoniou, M., Mignotte, V., Wall, L., Grosveld, F. (1988). The human beta-globin promoter: nuclear protein factors and erythroid specific induction of transcription. EMBO. J. 7, 4203-4212.

de Groot, R.P., den Hertog, J., Vandenheede, J.R., Goris, J., Sassone-Corsi, P. (1993) Multiple and cooperative phosphorylation events regulate the CREM activator function. EMBO. J. 12, 3903-3911.

de The, H., Chomienne, C., Lanotte, M., Degos, L., Dejean, A. (1990). The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor α gene to a novel transcribed locus. Nature. 347, 558-561.

Delatrre., O., Zuckman, J., Plougatel, B., Desmaze, C., Melot, T., Peter, M., Kovar, I., Joubert, P., de Jong, P., Rouleu, G., Aurias, A., Thomas, G. (1992). Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. Nature. 359, 162-165.

Delegeane, A.M., Ferland, L.H., Mellon, P.L. (1987). Tissue-specific enhancer of the human glycoprotein hormone a-subunit gene: dependence on cyclic AMP-inducible elements. Mol. Cell. Biol. 7, 3994-4002.

Delwel, R., Funabiki, T., Kreider, B.L., Morishita, K., Ihle, J.N. (1993) Four of the seven zinc fingers of the Evi-1 myeloid transforming gene are required for sequence-specific binding to GA(C/T)AAGA(T/C)AAGATAA. Mol. Cell. Biol. 13, 4291-4300.

DePamphilis, M.L (1988). Transcriptional elements as components of origin of eukaryotic DNA replication. Cell. 52, 635-638.

DePamphilis, M.L. (1993). How transcription factors regulate origins of DNA replication in eukaryotic cells. Trends. Cell. Biol. 3, 161-167.

Deutsch, U., Gruss, P. (1991). Semin. Dev. Biol. 2, 413-424.

Dexter. T.M., Garland, J., Scott, D., Scholnick, E.M., Metcalf, D. (1980). Growth of factor-dependent haemopoietic precursor cell lines. J. Exp. Med. 152, 1036-1047.

Donaldson, L.W., Peterson, J.M., Graves, B.J., McIntosh, L.P. (1994). Secondary structure of the ETS domain places murine Ets-1 in the superfamily of winged helix-turn-helix DNA-binding proteins. Biochemistry. 33, 13509-13516.

Donze, D., Townes, T.M., Bieker, J.J. (1995). Role of erythroid kruppel-like factor in human γ to β–globin gene switching. J. Biol. Chem. 270, 1955-1959.

Dressler, G.R., Douglass, E.C. (1992). Pax-2 is a DNA binding protein expressed in embryonic kidney and Wilms' tumour. Proc. Natl. Acad. Sci. 89, 1179-1183.

Dressler, G.R., Wilkinson, J.E., Rothenpierler, U.W., Patterson, L.T., William-Simons, L., Westphal, H. (1993) Deregulation of Pax-2 expressin in transgenic mice generates sever kidney abnormalities.

Dube, I.D., Kamel-Reid, S., Yuan, C.C., Lu, M., Wu, X., Corpus, G., Raimondi, S.C., Crst, W.M., Carroll, A.J., Minowada, J., Baker, J.B. (1991). A novel human homeobox gene lies at chromosome 10 breakpoint in lymphoid neoplasias with chromosomal translocation t(10;14). Blood. 78, 2996-

Durrin, L.K., Mann, R.K., Kayne, P.S., Grunstein, M. (1991). Yeast histone H4 N-terminal sequence is required for promoter activation *in vivo*. Cell. 65, 1023-1031.

Dynan, W.S. (1986). Trends. Genet. 2, 196-197.

Ebert, T., Bander, N.H., Finstad, C.L., Ramsawak, R.D., Old, L.J. (1990). Establishment and characterization of human renal cancer and normal kidney cell lines. Cancer Res. 50, 5531-5536.

Ekblom, P. (1981). Determination and differentiation of the nephron. Med. Biol. 59, 139-160.

Elgin, S. C. R. (1988). The formation and function of DNAse1 hypersensitve sites in the process of gene activation. J. Biol. Chem. 263, 19259-19262.

Emerson, B.M., Felsenfeld, G. (1984). Specific factor conferring nulease hypersensitivity at the 5' end of the chicken adult β-globin gene. Proc. Natl. Acad. Sci. USA. 81, 95-99.

Emili, A., Greenblatt, J., Ingles, J.C. (1994). Species-specific interaction of the glutamine-rich activation domains of Sp1 with the TATA box-binding protein. Mol. Cell. Biol. 14, 1582-1593.

Evans, T., Felsenfeld, G., Reitman, M. (1990). Control of globin gene transcription. Ann. Rev. Cell. Biol. 6, 95-124.

Faisst, S., Meyer, S. (1992). Compilation of vertebrate-encoded trnascription factors. Nucl. Acid. Res. 20, 3-26.

Fascher, K.D., Schmitz, J., Horz, W. (1990). Role of trans-activating proteins in the generation of actin chromatin at the PHO5 promoter in S. cerevisiae. EMBO. J. 9, 2523-2528.

Faulkner, F.G., Zachau, H.G. (1984). Correct initiation of an immunoglobulin k gene requires an upstream fragment containing conserved sequence elements. Nature. 310, 71-74.

Feinstein, E., Cimino, G., Gale, R.P., Canaani, E. (1992). Initiation and progression of chronic myelogenous leukaemia. Leukaemia. Suppl. 1, 37-43.

Felsenfeld, G., McGhee, J.D. (1986). Structure of 30nm fiber. Cell. 44, 375-377.

Felsenfeld, G. (1992). Chromatin as an essential part of the transcriptional mechanism. Nature. 355, 219-224.

Feng, W.C., Southwood, C.M, Bieker, J.J. (1994). Analysis of b-thalasemia mutant DNA interactions with erythroid kruppel-like factor (EKLF), an erythroid cell-specific transcription factor. J. Biol. Chem. 269, 1493-1500.

Fichelson, S., Dreyfus, F., Berger, R., Melle, J., Bastard, C., Miclea, M., Gisselbrecht, S. (1992). Evi-1 expression in leukaemic patients with rearrangements of thew 3q25-q28 chromosomal region. Leukaemia 6, 93-99.

Foley, K.P., Engel, J.D. (1992). Individual stage selector element mutations lead to reciprocal changes in β- vs ε-globin gene transcription: genetic confirmation of promoter competition during globin gene swithching. Genes. Dev. 6, 730-744.

Fong, T.C., Emerson, B.M. (1992). The erythroid-specific protein cGATA-1 mediates distal enhancer activity through a specialised b-globin TATA box. Genes. Dev. 6, 521-532.

Forman, B.M., Umesono, K., Chen, J., Evans, R.M. (1995). Unique response pathways are established by allosteric interactions among nuclear hormone receptors. Cell. 81, 541-550.

Frankel, A.D., Pabo, C.O. (1988). Fingering too many proteins. Cell. 53, 675.

Fraser, P., Pruzina, S., Antoniou, M., Grosveld, F. (1993). Each hypersensitive site of the human β-globin locus control region confers a different developmental pattern oof expression upon globin genes. Genes. Dev. 7, 106-113.

Freemont, P.S. (1993). The RING finger. A novel protein sequence motif related to the zinc finger. Ann. N.Y. Acad. Sci. 684, 174-192.

Freyd, G., Kim, S.K., Horvittz, H.R. (1990). Novel cysteine-rich motif and homeodomain in the product of the Caenorhabditis elegans cell lineage gene lin-11. Nature. 344, 876-879.

Fujii, T., Pichel, J.G., Taira, M., Toyama, R., Dawid. I.B., Westphal, H. (1994). Expression patterns of murine *lim* class homeobox gene lim1 in the developong brain and excretory systems. Dev. Dynam. 199, 73-83.

Funabiki, T., Kreider, B.L., Ihle, J.N. (1994). The carboxyl domain of the zinc fingers of the Evi-1 myeloid transforming gene binds a consensus sequence of GAAGATGAG. Oncogene. 9, 1575-1581.

Gasser, S.M., Laemelli, U.K. (1986). The organisation of chromatin loops: characterisation of a sacffold attachment site. EMBO. J. 5, 511-518.

