TRANSCRIPTION AND TRANSFORMATION BY PAPOVAVIRUSES

HARIKRISHNA NAKSHATRI, B.V.Sc.

TRANSCRIPTION AND TRANSFORMATION BY PAPOVAVIRUSES

by

© Harikrishna Nakshatri, B.V.Sc.

A thesis submitted to the
School of Graduate Studies
in partial fulfillment of the
requirements of the degree of
Doctor of Philosophy

Faculty of Medicine
Memorial University of Newfoundland
August, 1990

St. John's **Newfoundland**

This thesis is dedicated in memory

of my late grandparents

Smt Annapurna

and

Shri Narayana Nakshatri

ABSTRACT

The Regulation of eucaryotic gene expression and cellular transformation have been a central focus of research in the last two decades. Studies on DNA tumor viruses have provided information not only on viral pathogenicity but also on general eucaryotic gene regulation, as they depend on host cell machinery for their transcription. I have examined some aspects of transcriptional control in six DNA tumor viruses belonging to the papovaviridae family and the results are discussed in Section 1. Additionally, I have studied cellular transformation induced by one of these viruses the results of which are presented in Section 2.

The *cis*-acting regulatory elements of a gene required for accurate, efficient and cell type specific expression are classified into promoters and enhancers. To study the role of enhancers in cell type specific expression, I have performed comparative analysis of the SV40, BK and JC enhancers in differentiated and undifferentiated embryonal carcinoma (EC) cells. In transient transfection assays, transcription of a reporter gene was not activated by any of the viral regulatory elements in undifferentiated EC cells. However, transcriptional activation of the reporter gene was observed in retinoic acid (RA) differentiated neuronal cell types by all three regulatory elements. Moreover, SV40 and BK but not JC regulatory elements demonstrated activity in

dimethylsulfoxide (DMSO)-differentiated muscle cell types. To correlate *in vivo* activity with the binding of transcription factors, I performed DNaseI footprinting experiments. For SV40, the region containing the transcription factor Sp1 binding motif was protected in undifferentiated and differentiated cell types. However, in both RA- and DMSO-differentiated cells, additional protection at regions corresponding to P, SphI and SphII motifs was detected. The JCV enhancer demonstrated three retinoic acid differentiated cell type specific footprints, each containing sequences with homology to the nuclear factor-1 (NF-1) binding motif. With the BKV enhancer, a GC rich region was protected in all three cell types. Additional protection in four regions, each containing sequences with homology to the NF-1 motif, was observed in the two differentiated cell types. Further experiments suggest that the factor(s) interacting with NF-1 motif containing regions are different in the two differentiating cell types. Overall, the results demonstrate clear correlation of *in vivo* activity of the regulatory elements of the three viruses with *in vitro* DNA-protein interaction.

To analyze epitheliotropism of human papillomavirus types 11, 16 and 18, I performed similar studies in four epithelial cervical carcinoma (C33A, HeLa, SiHa and CaSki) and one fibroblast (143B) cell lines. All three viral enhancers demonstrated varying degrees of activity in C33A, HeLa and SiHa but not in CaSki and 143B. By DNaseI footprinting, I have

identified seven, nine and five footprints on the HPV 11 , 16 and 18 enhancers, respectively. NF-1 motifs were present in five, six and one protected regions of the HPV 11, 16 and 18 enhancers, respectively. Sequences homologous to LVc, and ØAP3 (HPV 11); AP1 (HPV 16 and HPV 18) and EFII (HPV 18) were also observed in a few protected regions. *In vitro* transcription-oligonucleotide competition and deletion analyses the of HPV 11 enhancer revealed that one of the NF-1 motif containing regions acts as a negative control element. One sequence motif of HPV 16, for which UV cross-linking studies revealed interaction with four protein molecules, is a strong modulator of HPV 16 enhancer function *in vivo* and shares 100% homology to a sequence motif, GTTTTAA in the tissue-specific enhancer of the c-*mos* oncogene. One sequence motif of the HPV 18 enhancer has three repeats of a TTTTA sequence contained within the c-*mos* sequence motif and interacts with at least four different polypeptides, as judged by UV cross-linking experiments.

Previous studies have indicated that tumor (T)-antigens of BK and SV40, although structurally related, possess different transformation potentials. To understand these differences, I have examined the role of BKV tumor antigens in the maintenance of transformation and have identified the domain of the T-antigen gene essential for the transformed phenotype. BKV DNA-transformed BHK 21 and NIH 3T3 cells expressing antisense BK T-antigen RNA lose their ability to grow in soft

agar, indicating the need for continued expression of T-antigen for the maintenance of the transformed phenotype. Experiments using translation termination linker insertion and deletion mutagenesis of BKV T-antigen demonstrate that amino acids 356 to 384 are essential for transformation. Although BKV T-antigen shares 100, 85, and 80% amino acid homology with the SV40 T-antigen for the nuclear localization signal, DNA-binding domain and p53-binding domain, respectively, the transformation domains of BKV and SV40 T-antigens share only 54% homology. In addition, the BKV T-antigen lacks a substantial portion of the ATPase domain present in the SV40 T-antigen. Further results indicate the dispensability of the remaining portion for transformation by this protein. I suggest that the differences in the amino acids in the identified transformation domains together with the differences in the ATPase domains may account for the differences in the transformation potentials of the two proteins.

ACKNOWLEDGEMENTS

I am indebted to Drs. Alan and Mary Pater for their supervision, help and encouragement. I thank them for providing laboratory space and materials. My sincere thanks to Dr. Banfield Younghusband for his advice and valuable comments on the work. I am also grateful to Dr. Chet Michalski for his support.

I wish to thank Ms. Glenys Hughes for technical assistance and editorial comments on this thesis. I am grateful to Christine Kisaka for help in proof-reading this thesis. My sincere thanks to David Hyslop, Anne Roche, Michelle Power for their technical assistance and encouragement. The company of my colleagues and friends, Trudy Marshall, W. Guan, Humphery Gardner, Kreg Stone, Krishna Bhat, Nagendra Prasad, Narasimhaswamy, Rakesh Mittal, Kotlo Umeshkumar, Kasinadhuni Satyanarayana, Christine Kisaka, Markandeya Jois, Jennifer, Vidya, Tara, Shrinivas, Ramdas, Bob Porter, Sulaman Afwork, Timothy Churchill, Jacques DeSousa and Margaret Reid, is very much appreciated. My sincere thanks to Kimberley Ross for typing the thesis and to Pah Baldeh for his assistance with the graphics programme.

The blessings of parents, my brothers and sisters-in-law and good wishes of my sister were the encouragement during this study. I am grateful to Drs. Narayan, B.P. Hegde, D.R.L. Shetty and A.V. Rai of Bangalore, Dr. Antanio Arpon of

Philippines and Mr. Ali Al. Komaie of Saudi Arabia for the support.

I thank Drs. H. Hamada, M.L. Brietman, A.J. Levine for providing certain materials used in this study. I gratefully acknowledge the financial support provided by the Faculty of Medicine and the School of Graduate Studies, MUN, the Cancer Research Society, Inc., Montreal, Quebec, and the National Cancer Institute of Canada.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BHK	=	Baby hamster kidney cells
bp	=	base pairs
BPV-1	=	Bovine papillomavirus type-1
BRK	=	Baby rat kidney cells
BRL	=	Bethesda Research Laboratories
BSA	=	Bovine serum albumin
CAT	=	Chloramphenicol acetyl transferase
CBP	=	CCAAT binding protein
cDNA	=	complementary DNA
CIP	=	Calf intestinal phosphatase
cpm	=	Counts per minute
CTF	=	CCAAT transcription factor
DBM	=	Diazobenzyloxymethyl paper
DMSO	=	Dimethylsulfoxide
DNA	=	Deoxyribonucleic acid
DNaseI	=	Deoxyribonuclease I
DTT	=	Dithiothreitol
E	=	Early region
E1A	=	Early region 1A
EC	=	Embryonal carcinoma
EDTA	=	Ethylenediaminetetra-acetic acid
EE	=	Early-early
EEP	=	Early-Early promoter
EES	=	Early-Early start site

EGTA	=	Ethyleneglycol-bis-(β-aminoethyl ether)N,N'-tetra-acetic acid
FCS	=	Foetal calf serum
HaPV	=	Hamster papovavirus
HEPES	=	H-(2-hydroxyethyl)-1-piperazinuthanesulfonic acid
HPVs	=	Human papillomaviruses
IgH	=	Immunoglobulin
kb	=	kilobasepairs
KD	=	kilodaltons
L	=	Late region
LCR	=	Long control region
LE	=	Late-Early
LES	=	Late-Early start site
LPV	=	Lymphotropic papovaviaruses
LTR	=	Long terminal repeat
MEM	=	Minimum essential medium
MoMuLV	=	Moloney murine leukaemia virus
MOPS	=	Morpholinopro-panesulfonic acid
NCR	=	Non-coding region
NEN	=	New England Nuclear
neo	=	neomycin
NF-1	=	Nuclear factor 1
NP40	=	Nonidet P-40
ORF	=	Open reading frame
PBS	=	Phosphate buffered saline
PML	=	Progressive multifocal leucoencephalopathy

PMSF	=	Phenylmethylsulfonyl fluoride
PVF	=	Papillomavirus factor
RA	=	Retinoic acid
RNA	=	Ribonucleic acid
rNTP	=	Riboxynucletide triphosphates
rpm	=	Revolutions per minute
RSV	=	Rous sarcoma virus
SDS	=	Sodium dodecyl sulfate
snRNA	=	small nuclear RNA
SSC	=	Sodium chloride-sodium citrate buffer
SV40	=	Simian vacuolating virus 40
T-antigen	=	Tumor antigen
TE	=	Tris-EDTA
TIF	=	Trans Inducing factor
TLC	=	Thin layer chromatography
TPA	=	12-O-tetradecanoylphorbol-13-acetate
U	=	units
URR	=	Upstream regulatory region
UV	=	Ultraviolet
VDRE	=	Vitamin D3 responsive element

SECTION 1

CELL-TYPE SPECIFIC TRANSCRIPTIONAL REGULATION

CHAPTER 1

INTRODUCTION

The last two decades have witnessed rapid growth in molecular biology which has revolutionized our understanding of the basic functioning of eucaryotic cells. The control of normal and oncogenic cell growth and gene expression have received considerable attention. Some of the major discoveries in these areas can be attributed to virus research, such as Schaffner's "enhancers" (Banerji *et al.* 1981) and Varmus and Bishop's "oncogenes" (Stehelin *et al.* 1976), which inspired many investigators to use viruses to study the molecular basis of gene expression and cancer.

The first section of this thesis is a preliminary survey of six papovaviruses with regard to cell-type specific transcriptional regulation. The second section is on cellular transformation by one of these viruses. This chapter presents a review of the literature on recent developments in the area of transcriptional regulation. This includes a brief description of transcriptional regulatory elements, functional domains of transcription factors, mechanisms of transcriptional activation and repression, tissue specific gene expression and structure-function relationships between cis-acting elements and trans-acting factors. Additionally, the molecular biology of the viruses used in this study is briefly described.

1.1 Transcriptional Regulation

The initiation of transcription is a primary control point in the regulation of gene expression. The theoretical model proposed by Britten and Davidson (1969) emphasized the importance of transcription initiation and laid the experimental framework for its molecular dissection. Although the terminologies used in this model differ from those in current usage, the basic elements of the model have been accepted.

This model proposes that "the gene is a region of the genome with a narrowly definable or elementary function. It need not contain information specifying the primary structure of a protein". There are five important elements in this model: producer genes, receptor genes, activator RNA, integrator genes and sensor genes. As described below, a regulatory network links these elements providing clues to the nature of gene regulation.

A producer gene is defined as a region of the genome which is transcribed to yield a template RNA molecule or other species of RNA molecule except those involved directly in genomic regulation. A receptor gene is a sequence linked to a producer gene and causes the transcription of that producer gene after forming a sequence specific complex with activator RNA. Activator RNAs are the products of integrator genes. DNA sequence elements linked to the producer gene that are required for the stimulation of transcription in response to

extracellular stimuli are called sensor genes. By analogy, receptor genes, activator RNA and sensor genes are currently called regulatory elements, transcription factors, and inducible enhancers, respectively.

DNA sequences in the regulatory region contain at least two elements that function to facilitate initiation of transcription. One element specifies the site of initiation and another governs the efficiency of initiation. The term "promoter" was first proposed by Epstein and Beckwith (1968) for the site of transcription initiation, and is described as an initiating element indispensable for the expression of bacterial structural genes. The presence of such an element in eucaryotes was firmly established by Corden *et al.* (1980). The second element, termed "enhancer", was first identified in SV40 and is defined as short, often tandemly repeated sets of nucleotides which activate transcription relatively independent of their position, orientation, and to a lesser degree distance from the target promoter (Banerji *et al.* 1981).

1.1.1 Promoters as sites for assembly of the pre-initiation complex

Early studies involving deletion mutational analysis of sequences 5′ to the ovalbumin and adenovirus major late gene coding sequences established that promoters contain a stretch of an A/T rich sequence called the TATA box which is involved

in the accuracy of transcription initiation (Gannon *et al.* 1979; Corden *et al.* 1980). Analysis of many other genes showed that this element is generally located 25-30 base pairs (bp) upstream from the beginning of the transcribed sequences and can be divided functionally into two distinct classes (Flavell, 1980; Breathnach and Chambon, 1981; Struhl, 1986). Sequences TATAAA form the regulatory TATA (Tr) elements while related sequences (for example, TATACA) form constitutive TATA (Tc) elements. Activity from regulated TATA elements usually requires the presence of one or more upstream regulatory elements located about 20-70 bp further upstream. Several upstream elements, notably CCAAT boxes and GGGCGG boxes, have been identified (reviewed by Dynan and Tjian, 1985). Some promoters contain elements that stimulate transcription in response to heavy metals and heat shock (Karin *et al.* 1984; Pelham, 1982). Several regulatory pathways, including those induced by serum, adenovirus E1A, or forskolin, are known to be mediated through the TATA box and an upstream element (Simon *et al.* 1988; Williams and Morimoto, 1990). Not all eucaryotic promoters contain a TATA box. In such cases sequences other than TATA determine the accuracy of initiation. For example, in the case of the lymphocyte-specific terminal deoxynucleotidyltransferase gene, the transcription initiation site itself functions as a TATA element (Smale and Baltimore, 1989).

Purification of RNA polymerase II, an enzyme responsible

for mRNA synthesis, and subsequent development of *in vitro* transcription systems initiated the fine functional analysis of TATA boxes and upstream regulatory elements (Schwartz and Roeder, 1975; Weil *et al.* 1979). It was observed that purified eucaryotic RNA polymerase II cannot initiate transcription unless the system is supplemented with crude nuclear extracts (Weil *et al.* 1979; Manley *et al.* 1980). Chromatographic fractionation of these extracts and indirect depletion assays with functional TATA homology sequences have enabled the identification of at least four factors that are required for accurate initiation from a minimum promoter containing only a TATA box and a start site (Davidson *et al.* 1983; Matsui *et al.* 1980; Samuels *et al.* 1982,. The factor TFIID (also known as DB or BTF-1) contains a protein that specifically binds to the TATA element (Sawadgo and Roeder, 1985). The amino acid sequences of the human TFIID, deduced from a cloned cDNA, reveals a basic amino acid rich C-terminal DNA binding domain and a glutamine rich N-terminal transactivation domain (Peterson *et al.* 1990; Kao *et al.* 1990). The C-terminal domain is highly conserved among the yeast, drosophila and human TFIID genes (Horikoshi *et al.* 1989; Peterson *et al.* 1990). Recombinant TFIID, synthesized either in bacteria or in a vaccinia virus expression vector, can bind the TATA box independently of other factors and can stimulate transcription upon addition of RNA polymerase II and the other general transcription factors TFIIA, TFIIB, TFIIE-

F (Peterson *et al.* 1990). The C-terminal DNA binding domain of TFIID alone is sufficient to confer basal level activity to a promoter containing only a TATA box and the other components of the in vitro system. However, as discussed later, full length TFIID is required for upstream activators induced transcription.

RAP 30/74, a heterodimeric general transcription factor, is believed to be identical to TFIIF, and cloned RAP30 shows DNA-helicase activity (Sopta *et al.* 1989). In an elegant piece of work, Buratowski *et al.* (1989) presented evidence for the sequential binding of TFIID, TFIIA, TFIIB, RNA polymerase II, and TFIIE-F to the TATA box. This sequential binding of general transcriptional factors is facilitated by the cooperative interaction of upstream element-binding factors and TFIID. This has been demonstrated with the adenovirus E4 promoter. DNase I footprinting analysis revealed a cooperative interaction between the upstream promoter binding factor ATF and TFIID when both are bound simultaneously to the adenovirus E4 promoter; these interactions in turn facilitate promoter recognition by RNA polymerase II, TFIIB, and TFIIE-F (Horikoshi *et al.* 1988).

Two models have been proposed to account for how such an interaction of upstream factors may facilitate accurate initiation. The first model predicts that the direct recruitment of a general transcription factor by an upstream activator facilitates the assembly of a pre-initiation complex. The

second model proposes that the activator enhances a step following assembly of the general transcription factor into a pre-initiation complex (Lillie and Green, 1989; Buratawaski *et al.* 1989).

TFIID may not always be the target for upstream factors as interactions between highly purified yeast TFIID and upstream factor Sp1 have not been detected (Schmidt *et al.* 1989). In this case, Sp1 may interact indirectly with TFIID through a "coactivator" protein.

1.1.2 Enhansons as functional units of enhancers

First identified in SV40, enhancers have been the target of extensive investigation. Saturation and deletion mutagenesis studies established that enhancers are composed of multiple sequence motifs of approximately 10 bases called enhansons (Ondek *et al.* 1988). Gel retardation and DNase I footprinting assays provided evidence for the interaction of trans-acting factors with these enhansons (Sassone-Corsi *et al.* 1985; for review see Maniatis *et al.* 1987; Jones *et al.* 1988). For optimum enhancer activity, a hierarchy in the functional organization of enhansons is required (Fromental *et al.* 1988; Dynan, 1989). The first level in the hierarchy is the proto-enhancer which is derived from a combination of enhansons.

At least four categories of proto-enhancers have been described. The first type consists of tandem duplication of

an enhanson in which a cooperative interaction of a factor(s) to adjoining enhansons generates synergistic activity. In this type, the distance between each enhanson is critical for synergistic effect. In the second type, cooperative activity is generated by the interaction of two different factors with adjoining, unrelated enhansons. The third group forms tissue-specific enhancers, where enhansons can bind only to tissue specific factors. Inducible enhancers constitute the fourth type of proto-enhancer, where the enhanson is active only in response to extracellular stimuli (Fromental *et al*. 1988). Different combinations of these proto-enhancers generate an enhancer. Thus, the strength and specificity of the enhancers are determined by individual enhansons, their organization into proto-enhancers, and the type of proteins that interact with the enhansons.

1.1.3 Trans-acting factors interact with DNA through specific DNA-binding domains

The advent of modern techniques enabled purification and cloning of genes encoding transcription factors. Although the exact mechanism by which transcription factors recognize DNA is not known, the emerging indication is that many different proteins recognize DNA via common structural domains. X-ray crystallography and *in vitro* mutagenesis studies have permitted the identification of at least four types of DNA-binding domains. The helix-turn-helix motif is widely

distributed from bacteriophage lambda cro and repressor proteins to mammalian homeodomain containing proteins (Pabo and Sauer, 1984; for review see Struhl, 1989). The crucial structure consists of two alpha helices separated by a beta turn. One helix, called the recognition helix, directly contacts bases exposed in the major groove of the target DNA, while the other lies across the major groove, making non-specific contacts with DNA.

In eucaryotes, the helix-turn-helix structure was first discovered in yeast mating type locus products MATα1 and MATα2 and in Drosophila homeotic genes. DNA sequence analyses of several homeotic genes defined a 60 amino acid region, the homeobox, with more than 90% amino acid homology. Homeoboxes contain a pattern of hydrophobic amino acids consistent with the helix-turn-helix motifs.

The recent characterization of four proteins, three of them mammalian transcription factors, enabled the identification of a new class of homeobox containing proteins. These proteins, Oct-1, Oct-2, Pit-1 and unc-86, have a 60 amino acid homeobox with common sequence similarities that distinquish them from other known homeoboxes and a second region of sequence similarity spanning about 75 amino acids upstream of the homeobox. Because these regions are shared by Pit-1, two octamer-binding proteins (Oct-1, Oct-2) and unc-86, they have been designated as POU domains (for review see Robertson, 1988; Stern and Herr, 1988; Bodner *et al.* 1988; Herr *et al.*

1988).

Most homeobox containing proteins in this class bind to A/T rich sequences where both individual bases and amino acids at nine conserved positions determine the DNA binding specificity. For example, a two base change in the binding motif of pituitary-specific transcription factor Pit-1 to that of the lymphoid-specific factor Oct-2 switches gene expression from pituitary-specific to lymphoid-specific (Elsholtz *et al.* 1990). Similarly, the DNA-binding specificity of Drosophila homeobox proteins are determined by the amino acid at the ninth conserved position (Triesman *et al.* 1989).

The second type of DNA binding domain, zinc fingers, is composed of two distinct classes. (Evans and Hollenberg, 1988). The C2H2 class, where the cysteine and histidine pairs serve as tetrahedral coordination sites for zinc ions producing "finger" like loops in the protein, hence the name zinc finger, was first identified in the 5S RNA gene transcription factor TFIIIA. A similar zinc finger motif has been identified in the GC box-binding protein Sp1 (Kadonaga *et al.* 1987). In contrast to the C2H2 group, the Cx proteins have a variable number of conserved cysteine residues available for zinc chelation. Examples for this class include yeast transcription factor GAL4 which contains a cluster of six invariant cysteines (C6 family) and the steroid receptors containing four invariant cysteines (C4 family). The role of the cysteine-rich domain in determining DNA-binding

specificity has been probed initially by finger swap experiments and more recently by site-directed mutagenesis (Green and Chambon, 1987; for review see Berg, 1989). These experiments implicate the nonconserved amino acids between the cysteines of the first finger and the region immediately following them in specific DNA recognition.

Proteins of the third group contain four or five leucine residues spaced seven residues apart in a amphipathic alpha helical structure and are called leucine zippers (Landschulz *et al*. 1988). Site-directed mutagenesis and gluteraldehyde cross-linking experiments have been employed to study the role of individual leucine residues. The results indicate that the leucine residues are important for interdigitating two alpha helices, one from each monomer unit, providing the structural basis for dimer formation (Landschulz *et al*. 1989; Turner and Tjian, 1989; Gentz *et al*. 1989; Kouzarides and Ziff, 1989). Dimerization by a leucine zipper is not a random event: GCN4 and C/EBP form homodimers, whereas fos and jun participate in a heterodimeric complex (Landschultz *et al*. 1989; Turner and Tjian, 1989; Gentz *et al*. 1989; Abel and Maniatis, 1989). The structural basis for this selectivity is not yet understood.

In proteins that are involved in direct DNA interaction, the leucine zipper is always followed by a stretch of basic residues. Recent studies have indicated that the spacing between the basic region and the leucine zipper is invariant and essential for maintaining the ability to bind to DNA

(Ransone *et al.* 1990). Vinson *et al.* (1989) have proposed a "Scissor grip" model to explain DNA-binding by this class of proteins. According to the model, two proteins criss-cross at the junction of the leucine zipper and basic region, analogous to scissor grip, and the basic regions bind to directly abutted, dyad-symmetric half-sites on the DNA.

The fourth class of DNA binding domain, first identified in the transcription-replication factor CTF/NF-1, is characterized by a basic amino acid-rich alpha helical structure (Mermod *et al.* 1989). An amphipathic helix-loop-helix motif identified recently in the immunoglobulin enhancer-binding, Drosophila daughterless, MYC and MYOD proteins represents another new class of DNA-binding domain (Murre *et al.* 1989).

In spite of considerable progress in understanding the structure of DNA binding domains of proteins, the three dimensional structure of DNA-protein interaction for most of the proteins are not known. Generalization of a concept in this regard is complicated by the observation that some proteins interact with degenerate sequence motifs while in some other instances, different proteins interact with the same sequence. For example, a HeLa cell factor (TEF1) interacts with two different cis-regulatory elements the of SV40 enhancer (Davidson *et al.* 1988). Conversely, factors Oct-1 and Oct-2 interact with the same DNA sequence ATGCAAAT (Bohmann *et al.* 1987; Johnson and McKnight, 1989). This

intriguing aspect has been a focus of recent research. The emerging evidence indicates that DNA binding specificity in the latter case is achieved, at least in part, by protein-protein interactions (Kristei *et al*. 1989; Stern *et al*. 1989).

1.1.4 Transcriptional activation by transcription factors is mediated through distinct activator domains

Most transcription factors are bipartite structures having a well defined DNA binding domain and a transcription activation domain. At least three classes of activator domains which can function independently and cooperatively have been identified (Mitchell and Tjian, 1989; Ptashne, 1988). The first class consists of a simple stretch of amino acids with a significant negative charge and the capacity to form an amphipathic helix. The examples for this class are the yeast transcription factors, GAL4 and GCN4 (Ma and Ptashne, 1987; Hope *et al*. 1988). Sigler (1988) proposed that the negatively charged region of the transcription factors interacts with either the carboxy terminal heptamer neutral amino acid repeats of RNA polymerase II or a similar structure of transcription factors to stabilize the initiation complex.

There is evidence that some transcription factors interact with RNA polymerase II, in particular its tail - the phosphorylated carboxy-terminal repeat structure of the largest subunit (Moyle *et al*. 1989; Zehring *et al*. 1988). A column containing the activator GCN4 retains RNA polymerase

II (Brandl and Struhl, 1989). However, studies with mutants of GCN4 have indicated that the acidic domain is not required for interaction with RNA polymerase II. Instead, genetic and biochemical evidence suggests that these acidic regions might be involved in interactions with other TATA-binding factors (Horikoshi *et al.* 1988; Struhl *et al.* 1988).

Transcription activator domains of the second class contain a glutamine-rich region. This type of domain is present in Sp1, Antennapedia, Ultrabithorax and Zeste proteins (Courey and Tjian, 1988; Mitchell and Tjian, 1989). Like the acidic domains described above, the three dimensional structure imparted by glutamines is also essential for transcriptional activation (Courey *et al.* 1989). Although the mechanism of activation is not known, one possibility is that the amide moieties of the glutamine side chains are involved in hydrogen bonding to RNA polymerase II or some other component of the general transcriptional machinery (Courey and Tjian, 1988).

The third class of activator domain, identified first in CTF/NF-1, consists of regions rich in proline (Mermod *et al.* 1989). This domain is different from the replication domain present in the same factor and can activate transcription when fused to the DNA-binding domain of the heterologous transcription factor Sp1 (Mermod *et al.* 1989). It has been suggested that Omega loops of prolines participate in protein-protein interactions with other factors in transcription machinery.

Not all proteins binding to DNA activate gene transcription. Some proteins are involved exclusively in transcriptional repression. A few transcription activators can also function as repressors depending on the context of DNA-protein interactions. As described below, at least four different mechanisms have been proposed for transcription repression.

1.1.5 Transcription repression can be mediated by at least four mechanisms

Competition between positive and negative transcription factors, sequence-specific negative regulators, quenching, and squelching have been proposed as means of transcription repression (Ptashne, 1988; Levine and Manley, 1989).