Gardiner-Garden, M., Frommer, M. (1987). CpG islands in vertebrate genomes. J. Mol. Biol. 196, 261-282.

Gerster, T., Matthias, P., Thali, M., Firicny, J., Schaffner, W. (1987). Cell type specificity of the immunoglobulin heavy chain gene enhancer. EMBO. J. 6, 1323-1330.

Ghosh, S., Baltimore, D. (1990). Activation in vitro of NF-κB by phosphorylation of its inhibitor IκB. Nature. 344, 678-682.

Gilbert, S.F., Migeon, B.R. (1975). D-valine as a selective agent for normal human and rodent epithelial cells in culture. Cell. 5, 11-17.

Gill, G. (1992). Complexes with a common core. Curr. Biol. Vol. 2. 10, 565-567.

Gill, G., Tjian, R. (1992). Eukaryotic coactivators associated with the TATA box binding protein. Curr. Opin. Gen. Dev. 2, 236-242.

Gill, G., Pascal, E., Tseng, Z.H., Tjain, R. (1994). A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII110 component of the Drosophila TFIID complex and mediates transcription activation. Proc. Natl. Acad. Sci. USA. 91, 192-196.

Gonzalez,G.A., Montminy, M.R. (1989). Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of the CREB ar Ser 133. Cell. 59, 675-680.

Goodrich, J.A., Hoey, T., Thut, C.J., Admon, A., Tjian, R. (1993). Drosophila $TAF_{II}40$ interacts with both VP16 activation domain and the basal trasnscription factor TFIIB. Cell. 75, 519-530.

Goddrich & Tjian, P. (1994). Transcription factors IIE and IIH and ATP hydolysis direct promoter clearance by RNA polymerase II. Cell. 77, 145-156.

Gong, Q., Dean, A. (1993). Enhaner-dependent transcription of the ε-globin promoter requires promoter-bound GATA-1 and enhancer bound AP1/NF-E2. Mol. Cell. Biol. 13, 911-917.

Graves, B. J., Johnson, P.F., McKinght, S.L. (1986). Homologous recognition of a promoter dom,ain common to the MSV LTR and the HSV tk gene. Cell 44, 565-76.

Greenberger, J.S., Sakakeeny, M.A., Humphries, R.K., Eaves, C.J., Eckner, R.J. (1983). Demonstration of permanent factor-dependent mutipotential (erythroid/neutrophil/basophil) haematopoietic progenitor cell lines. Proc. Natl. Acad. Sci. USA. 80, 2931-2935.

Grobstein, C. (1957). Some transmisson characterisitcs of the tubule inducing influence on the mouse metanephrogenic mesenchyme. Exp. Cell. Res. 13, 575-587.

Gross, D.S., Garrard, W.T. (1988). Nuclease hypersensitive sites in chromatin. Ann. Rev. Biochem. 57, 159-197.

Grosveld, F., Blom von Assendelft, G., Greaves, D.R., Kollias, G. (1987). Position-independent high-level expression of the human β–globin gene in transgenic mice. Cell. 51, 975-85.

Gruss, P., Walther, C. (1992). Pax in development. Cell. 69, 719-722.

Gumcio, D.L., Rood, K.L., Gray, T.A., Riordan, M.F., Sartor, C.I., Collins, F.S. (1988). Nuclear protiens that bind the human gamma-globin gene promoter: alterations in binding produced by point mutations associated with hereditary persistence of faetal haemoglobin. Mol. Cell. Biol. 8, 5310-5322.

Gurente, L., Bermingham-McDonogh. (1992) Conservation and evolution of trnascripitonal mechanisms in eukaryotes. Trens. Genet. 8, 26-32.

Haber, D.A., Buckler, A.J., Glaser, T., Call, K.M., Pelletier, J., Sohn, R.L., Douglass, E.C., Housman, D.E. (1990). An internal deletion within an 11p13 zinc-finger gene contributes to the development of Wilms' tumour. Cell. 61, 1257-1269.

Hagiwara, M., Albert, A., Brindle, P., Meinkoth, J., Feramisco, J., Deng., .. (1992). Transcritpional attenuation following cAMP induction requires RP-1-mediated dephosphorylation of CREB at ser 133. Cell. 59, 675-680.

Ham, J., Steger, G., Yaniv, M. (1992). How do eukaryotic activator proteins stimulate the rate of transcription by RNA polymerase II ? FEBS lett. 307, 81-86.

Han, M., Kim, U.J., Kayne, P., Grunstein, M. (1988). Depletion of histone H4 and nucleosomes activates the PHO5 gene in Saccharomyces cereiviae. EMBO. J. 7, 2221-2228.

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

Harrison, S.C., Aggarwal, A.K. (1990). DNA recognition by proteins with the helix-turn-helix motif. Ann. Rev. Biochem. 59, 933-69.

Harrison, S.C. (1991). A structural taxonomy of DNA-binding domains. Nature. 353, 715-719.

Harrison, C.J., Bohm, A.A., Nelson, H.C.M. (1994). Crystal structure of the DNA-binding domain of the heat shock transcription factor. 263, 224-227.

Harvey, R.P., Robins, A.J., Wells, J.R.E. (1982). Independently evolving chicken histone H2B genes; identification of a ubiquitious H"B-specific element. Nucleic. Acids. Res. 10, 7851-7863.

Hayward, W.S., Neel, B.G., Astrin, S.M. (1981). Activation of a cellular *onc* gene by promoter insertion in ALV-induced lymphoid leukosis. Nature 290, 475-480.

He, G-P., Muise, A., Wu-Li, A., Ro, H-S. (1995). A eukaryotic transcriptional repressor with carboxypeptidases activity. Nautre. 378, 92-96.

Hebbes, T.R., Thorne, A.W., Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. EMBO. J. 7, 1395-1402.

Herr (1988). The POU domain: a large conserved region in the mammalian *pit-1*, *oct-1*, *oct-2*, and *Caenorhabditis elegans unc*-86 gene products. Genes. Dev. 2: 1513-1516.

Hernandez, N (1993). TBP, a universal eukaryotic transcription factor? Genes. Dev. 7, 1291-1308.

Herzinger, D., Koseki, C., Mikawa, T., Al-Awqati, Q. (1992). Metanephric mesenchyme contains multipotent stem cells whose fate is restricted after induction. Development. 114, 565-572.

Higurashi, M., Cole, R.D. (1992). The combination of DNA methylation and H1 histone binding inhibits the action of a retriction nuclease on plasmid DNA. J. Biol. Chem. 266, 8619-8625.

Hill, C.S., Thomas, J.O. (1990). Core histone DNA interactions in sea urchin sperm chromatin: The N-terminal tail of H2B interacts with linker DNA. Eur. J. Biochem. 187, 145-153.

Hirschhorn, J.N., Brown, S.A., Clark, C.D., Winston, F. (1992). Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. Genes. Dev. 6, 2288-2298.

Hoey, T., Weinzierl, R.O.J., Gill, G., VChen, J-L., Dynlacht, B.D., Tjian, R. (1993). Molecualr cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. Cell. 72, 247-260.

Hollenberg, S.M., Evans, R.M. (1998). Multiple and cooperative *trans*-activation domains of the human glucocorticoid receptor. Cell. 55, 899.

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

Horikoshi, N.Maguire, K., Kralli, A., Maldonada, E., Reinberg, D., Weinmann, R. (1991). Direct interaction between Adenovirus E1A protein and the TATA box binding transcription factor IID. Proc. Natl. Acad. Sci. 88, 5124-5128.

Hoyt, P.R., Bartholomew, B., Gamar, L.W., Wright, C.V.E., Potter, S., Ihle, J.N., Mucenski, M.L. (1996) manuscript submitted.

Inostroza, J.A., Marmelstein, F.H., Ha, I., Lane, W.S., Reinberg, D. (1992). Dr1, a TATA-binding protein associated phosphoprotein and inhibitor of class II gene transcription Cell. 70, 477-490.

Isseman, I.,Prince, R.A., Tugwood, J.D., Green, S. (1993). The retinoid X receptor enhances the function of the peroxisome proliferator activated receptor. Biochemie 75, 251-256.