According to the competition model, a repressor protein binds at or near the transcription start site and blocks the interaction of general transcription factors with the promoter. For example, the binding of SV40 T-antigen to its origin of replication prevents the assembly of a transcription initiation complex in the nearby early promoter and represses transcription. A related form of competition is observed in the case of the human glycoprotein alpha subunit gene where the positively acting CREB protein and negatively acting glucocorticoid receptor compete for binding to overlapping sequences (Akerblom *et al.* 1988).

The second "direct repression" model, predicts that a repressor binds to a defined site, perhaps at some distance

from the target promoter, and interferes with the formation or activity of the basal transcription complex. For example, transcription repression by NF-A3, an undifferentiated embryonal carcinoma cell line F9 specific factor, is due to direct interaction with an octamer motif of immunoglobulin gene promoters (Lenardo *et al.* 1989).

The "Quenching" mechanism proposes that repressor and activator proteins bind to separate nonoverlapping DNA sequence elements, yet the repressor interferes with the function of activator in some way. Drosophila homeobox proteins may use this mechanism to regulate gene expression during early embryonic development (Han *et al.* 1989).

Squelching, first proposed by Gill and Ptashne (1988), predicts that overexpression of an activator protein can paradoxically inhibit expression from genes devoid of binding sites for that factor by sequestering common general transcription factors. For example, the inhibitory role of GAL4 on transcription from the cyc1 gene regulatory elements lacking a GAL4 binding site is believed to be due to squelching (Gill and Ptashne, 1988). Glucocorticoid-mediated repression of oestrogen-responsive genes may also be due to squelching (Meyer *et al.* 1989).

What might be the physiological relevance of transcription repressors? An intriguing possibility is that cell growth and differentiation are regulated by transcriptional repression. Recent studies on the osteocalcin

gene favour such a possibility. Differentiation of primary bone cells by vitamin D3 is accompanied with induction of the osteocalcin gene. Induction of this gene by vitamin D3 is through a vitamin D3 responsive element (VDRE) located in the promoter region of the gene. Interestingly, VDRE also contains an AP1 responsive element and AP1 dominantly represses osteocalcin gene expression. Since AP1 is a nuclear oncogene and can promote cell growth, it is proposed that AP1 prevents differentiation of bone cells by blocking osteocalcin gene expression (Schüle *et al.* 1990). Additionally, transcriptional repression plays a significant role in developmental regulation of Drosophila (Han *et al.* 1989; Ohkuma *et al.* 1990). Finally, transcriptional repression is involved in tissue-specific gene expression, as described below.

1.1.6 Tissue-specific transcription factors, regulated protein-protein interaction, and DNA modifications are involved in tissue-specific gene expression

A central problem in gene regulation is to understand the mechanisms by which specific genes are expressed in a tissue-specific manner. Although the major control is at the level of chromatin organization (for review see Gross and Garrard, 1987), several additional regulatory mechanisms have been described. The genes for albumin, immunoglobulin kappa chain, and tyrosine amino transferase provide examples of three

different levels of regulation; they are described below.

Tissue-specific transcription factors play an important role in tissue-specific gene expression. Tissue-specific factors HNF1, Oct-2, and Pit-1 are required for liver-specific, lymphoid-specific, and pituitary-specific expression of albumin, immunoglobulin, and growth hormone genes, respectively (Lichsteiner and Schibler, 1989; Scheidereit *et al*. 1987; Bodner and Karin, 1987). This has been demonstrated in experiments involving expression of cloned cDNAs for these factors in non-expressing cells, and supplementation of purified factors to non-expressing cell extracts in an *in vitro* transcription assay (Lichesteiner and Schibler, 1989, Bodner and Karin, 1987; Muller *et al*. 1988). The tissue-specific factors need not necessarily be transcription activators as tissue specific repressors are also involved in tissue specific gene expression (Lenardo *et al*. 1989; Beggs *et al*. 1988; Boshart *et al*. 1990).

It has recently been appreciated that regulated protein-protein interactions play an important role in tissue-specific gene expression. This has been demonstrated with the immunoglobulin kappa chain-binding factor $NF_{\kappa}B$. Initial studies using gel retardation assays identified this protein only in B cells (Lenardo *et al*. 1987). However, subsequent studies proved that $NF_{\kappa}B$ is widespread and is restricted to the cytoplasm in non-expressing cells by its association with an inhibitory protein, IkB (Baeuerle and Baltimore, 1988).

Likewise, the tissue-specific function of transcription factor MyoD, a factor required for muscle-specific expression of the muscle creatine kinase gene, is also regulated by an inhibitory protein (Benezra *et al.* 1990).

The involvment of DNA modification by methylation in tissue-specific gene expression has been studied in great detail (for review see Cedar, 1988). Methylation inhibits gene expression by affecting the protein-DNA interactions required for transcription. Insight into this mode of regulation was obtained with studies on the liver-specific tyrosine aminotransferase gene. *In vitro* DNaseI footprinting of this gene's enhancer revealed that enhancer binding factors are present in both expressing hepatoma cells and non-expressing fibroblasts. However, interactions of the same factors *in vivo*, as judged by *in vivo* footprinting, was found to be restricted to hepatoma cells. Two lines of evidence indicated that methylation was responsible for this altered DNA-protein interaction. Firstly, by genomic sequencing it was shown that the enhancer region in fibroblasts was methylated while in hepatoma cells it was unmodified. Secondly, methylation of the cloned DNA *in vitro* completely abolished the binding of factors (Becker *et al.* 1987).

Several questions still remain to be answered. The role of post-translational modifications of transcription factors, particularly glycosylation, on tissue specific gene expression needs to be addressed. Herende' i *et al.* (1990) have described

a mechanism of promoter specificity in bacteriophage T4 which involves a non-DNA binding protein acting as a communicator between enhancer and promoter binding factors. It is not known whether coactivators of eucaryotes and communicator proteins of procaryotes are analogous in their function. In any case, further experiments are required to identify the function of coactivators and/or factors equivalent to prokaryotic communicator proteins in tissue specific gene expression.

1.1.7 Sliding, twisting, oozing, nuclear matrix association and looping models for gene regulation by proteins acting nearby and at a distance

The mechanism by which enhancers and the enhancer binding factors stimulate transcription is not clearly understood. At least five models have been proposed. The sliding model proposes that a protein binds to a specific DNA sequence and then slides along the DNA to another site where transcription is initiated. The main evidence for this model is from the SV40 enhancer which shows preference for a nearby promoter over a distant promoter for transcription (Wasylyk *et al.* 1983).

According to the less favoured twisting model, the conformation of DNA is altered in a localized area and regulatory proteins have an affinity for the altered DNA conformation. In another version, regulatory proteins possess

enzymatic activity which alter DNA conformation, allowing other proteins to bind and begin transcription (Ptashne, 1986). Although sequences that can form Z-DNA are observed in the SV40 enhancer, mutational analysis failed to support a role for this region in transcriptional regulation (Zenke *et al.* 1986).

The oozing model, which may be applicable to enhancers with closely placed promoters, proposes that the binding of a regulatory protein to its recognition sequence helps binding of another protein to adjacent sequences, which in turn helps another protein to bind until a series of proteins have worked themselves down to promoters.

Cockerill and Garrard (1986) suggest that enhancers, in conjunction with topoisomerase II and a nuclear matrix association region anchor DNA into the nuclear matrix and induce torsional stress characteristic of transcriptionally poised chromatin.

The looping model proposes that the enhancer binding-protein can interact with other regulatory proteins, potentially near the RNA start site, causing a loop in the DNA between these two protein-bound DNA sequences (Ptashne, 1986). An elegant study by Muller-Storm *et al.* (1989) supports this model. These investigators showed that the enhancer linked to a promoter in trans-configuration via an avidin-biotin bridge can stimulate transcription.

None of the above models can be conclusively rejected,

and given the dramatic impact of enhancers on transcription, it is conceivable that enhancers may function by more than one mechanism.

1.2 Papovaviruses

As mentioned previously, studies of cis-acting elements of viruses, particularly of the DNA tumor viruses, have made major contributions towards our understanding of gene expression. My interest was to study the cell type specific expression of cis-acting elements of DNA tumor viruses belonging to the papovavirus family. I briefly describe the molecular biology of these viruses below. My major emphasis here will be on the molecular mechanism of tissue tropism by these viruses.

The name, papovaviradae, derives from the initial letters of three members; papilloma virus, polyomavirus, and simian vacuolating agent (SV40) (Tooze, 1981). The members of this family are characterized by their small size, a non-enveloped virion, an icosahedral capsid, a double-standed circular DNA genome, and the nucleus as the site of multiplication. The members are divided into two subfamilies, the Polyomavirinae and Papillomavirinae, based on differences in the sizes of the virions (40 nm in Polyomavirinae versus 55 nm in Papillomavirinae), and sizes of the genomes (5000 versus 8000 bp). There are differences in subfamily specific antigens. Furthermore, nucleic acids of subfamilies do not cross-

hybridize. The pattern of transcription is also different in these subfamilies. For example, in polyomavirinae, open reading frames (ORFs) are located in both strands of DNA while the ORFs of papillomavirinae are in one strand (Fields, 1990).

1.2.1 Polyomavirinae

This subfamily includes viruses that can infect a variety of species, including birds, rodents, and primates. However, the members show a strict host range; infection of distantly related hosts is inefficient. Mouse polyomavirus, simian virus 40 (SV40), human polyomaviruses BKV and JCV, lymphotropic papovavirus (LPV), hamster papovavirus (HaPV) and Budgerigar fleding disease virus are the members of this subfamily (Fields, 1990).

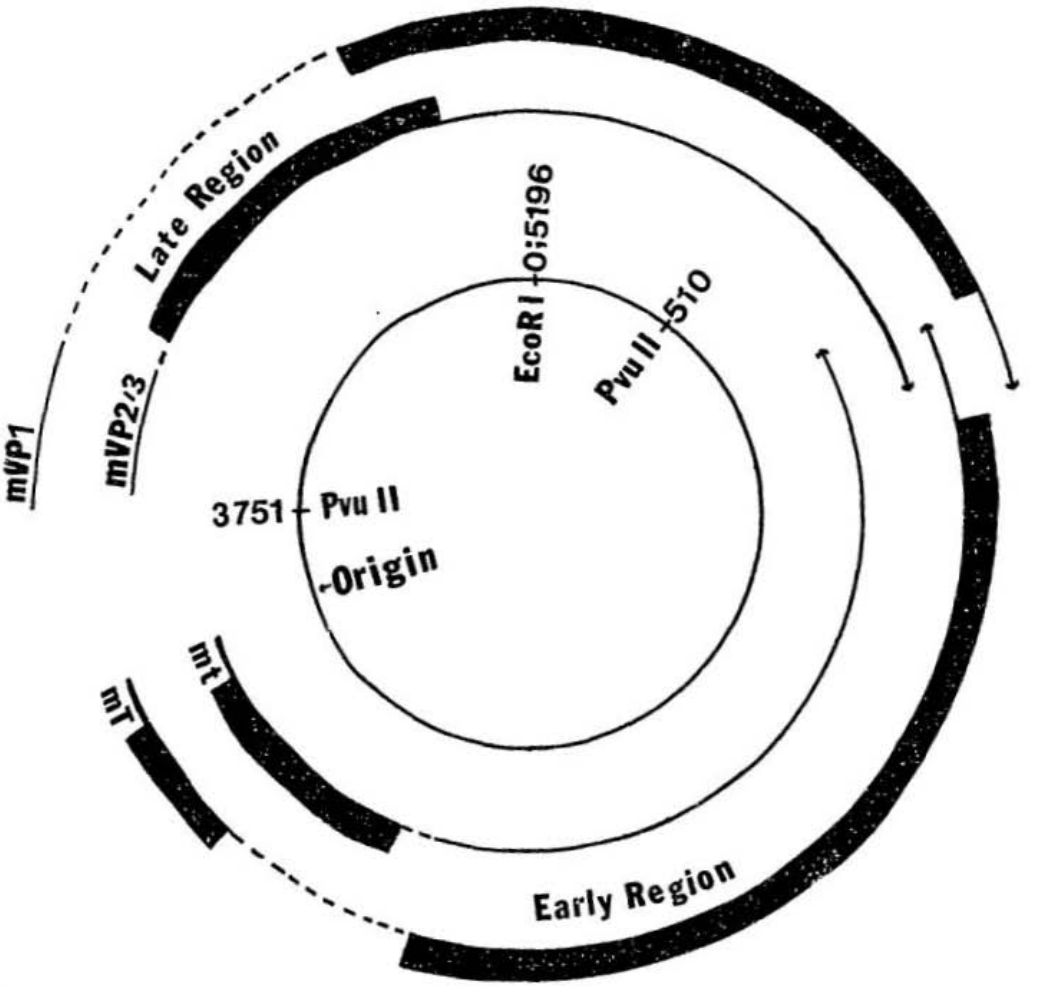
The mouse polyoma and monkey SV40 are the best characterized of the polyomavirinae. Since these viruses rely on host cellular machinery for their replication and transcription, they have served as useful tools for studying cellular processes. Although most studies on virion structure are limited to these two viruses, it is believed that other members of the family have a similar virion structure.

There are three types of virion particles in virus preparations: 1) Infectious virions, which contain viral DNA, 2) empty capsids, which lack viral DNA, and 3) pseudovirions, which contain fragments of cellular DNA. The virions contain three virus encoded proteins, referred to as VP1, VP2, and

VP3. The viral DNA in the virion is in the form of a minichromosome and is associated with four host cell histones, namely, H2A, H2B, H3 and H4. Adsorption of polyoma to a cell requires a cellular receptor and viral protein VP1. Studies with an enzyme that removes sialic acid from proteins suggest that the receptor for polyoma and JCV is sialidase sensitive while the SV40 receptor is sialidase resistant. Viruses attached to the cell surface are endocytosed, transported to the nucleus through a vesicle, and subsequently released into the nucleus by fusion of the vesicle into the nuclear membrane (Fields, 1990).

Transcription of viral genes occurs in at least two stages: early (before viral DNA replication) and late (after viral DNA replication). Organization of the early and late genes of BKV is presented in Figure 1.1 as an example. The number of proteins from early and late regions differs among the members of this subfamily. The early gene products, large T-antigen and small t-antigen are common to all members while polyoma and HaPV encode an additional middle T antigen. Similarly, VP1, VP2, and VP3 are common late gene products while SV40, JCV, and BKV encode an additional agnoprotein (Fields, 1990). Transcription of these genes is under the control of a regulatory element which also harbours the origin for viral DNA replication. The viral early proteins, mainly large T-antigen, play an important role in stage specific viral transcription (Fields, 1990). Since the regulatory

Figure 1.1 Physical map of BK virus. The genome is a double stranded circular DNA molecule of 5196 bp. the unique EcoRI restriction enzyme site is numbered 0/5196. Outer half circles represent transcripts arising from the genome where the 5′ non-translated region, translated region, introns and 3′ non-translated regions are represented by thin lines, thick lines, broken lines and thin lines with an arrow head, respectively. mT and mt indicates messages for large T-antigen and small t-antigen, respectively. Polycistronic messages for VP2 and VP3 (mVP 2/3) and VP1 (mVP1) are also indicated.


mVP1
mVP2/3
Late Region
EcoR I +0;5196
Pvu II +510
3751 – Pvu II
-Origin
mT
mT
Early Region

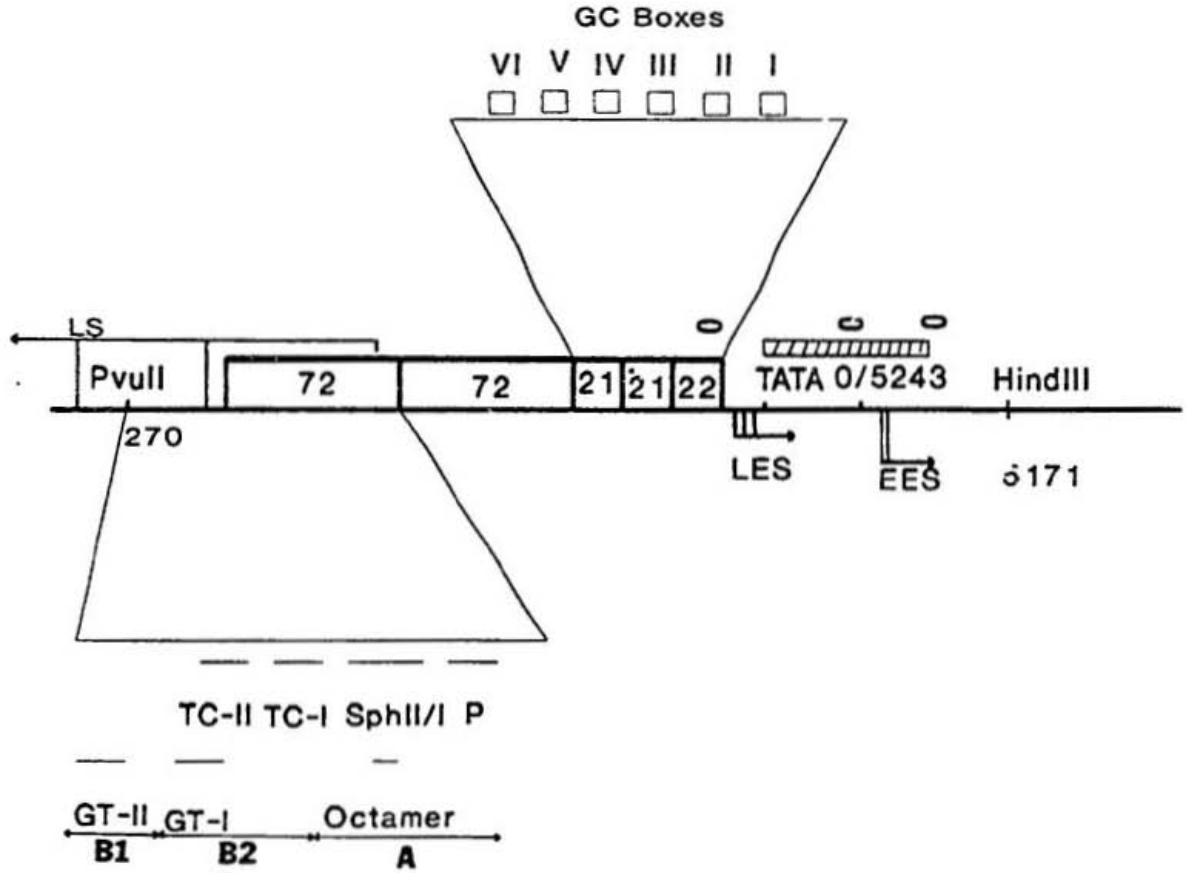
elements of SV40, JCV, and BKV are the topic of this study, detailed description of their structural organization is presented below.

1.2.1.1 SV40 transcription regulatory elements

As with most eucaryotic regulatory elements, SV40 regulatory elements contain a promoter and an enhancer both of which are essential for efficient early and late gene transcription (Fig. 1.2). The promoter and enhancer of this virus are active in a wide variety of cell types, as are the housekeeping genes of eucaryotic cells. In vivo and *in vitro* studies have shown that the SV40 early promoter region consists of two overlapping promoters which control initiation of transcription at the early-early (EE) and late-early (LE) start sites. The EE promoter (EEP) is composed of i) the TATA box sequence, which ensures accurate and efficient initiation of transcription from the EE start site (EES); ii) an upstream element, the 21-bp repeat region; and iii) the enhancer shown as 72 bp repeats in Figure 1.2. Both the upstream element and the enhancer are required for efficient transcription (Barrera-Saldana *et al*. 1985). The distance between the TATA box and the 21-bp repeat region is critical for accurate initiation. The 21-bp repeat region is composed of three imperfect 21 bp repeats and contains six copies of a hexanucleotide sequence 5′-GGGCGG-3′. Each of these hexa

Figure 1.2 Structural organization of SV40 regulatory elements. SV40 regulatory elements contain a TATA box (TATA), three imperfect 21 bp repeats (21), two 72 bp repeats (72) and 35 bp upstream of the second 72 bp repeat (towards late region). Early transcription start sites are indicated by arrows. Six GC-boxes (I-VI) present within three 21 bp repeats are also indicated. Sequences of 72 bp repeats and repeat upstream region are subdivided into three functional domains (A, B_1, and B_2). Both 72 bp repeats contain P, SphI, SphII, octamer, TC-I and TC-II motifs. Sequences upstream of the second repeat (towards late region) contain GT-II motif.

 = Origin of replication. EES = Early-Early start site. LES = Late-Early start site. LS = Late genes start site. ○ = T-antigen binding site.


GC Boxes
VI V IV III II I
LS
PvuII
72
72
21
21
22
TATA 0/5243
HindIII
270
LES
EES
5171
TC-II TC-I SphII/I P
GT-II GT-I Octamer
B1 B2 A

nucleotides interacts with one molecule of transcription factor Sp1 (Barrera-Saldana *et al.* 1985). Sp1 binds non-cooperatively; the sequences adjoining the hexanucleotides determine binding affinity. When the promoter is fully occupied, all Sp1 molecules lie on the same face of the helix (Barrera-Saldana *et al.* 1985; Courey *et al.* 1989). Interaction of Sp1 with GC-boxes I, II, and III is required for *in vitro* synthesis of early messenger RNA, whereas transcription in the late direction is mediated by binding of Sp1 to GC-boxes III, V, and VI (Barrera-Saldana *et al.* 1985). GC-boxes II and III are capable of interacting with another factor called LSF which is required for late gene transcription (Huang *et al.* 1990). In addition, transcription factor AP2 can also interact with GC-boxes both independently and concurrently with Sp1 (Mitchell *et al.* 1987).

Genetic analysis has shown that the two 72 bp repeats and sequences extending approximately up to 35 bp upstream from the second 72 bp repeat (towards the late region) in the SV40 regulatory region possess enhancer activity. Functional analyses have defined three domains, A, B1, and B2, that can compensate for one another (Fig. 1.2; Zenke *et al.* 1986; Ondek *et al.* 1987;). Studies involving site-directed mutagenesis and naturally occuring mutants (dcm mutants) have identified several enhansons. As presented in Figure 1.2, these are designated as P, SphI, SphII, Octamer, TC-I, TC-II, GT-I, GT-

II (Zenke *et al.* 1986; Ondek *et al.* 1987; Jones *et al.* 1988).

The P motif binds to transcription factor AP-1 and mutations that alter this motif have only a modest effect on enhancer activity in both HeLa and lymphoid cells (Jones *et al.* 1988). However, AP-1 binding sites are required for TPA and *ras* oncogene-mediated transcriptional activation (Angel *et al.* 1987; Imler *et al.* 1988).

Enhansons SphI, SphII, and GT-IIC can interact with the same cellular factor, TEF-I (Davidson *et al.* 1988). These enhansons are organized into the first type of proto-enhancer (see 1.1.2) where the cooperative interaction of factors generates synergistic activity. The junction of SphI and SphII motifs creates a second motif, the octamer motif, that is very similar to the octamer motif found in the immunoglobulin (IgH) promoter and enhancer, and in the promoters of the snRNA and histone H2B genes (Rosales *et al.* 1987; Jones *et al.* 1988). This enhanson is essential for expression of SV40 in lymphoid cells and interacts with one ubiquitous and three cell-type specific proteins (Schirm *et al.* 1987; Rosales *et al.* 1987).

Mutation of TC-II but not TC-I lowers the enhancer activity 25 fold in HeLa cells (Zenke *et al.* 1986). TC-II interacts with at least three proteins, including the ubiquitous factor KBF1/H2TF1 and the cell-type specific factors NFkB-like factor and AP2 (Macchi *et al.* 1989; Kanno *et al.* 1989). The GT-I enhanson, which along with TC-II forms

the enhancer C core, interacts with factors TEF-1 and AP3 (Davidson *et al*. 1988; Jones *et al*. 1988). This region of the enhancer appears to be important for enhancer function since mutations in the region have a deleterious effect on the enhancer function. Additionally, multimerized C core region, but not GT-I alone, is functional as an enhancer in a variety of cell lines tested (Schirm *et al*. 1987; Fromental *et al*. 1988).

GT-II is the most complex of all enhansons in SV40. It contains three overlapping motifs, GT-IIA, GT-IIB, and GT-IIC, capable of interacting with at least four factors (Xiao *et al*. 1987). Multimerized GT-II is active in HeLa, CV-1, and embryonal carcinoma (EC) cells but not in lymphoid cells (Nomiyama *et al*. 1987; Fromental *et al*. 1988; Schirm *et al*. 1987). GT-IIA mutants show a small decrease in enhancer activity which is more pronounced in MPC II and EC cells than in HeLa cells (Nomiyama *et al*. 1987). Similarly, GT-IIB, which does not have enhancer activity on its own, binds to a factor identical to that of the IgH enhancer AE3 sequence motif binding factor (Xiao *et al*. 1987). Most of the enhancer activity for the GT-II enhanson is derived from the GT-IIC motif (Fromental *et al*. 1988). Thus, the summation of various results suggests that the SV40 enhancer domain A contains enhansons P, SphI, SphII, and octamer, domain B1 contains GT-II, and B2 contains TC-I, TC-II, and GT-I. Schartz and Chalton (1990) propose that sterospecific alignment of these

enhancer motifs and the promoter is required for efficient initiation of transcription from the early promoter.

Transcription from the early genes is autoregulated by T-antigen. Two mechanisms for autoregulation have been proposed. The first is by sequence-specific interaction of T-antigen to three binding sites present downstream of the TATA box (towards the early gene) (DiMaio and Nathans, 1982). Although all these sites contain two or more repeats of the pentanucleotide binding signal GAGGC, they differ in function: binding to site I is primarily involved in repression of early transcription, whereas binding to site II is critical for the initiation of DNA replication (DiMaio and Nathans, 1982). The binding of T-antigen to these sites prevents the formation of the transcription initiation complex in the early promoter. The second proposed mechanism involves interaction of T-antigen with transcription factor AP2. This T-antigen-AP2 complex is incapable of interacting with the AP2 recognition site found within the enhancer and promoter (Mitchell *et al.* 1987). In the life cycle of the virus, these events are required to trigger the replication of viral DNA and allow the transcription of viral late genes.

Late gene transcription requires the late promoter which is very complex and includes the origin of replication, the three 21-bp repeats, the 72 bp repeats, and the region upstream from the 72 bp repeats beyond the major *in vivo* initiation site (Brady *et al.* 1982; Keller and Alwine, 1985;

Ayer and Dynan, 1988). Unlike most genes transcribed by RNA polymerase II, the late transcription unit does not contain a well defined TATA box. However, point mutational analysis has identified three regions with functional anology to the TATA box: a sequence 31 bp upstream of the cap site, the cap site itself, and a sequence 28 bp downstream from the cap site (Ayer and Dynan, 1988). The DNase I footprinting and *in vitro* transcription studies have identified at least two factors, AP4 and LSF, that are required for the late gene transcription (Mermod *et al.* 1988; Huang *et al.* 1990).

In summation, the various results discussed above indicate that the SV40 enhancer is the most complex among all studied so far and continues to be a source of information in eucaryotic gene regulation. My studies have focused on the transcription of SV40 in a pluripotent embryonal carcinoma cell line. The purpose was to identify differentiation induced change in the rate of transcription as well as DNA-protein interaction. The ultimate aim was to comparatively analyse SV40 regulatory elements with tissue-specific JCV regulatory elements and the widely active BK regulatory element.