Isreal., A., Le Bail, O., Hatat, D., Colette, J., Kieran, M., Logeat, F., Wallach, D., Fellous, M., Kourilsky, P. (1989). TNF stimulates expression of MHC class 1 genes by inducing an NF-κB-like enhancer binding activity which displaces constitutive factors. EMBO. J. 8, 3793-3800.

Jackson, P.D., Evans, T., Nickol, J.M., Felsenfeld, G. (1989). Developmental modulation of protein binding to the beta-globin gene regulatory sites within chicken erythrocyte nuclei. Genes. Dev. 3, 1860-1873.

Jaenish, R., Jahner, D., Nobis, P., Simon, I., Loher, J., Harbers, K., Grotkopp, D. (1981). Chromosomal position and activation of retroviral genomes inserted into the germ line of mice. Cell. 24, 519-529.

Jans, D.A. et al (1991). The p34cdc2-mediated phosphorylation at T124 inhibits nuclear import of SV-40 T antigen proteins. J. Cell. Biol. 115, 1203-1212.

Jantzen, H.M., Strahle, U., Gloss, B., Stewart, F., Schmid, W., Boshart, M., Miksicek, R., Schutz, G. (1987). Cooperativity of glucocorticoid response elements located far upstream of the tyrosine aminotransferase gene. Cell 49, 29-38.

Jarman, A.P., Higgs, D.R. (1988). Nuclear scaffold attahcment sites in the human globin gene complexes. EMBO. J. 7, 3337-3344.

Jat, P.S., Sharp, P.A. (1989). Cell lines established by a temperature sensitive simian virus 40 large T-antigen gene are growth restricted at the non-permissive temperature. Mol. Cell. Biol. 1672-1681.

Jat, P.S., Noble, M.D., Ataliotis, P., Tanaka, Y., Yannoutsos, N., Larsen, L., Kioussis, D. (1991). Direct derivation of conditionally immortal cell lines from an H-2Kb-tsA58 transgenic mouse. Proc. Natl. Acad. Sci. USA. 88, 5096-5100.

Johnson, E.M., Sterner, R., Allfrey, V.G. (1987). Altered nucleosomes of active nucleolar chromatin contain accessible histone H3 in its hyperacetylated forms. J. Biol. Chem. 262, 5943-5946.

Johnson, F..B. Krasnow, M.A. (1992). Differential regulation of transcription preintiation complex assembly by activator and repressor homeo domain proteins. Genes. Dev. 6, 2177-2189.

Johnson, L.M., Kayne, P.S., Kahn, E.S., Grunstein, M. (1990). Genetic evidence for an interaction between SIR3 and histone H4 in the expression of the silent mating loci in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA. 87, 6286-6290.

Jones, K.A., Yamamoto, K.R., Tjian, R. (1985). Two distinct transcription factos bind to the HSV thymidine kinase promoter *in vitro*. Cell. 42, 559-572.

Jordan, S.R., Pabo, C.O. (1988). Structure of the lambda comlex at 2.5 A resolution: details of the repressor-operator interactions. Science. 242, 893-899

Jost, J-P., Saluz, H-P., Pawlak, A. (1991). Esradiol down regulates the binding activty of an avian vitellogenin gene repressor (MDCP-2) and triggers a gradual demethylation of the mCpG pair of its DNA binding site. Nucleic Acids Res. 19, 57771-5775.

Karin, M. (1989). Complexities of gene regulation by cAMP. Trends. Genet. 5, 65-67.

Karlsson, O., Thor, S., Norberg, T., Ohisson, H., Edlund, T. (1990). Insulin gene enhancer binding protein lsl-1 is a member of a novel class of proteins containing both a homeo- and Cys-His domain. Nature. 344, 879-882.

Karpen, G.H. (1994). Position-effect variegation and the new biology of heterochromatin. Curr. Opin. Gen. Dev. 4, 281-291.

Karpov, V.L., Preobrazhenskaya, O.V., Mirzabekov, A.D. (1984). Chromatin structure of hsp 70 genes, activated by heat shock: selective removal of histones from the coding region and their absence from the 5' region.

Kaufman, L.S., Watson, J.C., Thompson, W.F. (1987). Proc. Natl. Acad. Sci. USA. 84, 1550-1554.

Kaye, J.S., Pratt-Kaye, S., Bellard, M., Dretzen, G., Bellard, F., Chambon, P. (1986). Steroid hormone dependence of four DNase 1-hypersensitive regions within the 7000bp 5' flanking segment of the ovalbumin gene. EMBO. J. 5,277--285.

Kellum, R., Schedl, P. (1991). A positional-effect assay for boundaries of higher order chromosomal domains. Cell. 64, 941-950.

Khana-Gupta, A., Lopingo, M.C., Savinelli, T., Zibello, T., Berliner, N., Perkins, A.S. (1996). Retroviral insertional activation of the *EVI1* oncogene does not prevent G-CSF-induced maturation of the mrine pluripotent myeloid cell line 32Dcl3. Oncogene 12, 563-569.

Kim, T.K., Roeder, R. (1993). Transcriptional activation in yeast by the proline-rich activation domain of human CTF-1. J. Biol. Chem. 268, 20866-20869.

Kingsley, C., Winoto, A. (1992). Cloning of GT box-binding proteins: a novel Sp1 multigene family regulating T-cell receptor gene expression. Mol. Cell. Biol. 12, 4251-4261.

Kissinger, C.R., Liu, B., Martin-Bianco, E., Kornberg, T., Pabo, C. (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: A framework for understanding homeodomain-DNA interactions. Cell. 63, 579-590.

Klehr, D., Mass, K., Bode, J. (1991). Scaffold-attached regions from the human interferon b domain can be used to enhance the stable expression of genes under the control of various promoters. Biochemistry. 30, 1264-1270.

Knezetic, J.A., Jacob, G.A., Luse, D.S. (1988). Assembly of RNA polymerase II preinintiation complexes before assembly of nucleosomes allows efficient intiation of transcrption on nucleosomal templates. Mol. Cell. Biol. 8, 3114-3121.

Ko, L.J., Engel, J.D. (1993). DNA-binding specificities of the GATA transcription factor family. Mol. Cell. Biol. 13, 4011-4022.

Konopka, J.B., Watanabe, S.M., Witte, O.N. (1984). An alteration of the human c-abl protein in K562 leukaemia cells unmasks associated tyrosine kinase activity. Cell. 37, 1035-1042.

Kreidberg, JA., Sariola, H., Loring, J.M., Maeda, M., Pelletier, J., Housman, D., Jaenisch, R. (1993). WT-1 is required for early kidney development. Cell. 74, 679-691.

Kreider, B.L., Orkin, S.H., Ihle, J.N. (1993) Loss of erythropoietin responsiveness in erythroid progenitors due to expression of the *Evi-1* myeloid-transforming gene. Proc. Natl. Acad. Sci. USA. 90, 6454-6548.

Kristie, T.M., Sharp, P.A. (1990). Interaction of the Oct-1 POU subdomains with specific DNA sequences and with the HSV α-*trans*-activator protein. Genes Dev. 4, 2383-2396.

Kurokawa, M., Ogawa, S., Tanaka, T., Mitani, K., Yazaki, Y., Witte, O.N., Hirai, H. (1995) The AML1/Evi-1 fusion protein in the t(3;21) translocation exhibits transforming activity of Rat1 fibroblasts with dependence on the Evi-1 sequence. Oncogene. 11, 833-840.

Laemmli, U.K., Kas, E., Poljak, L., Adachi, Y. (1992). Scaffold-associated regions: *cis*-acting determinants of chromatin structural loops and functional domains. Curr. Opin. Gen. Dev. 2, 275-285.

Lamb, P., McKnight, S.L. (1991). Diversity and specificity in trnascriptional regulation: The benefits of heterotypic dimerisation. Trends. Biochem. Sci. 16, 417-422.

Landsman, D., Bustin, M. (1993). A signiture for the HMG-1 box DNA-binding proteins. Bioessays. 15, 539-46.

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

LeBeau, M., Albain, K.S., Larson, R.A., Vardimam, J.W., Davis, E.M., Blough, R.R., Glomb, H.M., Rowley, J.D. (1986). Clinical and cytogenetical correlations of 63 pateints with therapy related myelodysplastic syndromes abd acute nonlymphocytic leukaemia: Furhter evidence for charactersitic abnormalites of chromosome no. 5 and 7. J.Clin. Oncol. 8, 325.