1.2.1.2 JC virus transcription regulatory elements

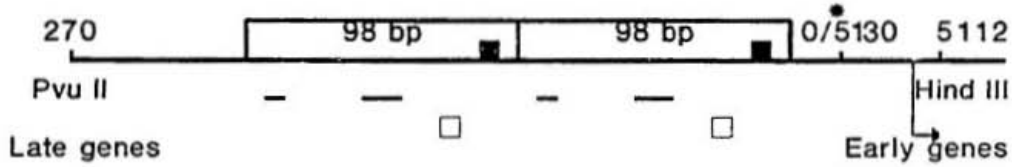
Infection of JCV is prevalent in the human population and under certain circumstances may lead to to the fatal brain disease, progressive multifocal leukoencephalopathy (Padgett

and Walker 1973; Padgett *et al.* 1976). In patients with this disease, JCV particles have been detected in brain cells of glial origin - namely, oligodendrocytes which maintain the myelin sheath, and at a much lower level in astrocytes (Dorries *et al.* 1979). In tissue culture, the virus can productively infect foetal glial cells and to a lesser extent, human uroepithelial cells, amnion cells, and endothelial cells (Padgett *et al.* 1977a; Beckemann *et al.* 1982; Takemoto *et al.* 1979a). The restricted host range and infectivity shown by JCV has prompted many investigators to use this viral DNA to study the factors that determine host range and tissue specific gene expression of viruses.

Viral regulatory elements that determine the host range have been identified by comparative analysis with the closely related SV40 and BKV regulatory elements and by deletion mutations (Kenney *et al.* 1984; Frisque *et al.* 1984; Feigenbaum *et al.* 1987). The host range is determined, at least, at two levels: cell type-specific transcription of early RNA and species-specific replication of the viral DNA in primate cells (Feigenbaum *et al.* 1987).

The cell type-specific transcription of early RNA synthesis requires two 98 bp repeats located on the late side of the origin of DNA replication (Fig. 1.3; Kenney *et al.* 1984). There is a 15 nucleotide AT-rich segment in each repeat that functions as a TATA box. The TATA box in the first repeat is required for accurate initiation whereas the TATA of the

Figure 1.3 JC virus regulatory element. Two 98 bp repeat region of the regulatory elements are shown. The transcription start site for early genes is indicated by an arrow.



p2A 120 homology

Myelin promoter homology

TATA box

Origin of replication

second repeat may have some regulatory role: the naturally occurring variant MAD-4 with a deletion of the second TATA box induces tumors that are distinct from those of prototype MAD-1 (Padgett *et al.* 1977b; Martin *et al.* 1985). The TATA box along with adjacent sequences confer an unusual DNA structure which is non-B, but is right-handed (Amirhaeri *et al.* 1987).

Unlike SV40 regulatory elements, very little is known about JCV regulatory elements. The main obstacle has been the lack of a suitable tissue culture system. However, results of a few studies listed below suggest that the 98 bp repeats contain signals responsible for the neurotropic expression of this virus. 1) There is a stretch of nucleotides within these repeats that shows significant homology to an 82-nucleotide element present within the introns of rat brain-specific precursor RNA molecules, p2A 120 (Kenney *et al.* 1984). 2) The promoter region of the tissue specific human myelin basic protein gene contains a 12 bp sequence which is also found in the 98 bp repeats (Kamholz *et al.* 1988). 3) JC regulatory elements can compete for nuclear factors that bind to myelin basic protein gene promoters (Tamura *et al.* 1988a). 4) The naturally occurring organ-specific variants and variants adapted to grow in primate cells (other than foetal glial cells) show marked rearrangements in the 98 bp repeats. In most of the rearranged regulatory elements the parental sequences are replaced by DNA sequences with homology to SV40 and adenovirus core enhancer elements (Loeber and Dorries.,

1988; Martin *et al.* 1985; Shinohara *et al.* 1989). 5) The heterokaryons generated between JCV-transformed hamster glial cells and mouse fibroblasts exhibit an extinction of large T-antigen. Since this decrease requires the viral regulatory elements, it could be mediated through a negative regulatory factor present in the non-expressing fibroblasts (Beggs *et al.* 1988). Thus, both cell type-specific positive and negative factors may be involved in JCV gene regulation. My studies are directed towards identifying the sequences that interact with factors present in both expressing and nonexpressing cells. Furthermore, a comparative analysis of JCV with SV40 and BKV, the latter two of which are active in a wide range of cell types, would permit a clear demonstration of differences between tissue-specific and widely active enhancers.

1.2.1.3 BK virus transcriptional regulatory elements

Like JCV, BKV is also distributed worldwide, infecting 70% of the population in early childhood (Tooze, 1981). Following this initial infection, the virus goes into a latent state, predominently in the kidney cells (Heritage *et al.* 1981). Several variants of this virus with different biological properties have been isolated, and sequence analyses have shown that they contain marked rearrangements in the regulatory region (Seif *et al.* 1979; Yang and Wu, 1979; Pater *et al.* 1983; Berg *et al.* 1988; Watanabe *et al.* 1984;

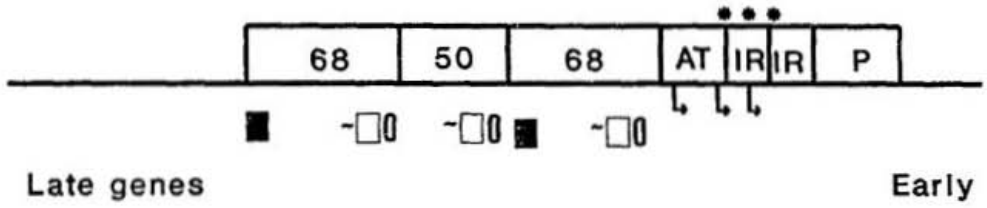
Pagnani *et al*. 1986).

The prototype BKV regulatory element contains a true palindrome of 17 bp followed by two sets of symmetrical sequences and a stretch of 20 AT residues, all of which form the replication origin. In addition, there is an inverted repeat and three direct 68 bp repeats (the middle one of which has an 18 bp deletion) (Fig. 1.4, Seif *et al*. 1979; Deyerle *et al*. 1989).

Rearrangement of the 68 bp repeats has generated variants with different biological activities. For example, mutant pm-522 having one 68 bp repeat unit and two sets of shorter 37 bp repeats in the regulatory region transforms rat and hamster cells more efficiently than does the wild type (Watanabe et al., 1984).In contrast, BK-IR, which has rearrangements in the first and second 68 bp repeats, transforms hamster cells with reduced efficiency and induces ependymomas in hamster at a lower frequency and with a longer latency period (Pagnani *et al*. 1986).

The regulatory elements of BKV have been characterized further. They are as follows. 1) These elements, like SV40 regulatory elements, have a constitutive enhancer-promoter activity in a wide range of cell types. Furthermore, human cells contain a homolog of these elements with similar properties (Rosenthal *et al*. 1983). 2) Unlike SV40 and polyoma, the BKV promoter extends downstream of the transcription initiation sites. Accurate initiation requires the inverted

Figure 1.4 BK virus regulatory elements. Palindromic sequences (P), inverted repeats (IR), AT rich region (AT) and enhancer repeats are shown. Early transcription start sites are indicated by arrows.



0 Adenovirus E1A enhancer core homology

□ SV40 enhancer core homology

■ Sp1 homology

~ NF1 homology

• • • Origin of replication

repeats, an AT-rich box and the first of the three 68 bp repeats (Fig. 1.4; Deyerle *et al*. 1987). The promoter and enhancers are overlapping elements, and are more homologous to those of JCV than SV40 and polyoma. 3) For efficient transcription from the early promoter, all three 68 bp repeats and a stretch of approximately 30 bp upstream of the third repeat (towards the late region) are required. Deletion studies have identified two putative enhansons bearing homology to the GT-II motif of SV40 and the adenovirus 5 E1A enhancer core (Deyerle *et al*. 1987). 4) The sequences within the 68 bp repeat units contribute to the species-specific differences (Watanabe and Yoshiika 1989; Deyerle and Subramani, 1989). For example, the wild type BKV enhancer functions positively on heterologous genes but negatively on its own gene in rat cells, whereas it is always positive in human cells for both heterologous and homologous genes. This property of the wild type enhancer is attributable to the sequences homologous to the SV40 enhancer core and the GT-I motif (Watanabe and Yoshiike, 1989). 5) The 18 bp deletion in the second repeat in the naturally occurring BKV optimizes early enhancer function at the expense of late promoter activity (Cassill and Subramani, 1989a). 6) The deletion of the 68 bp repeat units elevates the autoregulatory effect of BKV T-antigen (Deyerle and Subramani, 1989). 7) The late promoter of BKV requires a region that binds to Sp1, NF-1 protein

binding motifs, and a GC-rich sequence within each of the 68 bp repeats which has homology to the Sp1 binding site but is unable to bind to purified Sp1 (Cassill *et al*. 1989b). In addition, several elements in the enhancer confer cell type variability for late promoter activity. For example, a region with homology to the E1A enhancer core is necessary for activity of the late promoter in HeLa but not in CV1 cells. Similarly, the NF-1 site at the junction between the 68 bp repeat and the late region is dispensable in HeLa cells but not in CV-1 cells (Cassill *et al*. 1989b). 8) Transactivation of the late promoter by BK T-antigen requires neither a specific motif nor direct DNA binding (Cassill *et al*. 1989b). Taken together, these studies suggest that the repeat elements and adjoining sequences contain signals that are responsible for host range, tissue tropism, pathogenicity, and oncogenicity of BKV (Deyerle and Subramani, 1989; Small *et al*. 1986; Watanabe and Yoshiike, 1989; Sugimoto *et al*. 1989).

Although the above mentioned studies have noted sequences homologous to E1A enhancer core, SV40 enhancer core, Sp1 motif and NF-1 motif in the BKV enhancer, not much is known about the functional organization of these motifs and the factor(s) interacting with them. Since the BKV enhancer resembles SV40 in its activity profile, it is tempting to speculate that both enhancers are organized in a similar fashion. Evidence for this hypothesis is obtainable with experiments involving comparative analyses of SV40 and BKV enhancers with respect

to their expression and DNA-protein interactions.

1.2.2 Papillomavirinae

Over the past 5 years, there has been a virtual explosion in research on papillomaviruses for two main reasons. Firstly, there is compelling evidence for the role of papillomaviruses in the genesis of cancer. Secondly, papillomaviruses are proving to be the best vehicle to study keratinocyte differentiation and gene expression as these viruses interact specifically with terminally differentiated keratinocytes.

The members of this subfamily are widespread in nature and have been recognized primarily in higher vertebrates. They induce squamous epithelial cell tumors and fibroepithelial tumors in their natural hosts (Howley, 1990; Giri and Danos, 1986). The oncogenic potential of papillomaviruses was recognized in early studies on cottontail rabbit papillomaviurses (reviewed by Howley, 1990). More recent studies involving epidemiology and molecular biology have established the relationship between human papillomaviruses (HPVs) and certain human cancers, most notably cervical carcinoma. Consequently, a large body of information has been obtained on the molecular biology of certain genital associated HPVs. The following will be a brief discussion on the molecular biology of these viruses.

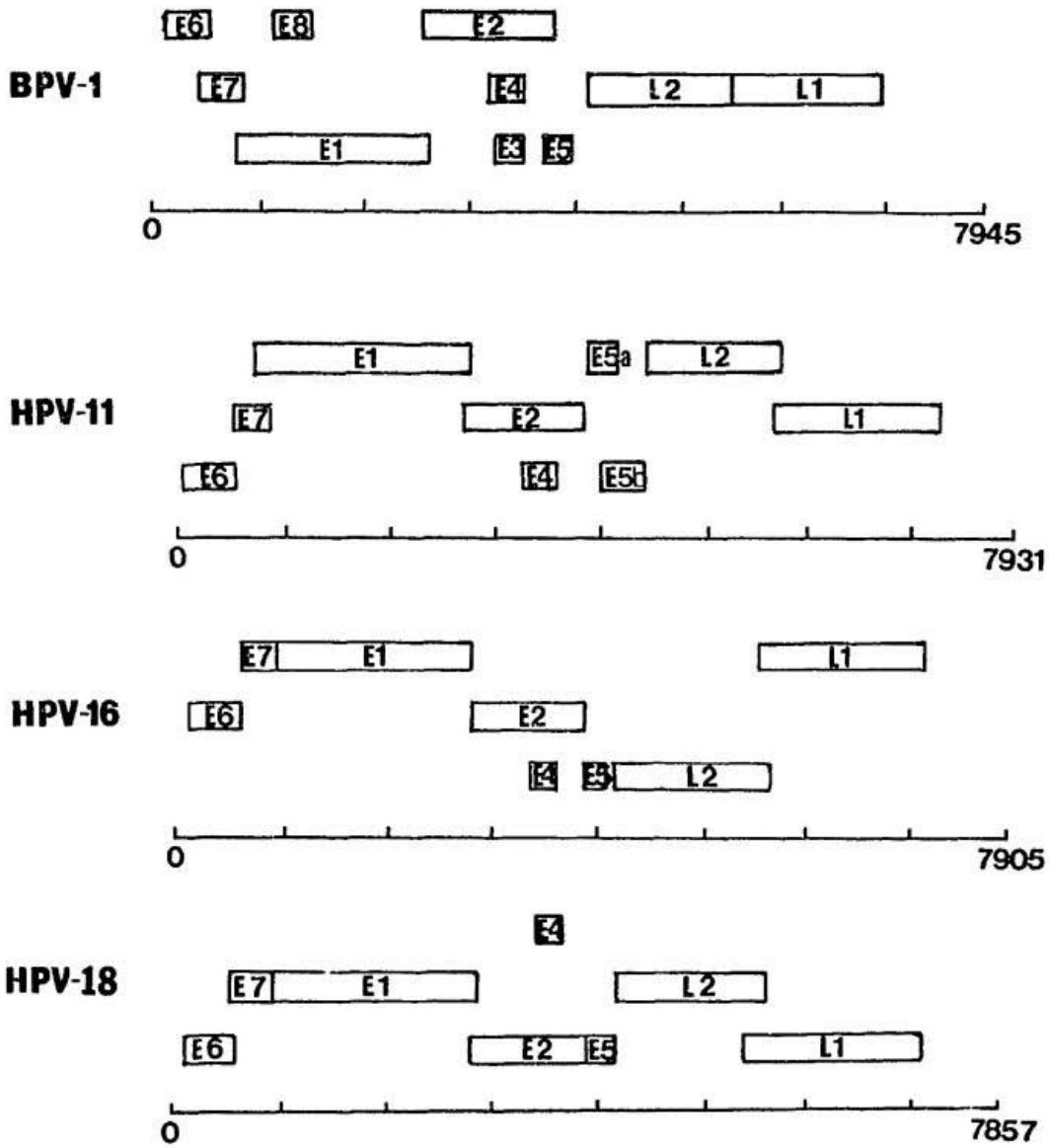
Papillomaviruses are small, nonenveloped, icosahedral DNA

viruses that replicate in the nucleus. The virion particles consist of a single molecule of double-stranded circular DNA of approximately 8000 bp in size that is associated with cellular histones (Howley, 1990). The virion capsid consists of two structural proteins encoded by the L1 and L2 genes of the viral DNA. The mode of viral entry is not known mainly due to the lack of success in propagating these viruses in culture.

Despite the fact that there is no *in vitro* tissue culture system for the virus propagation, viral DNA has been analysed extensively. First among such viruses studied was bovine papilloma virus type-1 (BPV-1) which is more readily obtainable because of the accessibility of BPV particles from lesions of infected animals. On the contrary, human papillomaviruses (HPVs) DNA have been isolated from tumor biopsies and are characterized with the help of cloning systems in bacteria. To date, 60 different types of HPVs have been isolated, with individual types showing less than 50% homology to all other types in liquid hybridization under stringent conditions (deVillers, 1989).

The molecular organization of HPVs is defined on a comparative basis with BPV-1 (Howley, 1990; Giri and Danos, 1986). Figure 1.5 compares the genomic organization of BPV1 with the three HPVs used in this study. The BPV-1 genome is functionaly divided into two domains: an early region (E) comprising 69% of the genome which is sufficient to transform

Figure 1.5 Comparison of the genomic organization of BPV1, HPV 11, HPV 16 and HPV 18. These are a linear maps of the circular genomes. The early (E) and late (L) region open reading frames are represented by open boxes. Most of the non-coding region (NCR) sequences are in the extreme right end (following L1 open reading frame).


BPV-1
E6
E8
E2
E7
E4
L2
L1
E1
E3
E5
0
7945
HPV-11
E1
E5a
L2
E7
E2
L1
E6
E4
E5b
0
7931
HPV-16
E7
E1
L1
E6
E2
E4
E5
L2
0
7905
HPV-18
E4
E7
E1
L2
E6
E2
E5
L1
0
7857

mouse fibroblasts, and a late region (L). The early region consists of seven open reading frames designated E1-E7 which encode proteins required for regulated episomal DNA replication (E1 ORF), enhancer activation or repression (E2 ORF), and cellular transformation (E5, E6 and E7 ORFs). The late region contains two open reading frames, L1 and L2, coding for structural polypeptides of the virions (Howley, 1990; Giri and Danos, 1986). A single region in each papillomavirus genome, referred to by several terms including the long control region (LCR), the upstream regulatory region (URR), and the non-coding region (NCR), contains signals necessary for RNA polymerase II-mediated transcription. Although the early region proximal portion of the NCR is highly conserved (especially among genital HPVs), homology studies have shown that the NCR is one of the most variable regions even among related HPVs (Giri and Danos, 1986). The 3' extremity close to the L1 ORF includes a domain which is GT-rich in genital HPVs and AT-rich in other viruses (Giri and Danos, 1986). This domain contains a polyadenylation signal for the transcripts of the L region and a sequence required for episomal replication (characterized only in BPV-1) (Lusky and Botchan, 1984; Howley, 1990).

As mentioned previously, papillomaviruses are strictly epitheliotropic and appear to be exclusively dependent on the environment provided by terminally differentiating keratinocytes for replication. Although the viruses can infect both stem cells in the epithelial basal layer and differentiating

keratinocytes, viral capsid production and virion assembly occur only in the upper differentiating layer (Giri and Danos, 1986). Thus, the papillomavirus life cycle may involve programmed early and late gene expression which is controlled by the state of differentiation of the keratinocyte. For example, early gene expression may occur in all layers of the epidermis whereas late gene expression is restricted to the upper most differentiating layer (Stoler *et al.* 1989). As a prerequisite for understanding this programmed gene expression and life cycle, many investigations have been performed with NCR elements of these viruses. Since my studies involve the NCRs of HPV types 11, 16 and 18, a review of the literature on these viruses with emphasis on the NCR is presented below.

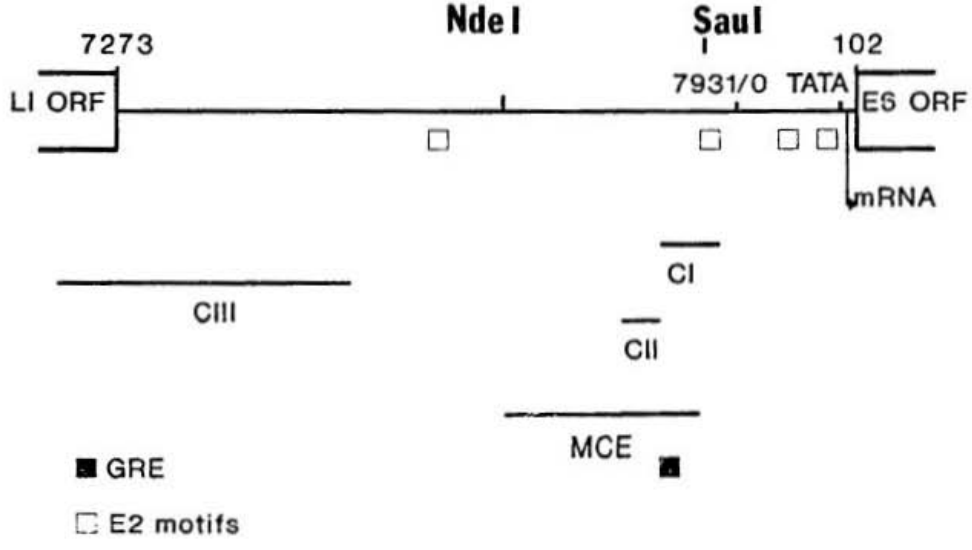
1.2.2.1 HPV 11, 16, and 18 NCR

Human papillomavirus type 11 DNA was originally cloned from a laryngeal papilloma, whereas types 16 and 18 DNA were from invasive cervical carcinomas (Gissmann *et al.* 1982; Seedorf *et al.* 1985; Boshart *et al.* 1984). Subsequent studies have identified HPV 11 DNA in condylomata acuminata and nasal inverting (Schneiderian) papillomas and HPV 16 and 18 DNA in cervical carcinoma cell lines (Gissmann *et al.* 1983; Respler *et al.* 1987 Pater and Pater, 1985). Although all three viruses show strict epitheliotropism, significant differences with regard to their sites of infection and degree of oncogenicity have been observed. For example, HPV 11 is

strongly associated with benign proliferative lesions, such as condylomata acuminata and laryngeal papillomas, which rarely progress into cancer whereas HPV 16 and 18 are associated with genital tract lesions and are risk factors for malignant progression (zur Hausen and Schneider, 1987). A major goal of some investigators is to understand the molecular basis for the above difference.

Like many other HPVs, the NCR of HPV 11 extends from the stop codon for the L1 ORF to the initiation codon for the E6 ORF (Dartmann *et al.* 1986). Cloning of these sequences upstream of enhancer-promoterless reporter genes has suggested that this region contains sequences for both enhancer and promoter function (Marshall *et al.* 1989). The central focus of study has been on the enhancer elements which can be functionally dissected into five independent units (Fig. 1.6). Three of these elements show constitutive activity while two others function in an inducible manner. A 269 bp fragment (nucleotides 7224-7493) located close to the L1 open reading frame forms one constitutive enhancer (CIII) which shows activity in different cell types including fibroblasts (Auborn *et al.* 1989; Steinberg *et al.* 1989). The second constitutive enhancer called CI (nucleotides 7761-7904) is active in the CV-1 cell line but not in C33A or HeLa cells (Hirochika *et al.* 1988; Chin *et al.* 1989). The third constitutive enhancer, CII, (nucleotides 7700-7769), is a tissue specific enhancer which is required for expression in cervical carcinoma cell

Figure 1.6 The noncoding region of HPV 11. MCE = minimal constitutive enhancer characterized in our laboratory (Marshall *et al.* 1989) and used in my studies. Arrow indicates transcription start site. CI, CII and CIII represent constitutive enhancers as described by Auborn *et al.* (1989), Hirochika *et al.* (1988) and Chin *et al.* (1989). GRE = Glucocorticoid responsive element. E2 motif = recognition sequence for E2 proteins. NdeI-SauI fragment indicated in the figure is used to the construct plasmid pT-1, as described in chapter 3


NdeI
SauI
7273
102
7931/0
TATA
LI ORF
E6 ORF
mRNA
CI
CIII
CII
MCE
GRE
E2 motifs

lines but is dispensable for expression in CV-1 cells (Chin *et al.* 1989). A 213 bp fragment (nucleotides 7657-7870) encompassing CI and CII has been designated as a minimal constitutive enhancer by our laboratory (Marshall *et al.* 1989).

One of the major discoveries of recent years in virus research is that some virally encoded proteins behave as transcriptional factors which influence the transcription from both viral and cellular genes. As described below, studies with papillomaviruses indicated that proteins encoded by the E2 ORF function as transcription factors. The number of proteins generated from this ORF varies from virus to virus.

The E2 ORF of HPV 11 encodes two proteins, one corresponding to the full length of the ORF (E2) while the other is a truncated protein containing only the carboxy-terminus of the ORF (E2-C) (Hirochica *et al.* 1988; Chin *et al.* 1988). These proteins are derived from four mRNAs, two of which are polycistronic and are generated by alternative splicing (Rotenberg *et al.* 1989). Although most of the structural analyses of these proteins are restricted to BPV1 E2 proteins, homology in amino acid sequences predicts a similar structure for HPV 11 E2. The amino terminal portion of these proteins contains an amphipathic helical structure with a relatively high negative charge, a characterstic of transcription transactivation domains, whereas the carboxy terminus

possesses a DNA binding domain (Howley, 1990; Ward *et al.* 1989). Since E2-C lacks the transactivation domain, it functions as a repressor. When both proteins are present in the same cell, inhibition of E2 transactivator function could be either due to competitive binding of E2-C to the responsive elements (see below), or to an inactive heterodimer of E2/E2-C at the binding sites. E2-C alone may repress E2-independent expression of the constitutive enhancer by steric hindrance with the binding of other transcription factors (Chin *et al.* 1988). E2 proteins of heterologous HPVs such as type 1, 6b, 16, 18, BPV-1, and cottontail rabbit papillomavirus can also induce expression from HPV 11 enhancer-SV40 promoter reporter genes (Hirochika *et al.* 1987).

The palindromic sequence 5′ ACCGN$_4$CGGT 3′ which is often repeated in the NCR of HPVs has been identified as an E2 ORF-dependent enhancer element. E2 proteins interact as dimers with these sequences. Affinity for the E2 protein is determined by the divergent sequences within the palindrome (Howley, 1990; Hirochika *et al.* 1988). As shown in Figure 1.6, a total of four such elements, two of them adjacent to the TATA box, have been identified in HPV 11 (Hirochika *et al.* 1988). This element can function independently of remaining sequences of the NCR as plasmids containing the SV40 promoter and two or more copies of this element function in an E2 dependent manner (Hirochika *et al.* 1988).

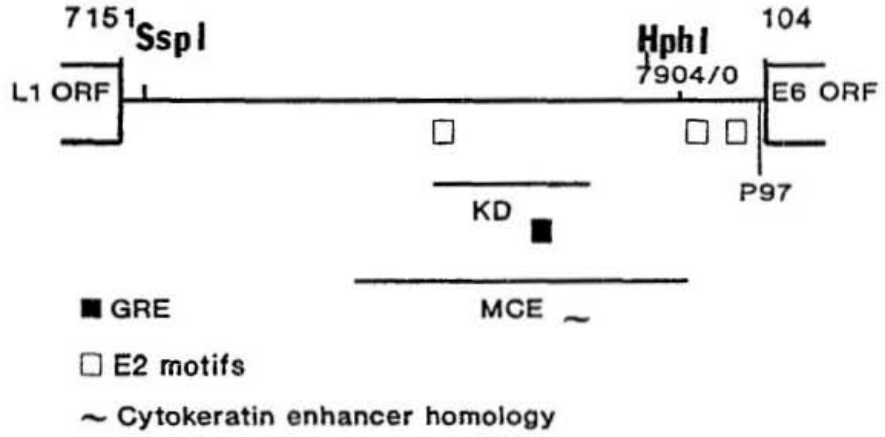
In addition to the E2 inducible enhancer, HPV 11 contains

sequences homologous to glucocorticoid responsive elements. Recent studies have indicated that these sequences are required for both dexamethasone and progesterone induced HPV 11 gene expression (Pater *et al.* 1988; Chan *et al.* 1989).

The NCR of HPV16 is 926 bp long and contains several features identical to the HPV 11 NCR (Seedorf *et al.* 1985). The transcription of early genes in HPV 16 is initiated from a promoter called P97 located approximately 50 bp upstream of the transcription start site. The promoter contains a Tr type TATA box without any well defined CCAAT or GC boxes (Seedorf *et al.* 1985).