Lee, D.Y., Hayes, J.J., Pruss, D., Wolffe, A.P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. Cell. 72, 73-84.

Lehtonen, E., Jalenko, H., Laitinen, L., Meittinen, A., Ekblom, P. Saxen, L. (1983). Differentiation of the metanephric tubules following a short induction pulse. Roux' Arch. Dev. Biol. 192, 145-151.

Leid, M., Kastner, P., Lyons, R., Nakshatri, H., Saunders, M., Zacharewski, T., Chen, J-Y., Staub, A, Garnier, J-M., Mader, S., Chambon, P. (1992). Purification, cloning and RXR indentity of the HeLa cell factor with which RAR or TR heterodimerises to bind target sequences effieciently. Cell 68, 377-395. Eratum: Cell 71(5).

Lewin, B. (1990). Genes IV. Oxford university press, New York

Lewin, B. (1994). Chromatin and gene expression: Constant questions but changing answers. Cell. 79,397-406.

Licht, J.D., Ro, M., English, M.A., Grossel, M., Hansen, U. (1993) Selective repression of transcription activators at a distance by the Drosophila kruppel protein. Proc. Natl. Acad. Sci. USA. 90, 11361-11365.

Lilley, D.M.J. (1992). DNA-protein interactions. HMG has DNA wrapped up. Nature. 357, 282-283.

Li, E., Bestor, T.H., Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 69, 915-926.

Lin , Y-S., Green, M.R. (1991). Mechanism of action of an acidic transcriptional activator *in vitro*. Cell 64, 971-981.

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

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

Llaing, H., Olejniczac, E.T., Mao, X, Nettesheim, D.G., Yu, L., Thompson, C.B., Feswick, S.W. (1994). The secondary structure of the ETS domain of human fli-1 resembles that of the helix-turn-helix DNA-binding motif of E.coli CAP. Proc. Natl. Acad. Sci. USA. 91, 11655-11659.

Look, A.T. (1994) Oncogenis role of "master" transcription factors in human leukaemias and sarcoma: a developmental model. Adv. Cancer. Res. 25-57.

Lugo, T.G., Witte, O. N. (1989). The BCR-ABL oncogene transforms RAT-1 cells and cooperates with v-myc. Mol. Cell. Biol. 9, 1263-1270.

Luisi, B.F., Xu, W.X., Otwinowski, Z., Freedman, L.P., Yamamoto, K.R. (1991). Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. Nature. 352, 497-505.

Ma, J., Ptsahne, M. (1987). Deletion anlaysis of GAL4 defines two transcriptional activating segments. Cell. 48, 847-853.

Madden, S.L., Cook, D.M., Morris, J.F., Gashler, A., Sukhatme, V.P., Rauscher III, F.J. (1991). Transcriptional repression mediated by WT1 Wilms' tumour gene product. Science. 253, 1550-1553.

Magrauer, G., Ekblom, P. (1991). Contrasting expression patterns of three members of the myc family of protooncogenes in the developing and adult mouse kidney. J. Cell. Biol. 112, 13-25.

Majumder, S., DePamphilis, M.L. (1995) A unique role of enhancers is revealed during early mouse development. Bioessays 17, 10 879-889.

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

Marmostein, R., Harrision, S.C. (1994). Crystal structure of a PPR1-DNA complex: DNA recognition by proteins containing a Zn_2-Cys_6 binuclear cluster. Genes. Dev. 8, 2504-2512.

Martin, D.I.K., Orkin, S.H. (1990). Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf 1. Genes Dev. 4, 1886-1898.

Matsugi, T., Morishita, K., Ihle, J.N. (1990) Identification, nuclear localisation, and DNA binding activity of the zinc finger protein encoded by the Evi-1 myeloid transforming gene. Mol. Cell. Biol. 10, 1259-1264.

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

McGhee, J.D., Felsenfeld, G. (1980). Nucleosome structure. Annu. Rev. Biochem. 49, 1115-1156.

McGhee, J.D., Wood, W.I., Dolan, M., Enjel, J.D., Felsenfeld, G. (1981). A 200 base pair region at the 5' end of the chicken adult beta-globin gene is accessible to nuclease digestion. Cell. 27, 44-55.

McGhinnis, W., Garber, J., Wirz, A., Kurooiwa, A., Gehring, W.J. (1984). A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. Cell. 37, 403-408.

McKay, D.B., Steitz, T.A. (1981). Structure of catabolite gene activator protein at 2.9 A resolution suggests binding to left-handed B-DNA. Nature. 290, 744-749.

Meehan, R., Lewis, J., Cross, S., Nan, X., Jeppesen, P., Bird, A. (1992). Transcriptional repression by methylation of CpG. J. Cell. Sci. 16, 9-14.

Meeker, T.C., Loeb, J., Ayres, M., Sellers, W. (1990). The human Pim-1 gene is selectively transcribed in different hemato-lymphoid cell lines in spite of a G+C-rich housekeeping promoter. Mol. Cell. Biol. 10, 1680-1688.

Merika, M., Orkin, S.H. (1995). Functional synergy and physical interactions of the erythroid trnascription factor GATA-1 and the Kruppel family proteins Spl and EKLF. Mol. Cell. Biol. 15, 2437-2447.

Merino, A., Madden, K., Lane, W.S., Champoux, J., Reinberg, D. (1993). DNA topoisomerase I is invovled in both repression and activation of transrciption. Nature. 365, 227-232.

Migliaccio, G., Migliaccio, A., Kreider, B., Rovera, G., Adamson, J. (1989). Selection of lineage restricited cell lines immortalised at different stages of heamopoietic differentiation from the murine cell line 32D. J. Cell. Biol. 109, 833-841.

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P.A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L.M., Ding, W. (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA-1. Science. 266, 66

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

Miller, I.J., Bieker, J.J. (1993). A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the kruppel family of nuclear proteins. Mol. Cell. Biol. 13, 2776-2786.

Mirkovitch, J., Mirault, M.E., Laemmli, U.K. (1984). Organisation of the higher-order chromation loop: specific DNA attachment sites on the nuclear scaffold. Cell. 39, 223-232.

Mitani, K., Ogawa, S., Tanaka, T., Miyoshi, H., Kurokawa, M., Mano, H., Yazaki, Y., Ohki, M., Hirai, H. (1994). Generation of the AML1-EVI-1 fusion gene in the t(3;21) (q26;q22) causes blastic crisis in chronic myelocytic leukaemia. EMBO. 504-510

Mitchell, P.J., Timmons, P.M., Hebert, J.M., Rigby, P.W.J., Tjian, R. (1991). Transcrip[tion factor AP-2 is expressed in neural crest cell lineages during mouse embryogenesis. Genes. Dev. 5, 105-119.

Miyoshi, H., Schimizu, K., Kozu, T., Maseki, N., Kameko, Y., Ohki, M. (1991). The t(8;21) breakpoints on chromosome 21 in acute myeloid leukamia clustered within a limited region of a novel gene AML1. Proc. Natl. Acad. Sci. USA. 88, 10431.

Moreau-Gaschelin, F., Tavitian A., Tambourin, P. (1988). Spi-1 is aputative oncogene invirally induced murine erythroleukaemias. Nature. 331, 277-280.

Morishita, K., Parker, D.S., Mucenski, M., Jenkins, N.A., Copeland, N.G., Ihle, J.N. (1988) Retroviral activation of a novel gene encoding a zinc-finger protein in IL-3-dependent myeloid leukaemia cell lines. Cell. 831-840.

Morishita, K., Parganas, E., Douglass, E.C., Ihle, J.N. (1990a) Unique expression of the human *Evi-1* gene in an endometrial carcinoma cell line: sequence of cDNAs and structure of the alternatively spliced transcripts. Oncogene. 5, 963-971.

Morishita, K., Parganas, E., Parham, D., Matsugi, T., Ihle, J.N. (1990b). The *Evi-1* zinc finger myeloid transforming gene is normally expressed in the kidney and in the developing oocytes. Oncogene. 5, 1419-1423.

Morishita, K., Parganas, E., Bartholomew, C., Sacchi, N., Valentiine, M.B., Raimondi, S.C., LeBeau, M.M., Ihle, J.N. (1990c). The human Evi-1 gene is located on chromosome 3q24-q28 but is not rearranged in three cases of acute nonlymphocytic leukamias containing t(3;5)(q25;34) translocations. Oncogene Res. 5, 221-231.

Morishita, K., Parganas, E., Matsugi, T., Ihle, J.N. (1992a). Expression of the *Evi-1* zinc finger gene in 32Dcl3 myeloid cells blocks granulocytes differentiation in response to granulocyte colony-stimulating factor. Mol. Cell. Biol. 12, 183-189.

Morishita, K., Parganas, E., Willman, C.L., Whitaker, M.H., Drabkin, H., Oval, J., Taetle, R., Valentine, M.B., Ihle, J.N. (1992b) Activation of the EVI-1 gene expression in human acute myeloid leukaemias by translocations spanning 300-400 kilobases on chromsome band 3q26. Proc. Natl. Acad. Sci. USA. 89, 3937-3941.

Morishita, K., Suzukawa, K., Taki, T., Ihle, J.N., Yokota, J. (1995) EVI-1 zinc finger protein works as a transcriptional activator via binding to a consensus sequence of GACAAGATAAGATAAN1-28CTCATCTTC. Oncogene 10, 1961-1967

Moser, M., Imhof, A., Pscherer, A., Bauer, R., Amselgruber, W., Sinowatz, F., Hofstadter, F., Schule, R., Buettner, R. (1995). Cloning and characterisation of a second AP-2 transcription factor: AP-2β. 121, 2779-2788.

Mucenski, M.L., Taylor, B.A., Ihle, J.N., Hartley, J.W., Morse III, H.C., Jenkins, N.A. (1988). Identifcation of a common site of viral integration site, Evi-1, in the DNA of AKXD murine myeloid tumours. Mol. Cell. Biol. 8, 301-308.

Mugrauer, G., Ekblom, P. (1991). Contrasting expression patterns of three members of the myc family of proto-oncogenes in the developing and adult mouse kidney. J. Cell. Biol. 112, 13-25.

Murray, E.J., Grosveld., F. (1987). Site specific demethylation in the promoter of human gamma globin gene does not alleviate methylation mediated repression. EMBO. J. 6, 2329-2335.

Nacheva, G.A., Guschin, D.Y., Preobrazhenskaya, O.V., Karpov, V.L., Ebralidse, K.K., Mirzabekov, A.D. (1989). Chnge in the pattern of histone binding to DNA upon transcriptional activation. Cell. 58, 27-36.

Nakajiama, N., Horikoshi, M., Roeder, R.G. (1988). Factors involved in specfic transcription by mammalian RNA polmerase II: purification, gene specificity, and TATA box-promoter interactions of TFIID. Mol. Cell. Biol. 8, 4028-4040.

Nichols, J., Nimer, S.D. (1992). Tanscription factors, translocations and leukaemia. Blood. 80, 2953-2963.

Nichols, M., Weih, F., Scmid, W., DeVack, C., Kowenz-Leutz, E., (1992). Phosphorylation of CREB affects its binding to high and low affinity sites: implications for cAMP-induced gene transcription. EMBO. J. 11, 3337-3346.

Nisson, P.E., Watkins, P.C., Sacchi, N. (1992). Transcriptionally active chimeric gene derived from the fusion of the AML1 gene and a novel gene on chromosome 8 in the t(8:21) leukaemic cells. Cancer Cytogenet. 63, 81.

Noble, M., Groves, A.K., Ataliotis, P., Ikram, Z., Jat, P.S. (1995). The H-2KbtsA58 transgenic mouse: a new toul for the rapid generation of novel cell lines. Transgenic Res. 4, 215-225.

Nucifora, G., Rowley, J.D. (1995). AML1 and the 8;21 and 3;21 translocations in acute and chronic myeloid leukaemia. Blood. 86, 1-14.

Nusslein-Volhard, C., Frohnhofer, H.G., Lehmann, R. (1987). Science. 238, 1675-1681.

Omichinski, J.G., Clore, G.M., Schaad, O., Felsenfeld, G., Trainor, C., Apella, E., Stahl, S.J., Gronenborn, A.M. (1993). NMR structure of a specific DNA complex Zn-containing DNA binding domain of GATA-1. Science. 261, 438-446.

O'Prey, J., Ramsey, S., Chambers, I., Harrison, P. (1993). Transcriptional up-regulation of the mouse cystolic glutathione peroxidase gene in erythroid cells is due to a tissue specific 3' enhancer containing functionally important CACC/GT motifs and binding sites for GATA and ets transcription factors. Mol. Cell. Biol. 13, 6290-6303.

Orkin, S.H. (1992). GATA-binding transcription factors in haemopoietic cells. Blood. 80, 575-581.

O'Shea, E.K., Klemm, J.D., Kim, P.S., Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science. 254, 539-544.

Otten, A., Saunders, M., McKnight, G. (1988). Molecular endocrinology. 2, 143-147.

Ozer, J., Moore, P.A., Bolden, A.H., Lee, A., Rosen, C.A., Lieberman, P.M. (1994). Molecular cloning of the small (γ) subunit of human TFIIA reveals functions critical for activated transcription. Genes. Dev. 8, 2324-2335.

Pabo, C.O., Lewis, M. (1982) The operator-binding domain of the λ repressor: structure and DNA recognition. Nature. 298, 447-443.

Pabo, C.O., Sauer, R.T. (1992). Transcription factors: Structural families and principles of DNA recognition. Annu. Rev. Biochem. 61, 1053-1095.

Parraga, G., Horvath, S.J., Eisen, A., Taylor, W.E., Hood, L., Young, E.T., Klevit, R.E. (1988) Zinc-dependent structure of a single-finger domain of the yeast ADR1. Science. 242, 1489-1492.

Parslow, T.G., Granner, D.K. (1982). Chromatin changes accompany immunoglobulin kaapa gene activation: a potential control region within the gene. Nature. 299, 449-451.

Parslow, T.G., Granner, D.K. (1983). Structure of the a nuclease-sensitive region inside the immunoglobulin kappa gene: evidence for a role in gene regulation. Nucleic. Acids. Res. 11, 4775-4792.

Parslow, T.G., Blair, D.L., Murphy, W.J. (1984). Sructure of the 5' ends of the immunoglobulin genes: a novel conserved sequence. Proc. Natl. Acad. Sci. USA. 81, 2650-2654.

Peterson, C.L., Herskowitz, I. (1992). Characterisation of the yeast SWI1, SWI2, SWI3 genes, which encode a global activator of transcription. Cell. 68, 573-583.

Patterson, L.T., Dressler, G.R. (1994). The regulation of kidney development: new insights from an old model. Curr. Opin. Gen. Dev. 4, 696-702.

Paul, D., Vasil, V., Vasil, I.K., Ferl, R.J. (1987). Proc. Natl. Acad. Sci. USA. 84, 799-803.

Paul, D., Hohne, M., Pinkert, C., Piasceki, N., Ummekmann, E., Brinster, R.L. (1988). Immortalized differentiated hepatocyte lines derived from trangenic mice harbouring SV40 T-antigen genes. Exp. Cell. Res. 175, 354-362

Pena, A., Reddy, C.D., Wu, S., Hickok, N., Reddy, E.P., Yumet, G., Soprano, D.R., Soprano, K.J. (1993). Regulation of the human orthinine decarboxylase expression by the c-Myc-Max protein complex. J. Biol. Chem. 268, 27277-27285.

Pena, A., Wu, S., Hickok, N., Soprano, D.R., Soprano, K.J. (1995). Regulation of the human decarboxylase expression following prolonged quiesence: a role for the c-Myc/Max protein complex. J. Cell. Physiol. 162, 234-245.

Perkins, A.S., Fishel, R., Jenkins, N.A., Copeland, N.G. (1991a) *Evi-1*, a zinc-finger proto-oncogene, encodes a sequence specific DNA-binding protein.

Perkins, A. S., Mercer, J.A., Jenkins, N.A., Copeland, N.G. (1991b). Patterns of *Evi-1* expression in embryonic and adult tissues suggests that *Evi-1* plays an important regulatory role in mouse development. Development. 111, 479-487.