Three independent enhancers, two of which are inducible, have been identified in the NCR of HPV16 (Fig. 1.7). The first inducible enhancer responds to the E2 ORF proteins of HPV 16 and BPV-1 (Cripe *et al.* 1987; Phelps and Howley, 1987). A total of three E2 binding palindromic sequences centered 40, 56, and 544 bp upstream of P97 have been identified (Fig. 1.7). The second inducible enhancer is homologous to the glucocorticoid responsive element. It can stimulate transcription of a linked gene in both glucocorticoid and a progesterone-dependent manner (Gloss *et al.* 1987; Pater *et al.* 1988a). Deletion mutational analyses have identified a 224 bp fragment which shows constitutive enhancer activity in uninfected human foreskin keratinocytes and in cervical carcinoma cell lines but not in human or animal fibroblasts or in a human breast cancer cell line (Cripe *et al.* 1987;

Figure 1.7 The noncoding region of HPV 16. MCE = minimal constitutive enhancer characterized in our laboratory (Marshall *et al.* 1989) and used in my studies. Transcription promoter p97 is also indicated. KD is the keratinocyte dependent enhancer described by Cripe *et al.* (1987). GRE = Glucocorticoid responsive element. E2 motif = recognition sequence for E2 proteins. SspI-HphI fragment was used to construct the plasmid pT-3, as described in chapter 3

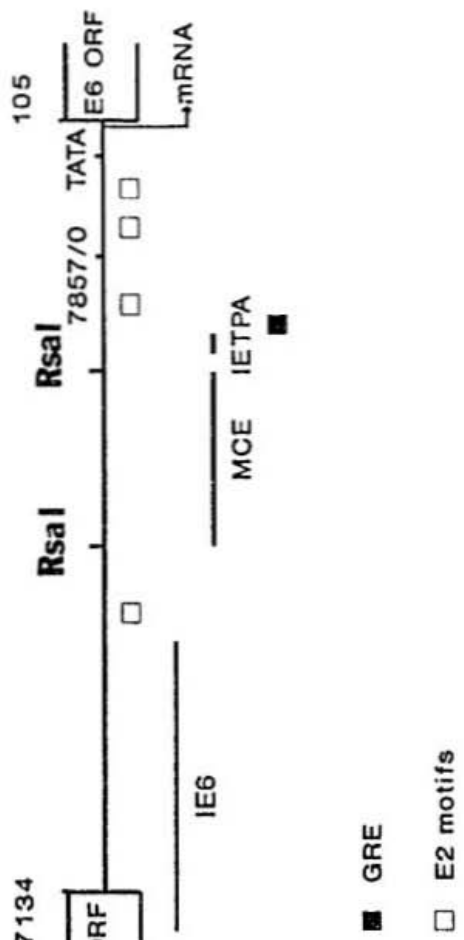

7151
SspI
HphI
7904/0
104
L1 ORF
E6 ORF
P97
KD
MCE ~
■ GRE
□ E2 motifs
~ Cytokeratin enhancer homology

Gloss *et al.* 1987; Marshall *et al.* 1989). The unique feature of this fragment is that it harbours a stretch of nucleotides 5′ TTTGGCTT 3′ which is also present in enhancer elements of human and bovine cytokeratin genes and human involucrin genes (Blessing *et al.* 1987). Thus, the current focus of research is to identify the importance of this homologous sequence and additional sequences in keratinocyte-dependent activity.

The NCR of HPV 18 is composed of 825 bp fragment with a typical TATA (Tr type) and CCAAT boxes and several other sequences conserved among papillomaviruses (Fig. 1.8; Thierry *et al.* 1987). Deletion mutagenesis of this fragment has identified the promoter, P105, which functions most efficiently in human epithelial cells (Thierry *et al.* 1987; Bernard *et al.* 1989). Promoter activity can be stimulated by SV40 large T-antigen and repressed by the adenovirus E1A gene product (Thierry *et al.* 1987).

The HPV 18 NCR contains four independent enhancers, three of which are inducible while the fourth is specific for epithelial cells. Four copies of the E2-binding palindromic sequences form the first type of inducible enhancer that responds to E2 proteins of HPV 18 and BPV-1. Two of these elements are found between the TATA box and the CCAAT box-like sequence and confer inducible activity in an enhancer configuration and repressor function in a promoter configuration for full length E2 products (Thierry and Yaniv, 1987; Bernard *et al.* 1989). Two other E2 motifs are located

Figure 1.8 The noncoding region of HPV 18. MCE = minimal constitutive enhancer characterized in our laboratory (Marshall *et al*. 1989) and used in my studies. I,6 = E6 inducible enhancer. IETPA = TPA inducible enhancer. Arrow indicates transcription start site. GRE = Glucocorticoid responsive element. E2 motif = recognition sequence for E2 proteins. RsaI-RsaI fragment was used to construct the plasmid pT-81, as described in chapter 3


7134
105
RsaI
RsaI
7857/0
TATA
E6 ORF
mRNA
L1 ORF
IE6
MCE
IETPA
GRE
E2 motifs

127 and 492 bp upstream of the E6 initiation codon and function as enhancers. The second inducible enhancer, IE6, is located approximately 500 bp upstream of the E6 cap site and is dependent on the E6 protein for its function (Gius *et al*. 1988). Recent studies have identified a third inducible enhancer (nucleotides 7839-7853) which responds to dexamethasone and progesterone (Chan *et al*. 1989).

The constitutive enhancer, located between 200 and 400 bpupstream of the E6 initiation codon, functions in a variety of cervical carcinoma cell lines, a neuroblastoma cell line, and a small number of fibroblast cell lines (Gius *et al*. 1988; Swift *et al*. 1987; Marshall *et al*. 1989). Deletion analyses of this enhancer have identified at least two functional domains that possess Z-DNA-forming capacity (Swift *et al*. 1987). Additionally, the constitutive enhancer activity is repressed by adenovirus E1A proteins (Swift *et al*. 1987) and activated by herpes simplex virus TIF and ICP0 genes (Gius and Laimins, 1989). A 30 bp fragment called IETPA, located 150 bp upstream of the cap site, contains sequences required for tumor promoter 12-0-tetradecanoylphorbol-13-acetate (TPA) mediated gene activation (Gius and Laimins, 1989). As with HPV 11 and HPV 16, current interest is focussed on identifying the DNA motifs in the enhancer and interacting cellular factors.

The presence of both constitutive and inducible enhancers

in HPV 11, 16 and 18 suggests that the transcription of these viruses is subject to complex regulatory pathways. It is possible that each of these enhancers are used either individually or in a combinatorial fashion for programmed gene expression during the viral life cycle and for malignant conversion of infected cells. Additionally, these enhancers may contain elements that determine the epitheliotropism of these viruses. One or more of the mechanisms described previously (Chapter 1.1.6) or an as yet undefined HPV specific mechanism may be responsible for epithelial cell type specific expression of these enhancers. In this regard, a comparative analysis of the HPVs in different cell types for their expression and DNA-protein interaction may provide enough clues for further experimentation. The use of HPV 11, 16 and 18 for this purpose should also provide a unique opportunity to probe the enhancers for their role in determining oncogenicity of the viruses.

1.3 Statement of Objectives

Transcriptional regulation of eucaryotic genes has been a focal point of research in molecular biology. Transcriptional control regions containing promoter and enhancer elements have been identified for various cellular genes. Cellular transcription factors that initiate and enhance transcription from control regions have been characterized using recombinant DNA technology and

biochemistry. Several models that explain DNA-protein interaction, transcription initiation, activation and repression, have been proposed. Information derived from these studies has been utilized to gain enhanced understanding of tissue-specific and inducible gene expression as well as oncogenesis.

Most of our knowledge on eucaryotic gene regulation is derived from studies on viruses. Although the principal interest in these studies has been to comprehend viral infectivity and pathogeniity, host cell dependence of these viruses has prompted many scientists to use them as model systems to study cellular processes such as gene transcription. Papovaviruses are of special interest since members of this family, although structurally related, manifest distinct tissue tropism. As mentioned previously, SV40 and BK virus regulatory elements are active in a wide range of tissues whereas JC virus regulatory elements are active only in cells of neuronal origin. Similarly, the HPV 11 regulatory elements are active in a wide range of cells as opposed to those of HPV 16 and 18.

The SV40 regulatory elements have been studied most extensively. Several enhancer motifs (enhansons), each displaying cell-specific activity, have been identified. Considering the results of these studies, I propose that the broad tissue tropism of SV40 regulatory elements is due to this complex array of enhansons that interact with distinct

cellular factors. Whether the regulatory elements of other viruses such as BKV, which has a broad range of tissue tropisms, have a similar complex organization as compared to the regulatory elements of viruses such as JCV, which have a narrow tissue tropism, is a subject of great interest. For this purpose, I have undertaken comparative studies of SV40, BK and JC regulatory elements and the results of these studies are presented in Chapter 2. Additionally, these studies were also directed towards developing a suitable system to analyse glial cell type specific JCV regulatory elements for which studies have been hampered by difficulties in obtaining convenient cell types.

Human papillomaviruses are interesting for two reasons. First, these viruses have a strict tropism for epithelial cells and require differentiation of epithelial cells for vegetative growth. Second, they can induce oncogenesis of epithelial cells. Epitheliotropism of these viruses is most probably due to epithelial cell type specific enhancer motifs present in the viral regulatory elements. Previous studies have located epithelial cell type specific enhancer elements but not enhancer motifs. To characterize such sequences and the cellular factors interacting with them, I have studied the enhancer elements of HPV 11, 16 and 18 and the results are presented in Chapter 3. These studies are aimed at determining the role of enhancers in the virus life cycle and oncogenesis.

CHAPTER 2

Activity And Enhancer Binding Factors For SV40, JC And BK Virus Regulatory Elements In Differentiating Embryonal Carcinoma Cells

2.1 Introduction

One of the central challenges in molecular virology is to understand the mechanisms of tissue specific gene expression. Studies with several viruses reaffirmed that viral regulatory elements play an important role in this process (Amati, 1985; Chatis *et al*. 1984; Hilberg *et al*. 1987; Jones *et al*. 1988; Laimins *et al*. 1982; Rockford *et al*. 1987). The differences in the infectivity as well as gene expression of closely related viruses have been attributed to variations in regulatory regions. For example, regulatory elements of the closely related SV40, BK and JC viruses show marked difference in activity: SV40 and BKV regulatory elements are active in a wide variety of cells whereas the JCV enhancer is active only in cells of neuronal origin (Chapter 1).

This study is aimed at identifying the elements within the SV40, JC and BK regulatory regions that contribute to the above mentioned phenomena. It is also aimed at developing a suitable system to analyse JC virus, for which studies have been hampered by difficulties in obtaining glial cells, the only cell type that supports JCV expression. To this end, the embryonal carcinoma cell line, P19, has been used. The P19

cell line was derived from a primary tumor developed on the testis of a mouse after transplantation of a mouse egg cylinder to this site (McBurney & Rogers, 1982). These cells can be induced to differentiate into a mixture of glial cells, neurons and astrocytes upon retinoic acid treatment and into cardiac and skeletal muscle cells upon dimethylsulfoxide (DMSO) treatment. As a first step, the activity of viral enhancers in undifferentiated, retinoic acid differentiated and DMSO differentiated cells was examined. Several deletion mutants of the JCV regulatory region were also tested for their activity. DNaseI footprinting and oligonucleotide competition studies were subsequently employed to correlate *in vivo* activity with *in vitro* DNA-protein interaction and to understand the nature of the DNA-protein interactions.

2.2 Materials and Methods

2.2.1 Materials

Restriction endonucleases were obtained from New England Biolabs, Bethesda Research Laboratories (BRL) or Boehringer Mannheim. T4 DNA ligase, S1 nuclease, calf intestinal phosphatase (CIP) and reverse transcriptase were purchased from BRL, BRL, Boehringer Mannheim, and Life Sciences, respectively. The T4 ligase 5X premix and some of the restriction endonuclease (10X REact) buffers were obtained from BRL. The remaining restriction enzyme buffers and the reverse tran-

scriptase premix were made according to manufacturers' recommendations or Maniatis *et al.* (1982).

The radioactive [^{14}C]chloramphenicol and α[^{32}P]dCTP were purchased from New England Nuclear and Amersham, respectively. Sephadex G-50 columns (Nick-column) were from Pharmacia. Thin layer chromatography (TLC) plates and X-ray films were purchased from Kodak. Acetyl coenzyme A was obtained from Pharmacia, P.L. Biochemicals and Sigma.

Flow laboratories supplied the tissue culture medium (alpha-MEM and Dulbecco's modified media), penicillin-streptomycin, phosphate buffered saline (PBS) and the trypsin-EDTA, while Bockneck Laboratories supplied the fetal calf serum (FCS). Retinoic acid and DMSO were obtained from Sigma and Baker, respectively. The P19 cell line was provided by H. Hamada.

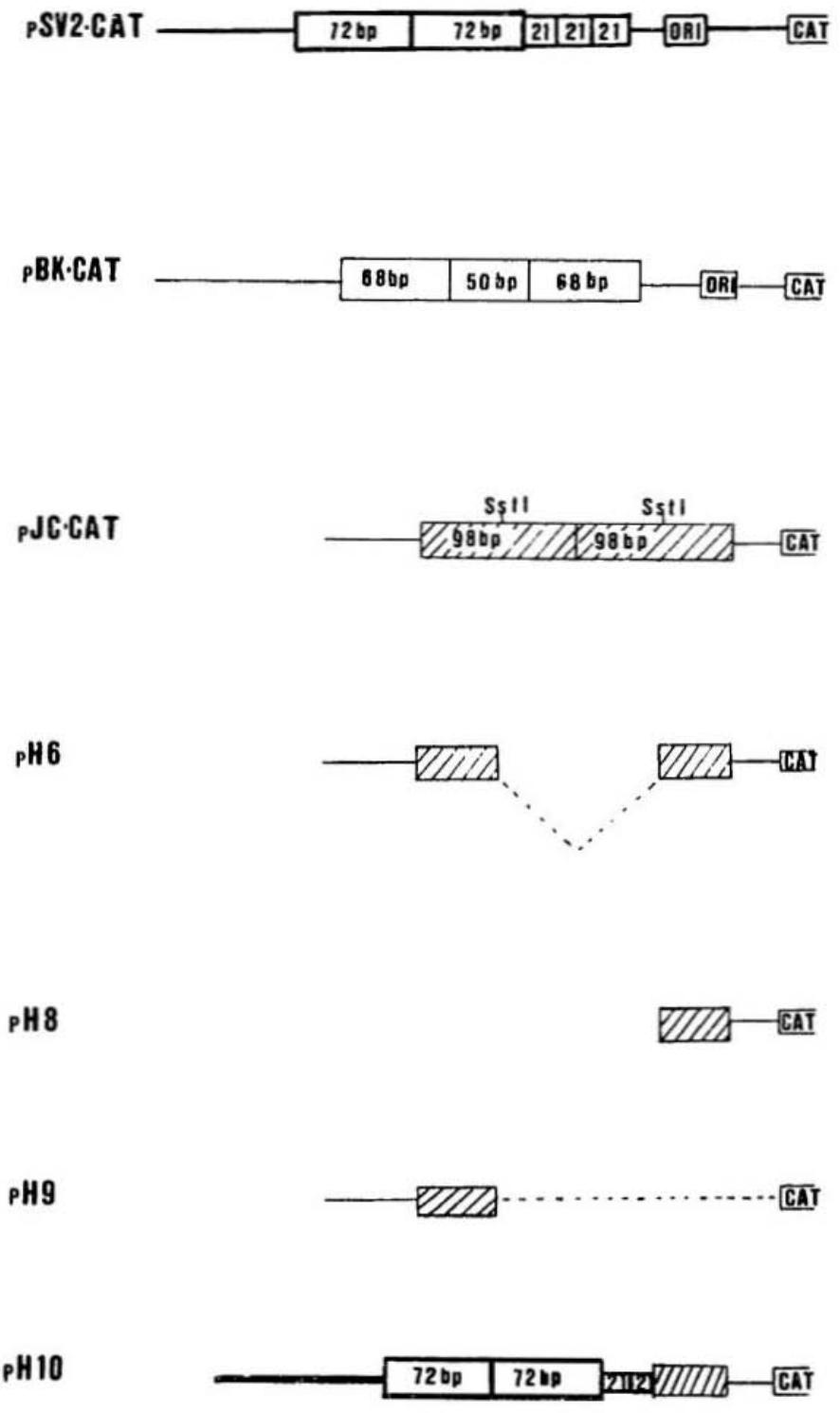
The plasmids pSV2-cat and pSV0-cat were kindly provided by B. Howard. Synthetic oligonucleotides were purchased from The Regional DNA Synthesis Lab, University of Calgary.

2.2.2 Construction of recombinant plasmids

Most of the papovaviral plasmid constructs are presented diagrammatically in Figure 2.1. Plasmids pSV0-cat, pSV2-cat and pBK-cat are as described previously (Gorman *et al.* 1982; Pater and Pater, 1988b).

Plasmid pSV2-cat (Gorman *et al.* 1982) contains the bacterial replication origin and the ampicillin resistance gene from

Figure 2.1 Structure of regulatory elements in recombinant plasmids. Restriction enzymes used in the construction of these plasmids are described in Materials and Methods. Repeat sequences of SV40, JC and BK enhancers are given as boxes. The internal deleted sequences of plasmids pH 6 and pH 9 are indicated by broken lines.


pSV2-CAT
72bp
72bp
21
21
21
ORI
CAT
pBK-CAT
68bp
50bp
68bp
ORI
CAT
pJC-CAT
SstI
SstI
98bp
98bp
CAT
pH6
CAT
pH8
CAT
pH9
CAT
pH10
72bp
72bp
21
12
CAT

pBR322, SV40 regulatory sequences, the bacterial chloramphenicol acetyltransferase gene and the SV40 poly adenylation signal. The pSV0-cat plasmid is derived from pSV2-cat by the removal of the entire enhancer and promoter regions of SV40 (Gorman *et al*. 1982).

Plasmid pBK-cat is derived from pSV0-cat by the insertion of the BK virus enhancer/promoter region (Yang and Wu, 1979) upstream of the CAT coding region in pSV0-cat (Pater and Pater, 1988).

Plasmid pJC-cat contains the HindIII-PvuII fragment (JCV MADI strain, nucleotides 5112-270; Frisque *et al*. 1984) of JC inserted into the HindIII site of plasmid pSV0-cat such that the CAT gene is expressed from the early JC promoter. Plasmid pH6 is a derivative of pJC-cat in which cleavage with restriction endonuclease SstI has deleted sequences from nucleotides 57 to 155 and subsequent religation has generated a plasmid with one 98 bp repeat and the region upstream of the enhancer. Insertion of the HindIII-SstI fragment (nucleotides 5112-55) and SstI-PvuII fragment (nucleotides 157-270) of JCV DNA such that the early promoter of JCV DNA is used to express the pSV0-cat CAT gene yielded plasmids pH8 and pH9, respectively. Plasmid pH10 has JCV sequences identical to that of pH8 and the additional Fok I fragment (nucleotides 150-424, Gorman *et al*. 1982) of pSV2-cat which contains the two 72 bp repeats of the SV40 enhancer.

Bacterial transformation, competent cells, plasmid DNA purification and other basic molecular biology methods were according to Maniatis *et al.* (1982).

2.2.3 Cell culture, induction of differentiation and transfection

The embryonal carcinoma cell line, P19, was maintained in alpha-MEM medium containing 10% foetal calf serum and cells were induced to differentiate as described previously (Rudnicki and McBurney, 1987). Briefly, undifferentiated cells were collected by trypsinization, resuspended in medium containing either 300 nM retinoic acid (RA) or 1% dimethylsulfoxide (DMSO) and seeded into bacterial culture plates. Fresh media with RA or DMSO was added after two days and after four days the cell aggregates were collected, washed in phosphate buffered saline (PBS), trypsinized briefly and plated into tissue culture plates. Eight hours after plating, DNA transfection using the calcium-phosphate precipitation procedure (Gorman *et al.* 1982) was performed. Briefly, five hundred microlitres of a $CaCl_2$/ DNA solution (62 μl of 2M $CaCl_2$ + DNA in 0.1 X TE + water) were added slowly with constant bubbling of air into 500 μl of 2X HBS (0.28 M NaCl; 0.05 M HEPES; 2.8 mM Na_2HPO_4, pH 7.1). In some experiments one ml of $CaCl_2$/DNA solution and one ml of 2X HBS were mixed at a time. The cloudy precipitate was then added evenly to the cells. Four hours after transfection the cells were carefully washed with serum-free medium and subsequently exposed to glycerol (1.5 ml of 15%

glycerol in 1% HBS) for three minutes in the case of undifferentiated cells and one minute for differentiated cells. The cells were then washed with serum-free medium and incubated with alpha MEM + 10% FCS for 48 hours.

2.2.4 CAT assays

Cellular extracts for assaying CAT enzyme were prepared 48 hours after transfection and assayed as described (Gorman *et al.* 1982). Briefly, cells were first washed thrice with 2 ml of PBS and then incubated at room temperature for 5 minutes in 1 ml per plate of Tris-EDTA-NaCl (0.04 M Tris HCl, pH 7.4; 1 mM EDTA; 0.15 M NaCl). The cells were collected by rubber policeman and pelleted by microfuging for two minutes at 4°. The pellet was then frozen at -70° until ready to assay. The frozen cells were thawed on ice and resuspended in 50 μl of 0.25 M Tris HCL (pH 7.8).

The cells were disrupted by three cycles freezing in liquid nitrogen for 5 minutes and thawing at 37° for 5 minutes. The cell debris was pelleted (5 minutes in a microfuge at 4°) and the supernatant was used for CAT assays. The standard assay reaction contained 1 μl of ^{14}[C]chloramphenicol, 4 μl of 4 mM acetyl coenzyme A, 10 μl of 1 M Tris HCl (pH 7.8) and 15 μl of extracts. Reaction was at 37° for 60 minutes and terminated by the addition of ethyl acetate. Ethyl acetate which separates chloramphenicol from the aqueous solution was collected in separate tubes and dried in a speed vacuum

(Savant). The samples were resuspended in 15 µl of ethyl acetate and spotted on silica gel thin layer chromatography plates. Acetylated and nonacetylated forms of chloramphenicol were separated by ascending thin layer chromatography (TLC) in chloroform:methanol (95:5). The TLC plate was treated with omnifluor enhance spray to increase sensitivity and was exposed against XAR-5 Kodak film at -70°. After autoradiography, quantitation of the enzyme reaction products was achieved by liquid scintillation counting of the acetylated and nonacetylated chloramphenicol excised from the chromatography plates. Data is expressed as the percentage of conversion of chloramphenicol to its acetylated form (% CAT activity) where 100% would be 0.25 units of CAT activity (nmole/minute at 37°).

2.2.5 Preparation of radiolabelled DNAs

The DNA fragment used as probe for SV40 was derived from the plasmid pSV2-cat by digesting with restriction enzymes NcoI and PvuII and subsequent electroelution of the 233 bp fragment (nucleotides 107-340; Gorman *et al.* 1982). The BKV enhancer contained in a HaeIII fragment (nucleotides 3186-3479; Yang and Wu, 1979) and the JCV enhancer in a HindIII-PvuII fragment (nucleotides 5112-270; Frisque *et al.* 1984) were first cloned into the XbaI site of pUC-19. Inserted enhancer fragments were then liberated by digestion with restriction enzymes SalI and SmaI and electroeluted from agarose gels. End-labelling was performed by using reverse transcriptase in the presence of

$\alpha[^{32}P]$dCTP, which fills in the NcoI (nucleotide 107) ends of the SV40 fragment and SalI end of the JC (nucleotide 5112) and BK (nucleotide 3479) enhancer fragments. End-labelled DNA was separated from non-incorporated nucleotides through Sephadex G-50 chromatography. When competitor DNA was used for DNaseI footprinting, they were the same fragments without radiolabel. The PvuII fragment of pUC-19 (nucleotides 306-638) was used as a nonspecific competitor in such experiments. When oligonucleotides were used for competition, complementary oligonucleotides were hybridized according to Kadonaga and Tjian (1986) and used directly in the binding reactions.

2.2.6 Nuclear extracts and DNaseI footprinting

Preparation of nuclear extracts and DNaseI footprints were by the procedures of Hennighausen and Lubon (1987). Briefly, cells from approximately 50 plates were used for each nuclear extract preparation. The cells were washed with PBS and collected by scrapping with a rubber policeman and subsequently pelleted by centrifugation for 10 minutes at 1800 rpm at 4°. The pelleted cells were resuspended in 5 volumes of 0.3 M sucrose in buffer A (10 mM HEPES-KOH, pH 7.9; 10 mM KCl, 1.5 mM $MgCl_2$, 0.1 mM EGTA, 0.5 mM dithiothreitol (DTT), 0.5 mM of protease inhibitor phenylmethylsulfonyl fluoride (PMSF), and 2 μg/ml each of antipain and pepstatin A. Cells were lysed by 8-12 strokes with a B pestle in a Dounce glass homogenizer (Knots) and 1-2 strokes in the presence of 0.3-

0.4% Nonidet P-40 (NP-40). The homogenate was then centrifuged at 2700 rpm for 10 minutes and the pelleted nuclei were washed twice in 0.3 M sucrose in buffer A without Nonidet P-40.

Nuclei were resuspended with a glass homogenizer (10 strokes, B pestle) in 2.5 pelleted nuclei volumes of 400 mM NaCl, 10 mM HEPES-KOH at pH 7.9, 1.5 mM $MgCl_2$, 0.1 mM EGTA, 0.5 mM DTT, 5% glycerol and 0.5 mM PMSF. The resuspended nuclei were stirred slowly for 30 minutes at 4° followed by centrifugation for 60 minutes at 37,000 rpm in a 75 Ti rotor. The supernatant was dialyzed for 4 hours at 4° against 50 volumes of 20 mM HEPES-KOH at pH 7.9, 75 mM NaCl, 0.1 mM EDTA, 0.5 mM DTT, 20% glycerol and 0.5 mM PMSF. Material which precipitated during dialysis was removed by centrifugation (20,000 rpm, 20 minutes in 75 Ti rotor, Beckman) and the supernatant was aliquoted and flash frozen in liquid nitrogen and stored at -70°.

The protein concentration was measured by the method of Bradford (1976). For DNaseI footprinting, nuclear extracts (10-15 μl volume depending on the protein concentration) were incubated in a total volume of 50 μl containing binding buffer (50 mM NaCl, 0.1 mM EDTA, 20 mM HEPES-KOH at pH 7.5, 0.5 mM DTT, 10% glycerol) and one microgram of poly (dI.dC). poly (dI.dC) for 15 minutes in ice. Probe was added (~40,000 cpm) and incubation continued for 10 minutes at 20°. $MgCl_2$ and $CaCl_2$ were added to a final concentration of 5 and 1 mM, respectively. This was immediately followed by pancreatic

deoxyribonuclease I (DNaseI). The DNaseI solution (2 μg/μl in 150 mM NaCl and 50% glycerol) was diluted with 25 mM NaCl, 10 mM HEPES-KOH at pH 7.5, and 0.5 mM DTT. All DNaseI footprinting experiments included a control of naked probe DNA (lanes designated as F in the relevant figures) digested at two different concentrations of DNaseI. The concentrations of DNaseI (2050 kunitz units/μg) were 2.5 and 1.5 ng for SV40 and BK and 2.0 and 1.0 ng for JC. Probes with protein received 320 ng in case of SV40 and BK and 240 ng for JCV. After 30 seconds of incubation with DNaseI at 20°, 100 μl of stop buffer (0.375% SDS, 15 mM EDTA, 100 mM NaCl, 100 mM Tris HCL at pH 7.6, 50 μl/ml sonicated salmon sperm DNA, and 100 μg/ml pronase) was added to the reaction mixture. Samples were incubated for 15 minutes at 37°, and 2 minutes at 90° followed by phenol-chloroform-isoamylalcohol extraction and ethanol precipitation. The reaction products were separated in polyacrylamide-8M urea sequencing gels and subjected to autoradiography.

2.3 Results

2.3.1 Expression of SV40, JC and BKV DNA regulatory elements in the EC cell line, P19

To assess the cell type specific expression of SV40, JC and BK regulatory elements and to examine the suitability of P19 cells for *in vitro* studies, I performed transient transfection assays. Recombinant plasmids containing the

bacterial chloramphenicol acetyl transferase (CAT) reporter gene (Gorman *et al*. 1982) and the SV40 (pSV2-cat), JCV (pJC-cat) and BKV (pBK-cat) [Fig. 2.1] regulatory elements were transfected into the undifferentiated, retinoic acid and DMSO differentiated EC cell line, P19. The enhancer-promoterless pSV0-cat was used as a negative control. CAT enzyme assays were used to quantitatively measure the transcriptional activity of the SV40, JC and BKV regulatory elements. The results are shown in Table 2.1. In undifferentiated P19 cells, none of the regulatory elements expressed significantly. However, differentiation by retinoic acid resulted in significant expression of the CAT gene from all three viral constructs. Increased levels of expression in the differentiated cells relative to the undifferentiated cells was not due to increased plasmid uptake since both cell types contained the same amount of plasmid as determined by spot blot hybridization of the Hirt extract from transfected cells. The levels of activity from pSV2-cat and pBK-cat were significantly higher than from pJC-cat. Both pSV2-cat and pBK-cat, but not pJC-cat were also expressed in DMSO differentiated cells (Table 2.1). These results indicate a wider range of tissue specificity for SV40 and BKV regulatory elements and a restricted cell specificity of JCV regulatory elements to cells of neuronal origin.

Table 2.1 The activity of SV40, JC and BK regulatory elements in undifferentiated, retinoic acid and DMSO differentiated embryonal carcinoma cell line, P19

	Cell Phenotype		
	P19-UD	P19-RA	P19-DMSO
pSV0-cat	0.1 (0.07-0.12)	0.1 (0.07-0.12)	0.1
pSV2-cat	0.4	85.8 (84-87)	58.4 (57-59)
pJC-cat	0.1	6.6 (6.0-7.2)	0.3 (0.25-0.35)
pH 6	ND	1.5	ND
pH 8	ND	1.3	ND
pH 9	ND	0.2	ND
pH 10	0.07	1.30	0.30
pBK-cat	1.2 (1.0-1.47)	372.00 (346-398)	69.6 (67-70)

P19-UD, P19-RA and P19-DMSO are undifferentiated and retinoic acid and DMSO differentiated P19 cells, respectively. The cells were transfected with 10μg of the indicated plasmids and 10 μg of pUC-19. Units are defined as the percentage of conversion of [^{14}C]chloramphenicol to acetylated chloramphenicol. Figures greater than 100% represent numbers for which diluted extracts had been used. The average of two experiments are presented. Range of activity is given in the brackets. pSV2-cat and pBK-cat transfected differentiated cell extracts were diluted 1 in 30 times and the activity is quantitated after considering the dilution factor. ND: Not determined.

2.3.2 Deletion analysis of the JCV DNA regulatory region

The regulatory elements of SV40 and BKV have been well characterized (Deyerle et al. 1987; Deyerle and Subramani, 1988; Herr and Clarke, 1986; Ondek *et al*. 1987; Schirm *et al*. 1987; Zenke et al. 1986). However, analogous studies have not been reported for the regulatory region of JCV. To investigate the role of the two repeat sequences of this region, I performed deletion analysis. Deletion mutant pH6 (Fig. 2.1), which contains only one perfect 98 bp repeat and repeat upstream sequences, has only 22% of wild type activity (Table 2.1). Deletion mutant pH8, which contains only a portion of the first repeat (nucleotides 5112-55; Frisque *et al*. 1984) encompassing the putative TATA box and CCAAT box, showed activity equivalent to mutant pH6 (19% wild type activity). Mutant pH9, which contains the second half of the second repeat (nucleotides 157-217) and repeat upstream sequences (nucleotides 218-270), did not have any significant activity (Table 2.1). To study whether the 72 bp repeat sequences of the SV40 enhancer can affect the expression of the JCV early promoter, I cloned these sequences just upstream of the putative JCV TATA and CCAAT boxes (as in plasmid pH8) to generate the plasmid pH10. This modification did not have any effect on JC promoter function (Table 2.1). Therefore, it would appear that the JCV promoter, represented by the TATA and CCAAT boxes, shows enhanced and cell specific activity only in

response to a duplication of these sequences but not in response to the generally very effective SV40 enhancer sequences.

2.3.3 DNaseI footprinting and competition analysis for viral regulatory elements in P19 nuclear extracts

The results presented above suggested that differentiation of EC cells might be correlated with quantitative and/or qualitative changes in cellular transcription factors. Therefore, DNaseI footprinting experiments were undertaken to compare the binding of transcription factors to the regulatory region of these viruses in various P19 cell types. DNaseI footprinting, also termed DNase protection assays, is a useful *in vitro* method to assess events that may be involved in the regulation of gene expression. With this method, direct interactions of proteins which may positively or negatively affect transcription by interacting with the sequences of the enhancers are evaluated. The procedure involves incubation of end-labelled DNA with nuclear proteins and subsequent partial cleavage of DNA with DNaseI. The region of DNA which interacts with proteins is inaccessible to DNaseI while the remaining portion of the DNA molecule is cleaved by DNaseI. The regions of DNA that interact with proteins are then identified by comparing the gel electrophoresis pattern of DNaseI cleaved naked DNA with that of DNase I cleaved DNA which has been incubated with proteins.

The following are the guidelines used throughout this thesis to define terms used to describe DNase I protection / competition assays results. All comparisons made consider only the intensity of bands in hypersensitive or protected/ competed regions relative to the intensity of bands in unaffected regions. For a completely protected region most of the following criteria must be met. First, the relative intensity of bands within the protected region (lanes B in the relevant figures) must be less than 30% of that of corresponding bands in the DNase I digested naked DNA (lane F in the relevant figures). Second, the boundaries of the protected region are defined by the bands above and below the protected regions in lanes B which are 80% or more intense as those in lanes F. Third, competition by a non-specific competitior should be no more than 30% of that by a specific competitior. Protectition should also be reproducible with different preperations of nuclear extracts. A protection is considered weak or partial if the protection is reproducible but fails to meet some of the above criteria. Competition by papovaviral DNA fragments is considered to be strong when 50% or more of the original relative band intensity is restored. Hypersensitive sites refer to band which are a minimum of 30% increased in intensity in the presence of protein extracts.

Table 2.2 provides the summary of the protected regions for all three viral regulatory elements in undifferentiated

Table 2.2 Summary of DNase I protection for SV40, JC and BK regulatory elements in undifferentiated and differentiated P19 cell line

	SV40		JC		BK	
	No. of protected regions	Homology	No.of protected regions	Homology	No. of protected regions	Homology
UD	1	Sp1	-	-	1	Sp1
RA	3	Sp1, P Sphl/Sphll	3	NF-1 E4TF-1	5	Sp1 NF-1
DMSO	3	Sp1, P Sphl/Sphll	-	-	5	Sp1 GTIIc

and differentiated cell extracts. Further discription of individual footprints is provided below.

SV40: Figure 2.2 illustrates footprint analysis of the SV40 regulatory region and Figure 2.3 shows the sequence of the protected regions. P19 undifferentiated cell extracts contain proteins which protect sequences within the 21 bp repeat region in the SV40 promoter (Fig. 2.2, UD, region III). The 21 bp repeat region contains several GC rich boxes, each of which are binding sites for the transcription factor Sp1 (reviewed by Dynan and Tjian, 1985). In addition to protection of the 21 bp repeat region, there was partial DNaseI protection, mainly in the second 72 bp repeat, of sequences interspersed between the P and SphI motifs (Fig. 2.2 dashed line; see Fig. 2.8 for more pronounced protection). The homologous SV40 DNA showed poor competition for this protection while no competition was observed from other DNAs (compare lane B with lanes 1-8, Fig. 2.2, UD). However, the Sp1 motif (region III) binding is competed by SV40 DNA, slightly by BKV DNA and not at all by JC and pUC-19 DNA.

Similar experiments with retinoic acid differentiated cell extracts demonstrated identical protection in both 72 bp enhancer regions as well as in the Sp1 binding sites (Fig. 2.2, RA; compare lanes F and B for regions I, II and III). The protected regions of I and II, located in the second and

Figure 2.2 DNaseI footprinting analysis of SV40 regulatory elements. Twenty-five μg of protein from undifferentiated (UD), retinoic acid differentiated (RA) and DMSO differentiated (DM) P19 nuclear extracts were used. Chemical cleavage of purines (Maxam and Gilbert, 1980) is in lane A/G. DNA treated with DNaseI in the absence and presence of nuclear extracts are in lanes F and B, respectively. Lanes 1 to 8 represent competition experiments where the same amount of protein and probe were incubated, along with cold competitor DNA, before DNaseI digestion. Competitors were: SV40 in lanes 1 and 2, JC in lanes 3 and 4, BK in lanes 5 and 6, and pUC-19 in lanes 7 and 8. Amounts of competitors were 4 ng (10 fold excess) and 20 ng in odd numbered and even numbered lanes, respectively. Protected regions are shown as brackets and numbered I, II and III. Dashed lines in UD designate regions of partial protection. The autoradiograms containing region III protection are from more briefly run gels of the same samples to allow resolution on the top and compactness on the bottom of the figure.


UD
RA
DM
A-G F B 1 2 3 4 5 6 7 8 F
I
II
III

Figure 2.2A DNaseI footprinting of SV40 regulatory elements with an extract from a different set of DMSO differentiated cells. Competitors and description of lanes are as in Figure 2.2.

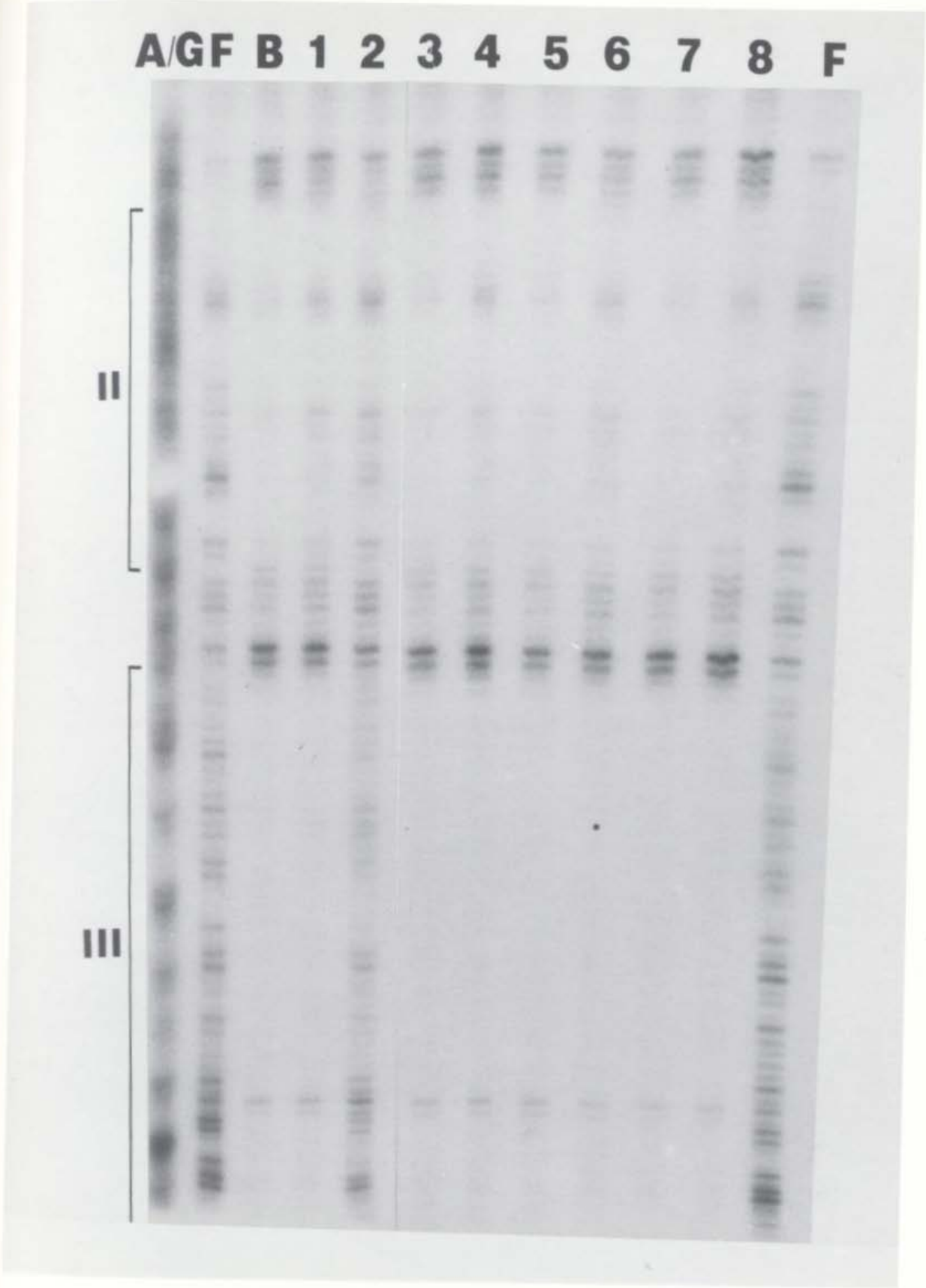
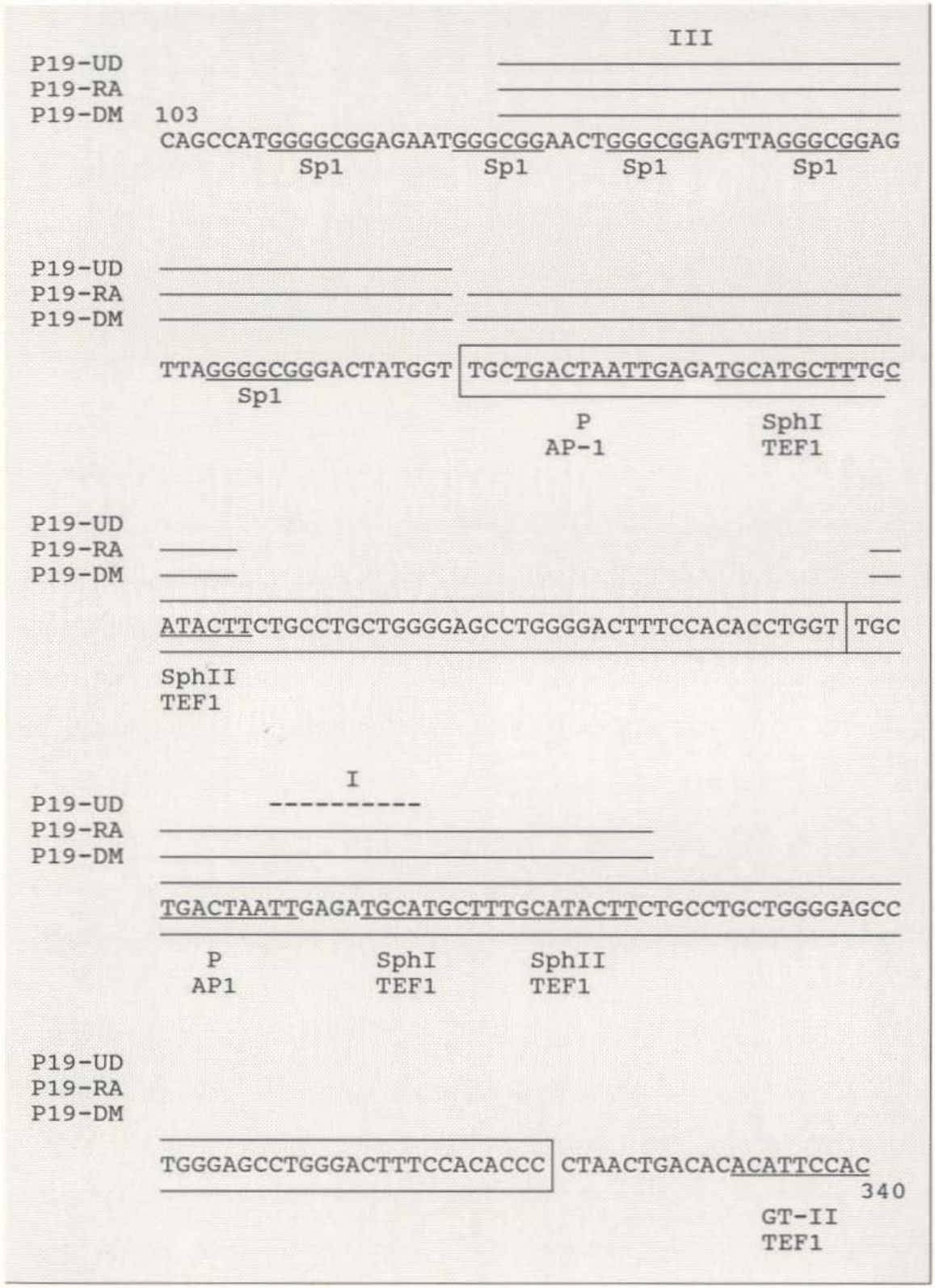

A/GF
B
1
2
3
4
5
6
7
8
F
II
III

Figure 2.3 DNaseI protected nucleotides of SV40 in all three cell types. Protected sequences are overlined and the numbers above the lines are the same as in Figure 2.2. P19-UD, P19-RA and P19-DM are undifferentiated, RA-differentiated and DMSO-differentiated P19 cell extracts, respectively. Consensus sequences for known transcription factors are underlined with the names of the transcription factors indicated below. The two 72 bp repeats are boxed.


III
P19-UD
P19-RA
P19-DM
103
CAGCCATGGGGCGGAGAATGGGCGGAACTGGGCGGAGTTAGGGCGGAG
Sp1
Sp1
Sp1
Sp1
P19-UD
P19-RA
P19-DM
TTAGGGGCGGGACTATGGT
TGCTGACTAATTGAGATGCATGCTTTGC
Sp1
P
AP-1
SphI
TEF1
P19-UD
P19-RA
P19-DM
ATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACCTGGT
TGC
SphII
TEF1
I
P19-UD
P19-RA
P19-DM
TGACTAATTGAGATGCATGCTTTGCATACTTCTGCCTGCTGGGGAGCC
P
AP1
SphI
TEF1
SphII
TEF1
P19-UD
P19-RA
P19-DM
TGGGAGCCTGGGACTTTCCACACCC
CTAACTGACACACATTCCAC
340
GT-II
TEF1

first 72 bp repeats, respectively, are identical and span P and SphI, SphII motifs (Fig. 2.3; regions I and II). It has been reported that HeLa cell factors AP1 and TEF-1 bind P and SphI/SphII motifs, respectively (Lee *et al.* 1987; Davidson *et al.* 1988). SV40 DNA competed for binding in all regions while JC, BK and pUC DNAs did not compete for any of the protected sites (Fig. 2.2).

The footprints in DMSO differentiated cell extracts, presented in Figure 2.2, DM, show the protection in the Sp1 motifs which is identical to that of the other two cell types. The protection in the 72 bp repeat regions is partial and may encompass the P motif and sequences interspersed between P and the SphI motif. However, the protection differed from experiment to experiment when proteins from independent nuclear extract preparations were used. In a separate experiment shown in Figure 2.2A, the P, SphI and SphII motifs are protected and are almost identical to retinoic acid differentiated cell footprints. Some difference in competition was observed. BKV DNA and to a lesser extent JCV DNA, both of which did not compete in retinoic acid differentiated cells, competed for SV40 protection in DMSO differentiated cell extracts (compare Fig. 2.2RA to 2.2A). This difference among the two differentiated cell extracts could be due to quantitative as well as qualitative difference in factors. Whatever the case, the phenomena which produce these *in vitro* results may be involved in the induction of expression in differentiated cells (Table 2.1).

JC: DNaseI footprints of JCV DNA are shown in Figure 2.4 and summarized in Figure 2.5. In undifferentiated cell extracts (P19 UD), there was partial protection principally for the TATA box. I next examined the DNaseI footprint pattern in retinoic acid differentiated cell extracts (Fig. 2.4, P19 RA). At least three completely protected regions and two partially protected regions, including the TATA box, were detected. All three completely protected regions contained sequences homologous to motifs for CCAAT binding protein, CTF/NF-1 (Jones *et al.* 1985; Jones *et al.* 1987). Regions II and III are located within the two 98 bp repeats. The third protected region, region I, also shares homology to the E4TF1 binding site (Jones *et al.* 1988). The DNaseI footprints showing cell type specific protein - JCV DNA complexes correlates with JCV regulatory region activity which is observed only in retinoic acid differentiated cells.

Some of the factors that bind to JCV regulatory elements can also bind to BKV regulatory elements. This was evident from competition experiments in which JCV and BKV DNA competed for binding but SV40 and pUC did not (Fig. 2.4). This may suggest a closer relationship between JCV and BKV transcriptional control mechanisms than between those of JCV and SV40. Only the TATA box region was protected in DMSO-differentiated cell extracts (Fig. 2.4, P19 DM), which correlates with lack of activity of JCV regulatory elements in these cells. Fairly uniform protection was observed for the TATA

Figure 2.4 DNaseI footprinting analysis of JC regulatory elements. End-labelled JCV DNA was subjected to DNaseI footprinting with 25 μg of protein of nucelar extracts from undifferentiated (P19 UD), retinoic acid differentiated (P19 RA) and DMSO differentiated (P19 DM) cells. Samples applied to the lanes are as follows. Chemical cleavage of purines is in lane A+G. DNA treated with DNaseI in the absence and presence of nuclear extract is in lanes F and B, respectively. Lanes 1 to 8 represent competition experiments where the same amount of protein and probe were incubated with cold competitior DNA before DNaseI digestion. Competitors were: SV40 in lanes 1 and 2, JC in lanes 3 and 4, BK in lanes 5 and 6 and pUC-19 in lanes 7 and 8. In odd and even numbered lanes 4 ng (10-fold excess) and 20 ng, respectively, of competitor DNAs were used. Three completely protected regions (I, II, III) and one partially protected region (bracked with broken line), were detected in retinoic acid differentiated cell nuclear extracts.

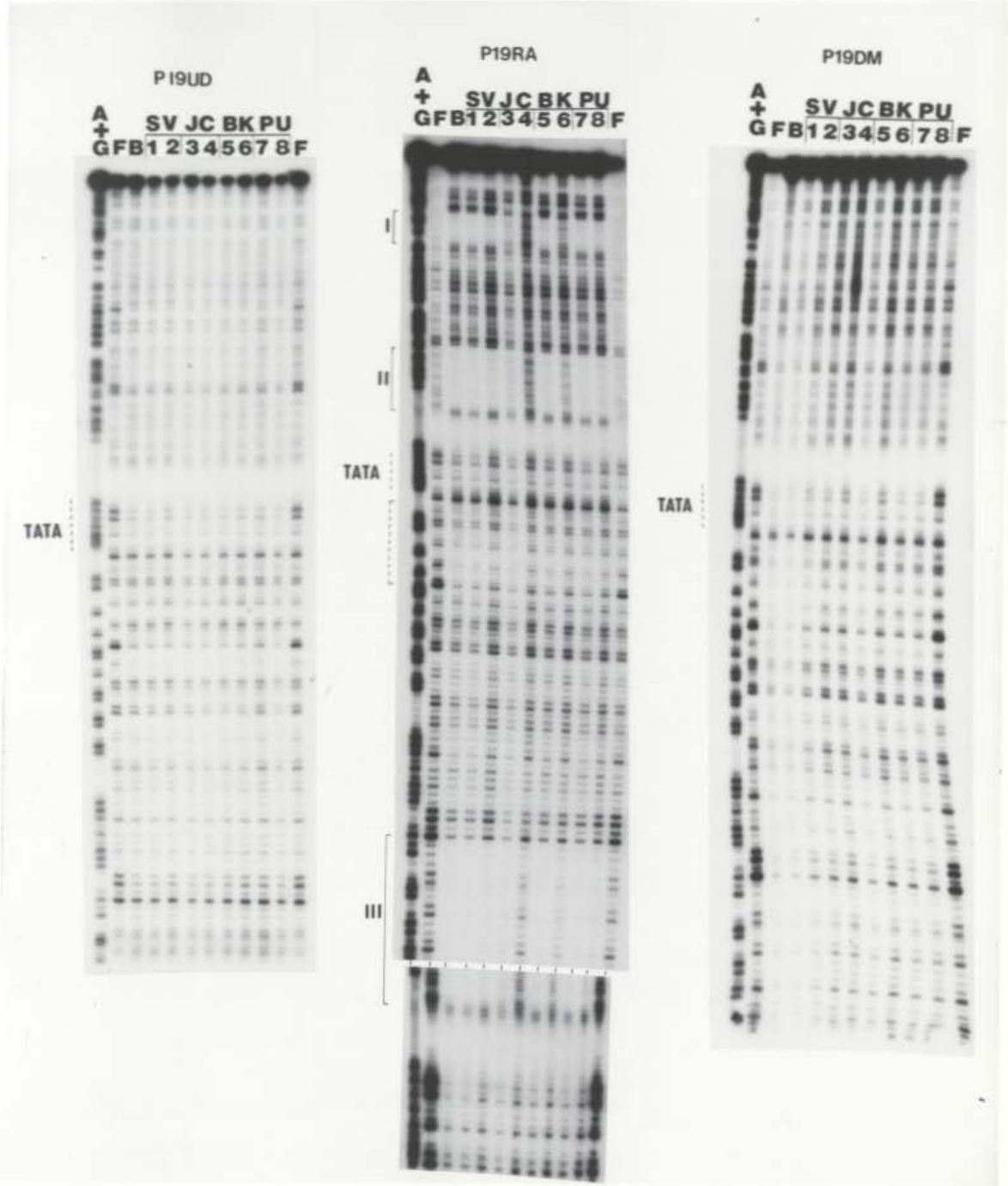
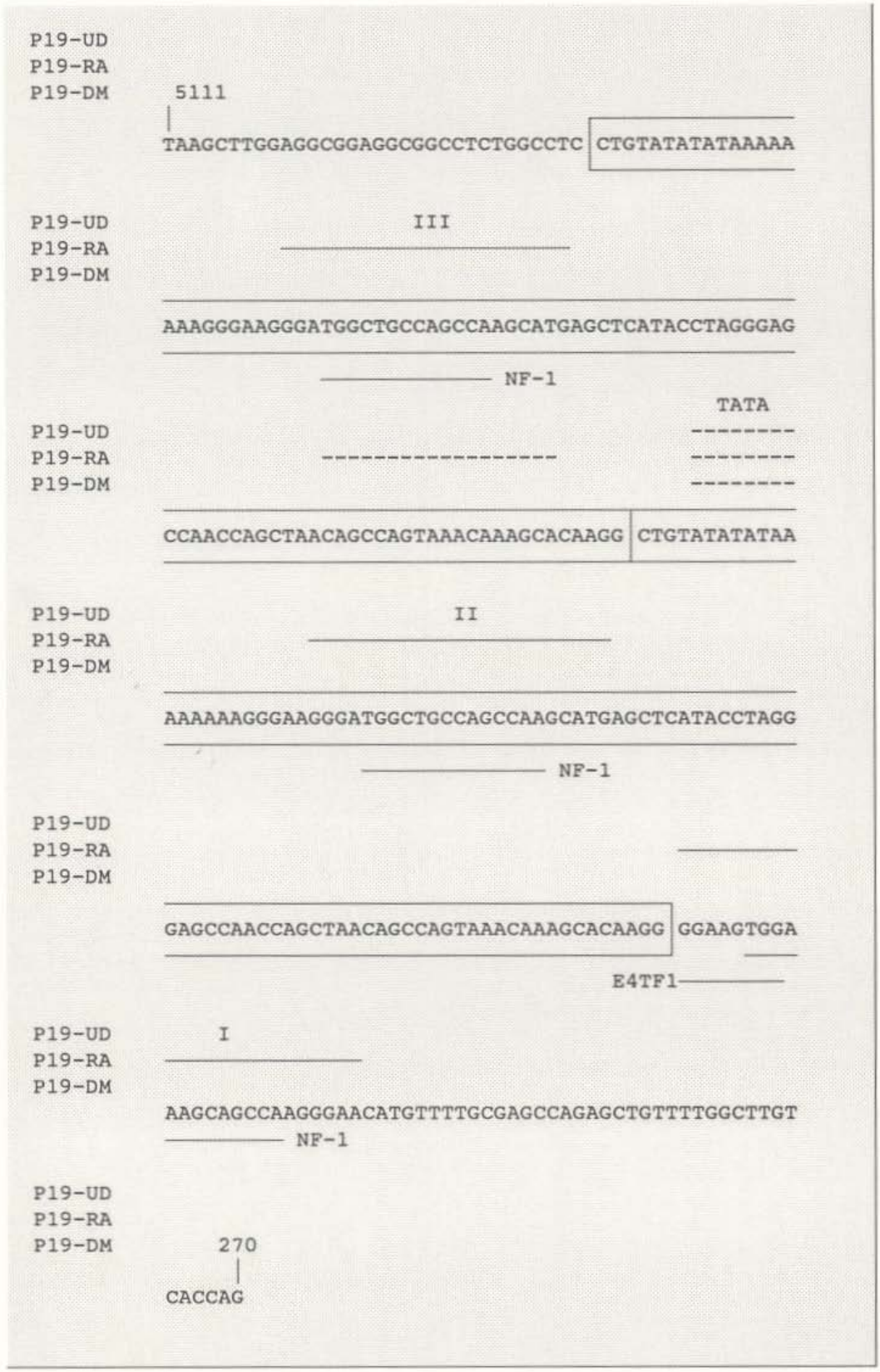

P19UD
A + G
F
B
SV 1
JC 2
BK 3
PU 4
5
6
7
8
F
TATA
P19RA
I
II
III
TATA
P19DM
TATA

Figure 2.5 DNaseI protected nucleotides of JC in all three cell types. Protected sequences are overlined and the numbers above the lines are the same as in Figure 2.4. P19-UD, P19-RA and P19-DM represents undifferentiated, retinoic acid differentiated and DMSO differentiated cell extracts, respectively. Consensus sequences for known transcription factors are underlined with the names of the transcription factors indicated below. Two 98 bp repeat regions are boxed.