Perkins, A.S., Kim, J.H. (1996). Zinc-fingers 1-7 of Evi-1 fail to bind the GATA motif by itself but require the core site GACAAGATA for binding. J. Biol. Chem. 271, 1104-1110.

Peterson, C.L., Herskkowitz, I. (1992). Characterisation of the yeast *SWI1, SWI2, SWI3,* which encode a global activator of transcription. Cell. 68, 573-583.

Philipsen, S, Talbot, S., Fraser, P., Grosveld, F. (1990). The β-globin dominant control region: hypersensitive site 2. EMBO. J. 9, 2159-2167.

Phi-Van, L., von Kries, J.P., Ostertag, W., Stratling, W.H. (1990). The chicken lysozyme 5' matrix attachment region increases transcription from a heterologous promoter in heterologous cells and dampens position effects on the expression of transfected genes. Mol. Cell. Biol. 10, 2302-2307.

Piedagnel, R., Prie, D., Cassingena, R., Ronco, P., Lelong, B. (1994). SV40 large-T oncogene inhibits transcription of the perlecan-related proteoglycans but stimulates hyaluronan synthesis in a temperature sensitive renal tubule pricipal cell line. J. Biol. Chem. 269, 17469-17476.

Piette, J., Yaniv, M. (1987). Two diffrent factors bind to the a-domian of the polyoma virus enhancer, one of which also interacts with the SV40 and c-fos enhancers. EMBO. J. 6, 1331-1337.

Plumb, M.A., Lobenenkov, V.V., Nicholas, R.H., Wrigh, C.A., Zavou, S., Goowin, G.H. (1986). Characterisation of chicken erythroid nuclear proteins which bind to the nuclease hypersensitive regions upstream of the betaA- and betaH-globin genes. Nucleic. Acids. Res. 14, 7675-7693.

Plumb, M., Goodwin, G.H. (1988). Detection of sequence-specific protein-DNA interactions by the DNA footprinting technique. In J.M. Walker (ed.), Methods in molecular biology, vol. 4. Human Press, Clifton, N.J.

Pruzina, S.O., Hanscombe, O., Whyatt, D., Grosveld, F., Philipsen, S. (1991). Hypersensitive site 4 of the human β-globin locus control region. Nucleic. Acids. Res. 19, 1413-1419.

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

Pugh , B.F., Tjian, R. (1991). Transcription from a TATA-less promoter requires a multi-subunit TFIID complexes. Genes. Dev. 5, 1935-1945.

Pulverer, B.J., Kyriakis, J.M., Avruch, J., Nikolakaki, E., Woodgett, J.R. (1991). Phosphorylation of c-jun mediated by MAP kinases. Nature. 353, 670-674.

Pulverer, B.J., Hughes, K., Frnklin, C.C., Kraft, A., Woodgett, J.R. (1992). Dual modes of c-jun phosphorylation by phorbal esters in U937 leukaemic cells. Oncogene.

Quinn, J., Fyerberg, A.M., Ganster, R.W., Schmidt, M.C., Peterson, C.L. (1996). DNA-binding properties of the yeast SWI/SNF complex. Nature. 379, 844-847.

Rabbits, T.H. (1994). Chromosomal translocations in human cancer. Nature. 372, 143-149.

Raich, N., Clegg, C.H., Grofti, J., Romeo, P-H., Stamatoyannopoulos, G. (1995). GATA1 and YY1 are developmental repressors of the human ε-globin gene. EMBO. J. 14, 801-809.

Rauscher III, F.J., Morris, J.F., Tournay, O.E., Cook, DM., Curran, T. (1990). Binding of the Wilms' tumour locus zinc-finger protein to the EGR-1 consensus sequence. Science. 250, 1259-1262.

Reddy, P.M., Shen, C.J. (1991). Protein-DNA interactions in vivo of an erythroid-specific human b-globin locus enhancer. Proc. Natl. Acad. Sci. USA. 88, 8676-8680.

Richard-Foy, H., Hager, G.L. (1987). Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. EMBO. J. 6, 2321-2328.

Robinson, M.O., Zhou, W., Hokom, M., Danilenko, D.M., Hsu, R-Y., Atherton, R.E., Xu, W., Mu, S., Saris, C.J.M., Swift, S., Elliot, G., Castillo, J.D., Hunt, P., Bosselman. R.A. (1994). Proc. Natl. Acad. Sci. USA. 91, 12798-12802.

Roeder, R.G. (1991). The complexities of eukaryotic transcripiton intiation: regulation of preinitiation complex assembly. Trends Biochem. Sci. 16, 402-408.

Ron, D., Habener, J.F. (1992). CHOP, a developmentally regulated nuclear protein that dimerizes with transcription factors C/EBP and LAP and functions as a dominant-negative inhibito of gene transcription. Genes. Develop. 6, 439-453.

Rothenpieler, U.W., Dressler, G.R. (1993). Pax-2 is required for mesenchyme-to-epithelium conversion during kidney development. Development. 119, 711-720.

Rosales, R., Vigneron, M., Macchi, M., Davidson, I., Xiao, J.H. (1987). EMBO. J. 6, 3015-3025.

Rowley, J.D. (1973). Idnetification of a translocation with quinacrine flouresence in a patient with acute leukaemia. Annal. Genet. 16, 109.

Rowley, J.D. (1990). Recurring chromosome abnormalities in leukaemia and lymphoma. Semin Hematol. 27, 122.

Roy, A.L., Malik, S., Meiisterenst, M., Roeder, R.G. (1993). An alternative pathway for transcription initiation involving TFII-I. Nature. 365, 355-359.

Rubin, C.M., Larson, R.A., Anastasi, J., Winter, J.N., Thangavelu, M.., Vardiman, J., Rowley, J.D., LeBeau, M.M. (1990). t(3;21)(q26;q22): a reccuring chromosomal abnormality in therapy related myelodysplastic sundrome and acute myeloid leukaemia. Blood. 76, 2594-2598.

Russel, M., Thompson, F., Spier, C., Taetle, R. (1993). Expression of the EVI1 gene in chronic myelogenous leukaemia in blast crises. Leukaemia. 7, 1654-1657.

Russel, M., List, A., Greenberg, P., Woodward, S., Glinsmannn, B., Parganas, E., Ihle, J.N., Taetle, R. (1994). Expression of EVI1 in myelodysplastic syndromes and other haematologic malignancies without 3q26 translocations. Blood. 84, 1243-1248.

Rutherford, T., Nienhuis, A.W. (1987) Human globin gene prmoter sequences are significant for specific expression of a hybrid gene transfected into tissue culture cells. Mol. Cell. Biol. 7, 398-402.

Ryan, G., Steele-Perkins, V., Morris, J.F., Rauscher III, F.J., Dressler, G.R. (1995). Repression of *Pax-2* by WT1 during normal kidney development. Development. 121, 867-875.

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). Molecular cloning: a laboratory manual, 2nd edition. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, N.Y.

Sanchez-Garcia, I., Rabbitts, T.H. (1994). The LIM domain: a new structural motif found in zinc-finger-like proteins. Trends. Gen. vol 10. 9, 315-320.

Sassone-Corsi, P. (1995). Transcription factors responsive to cAMP. Annu. Rev. Cell. Dev. Biol. 11, 355-377.

Saxen, L., Salonen, J., Ekblom, P., Nordling, S. (1983). DNA synthesis and cell generation cycle during determination and differentiation of the metanephric mesenchyme. Dev. Biol. 98, 130-138.

Schild, C., Claret, F.X., Wahu, W., Wolffe, A.P. (1993). A nucleosome-dependent static loop potentiates estrogen-regulated transcription from xenopus vitellogenin B1 promoter *in vitro*. EMBO. J. 12, 423-433.

Schule, R., Muller, M., Otsuka-Murakami, H., Renkawitz, R. (1988). Co-operativity of the glucocorticoid receptor and the CACCC-box binding factor. Nature. 332, 87-90.

Scott, M.P., Weiner, A.J. (1984). Structural relationships among genes that control development: sequence homology between Antennnapedia, Ultrabithorax, and fuzi tarazu loci of Drosophila. Proc. Natl. Acad. Sci. USA. 81, 4155-4119.

Shawab, M., Almer, L. (1990). Genes Chromosomes and Cancer. 1, 83-93.

Senear, A.W., Palmiter, R.D. (1983) Expression of the mouse metallogenin-1 gene alters the nuclease hypersensitivity of its 5' regulatory region. Cold Spring Harbour Symp Quant Biol. 47, 539-547.

Seth, A., Ascoine, R., Fisher, R.J., Mavrothalassitis, G.J., Bhat, N.K., Papas, T.S. (1992). The ets gene family. Cell Growth. Differentiation. 3, 327-334.

Seto, E., Shi, Y., Shenik, T. (1991). YY1 is an initiatior sequence-binding protein that directs and activates transcription *in vitro*. Nature. 354, 241-245.

Shapiro, D.N., Sublett, J.E., Li, B., Downing, J.R., Naeve, C.W. (1993). Fusion of PAX3 to a member of the forkhead family of transcription factors in human alveolar rhabdomyosarcoma. Cancer. Res. 53, 5108-5112.

Shen-Ong, G.L.C., Morse, H.C., Potter, M., Mushinski, J.F. (1986). Two modes of c-myb activation in virus-induced mouse myeloid tumours. Mol. Cell. Biol. 6, 380-392.

Shi, Y., Seto, E., Chang, L., Shenk, T. (1991). Trabscription repression by YY1, a human-related protein, and relief of repression by adenovirus E1A protein. Cell. 67, 377-388.

Shirakawa, F., Mizel, S.B. (1989). In vitro activation and nuclear translocation of NF-kB catalysed by cyclic AMP-depenent protein kinase and protein kinase C. Mol. Cell. Biol. 9, 2424-2430.

Sippel, A., Saueressig, H., Winter, D., GerwL, t., Faust, N., Hecht, A., Bonifer, C. (1992). The regulatory domain orgainsiation of eukaryotic genomes: implications for stable gene transfers. in Transgeneic animals. editied by Grosveld F, Kollias, G. London: academic press; 1-26

Singh, H., Sen, R., Baltimore, D., Sharp, P.A. (1986). A nuclear factor that binds to a conserved sequence motif in transcriptional contro, elements of immunoglobin genes. Nature. 319, 154-158.

Smale, S.T., Schmidt , M.C., Berk, A.J., Baltimore, D. (1990). Transcriptional activation by Sp1 as directed through TATA or Initiator: Specific requirement for mammalian transcription factor IID. Proc. Natl. Acad. Sci. 1990. 87, 4509-4513.

Sola, B., Simon, D., Mattei, M-G., Fichelson, S., Bordereaux, D., Tambourin, P.E., Guenet, J-L., Gisselbrecht, S. (1988) Fim-1, Fim-2/c-fms, and fim-3, three common integration sites of friend murine leukaemia virus in myeloblastic leukaemias, map to mouse chromosome 13, 18, and 3, respectively. J Virol. 62, 3973-3978.

Sowadago, M., Roeder, R.G. (1985a). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell. 43, 165-175.

Sowadago, M., Roeder, R.G. (1985b) Factors involved in specific transcription by RNA polymerase II. Analysis by a repid and quatative *in vitro* assay. Proc. Natl. Acad. Sci. USA 82, 4394-4398..

Stanton, B.R., Perkins, A.S., Tessarolla, L., Sassoon, D.A., Parada, L.F. (1992). Loss of N-myc function results in embryonic lethality and failure of the epithelium component of the embryo to develop. Genes. Dev. 6, 2235-2247.

Stark, G.R. (1993). Regulation and mechanism of mammalian gene amplification. Adv. Cancer. Res. 87-111.

Staudt, L.M., Singh, H., Sen, R., Wirth, T., Sharp, P.A. (1986). A lymphoid-specific protein binding to the octomer motif of immunoglobulin genes. Nature 323, 640-643.

Steif, A., Winter, D.M., Stratling, W.H., Sippel, A.E. (1989). A nuclear DNA attachment element mediates elevated and position-independent gene activity. Nature 341, 343-345.

Steiner, C., Muller, M., Baniahmad, A., Renkawitz, R. (1987). Lysozyme gene activity in chicken macrophages is controlled by positive and negative regulatory elements. Nucleic. Acids. Res. 15, 4163-4178.

Stauss, E.C., Orkin, S. (1992). *In vivo* protein-DNA interactions at the hypersensitive site 3 of the human β-globin locus control region. Proc. Natl. Acad. Sci. USA. 89, 5809-5813.

Stringer, K.F., Ingles, C.J., Greenblat, J. (1990). Direct and selective binding of an acidic trnascriptional activation domain to the TATA-box factor TFIID. Nature. 345, 783-786.

Su, W., Jackson, S., Tjian, R., Echols, H. (1991). DNA looping between sites for transcription activation: self-association of DNA-bound Sp1. Genes. Dev. 5, 820-826.

Superti-Furga, G., Barberis, A., Schaffner, G., Busslinger, M. (1988). The -117 mutation in Greek HPFH affets the binding of three nuclear factors to the CCAAT region of the gamma-globin gene. EMBO. J. 7, 3099-3107.

Suzukawa, K., Parganas, E., Gajjar, A., Abe, T., Takahsi, S., Tani, K., Asano, S., Asou, H., Kamada, N., Yokata, J., Morishita, K., Ihle, J.N. (1994). Identification of a breakpoint cluster region 3' of the ribophorin I gene at 3q26 associated with the transcriptional activation of the EVI1 gene in acute myelogenous leukaemias with inv (3)(q21q26). Blood. 84, 2681-2688.

Talbot, D., Collis, P., Antoriou, M., Vidal, M., Grosveld, F., Greaves, D.R. (1989). A dominant control region from the human β-globin locus cinferring integration site-independent gene expression. Nature 338, 352-55.

Talbot, D., Philipson, S., Fraser, P., Grosveld, F. (1990). Detailed analysis of the site 3 region of the human β-globin dominant control region. EMBO. J. 9, 2169-2178.

Tanaka, M, Clouston, W.M., Herr, W. (1994). The Oct-2 glutamine-rich and proline-rich activation domains can synergise with each other or duplicates of themselves to activate domain. Mol. Cell. Biol. 14, 6046-6055.

Tanaka, M., Herr, W. (1994). Reconstitution of transcriptional activation domains by reiteration of short peptide segments reveals the modular organisation of a glutamine-rich activation domain. Mol. Cell. Biol. 14, 6056-6067.

Tanaka, T., Nishida, J., Mitani, K., Ogawa, S., Yazaki, Y., Hirai, H. (1994). Evi-1 raises AP-1 activity and stimulates c-fos promoter transactivation with dependence no the second zinc finger domain. 269, 24020-24026.

Tanaka, T., Mitani, K., Kurokawa, M., Ogawa, S., Tanaka, K., Nishida, J., Yazaki, Y., Shibata, Y., Hirai, H. (1995) Dual functions of the AML1/Evi-1 chimeric protein in the mechanism of leukaemogenesis in t(3;21) leukaemias. Mol. Cell. Biol. 15, 2383-2392.

Tasset, D., Tora, L., Fromental, C., Scheer, E., Chambon, P. (1990). Distinct classes of transcriptional activating domains function by different mechanisms. Cell. 62, 1177-1187.

Tate, P.H., Bird, A.P. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. Curr. Opin. Genet. Dev. 3, 226-231.

Taub, M., Chuman, L., Saier, M.H., Sato, G. (1979). Growth of Madin-Darby canine kidney epithelial cell (MDCK) line in hormone-supplemented, serum- free. Proc. Natl. Acad. Sci. USA. 76, 3338-3342.

Taub, R., Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S., Leder, P. (1982). Translocation of the c-myc gene into the immunglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc. Natl. Acad. Sci. USA. 79, 7837-7841.

Theisen, M., Steif, A., Sippel, A.E. (1986). EMBO. J. 5, 719-724.

Theveny, B., Bailly, A., Rauch, C., Rauch, M., Delain, E., Milgrom, E. (1987). Association of DNA-bound progesterone receptors. Nature. 329, 79-81.

Tsai, S.Y., Tsai, M.J., O'Malley, B.W. (1989). Cooperative binding of steroid hormone receptors contributes to transcription synergism at target enhancer elements. Cell. 57, 443-448.