P19-UD
P19-RA
P19-DM
5111
TAAGCTTGGAGGCGGAGGCGGCCTCTGGCCTC CTGTATATATAAAAA
P19-UD
P19-RA
P19-DM
III
AAAGGGAAGGGATGGCTGCCAGCCAAGCATGAGCTCATACCTAGGGAG
NF-1
TATA
P19-UD
P19-RA
P19-DM
CCAACCAGCTAACAGCCAGTAAACAAAGCACAAGG CTGTATATATAA
P19-UD
P19-RA
P19-DM
II
AAAAAAGGGAAGGGATGGCTGCCAGCCAAGCATGAGCTCATACCTAGG
NF-1
P19-UD
P19-RA
P19-DM
GAGCCAACCAGCTAACAGCCAGTAAACAAAGCACAAGG GGAAGTGGA
E4TF1
P19-UD
P19-RA
P19-DM
I
AAGCAGCCAAGGGAACATGTTTTGCGAGCCAGAGCTGTTTTGGCTTGT
NF-1
P19-UD
P19-RA
P19-DM
270
CACCAG

box in all three cell types indicating that all the three cell types contain the same level of general transcription factors.

BK: Figure 2.6A and B represent the DNaseI footprints for BKV DNA. Results are summarized in Figure 2.7. In undifferentiated P19 cell extracts, only one protected region was detected. This region appears to bind to an abundant protein since poor competition was observed and only for higher concentrations of BKV and SV40 DNA (Fig. 2.6A, UD). This protected region spans a GC-rich sequence and has homology to the transcription factor Sp1 binding motif and to a transcription repressor protein binding site (Kageyama and Pastan, 1989).

The footprint pattern was completely different in retinoic acid differentiated cell extracts (Fig. 2.6A, RA), in which BKV regulatory elements function with greater efficiency (Table 2.1). A total of five protected regions were detected and four of them designated I, II, IV and V in Figure 2.6A (RA), have homology to the binding site of transcription factor CTF/NF-1. The other protected region (III) is the GC rich region which was also protected in undifferentiated cells. Interestingly, JCV regulatory elements competed for the four protected regions (I, II, IV and V), which contain a sequence homologous to the CTF/NF-1 binding site while SV40 regulatory elements competed, although poorly, only for region III which has homology to the Sp1 binding site.

Figure 2.6A DNaseI footprint analysis of BK regulatory elements. The end-labelled BKV DNA was as described in Methods. Twenty-five μg of protein from undifferentiated (UD), retinoic acid differentiated (RA) and DMSO differentiated (DM) P19 nuclear extracts were used. Chemical cleavage of purines was used for the sample in lane A/G. DNA treated with DNaseI in the absence and presence of nuclear extracts are in lanes F and B, respectively. Lanes 1 to 8 represent competition experiments where the same amount of protein and probe were incubated, along with cold competitor DNA, before DNaseI digestion. Competitors were: SV40 in lanes 1 and 2, JC in lanes 3 and 4, BK in lanes 5 and 6, and pUC-19 in lanes 7 and 8. Amounts of competitors were 4 ng (10-fold excess) and 20 ng in odd numbered and even numbered lanes, respectively. Protected regions are shown as brackets and numbered I-V.

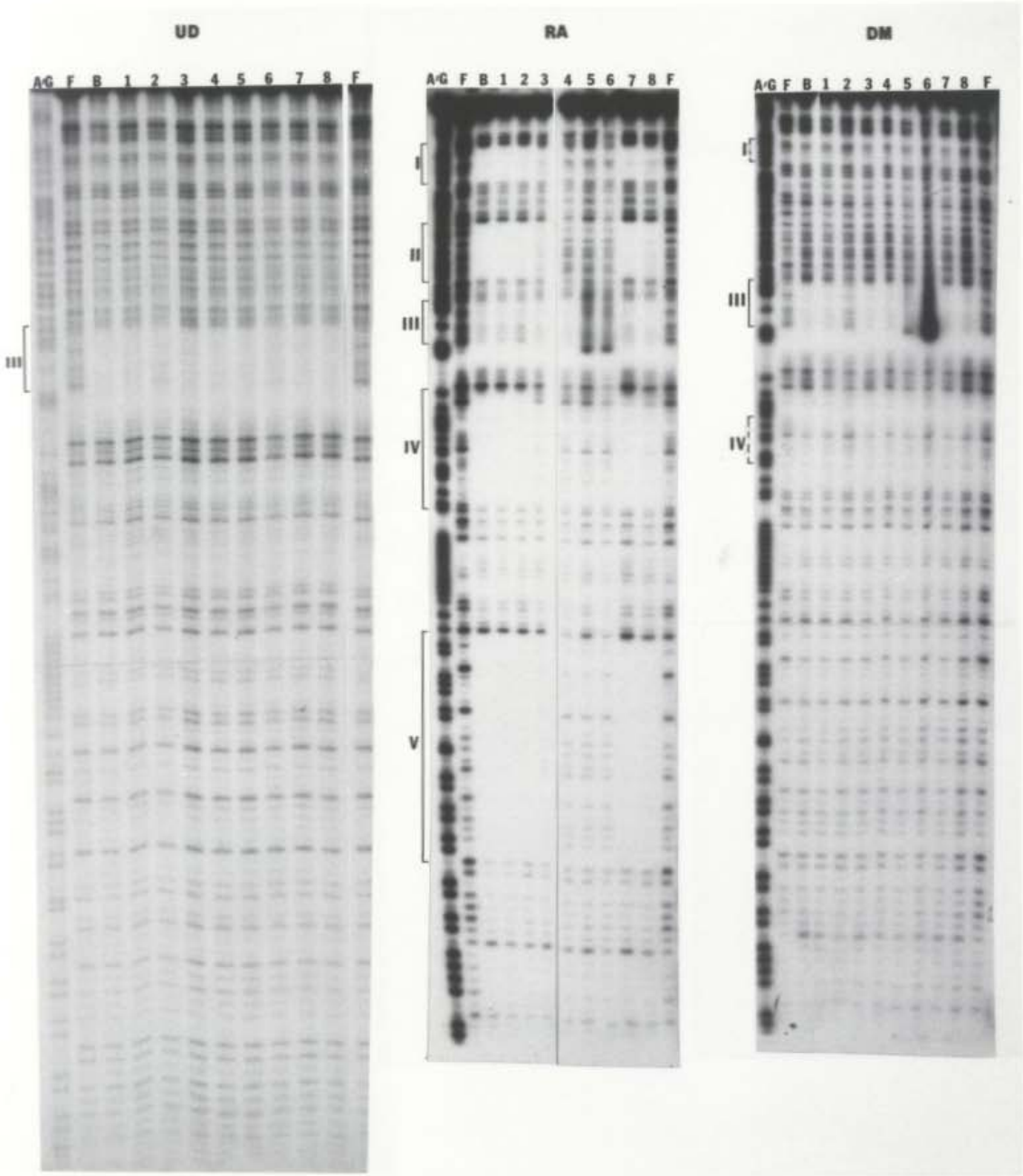

UD
A-G F B 1 2 3 4 5 6 7 8 F
III
RA
A-G F B 1 2 3 4 5 6 7 8 F
I
II
III
IV
V
DM
A-G F B 1 2 3 4 5 6 7 8 F
I
III
IV

Figure 2.6B DNaseI footprinting of BKV regulatory elements with 50 μgs of undifferentiated (UD) and DMSO (DM) differentiated cell nuclear extracts. Competitors and description of lanes are as in Figure 2.6A.

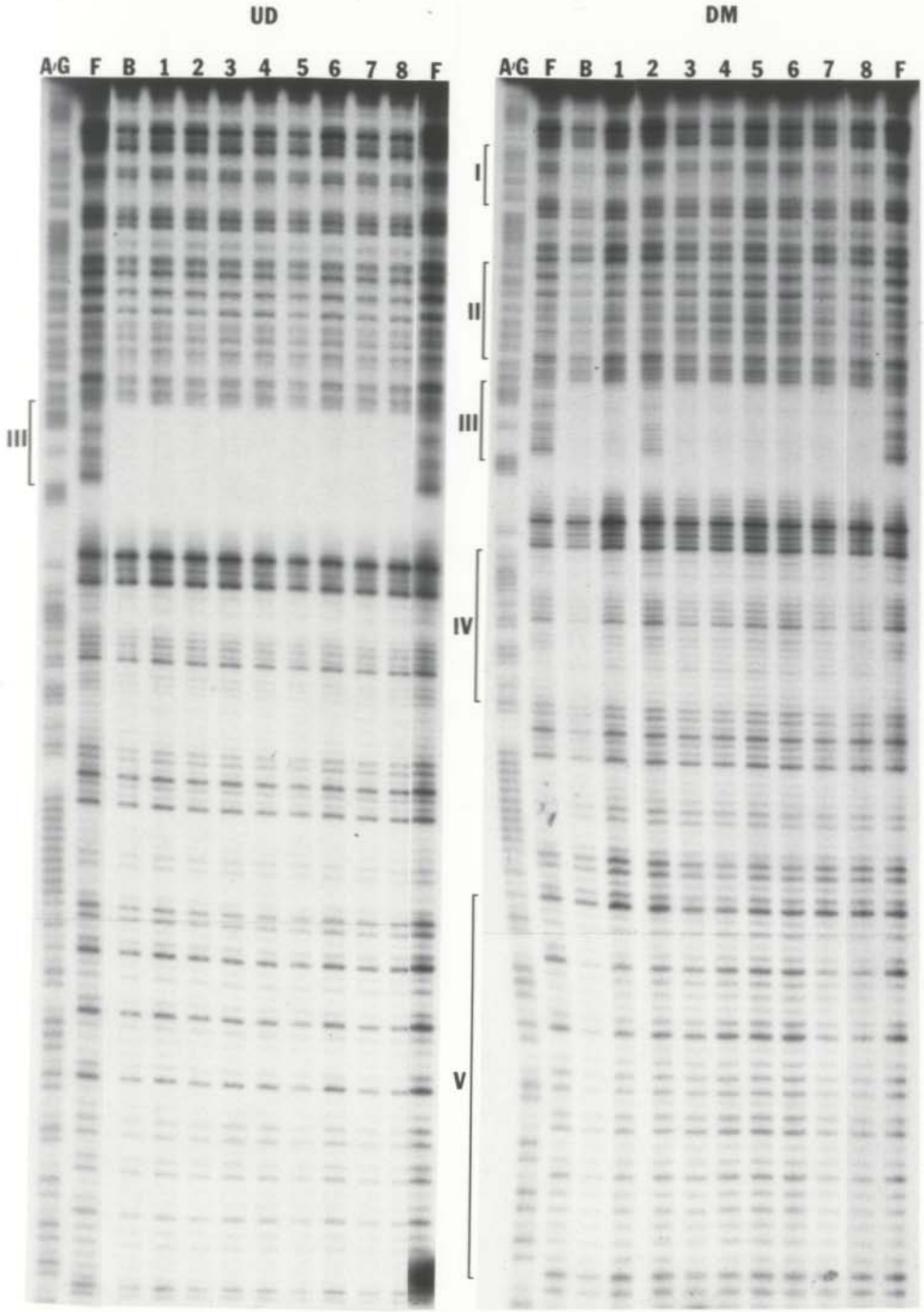
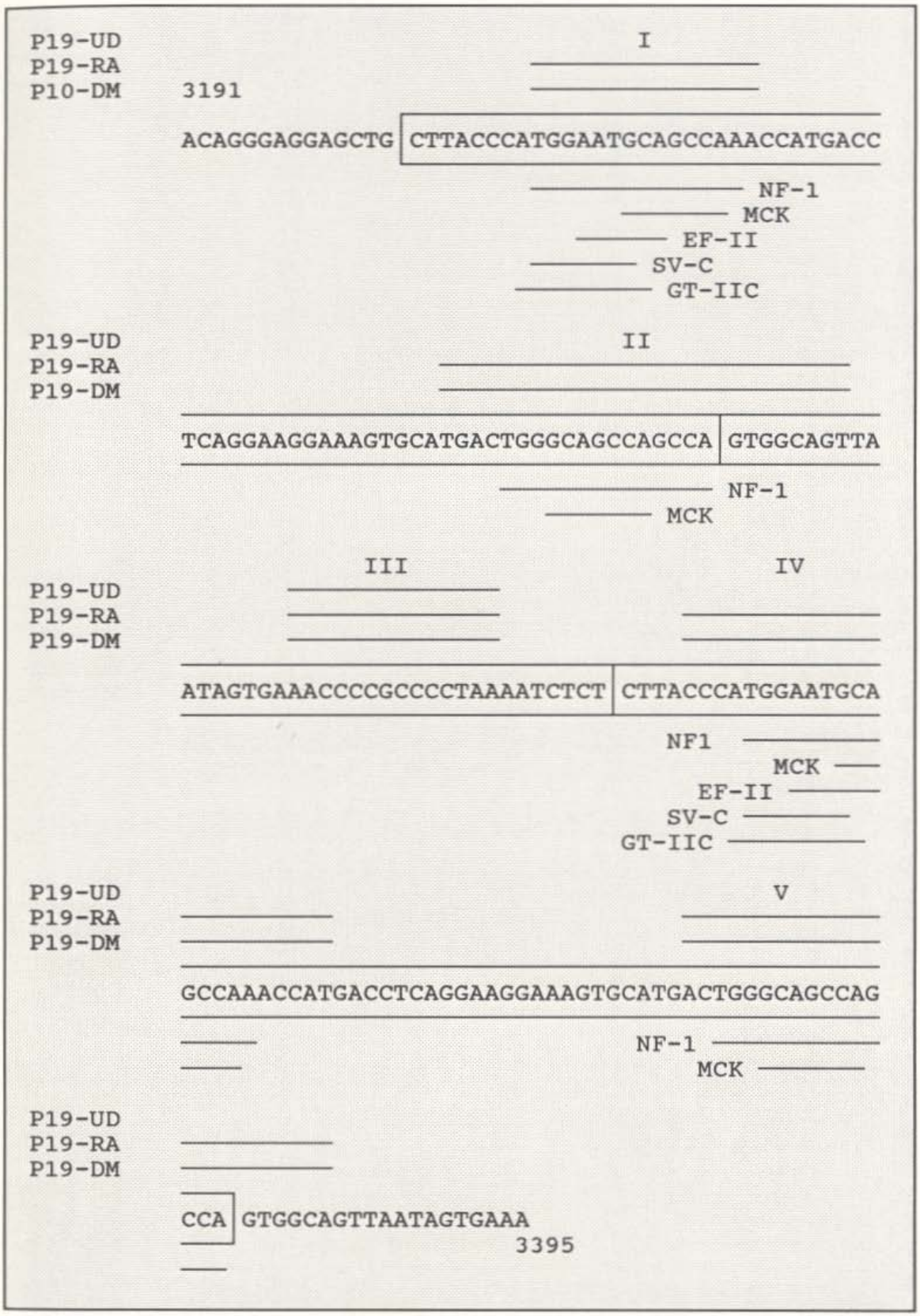

UD
DM
A/G F B 1 2 3 4 5 6 7 8 F
A/G F B 1 2 3 4 5 6 7 8 F
I
II
III
III
IV
V

Figure 2.7 DNaseI protected nucleotide sequences of BK regulatory elements in all three cell types. Protected sequences are overlined and the numerals above the lines are the same as in Figure 2.6A and B. P19-UD, P19-RA and P19-DM represent undifferentiated, retinoic acid differentiated and DMSO differentiated P19 cells, respectively. Consensus sequences in the corresponding protected regions for known transcription factors are underlined with the names of the transcription factors indicated. Repeat regions of the enhancer are boxed.


P19-UD
P19-RA
P10-DM
I
3191
ACAGGGAGGAGCTG
CTTACCCATGGAATGCAGCCAAACCATGACC
NF-1
MCK
EF-II
SV-C
GT-IIC
P19-UD
P19-RA
P19-DM
II
TCAGGAAGGAAAGTGCATGACTGGGCAGCCAGCCA
GTGGCAGTTA
NF-1
MCK
III
IV
P19-UD
P19-RA
P19-DM
ATAGTGAAACCCCGCCCCTAAAATCTCT
CTTACCCATGGAATGCA
NF1
MCK
EF-II
SV-C
GT-IIC
P19-UD
P19-RA
P19-DM
V
GCCAAACCATGACCTCAGGAAGGAAAGTGCATGACTGGGCAGCCAG
NF-1
MCK
P19-UD
P19-RA
P19-DM
CCA
GTGGCAGTTAATAGTGAAA
3395

For the DMSO-differentiated cells, the cell system in which the BKV enhancer was active but to a lesser extent than in retinoic acid differentiated cells, sequences in regions I, III and IV were protected. The protection in regions I and IV was very weak compared to the protection in retinoic acid differentiated cells. These differences could be due to quantitative and/or qualitative difference in the transcription factors in each cell type. To test these possibilities, I performed footprints with double the amount of protein from DMSO-treated cells (Fig. 2.6B). Significantly, protection of the regions I and IV was more complete. In addition, there were two more protectedregions corresponding to regions II and V of retinoic acid differentiated cells. This suggests that these two regions are low affinity binding sites for the protein present in DMSO differentiated cells, which also demonstrates the unique nature of the factors binding to the NF-1 site in the two differentiated cell types.

An interesting shared feature of the regions I, II, IV and V was revealed by sequence comparison of the protected regions of the BKV enhancer with known enhancer motifs. This has revealed that the sequence 5′ GCAGCCA 3′ of regions I, II, IV and V, all of which were protected in DMSO-treated cells are also present in the κ motif of the muscle creatine kinase gene enhancer (Lassar *et al.* 1989; Muller and Wold 1989).

Competition was observed with all three viral regulatory elements in the order BKV > SV40 > JC DNA. Another region

protected in DMSO differentiated cell extracts is region III. Note that, for the lower protein concentration of Figure 2.6A, the competition with BKV sequences resulted in a dramatic increase in the cleavage of this region by DNaseI to the degree where DNaseI digestion exceeds that for DNA in the absence of cell extracts. An additional protected region, protected significantly in the RA panel (Fig. 2.6A), weakly for DM and not for UD, was observed below region V. However, competition for protection was not apparent for any of the four DNA competitors.

2.3.4 Oligonucleotide competition studies

Once an overall evaluation of DNA-protein interaction has been made by DNaseI footprinting, more refined analyses are possible using oligonucleotide competition. This method provides preliminary data on the nature of protein interactions with repeated sequence motifs of the enhancer and the influence of neighbouring sequences of a motif on the DNA-protein interaction. The technique is very similar to footprinting and uses chemically synthesized oligonucleotides which are the same as or mutant forms of DNA sequences of interest. Competition by oligonucleotides for a protected region can be complete or weak. The pattern of competition is influenced by the abundance of specific factors and quantity of oligonucleotide used for competition. Additionally, in some cases, DNA binding proteins do not bind to fragments that

include only the footprinted region: sometimes relatively non-specific contacts on either side of the footprint are also required. Footprints requireing additional sequences on the boundaries of the footprinted region are competed weakly by the oligonucleotides.

When oligonucleotides corresponding to the SV40 P motif and the SphI and SphII motifs (Fig. 2.8; lanes 1 and 2, respectively) were used for these experiments, each independently competed for protection of regions I and II of SV40 DNA. Competition of the entire protected region, which is composed of P, SphI and SphII motifs, by both oligonucleotides suggests a cooperative interaction between factors binding to these regions. Although interaction of the same factor(s) to both P and SphI/SphII motifs can not be ruled out, available literature does not support such a possibility (Jones *et al.* 1988; Davidson *et al.* 1988). A control oligonucleotide representing a sequence present in the JC enhancer (lane J) produced no competition, demonstrating the specificity of competition.

All the protected regions in the JCV enhancer contain the GCCAA NF-1 binding motif. To test whether the same factor(s) binds to these different regions, I performed competition experiments with oligonucleotides corresponding to regions I and II (Fig. 2.5). Region I oligonucleotide (Fig. 2.9 lane 4) competed effectively for all three regions while competition by region II oligonucleotide was poor for all three regions

Figure 2.8 Oligonucleotide competition experiments for SV40. Competition was with 100 ng (3200 to 10,000-fold excess) of the indicated oligonucleotides. DNaseI digestion in the absence and presence of protein and without any competitor is in lanes F and B, respectively. Oligonucleotides corresponding to the SV40 P motif (SV40, lane 1, 5′ CTGACTAA 3′; Fig. 2.3) and SphI, SphII motifs (SV40, lane 2; 5′ TGCATGCTTTGCATACT 3′) and control oligonucleotide (lane J; 5′ AAGGGAAGGGATGG 3′) were incubated with nuclear extract prior to addition of probe and DNaseI digestion. Competition experiments were with nuclear extracts from undifferentiated cells (UD) and differentiated cells (RA and DM).