Vainio, S., Lehtonen, E., Jalkanen, M., Bernfield, M., Saxen, L. (1989). Epithelial-mesenchyme interactions regulate the stage-specific expression of a cell surface proteoglycan, syndecan, in the developing kidney. Dev. Biol. 134, 382-391.

van Assendeflt, G., Hanscombe, O., Grosveld, F., Greaves, D.R. (1989). The b-globin dominant control region activates homologous and heterologous promoters in a tissue specifc manner. Cell. 56, 969-977.

van der Vliet, P.C. Verrijzer, C.P. (1993). Bending of DNA by transcription factors. Bioessays. 15, 25-32.

Verrijezer, C.P., Yokomori, K., Chen, J-L, Tjian, R. (1994). Drosophila TAFII150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA. Science 264, 933-941.

Vestweber, D., Kemler, R., Ekblom, P. (1985). Cell-adhesion molecule uvomorulin during kidney development. Dev. Biol. 112, 213-221.

Udvardy, A., Maine, E., Schedl, P. (1985) The 87A7 chromomere. Identifcation of a novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. J. Mol. Biol. 185, 341-358.

Walters, M.C., Fiering, S., S., Eidmiller, J., Magis, W., Groudine, M., Martin, D.I.K. (1995). Enhancers increase the probability but not the level of gene expression. Proc. Natl. Acad. Sci. USA. 7125-7129.

Walters, M.C., Magis, W., Fiering, S., Eidmiller, J., Scalzo, D., Groudine, M., Martin, D.I.K. (1996). Transcriptional enhancers act in Cis to suppress positional-effect variegation. Genes. Dev. 10, 185-195.

Warren, R., Carroll, S. (1995). Homeotic genes and diversification of the insect body plan. Curr. Opin. Genet. Dev. 5, 459-465.

Waterborg, J.H., Matthews, H.R. (1984). Patterns of histone acetylation in Phsarum polycephalum H2A and H2B actylation is functionally distinct from H3 and H4 activation.. Eur. J. Biochem. 142, 329-333.

Way, J.C., Chalfie, M. (1988). mec-3 a homeobox-containing gene that specifies differentiation ofthe touch receptor neurons in C.elegans. Cell. 54, 5-16.

Weinzierl, R., Dynlacjt, B.D., Tjian, R. (1993). Largest subunit of Drosphila transcription factor IID directs assmebly of a complex containing TBP and a coactivator. Nature. 362. 511-517.

White., J.H., Brou, C., Wu, J., Burton, N., Egly, J.M., Chambon, P. (1991). Evidence for a factor required for transcriptional stimulation by the chimeric acidic activator GAL4-VP!6 in HeLa cell extracts. Proc. Natl. Acad. Sci. USA 88, 7674-7678.

White, J.H., Jackson, S.P., Rigby, P.W.J. (1992). A role for TATA-box binding protein component of transcription factor-IID complex as a genereal RNA polmerase III ranscription factor. Proc. Natl. Acad. Sci. USA. 89, 1949-1953.

Whitlock, J.P., Simpson, R.T. (1977). Localisation of the sites along nucleosomal DNA which interact with the NH2-teminal histone region. J. Biol. Chem. 252, 6516-6520.

Williams, D.A., Rosenblatt, M.F., Beier, D.R., Cone, R.D. (1988). Generation of murine stromal cell lines supporting hematopoietic stem cell proliferation by use of recombinant retrovirus vectors encoding simian virus 40 large T antigen. Mol. Cell. Biol. 8, 3864-3871.

Wilson, P.D., Anderson, R.J., Breckon, R.D., Nathrath, W., Schrier, R.W. (1987). Retention of differentiated characterisitics by cultures of defined rabbit kidney epithelia. J. Cellular. Physiology. 130, 245-254.

Wise, T.L., Harris, M. (1988). Deletion and hypermethylation of thymidine kinase gene in V79 Chinese hamster resistant to bromodeoxyuridine. Somatic. Cell. Mol. Genet. 14, 567-581.

Wolberger, C., Vershon, A.K., Liu, B., Johnson, A.D., Pabo, C.O. (1991). Crystal structure of a MAT a2 homeodomain-operator complex suggests a general model of homoeodomain-DNA interactions. Cell. 67, 517-528.

Wolf, S.F., Migeon, B.R. (1985) Clusters of CpG dinucleotides implicated by nuclease hypersensitivity as control elements of housekeeping genes. Nature (London) 314, 467-469.

Wolffe, A.P., Pruss, D. (1996). Deviant nucleosomes: the functional specialization of chromatin. Trends. Genet. 12, 58-62.

Workman, J.L., Roeder, R.G. (1987). Binding of the transcription factor TFIID tothe major late promoter during *in vitro* nucleosome assembly potentiates subsequent initiation by RNA polymerase II. Cell. 51, 613-622.

Workman, J.L., Abmayr, S.M., Cromlish, W.A. Roeder, R.G. (1988). Transcriptional regulation by the immediately early protein of pseudorabies virus during in vitro nucleosome assembly. Cell. 55, 211-219.

Workman, J.L., Roeder, R.G., Kingston, R.E. (1990). An upstream transcription factor, USF (MLTF), facilitates the formation of preintiation complexes during in vitro chromatin assembly. EMBO. J. 9, 1299-1303.

Wu, C., Bingham, P.M., Livak, K.J., Holmgren, R., Elgin, S.C. (1979a). The chromatin structure of specific genes: I. Evidence for higher order domains of defined DNA sequence. Cell. 16, 797-806.

Wu, C., Wong, Y.C., Elgin, S.C. (1979b). The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activty. Cell. 16, 807-14.

Wu, C. (1980). The 5' ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNAse1. Nature 286, 854-60.

Wu, C., Wilson, S., Walker, B., Dawid, I., Paisley, T. (1987). Purification and properties of Drosophila heat shock activator protein. Science. 238, 1247-1253.

Wu, S., Pena, A., Korez, A., Soprano, D.R., Soprano, K.J. (1996). Overexpression of Mxi1 inhibits the induction of the human ornithine decarboxylase gene by the Myc/Max protein complexes. Oncogene. 12, 621-629.

Xiao, J.H., Davidson, I., Macchi, M., Rosales, R., Vigneron, M., Staub, A., Chambon, P. (1987). In vitro binding of several cell-specific and ubiquitous nuclear proteins to the GT-1 motif of the SV40 enhancer. Genes. Dev. 1, 794-807.

Xiao, H., Pearson, A., Coulombe, B., Traut, R., Zhang, S., Reiger, J.L., Triezenberg, S.J., Reinberg, D., Flores, O., Ingles, C.J., Greenblatt, J. (1994). Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53. Mol. Cell. Biol. 14, 7013-7024.

Xu, X., Prorock, C., Ishikawa, H., Maldonado, E., Ito, Y., Gelinas, C. (1993). Functional interaction of the v-Rel and c-Rel oncoproteins with the TATA-binding protein and association with transcription factor-IIB. 13, 6733-6741.

Yamamoto, K.R. (1985). Steriod receptor regulated transcription of specific gene networks. Ann. Rev. Genet. 19, 209-252.

Yanai, N., Satoh, T., Kyo, S., Abe, K., Suzuki, M., Obinata, M. (1991a). A tubule cell line estabalished from transgenic mice harbouring temperature-sensitive Simian virus 40 large T-antigen gene. Jpn. J. Cancer. Res. 82, 1344-1348.

Yanai, N., Suzuki, M., Obinata, M. (1991b). Hepatocyte cell lines estabalished from transgenic mice harbouring temperature-sensitive Simian virus 40 large T-antigen gene. Exp. Cell. Res. 197, 50-56.

Yanai, N., Obinata, M. (1994). Apoptosis is induced at nonpermissive temperature by a trnasient increase in p53 in cell lines immortalised with temperature-sensitivie SV40 large T-antigen gene. Exp. Cell. Res. 221, 296-300.

Zaret, K.S., Yamamoto, K.R. (1984). Reversible and persisent changes in chromatin structure accompany activation of a glucocorticoid enhancer element. Cell. 38, 29-38.

Zawel, L., Reinberg, D. (1995) Common themes in assembly and function of eukaryotic transcription complexes. 64, 553-561.