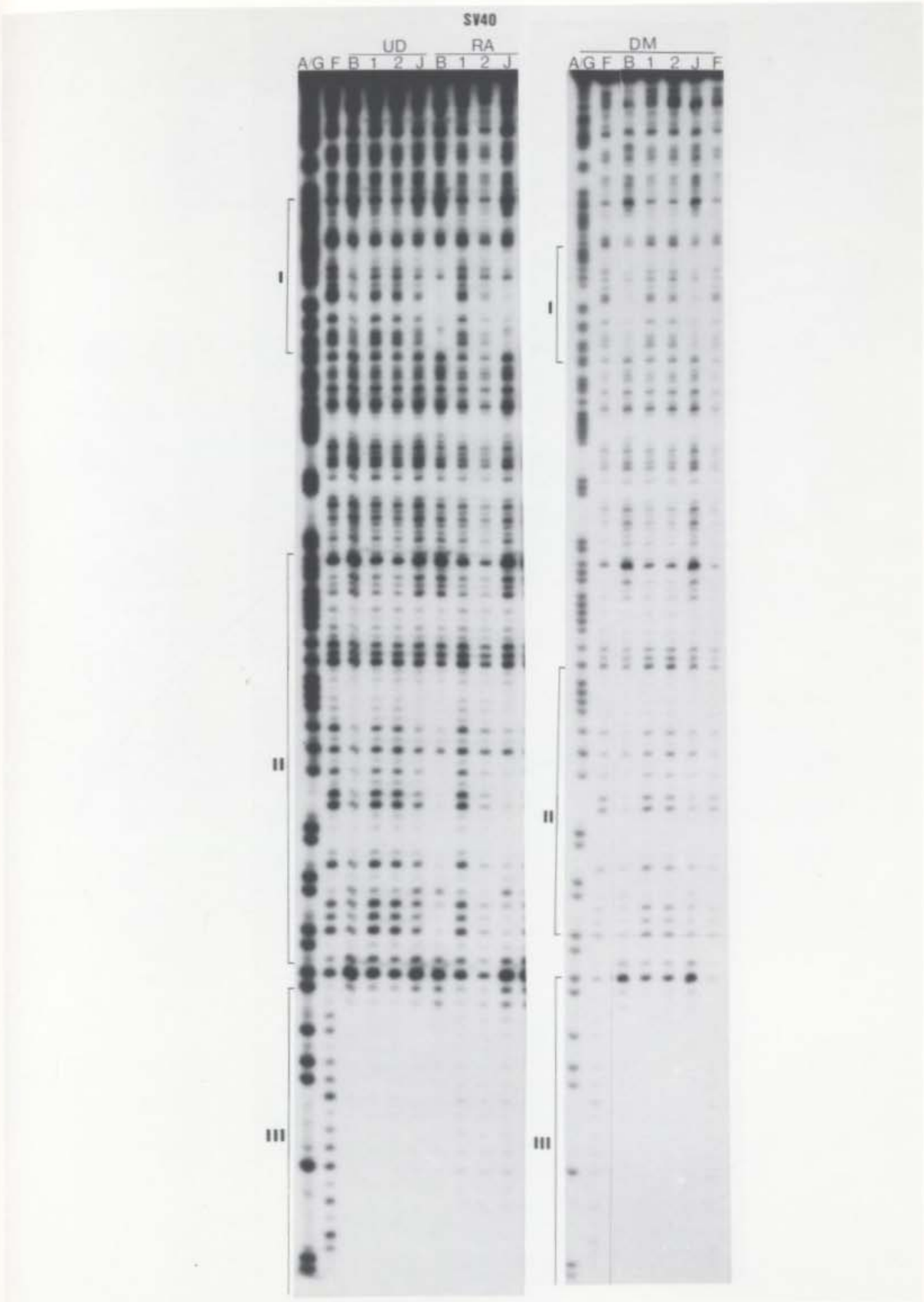
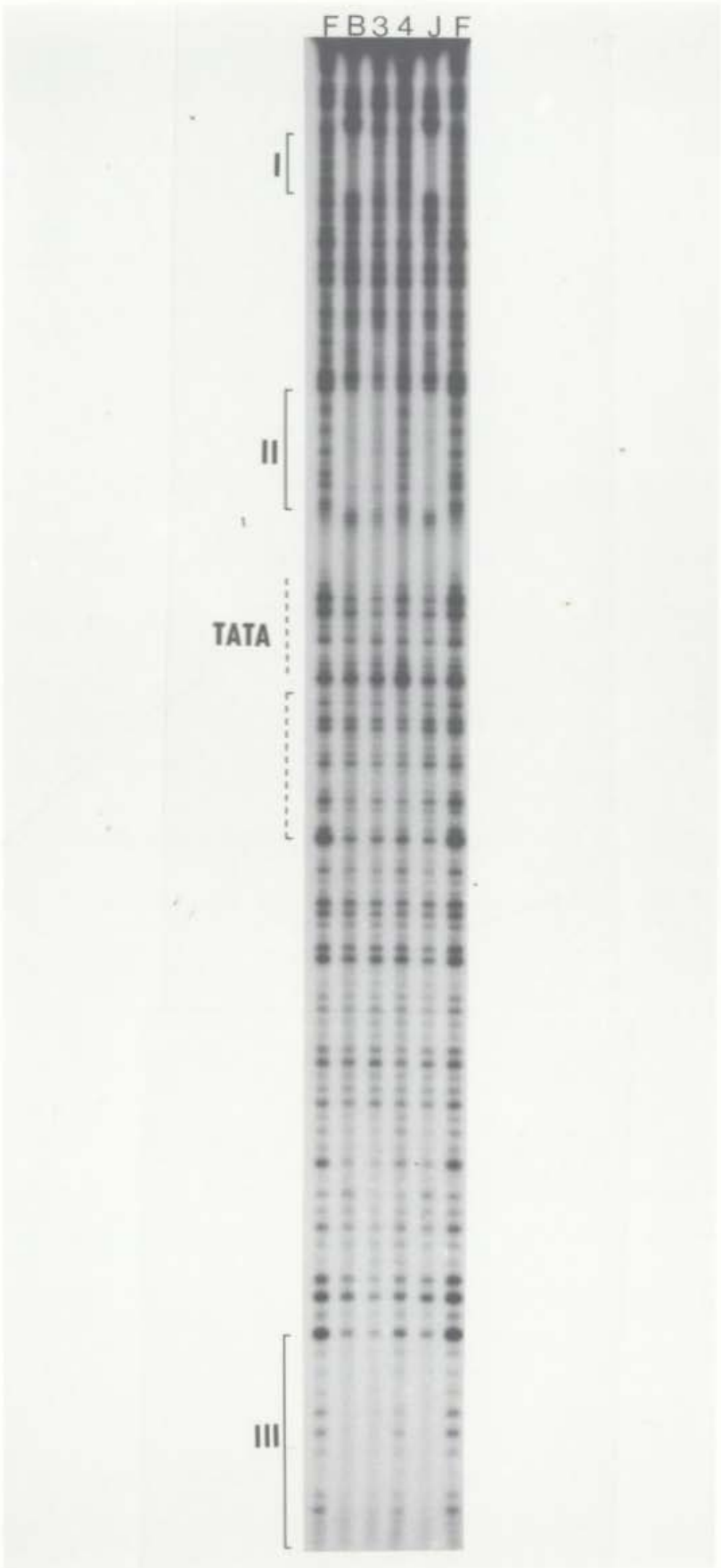

SV40
UD
RA
DM
A G F B 1 2 J B 1 2 J
A G F B 1 2 J F
I
II
III
I
II
III

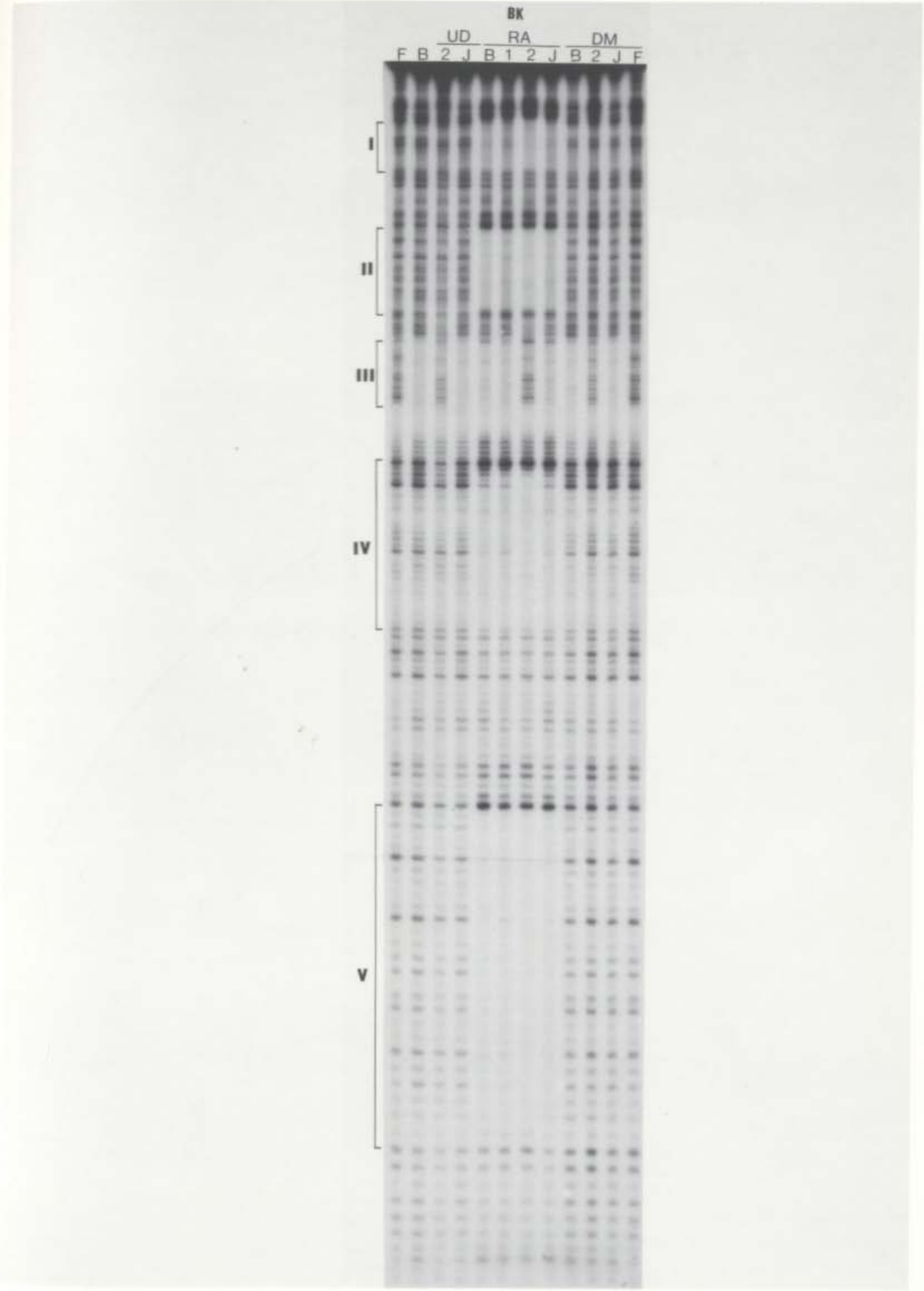
Figure 2.9 Oligonucleotide competition experiments for JC. Competition was done using 100 ng of the oligonucleotides corresponding to region II (lane 3; 5'TGGCTGCCAGCCAA 3') and region I (lane 4; 5'AAGGGGAAGTGGAAAGCAGCCAA 3') of JC enhancer (Fig. 2.5). DNaseI digestion in the absence and presence of cell extracts without any competitor are in lanes F and B, respectively. In lane J, a control oligonucleotide of sequence 5' AAGGGAAGGGATGG 3' was used. Competition was performed in retinoic acid differentiated cell extracts. Competition with a region I oligonucleotide was strong compared to competition with region II oligonucleotide.


F B 3 4 J F
I
II
TATA
III

(Fig 2.9, compare lanes B with 3 and 4). This suggests interaction of the same factor(s) to all three regions, possibly with different affinities. Apparent competition for all regions by these oligonucleotides could also be due to interference with cooperative interactions. This is in agreement with the results for JC in other neurogenic cells as reported by others (Amemiya *et al.* 1989; Tamura *et al.* 1988). As with the SV40 studies, the control oligonucleotide did not affect any protected regions (Fig. 2.9, JC; compare lane B with J).

For the BKV enhancer, oligonucleotides corresponding to region II and region III were tested. The oligonucleotide for region II competed partially for regions I, II, IV and V (Fig. 2.10, BK; RA; lane 1). It is possible that all these four protected regions bind to the same factor(s) which in some cell types could most likely be members of the NF-1 family. Region III oligonucleotide competed only for the homologous region in all three cell types, suggesting that the factor(s) binding to this region is unique and that binding is not cooperative with factors binding with neighbouring regions. It is interesting to note that the opposite was observed: competition with this oligonucleotide increased the degree of protection in the region IV proximal side in retinoic acid cell extracts (Fig 2.10; BK; RA; lane 2), suggesting that the interaction of a protein(s) at region III has a negative influence on protein binding to region IV, possibly through steric hinderance.

Figure 2.10 Oligonucleotide competition experiments for BK. Compeition was done using 100 ng of the oligonucleotide corresponding to region II (lane 1; 5'TGGGCAGCCAGCCA 3') and region III (lane 2; 5'AACCCCGCCCCTA 3') of the BKV enhancer (Fig. 2.7). DNaseI digestion in the absence and presence of cell extracts without any competitors are in lanes F and B, respectively. In lane J, a control oligonucleotide of sequence 5' AAGGGAAGGGATGG 3' was used. Competition was performed in undifferentiated (UD), retinoic acid differentiated (RA) and DMSO differentiated (DM) cell extracts.


BK
UD
RA
DM
F B 2 J B 1 2 J B 2 J F
I
II
III
IV
V

2.4 Discussion

There has been a controversy regarding activity of the SV40 enhancer in undifferentiated EC cells. Earlier studies found no activity of this enhancer in F9, P19 and PCC4 cells (Gorman *et al.* 1985). Subsequent studies have shown that the enhancer is active but at a lower level than in differentiated cells and that the differentiation by retinoic acid results in ~30 fold increase in activity (Nomiyama *et al.* 1987; de Groot *et al.*, 1990). However, the data presented in Table 2.1 suggests that the SV40 enhancer is inactive in undifferentiated cells and that the differentiation by retinoic acid leads to ~200 fold activation. The presence of an undifferentiated cell type specific repressor(s) and the lower abundance of the necessary transcription factors have both been proposed as explanations for this lower level of activity (Gorman *et al.* 1985; Sleigh 1987; Sleigh *et al.* 1987; Fromental *et al.* 1988).

Although the correlation of enhancer activity *in vivo* with *in vitro* DNA-protein interactions could be inferential, there is significant supporting evidence that such a correlation is relevant, particularly for the SV40 enhancer. Unlike cellular genes such as the tyrosine aminotransferase gene and the muscle creatine kinase gene for which the *in vivo* activity does not correlate with the results of *in vitro* DNA-protein interactions, most of the results published to date do show a definite correlation of SV40 enhancer *in vivo* activity with *in vitro*

DNA protein interactions (Becker *et al.* 1987; Muller & Wold, 1989; Zenke *et al.* 1986; Xiao *et al.* 1987; Rosales *et al.* 1987; Fromental *et al.* 1988). Considering these observations, I interpret from my results (Fig. 2.2) that the increase in SV40 enhancer activity after the differentiation of P19 cells is due to the selective increase in the level of and/or activation of transcription factors. Thus, the complete protection of SphI and SphII motifs in retinoic acid differentiated cells may be attributed to a selective increase in factors binding to this motif. A common HeLa factor, TEF-1, can bind to both of these motifs and to the unrelated GT-II motif (Davidson *et al.* 1988). It has been observed that mutation of the GT-II, SphI and SphII motifs drastically reduce SV40 enhancer activity in F9 differentiated cells (Nomiyama *et al.* 1987). Moreover, the tetramer of two closely opposed GT-II motifs and the SphI/SphII motif, show activity in retinoic acid differentiated F9 cells (Fromental *et al.* 1988). If a similar mechanism is operative in retinoic acid differentiated P19 cells, the activity as well as protection of the SphI and SphII motifs of the SV40 enhancer could be due to TEF-1.

The protection observed for the region corresponding to the P motif in retinoic acid differentiated cells could be due to the activation of additional transcription factors. Transcription factors AP1 and PEA1 have been observed to bind to this motif (Jones *et al.* 1988). Differentiation induced activation of PEA1 proteins in F9 cells and their binding to

the polyomavirus enhancer has been reported (Kryszke *et al.* 1987). Transcription factor AP1, encoded by members of the c-*jun* oncogene family, can interact with DNA efficently only in association with c-*fos* protein (Chiu *et al.* 1988; Sassone-Corsi *et al.* 1988; Jones, 1990). Thus, the increase in DNA binding activity at the P motif may involve activation of either AP1 or c-*fos*. A recent report indicates that c-jun but not c-fos is induced after retinoic acid treatment of P19 cells (deGroot *et al.* 1990). Thus, the approximately 200-fold increase in activity of pSV2-cat after retinoic acid differentiation could be due to induction of AP1 in conjunction with the increase in the amount of TEF-1-like factors.

The footprints of SV40 DNA with DMSO differentiated cell extracts varied from experiment to experiment. For example, the footprints presented in Figure 2.2, DM shows protection in P motifs and sequences interspersed between the P and SphI motifs. Further extension of protection covering the SphII motif was obtained when footprints were performed with another batch of nuclear extracts. These differences could be due to variation in differentiation efficiency. TEF-1-like factors could be responsible for the protection.

The protected region in DMSO differentiated cell extracts is competed by both SV40 and BKV DNA. It is interesting to note that SV40 DNA competed for BKV footprints in this cell type (see below). Further discussion of these observations is presented below in the discussion of the BKV footprints.

The JCV regulatory elements, unlike those of SV40 and BKV which are active in both retinoic acid and DMSO differentiated cells, is active only in retinoic acid differentiated cells (Table 2.1). Retinoic acid differentiation gives rise to neurogenic cells in which JCV is active. Thus, our studies, along with the results from Kenney *et al.* (1984), provide more conclusive evidence for the restricted neurotropic cell specificity of this regulatory element. DNaseI footprint analysis demonstrated a specific DNA-protein complex in retinoic acid differentiated cells but not in undifferentiated or DMSO-differentiated cells. Based on the sequences of the JCV enhancer that are protected in retinoic acid differentiated cell extracts, it appears that one of the factors belongs to the NF-1 protein family (Jones *et al.* 1987) and the other bears homology to the recently identified adenovirus E4 promoter binding protein, E4TF1 (Jones *et al.* 1988). Similar protection has been observed by others with extracts from glial cells and purified NF-1 (Amemiya *et al.* 1989; Tada *et al.* 1989; Tamura *et al.* 1988a). NF-1 proteins vary between different cell types and bind to related sequences with different affinities (Jones *et al.* 1987; McKnight and Tjian, 1986; Gil *et al.* 1988; Paonessa *et al.* 1988; Santoro *et al.* 1988). Since deletion mutation analysis (Table 2.1) indicated that mutants with only a TATA box and NF-1 motif sequences retain 25% of the activity of the entire JCV regulatory region and that the level of this activity is equivalent to the activity of one repeat, it

appears that NF-1 motif binding protein provides much of the JCV regulatory element function. It is possible that the NF-1 binding to the JCV enhancer in retinoic acid differentiated cells could be different from the NF-1 in nonexpressing cells, as suggested by others (Khalili *et al.* 1988). It has been reported that the degree of protection afforded to a motif corresponding to region I in JCV expressing glial cells and nonexpressing HeLa cells differ markedly, with complete protection only in glial cells (Amemiya *et al.* 1989). This complete protection could be due to the interaction of a second factor, present only in glial cells, along with NF-1. The factor(s) binding to the E4TF1 motif of this region in differentiated cells is a reasonable candidate. Affinity purification and UV crosslinking are required to examine this possibility. It is interesting to note that the same NF-1 motif binding protein can also bind to the BKV enhancer, since BKV DNA could compete for this factor *in vitro* (Fig. 2.4). Competition appeared to occur only between these two human papovaviruses, as SV40 showed no competition.

For the BKV enhancer, while the *in vivo* expression was closely parallel to that of SV40, DNaseI protection patterns suggested unique transcription factors were involved. As shown in Figure 2.6, retinoic acid induced differentiation produced five protected regions. Four of these regions contain sequences with homology to the NF-1 motif. Sequence $TGGN_7CCA$ which constitutes the most favourable binding site for CTF/NF-

1 is present only in region I and IV. NF-1 motif of region II and V contained the sequence $TGGN_8CCA$. As mentioned previously, NF-1 motifs have been shown to bind to a family of transcription factors (Gil *et al.* 1988; Paonessa *et al.* 1988; Santoro *et al.* 1988). Thus, it is an interesting possibility that the factor(s) binding to these motif(s) are also retinoic acid and DMSO-differentiated cell type specific. This is supported by my preliminary results with gel retardation assays which indicated that oligonucleotides containing NF-1 motifs elicit different gel shift patterns for retinoic acid differentiated P19 extracts and HeLa extracts.

Grinnell *et al.* (1988) reported that distinct factors bind to individual BKV NF-1 sites. Their results indicate that, in BKV P2 strain, region II and V are protected only in only HeLa cells but not in 293 and MK-2 cell extracts whereas regions I and IV were protected in all three cell types. Further, from results of *in vivo* competition studies, they suggested that factors binding to regions II and V are distinct from those for regions I and IV and act as negative regulators of transcription. However, Markowitz and Dynan (1988) were able to affinity purify proteins NF-BK from HeLa cells using oligonucleotides corresponding to region II and to show that the same factor(s) binds to all NF-1 motifs of BK (Dunlop), BK (WW) and BK (MM) strains. Factors potentially influencing the disparate results include 1) BKV strain specific differences in the sequences adjoining the NF-1 motifs, 2)

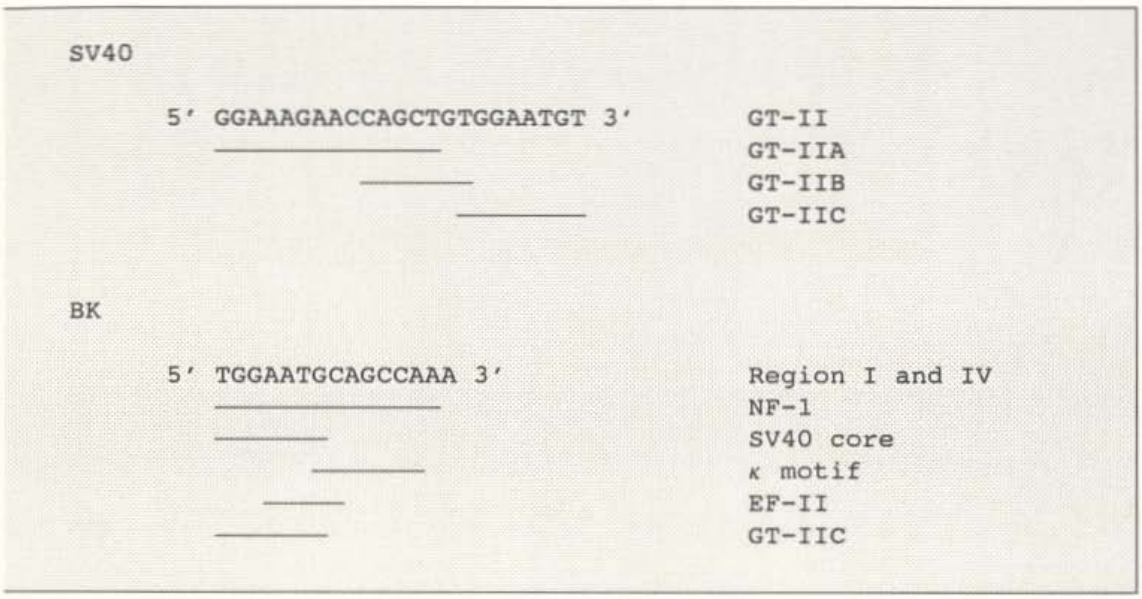
different affinities of factors for the four NF-1 motifs, 3) cell type specific differences in NF-1 factors, 4) differences between factors binding *in vitro* and *in vivo*, and 5) inability to copurify, by affinity chromatography, noncovalently associated protein species. Since the oligonucleotide competition indicated interaction of the same or similar factor(s) and there is good correlation in activity and DNA-protein interaction in retinoic acid differentiated cells, it is likely that all four NF-1 motif-containing regions interact with the same or similar positive regulatory factor(s) in this cell type.

As alluded to earlier, DNaseI footprints of BKV DNA in DMSO differentiated cells again showed protection but with a different pattern of regions protected as compared to those in retinoic acid differentiated cells. At a lower level of protein, mainly sequences corresponding to region I and region IV were protected while at a higher protein concentration additional protection for regions II and V was also observed (Fig. 2.6A; B, respectively). The protein concentration dependent variation in footprints suggests that some of the quantitative differences in *in vivo* expression in the retinoic acid and DMSO differentiated cells is due to quantitative differences in levels of NF-1 site and region III binding factors.

The BKV protected regions in DMSO differentiated cell extracts were competed by all three regulatory elements in the

order BKV > SV40 > JC DNA. The competition by SV40 is interesting as competition is observed only in DMSO differentiated cell extracts but not in retinoic acid differentiated cell extracts. Sequence comparison revealed that the three sequences in protected regions of BKV enhancer show considerable homology to the SV40 enhancer. They are located in regions I, II, IV and V. One of them is TGGGC of regions II and V which is identical to the LSF factor binding motif of SV40 (Wingender, 1988; Huang *et al*. 1990). The second is TGGAATG of regions I and IV which differs from the SV40 enhancer core by one nucleotide (TGGAAAG; Fig. 2.11; Angel *et al*. 1987). The third region has the sequence GTGGAATGT, which differs from the SV40 GT-IIC motif by two bases (underlined). As mentioned earlier, the GT-IIC motif and the SphI and SphII motifs all bind to the same factor, TEF-1 (Davidson *et al*. 1988). This observation allows the following interpretation: (1) Protection at the SphI and SphII motifs of SV40 and the GT-IIC motif homology region of BKV in DMSO differentiated cells could be due to binding of the same factor which is likely TEF-1. (2) Competition by BKV for SV40 footprints and vice versa in DMSO differentiated cells could be due to this shared factor(s). (3) Inability of BKV to compete for SV40 and vice versa in retinoic acid differentiated cells may be

Figure 2.11 Organization of SV40 GT-II motifs and BKV region I and IV



SV40: Sequence of GT-II enhanson is shown. GT-IIA, GT-IIB and GT-IIC motifs of GT-II enhanson is indicated by a line under the sequence. The data is adapted from Fromental *et al.* (1988).

BK: Sequences within regions I and IV are shown. NF-1, SV40 core, κ motif, EF-II and GT-IIC homologous sequences are underlined.

due to the presence of NF-1 family factors which bind to BKV preferentially over TEF-1. (4) BKV enhancer activity in retinoic acid differentiated cells may be due to the NF-1 family of factors whereas in DMSO differentiated cells the activity is due to TEF-1 like factors.

Apart from the sequences in the enhancer region, as described above, the TATA box of BKV may also play a role in DMSO differentiated cell type expression. Recent studies have established that the muscle specific expression of the myoglobin gene requires a muscle specific enhancer and a TATAAAA type TATA box (Wefald *et al.* 1990). Interestingly, the BKV promoter also contains the same TATAAAA type box. Further studies involving point mutation analysis should be helpful in identifying the importance of the motifs mentioned above and the TATA box in DMSO differentiated cell type specific expression of the BKV enhancer.

Previously published studies (see Chapter 1.2.1.1) and the present report suggest that the wide range activity of the SV40 enhancer is due to a complex cooperative and hierarchical organization of enhansons in the SV40 enhancer. I can draw a similar conclusion for the BKV enhancer which, like that of SV40, is active in a variety of cells. For example, the organization of the GT-II enhancer of SV40 and BKV regions I and IV show some interesting correlations (Fig. 2.11). The GT-II region of SV40 contains three overlapping motifs, GT-IIA, GT-IIB and GT-IIC. GT-IIA and GT-IIB bind ubiquitous

factors, whereas GT-IIC binds a cell type specific factor (Xiao *et al.* 1987). BKV regions I and IV contain sequences homologous to the NF-1 motif, SV40 core, GT-IIC, muscle creatine kinase κ motif and to an EF-II motif of Rous Sarcoma Virus (RSV) LTR (TATGCA in RSV LTR and AATGCA in BK; Sealey and Chalkley, 1987; Wingender, 1988; Fig. 2.11). The NF-1 motif, EF-II and SV40 core motif binding proteins are ubiquitous, whereas κ motif binding proteins are functionally active only in muscle cells (Ondek *et al.* 1988; Schirm *et al.* 1987; Lasser *et al.* 1989; Muller and Wold, 1989; Benezra *et al.* 1990). Thus, the expression of the BKV enhancer in great variety of cell types could be due to its ability to interact with both ubiquitous and cell type specific factors. This type of complex organization appears to be restricted to BK and SV40 as JCV protected regions showed homology only to NF-1 and E4TF1 motifs. This also provides an explanation for the restricted activity of JCV regulatory elements.

The strong tissue specificity of JCV regulatory elements and wider activity of BKV regulatory elements in tissue culture show good correlation with the results from clinical investigations. For example, JCV is associated with progressive multifocal leucoencephalopathy (PML), a degenerative subacute demyelinating disease resulting from infection of oligodendrocytes. In these patients, JCV is present in mononuclear cells in the bone marrow during latency, and virus-infected lymphocytes enter the perivascular space of the brain and

infect glial cells (Houff *et al.* 1988). Extensive viral replication is observed only in glial cells of brain of these patients (Padgett and Walker, 1976; Tooze, 1981). Since any primate cells can support viral DNA replication when the viral early protein, the T-antigen, is provided (Feigenbaum *et al.* 1987), glial cell type specific viral replication observed in PML patients could be due to glial cell type specific expression of T-antigen.

In contrast to JCV, BKV is associated with several lesions. For example, BKV has been implicated in cases of ureteral stenosis in recepients of renal transplants, hemorrhagic cystitis in recipients of bone marrow transplants and with an occasional case of cystitis in immunologically normal children. Furthermore, episomal BKV DNA has been isolated from tonsils (Goudsmit *et al.* 1982), kidneys (Chesters *et al.* 1983) and insulinomas (Caputo *et al.* 1983). These results suggest that BKV replicates in a variety of organs. Since, like JCV replication, BKV replication requires viral T-antigen, the ability of BKV to infect and replicate in a variety of tissues could be attributed to the wide range of activity of its regulatory elements.

CHAPTER 3

Ubiquitous And Cell-Type Specific Protein Interactions With Human Papillomavirus Type 11, 16 and 18 Enhancers

3.1 Introduction

Human papillomaviruses (HPVs) are small double-stranded DNA viruses which cause epithelial tumors or warts in humans. HPV types 6 and 11 are associated with benign lesions of external genitalia, while the HPV types 5, 8, 16, 17, 18, 20, 30, 31, 33, 35, 38, 39, 41, 45, 48, 51, 52 and 56 are associated with cervical lesions which can progress to malignant carcinomas (zur Hausen, 1986; Lancaster and Olson, 1982; Giri and Danos, 1986; deVilliers, 1989). Normal life cycle of the HPVs begins with the infection of the basal layer (Stratum basale) of the epithelium. The viral replication occurs in the stratum spinosum which is followed by maturation of viral particles in the stratum granulosum. The mature viral particles are, however, seen only in the uppermost layer of differentiated cells, the stratum corneum, of the the epidermis. In benign tumors, it is assumed that virus infection leads to transformation of one or more basal cells which results in increased cell proliferation and papilloma formation. Conversion of papilloma to carcinoma occurs after a long latency period. This process is usually accompanied by the integration of viral DNA into the host cell genome.

As for a number of other HPV types, HPV 11, 16 and 18 have been found in human epithelial tissues of diverse origin but

virus growth and expression is most favourable in epithelial cells of the female and male genital region. Functional analysis of their genomes have provided evidence for cervical epithelial cell specific promoter and enhancer elements which control gene expression (Chapter 1.3). A hypothesis proposed by many investigators is that the epitheliotropism of the HPV enhancers is mediated by epithelial cell-type specific transcription factors which activate transcription by sequence specific interactions (Cripe *et al.* 1987; Marshall *et al.* 1989; Garcia-Carranaca *et al.* 1988).

This study is aimed at testing the above hypothesis. For this purpose, constitutive enhancers of HPV 11, 16 and 18 were tested for their expression in four epithelial cervical cell lines and a non-cervical fibroblast cell line. The expression of all three enhancers was restricted to epithelial cells. DNaseI footprinting, performed as a measure of DNA-protein interaction, revealed seven, nine, and five footprints for HPV 11, 16 and 18, respectively. Cell-type variability in footprints was also obtained. Some DNaseI protected regions have been characterized further by oligonucleotide competition, gel retardation assays, *in vitro* transcription-competition, UV cross-linking and deletion analysis.

3.2 Materials and Methods

3.2.1 Materials

All materials are as described in Chapter 2.2.1.

3.2.2 Construction of recombinant plasmids

The plasmid pSV2-cat is described in Chapter 2.2.2. The plasmid pA10-cat containing the 21-bp repeated sequences and TATA box of SV40 coupled to the CAT gene was generated by joining the SphI/BamHI fragment of pBR322 plasmid with the SphI/BamHI (CAT containing) fragment of pSV2-cat (Laimins *et al.* 1982). At the SphI site, a BglII linker has been added. This site has been used to construct HPV-enhancer plasmids. HPV 11 enhancer-pA10-cat (pT-1), HPV 16 enhancer-pA10-cat (pT-3) and HPV 18 enhancer-pA10-cat (pT-81) were constructed as described in Marshall *et al.* (1989). Briefly, plasmids pT-1, pT-3 and pT-81 contain the 213 bp NdeI-SauI fragment of HPV 11 (nucleotides 7657-7870, Dartmann *et al.* 1985), the 554 bp SspI-HphI fragment of HPV 16 (7224-7778, Seedorf *et al.* 1985) and the 230 bp RsaI fragment of HPV 18 (nucleotides 7508-7738, Cole and Danos, 1987), respectively, in pA10-cat. Plasmid pT-1Δ was constructed by inserting the BalI-SauI fragment of HPV 11 (nucleotides 7718-7870, Dartmann *et al.* 1985) into the BglII site of pA10-cat vector in same orientation as in pT-1.

3.2.3 Cell culture and transfection

All cells were maintained in Dulbecco's modified Eagles medium with 10% fetal calf serum. The calcium phosphate

precipitation method, as described in Chapter 2.2, was used for transfection.

3.2.4 CAT assays

CAT assays are as described in Chapter 2.2.

3.2.5 Preparation of probes

The enhancer fragments of HPV 11 (nucleotides 7657-7870), HPV 16 (nucleotides 7400-7878) and HPV 18 (nucleotides 7400-7738) were first cloned into the XbaI site of pUC-19. The derivative plasmids were digested with SalI and SmaI in the case of HPV 11 and 18 and SphI and SalI in the case of HPV 16. The enhancer fragments were separated on agarose gels and electroeluted. For DNaseI footprinting, end-labelling at the SalI site was by reverse transcriptase using $\alpha[^{32}P]$dCTP. For labelling the non-coding strand, BamHI and PstI restricted fragments were used. For gel retardation, the oligonucleotides were end-labelled either with $\gamma[^{32}P]$-ATP by polynucleotide kinase (HPV 11) or by nick-translation in the presence of $\alpha[^{32}P]$dCTP and 5-bromodeoxyuridine (HPV 16 and HPV 18).

3.2.6 Nuclear extracts and DNaseI footprinting

Nuclear extracts and DNaseI footprinting assays were performed as described in Chapter 2.2. The amount of DNaseI used for digestion of HPV 11 DNA without any protein extracts was 4.5 ng or 3.5 ng and for DNA with protein the amount was 480 ng. HPV 16 and HPV 18 DNA without protein was digested

with 3.5 or 2.5 ng while with protein, the DNaseI concentration was 300 ng. In competition experiments, the same enhancer fragments as the probes were used. A 323 bp PvuII fragment of pUC-19 DNA served as control. Preparation of double stranded oligonucleotides and the oligonucleotide in control footprints was the same as in Chapter 2.2

3.2.7 Gel retardation and UV cross-linking

Gel retardation experiments were performed as described by Chodosh *et al.* (1988) with some modifications. Breifly, nuclear extracts were incubated with approximately 20,000 cpm of an end-labelled DNA (HPV 11) or 120,000 cpm of nick-translated DNA (HPV 16 and 18) in the presence of 7.5 μg of poly (dI.dC).poly (dI.dC) in a final volume of 15 μl. Binding reactions contained 12% glycerol, 12 mM HEPES-NaOH (pH 7.9), 4 mM Tris HCl (pH 7.9), 1 mM EDTA, 0.6 mM DTT, 300 μg/ml bovine serum albumin and 60 mM KCl. Incubations were carried out at 30° for 30 min. Gels were run at 4° with constant circulation of buffer.

For UV cross-linking, the binding reaction, as above, was irradiated under a Fotodyne UV lamp (maximum emisson wave-length, 310 nm; maximum intensity 7,000 $\mu W/cm^2$) for 20 minutes prior to loading gels. To determine the molecular weight sizes of interacting proteins, protein-DNA complexes in gel shift bands were eluted, heated at 85° for 10 minutes in sodium dodecyl sulfate polyacrylamide gel electrophoresis sample

buffer and then subjected to electrophoresis through 10% polyacrylamide-sodium dodecyl sulfate gels (Maniatis *et al.* 1982) along with ^{14}C-labelled molecular weight protein markers (lysozyme, 14,300; carbonic anhydrase, 30,000; ovalbumin 46,000; bovine serum albumin, 69,000; phosphorylase b, 92,500; myosin, 200,000; Amersham).

3.2.8 *In vitro* transcription

In vitro transcription reactions were according to Hai *et al.* (1988) with some modifications. Briefly, fifty μg protein in 20 μl of HeLa nuclear extract was incubated with 0.15 pmol of plasmid pT-1 in a 40 μl reaction containing 12.5 mM HEPES-NaOH (pH 7.9), 7.5 mM $MgCl_2$, 60 mM NaCl, 0.125 mM EDTA, 0.3 mM DTT, 12.5% glycerol, and 0.5 mM each of four ribonucleoside triphosphates. The reaction mixtures were incubated at 30° for 1 hr followed by treatment with DNaseI for 10 minutes. After DNaseI digestion, 10 μg of pronase in 60 μl of water and 100 μl of solution containing 200 mM Tris HCl (pH 7.5), 25 mM EDTA, 300 mM NaCl, and 2% SDS was added. After 30 min incubation at 37°, 500 cpm of non-specific end-labelled DNA was added and the transcripts were purified by phenol-chloroform-isoamyl alcohol followed by ethanol precipitation. Transcripts were quantitated by primer extension as described by Phelps and Howley (1987). A 20 base oligonucleotide complementary to a portion of the 5′ end of the CAT gene (5′ TCCATTTTAGCTTCCTTACG 3′) was used. The primer was 5′ end-

labelled with polynucleotide kinase (in 50 mM Tris HCl, pH 9.0, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA, 0.08 μg of primer DNA, and 5U of kinase for 1 hr at 37°). After heat inactivation at 70° for 10 min, the end-labelled primer (4 ng) was annealed to RNA by mixing in 100 mM NaCl, 20 mM Tris HCl, 0.1 mM EDTA and heating to 90° for 3 min. This mixture was then incubated at 55° for 10 min and cooled slowly to room temperature. Thirty μl of 2X primer extension buffer (100 mM Tris HCl, pH 7.5, 200 μg of BSA per ml, 6 mM $MgCl_2$, 150 mM KCl, 100 μg of actinomycin D per ml, 1 mM deoxyribonucleotide triphosphates) were added to the above reaction and the total volume was increased to 60 μl by the addition of water. One hundred units of cloned murine leukemia virus reverse transcriptase (BRL) were added to the reaction. Incubation was at 37° for 1 hr. The reaction was terminated by the addition of EDTA to 25 mM, followed by phenol-chloroform extraction and ethanol precipitation. The primer extension products were fractionated through an 8% polyacrylamide urea gel. The gel contained 8% polyacrylamide (20:1 ratio of acrylamide to bisacrylamide) and 8M urea in 1 X TBE (89 mM Tris-Borate, 2 mM EDTA, pH 8.3). Electrophoresis was in 1 x TBE.

3.3 Results

3.3.1 HPV 11, 16 and 18 enhancer activities in epithelial and fibroblast cell lines

Transient assays with plasmids, in which the CAT reporter gene is expressed under the regulation of the SV40 promoter and HPV enhancers were used to measure the enhancer function *in vivo*. Assays were performed with the four cervical carcinoma cell lines, C33A, HeLa, SiHa and CaSki, which contain and express, respectively, no HPV, HPV 18, HPV 16 and HPV 16 DNA (Pater and Pater, 1988a) and one fibroblast cell line, 143B. The activity of the enhancerless-SV40 promoter containing pA10cat, from which the HPV 11 enhancer plasmid pT-1, the HPV 16 enhancer plasmid pT-3 and the HPV 18 enhancer plasmid pT-81 were derived, was taken as the basal level of activity. The activities of the HPV enhancers in different cell lines are presented in Table 3.1. To allow better comparisons between different cell types and to avoid ambiguity due to differences in plasmid uptake, the SV40 enhancer-promoter construct pSV2cat was included as a positive control. In relation to pSV2cat, HPV 11 showed the highest level of activation in C33A cells, slightly lower levels of expression in HeLa and SiHa cells and near basal levels of expression in CaSki and 143B cells. The HPV 16 enhancer construct T-3 showed the highest level of activity in SiHa cells, followed by HeLa and C33A cells. No increase in activity over pA10-cat levels was observed in CaSki or 143B cells. The HPV 18 enhancer construct T-81 showed higher activities in HeLa and SiHa cells and slightly lower activity in C33A cells. As for the HPV 11 and 16 enhancers, the HPV 18 enhancer did not increase pA10-

Table 3.1 Activity[a] of HPV 11, HPV 16 and HPV 18 constitutive enhancers in cervical and fibroblast cell lines

Plasmids	C33A(-)	HeLa (HPV 18)	SiHa (HPV 16)	CaSki (HPV 16)	143B(-)
pA10cat	1.0 (0.4)	1.0 (0.1)	1.0 (0.7)	1.0 (0.3)	1.0 (0.3)
pSV2cat	893 (357)	160 (15.80)	30 (20.7)	13 (4.0)	23 (6.9)
T-1 (HPV 11)	750 (298.9)	114 (11.40)	21 (14.6)	2.0 (0.5)	1.0 (0.3)
T-3 (HPV 16)	4.0 (1.7)	10.0 (1.0)	11 (7.9)	1.0 (0.3)	1.0 (0.3)
T-81 (HPV 18)	35 (14.1)	19 (1.9)	4.5 (3.0)	1.0 (0.3)	1.0 (0.3)

[a] 10 μg of test plasmid was used along with 10 μg of pUC-19 DNA. Cellular extracts were prepared 48 h post-transfection. The CAT activity of pA10cat was normalized to 1.0 and fold activation above this level by different enhancers is presented. Numbers in brackets are the actual percentage conversion of substrate. Results are averages of duplicate experiments. Type of HPVs integrated into the genome of cervical carcinoma cell lines are also indicated.

cat activity in CaSki and 143B cell lines. The activity of pSV2-cat is also low in CaSki and 143B in comparison to other cell types. The lower level activity of HPV 16 enhancer in CaSki is surprising as the integrated HPV 16 DNA is expressed in this cell line. One possible explanation for the reduced expression is that CaSki contains fewer transcription activation factors. This is supported by the DNaseI footprinting results of Fig. 3.6A and B, in that protection is weaker and competition by the various HPV enhancer fragments is much more pronounced, even at low concentrations, for CaSki nuclear extracts as compared with the other nuclear extracts. Also, the results of my gel retardation assays (Fig. 3.3 and 3.14) have shown relatively low levels of factors binding to oligonucleotides corresponding to some of the protected regions. Taken together, these results document cell-type variability of HPV 11, 16 and 18 enhancer function even though all three HPV expressing cell lines were derived from human cervical cancers and were of epithelial origin.

3.3.2 Protein-binding motifs within the constitutive enhancer of HPV 11

The footprints for HPV 11 are shown in Figure 3.1 and the protected sequence motifs are summarized in Table 3.2. There were five protected regions for C33A (Fig. 3.1A), six for SiHa cell extracts (Fig. 3.1C) and seven for HeLa, CaSki and 143B cell extracts (Fig. 3B, D and E, respectively). Protected

Figure 3.1 DNaseI footprinting analysis of the HPV 11 constitutive enhancer. An asymmetrically labelled HPV 11 enhancer fragment was subjected to DNaseI footprint analysis using 20 μg of nuclear extract protein from C33A (A), HeLa (B), SiHa (C), CaSki (D) and 143B (E) cell lines. Chemical cleavage of purines are in lane A+G. DNA treated with DNaseI in the absence of nuclear extract is in lane F, and in the presence of nuclear extract in lane B. Lanes 1 to 8 represent competition experiments where the same amount of protein and probe were incubated together with cold competitor DNA before DNaseI digestion. Competitor DNAs were: HPV 11 enhancer in lanes 1 and 2, HPV 16 enhancer in lanes 3 and 4, HPV 18 enhancer in lanes 5 and 6 and pUC-19 DNA in lanes 7 and 8. Competitor DNAs used were 10 ng (5-fold excess) for odd numbered lanes and 50 ng for even numbered lanes. Protected regions are designated I to VII.

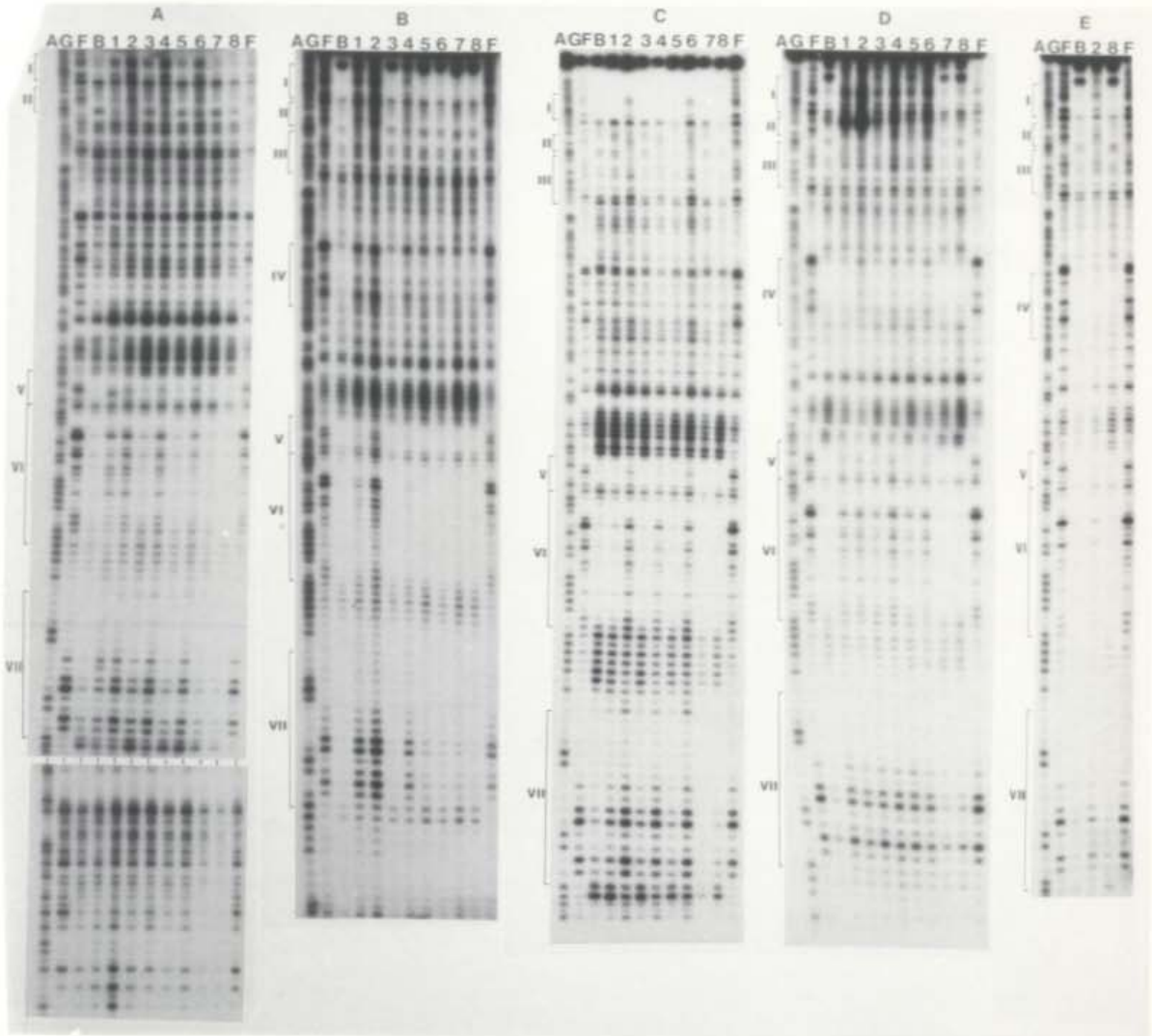

A
AGFB1 2 3 4 5 6 7 8 F
I
II
V
VI
VII
B
AGFB1 2 3 4 5 6 7 8 F
I
II
III
IV
V
VI
VII
C
AGFB1 2 3 4 5 6 7 8 F
I
II
III
V
VI
VII
D
AGFB1 2 3 4 5 6 7 8 F
I
II
III
IV
V
VI
VII
E
AGFB2 8 F
I
II
III
IV
V
VI
VII

Table 3.2 Sequences and locations of HPV 11 protected regions and cell lines for which protection is observed

Protected regions[a]	Sequences[b]	Nucleotide number[c]	Cell lines				
			C33A	HeLa	SiHa	CaSki	143B
I	ATGTTGTGTGCCAAGGT	7660-7676	+	+	+	+	+
II	TATTGCCCTGCCAA[d]	7680-7693	+	+	+	+	+
III	TATCTTGCCAACA	7695-7707	-	+	+	+	+
IV	TATTGCATGACTAATGT[e]	7729-7745	-	+	-	+	+
V	TGTTG	7771-7775	+	+	+	+	+
VI	TGGATTGCAGCCAAAGGTT[f]	7776-7794	+	+	+	+	+
VII	TTTGGCTTCTAGCTGA	7804-7819	+	+	+	+	+

[a] Protected regions summarized here and shown in Fig. 3.1 are identical.

[b] 5' GCCAA 3' motifs are underlined, LVc homology is represented by hatched box and øAP3 homology is represented by broken lines. Sequence ATTGC which is present in HPV 16 and 18 is indicated by wavy lines.

[c] Nucleotide numbers are according to Dartmann et al. (1985).

[d,e,f] Competition experiments were performed using synthetic oligonucleotides identical to these sequences.

regions I, III, IV, V plus VI and VII correspond to fp1.11, 2.11, 3.11, 4.11 and 5.11, respectively, of Gloss *et al.* (1989b) It appears that proteins interacting with region I, II and III in HeLa, CaSki and 143B interact with each other as the sequences between these regions are also partially protected. Gloss *et al.* (1989b) report part of region II and all of region III as a single protected region (fp2.11) in HeLa cells. However, the observation that region II is protected in all cell lines and region III is protected in all cells except C33A suggests that region II and III are separate footprints. All protected regions were competed by the HPV 11 DNA but not by the nonspecific competitor 323 basepair PvuII fragment of pUC-19 DNA.

Two sequence motifs are found in more than one protected region. The first is a sequence ATTGC (indicated by wavy lines; Table 3.2) of regions II, IV and VI which differs from prototype CBP protein binding site by one nucleotide (ATTGG) and has been shown to bind CBP with higher efficiency (Graves *et al.* 1986). The second motif, GCCAA (underlined in Table 3.2), is present in protected regions I, II, III, VI and VII and has also been identified as NF-1 binding site by Gloss *et al.* (1989b). NF-1 constitutes a large family of factors, few of which have been cloned. The first is NF-1/CTF which binds as a dimer with high affinity to sequences that contain the inverted repeat $TGGN_7CCA$ (Jones *et al.*, 1987; Gournavi *et al.*, 1990). It also binds as a dimer to "half sites" that contain

only one copy of TGG. Santoro *et al.* (1988) isolated three cDNAs for the NF-1/CTF protein from human HeLa cells. All of the mRNAs were derived from a single gene by alternative splicing and differed by the presence or absence of an insertion in the middle of the protein and the presence of various COOH termini. Another member of the NF-1/CTF family is the TGGCA-binding protein, the cDNA of which has been isolated from rat liver (Paonessa *et al.*, 1988). Among the protected regions of HPV 11, a high affinity NF-1/CTF site ($TGGN_7CCA$) was observed in region VI, half sites in regions I, II, III and VII, and a TGGCA motif in regions I, II and III.

Cell-type variability was observed in region IV: this region is protected in HeLa, CaSki and 143B cells but only partially in C33A and weakly in SiHa cell lines. However, sequences above and below this protected region showed increased cleavage by DNaseI in all cell types, a characteristic commonly observed at boundaries of DNA-protein interaction, suggesting that C33A and SiHa contain either a low level or a unique factor(s) interacting with this region (Fig. 3.1). For region II, a difference in the sequences protected was observed for CaSki. This altered protection pattern could also be related to the uniquely low level of activity of the enhancer in CaSki (Table 3.1) as compared with the other three cervical carcinoma cell lines.

A computer search of different HPV enhancer elements showed distinct blocks of homology among the enhancer sequences

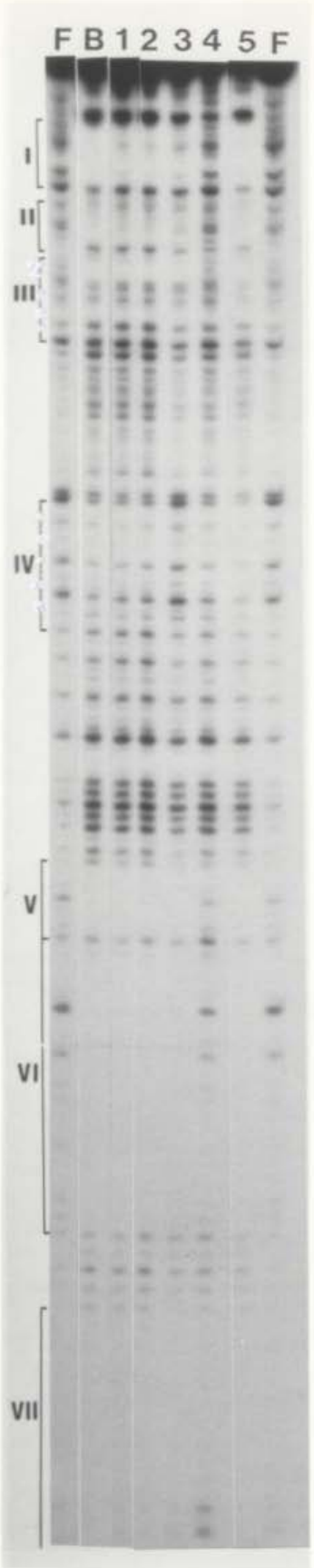
of HPV types 11, 16 and 18. This suggested that these three HPVs, which share cell specificity for epithelial tissues of the reproductive system, might also share similar mechanisms in enhancer activation. To this end, I performed cross competition experiments using unlabelled HPV 16 and 18 enhancer elements (Fig. 3.1). Competitors were the 415 bp fragment of HPV 16 (nucleotide 7463-7878; Seedorf *et al.* 1985) and the 230 bp RsaI fragment of HPV 18 (nucleotides 7508-7738; Cole and Danos, 1987). Both enhancers were able to compete for the NF-1 binding sites of HPV 11 enhancers, albeit at generally reduced efficiency when compared with the HPV 11 enhancer. There is an interesting negative correlation between the type of HPV competitor DNA and the type of cervical cell line. HPV 16 and 18 compete equally well for NF-1 sites in C33A cells, which contain neither type of HPV DNA, while HPV 16 competes better than HPV 18 in the HeLa cells, which contains HPV 18 DNA and the opposite is true for SiHa, which contains HPV 16 DNA. The correlation might be fortuitous or it might be due to the endogenous HPV type specific DNA, RNA or proteins and/or virally induced cellular proteins. In addition to the GCCAA motif containing regions, competition of HPV 16 and HPV 18 was also observed for the other two regions but at a generally reduced efficiency. This could be due to the sequence ATTGC which is shared among all three viruses. Alternatively, competition could be due to redundancy in enhancer binding factors or due to the depletion of some of the common "piggy-

back" transcription factors (factors that interact with DNA bound proteins; Shaw *et al.* 1989) as the competitors used were of whole length enhancers.

3.3.3 Oligonucleotide competition studies of DNaseI protected regions of HPV 11

Further analysis of the HPV 11 enhancer was performed by oligonucleotide competition using synthetic oligonucleotides corresponding to region II, region II with a mutation of the GCCAA motif to GGGAA, region IV and region VI. The results of the competition studies are presented in Figure 3.2. The protection of region II was reduced slightly only at high concentration of oligonucleotide (250 ng) corresponding to the homologous region (Fig. 3.2, lane 1). At the same concentration, mutant region II oligonucleotide competed with the same efficiency as the wild type. It is not clear why mutation of the GCCAA motif did not affect the binding of factors (see Discussion). However, competition by these oligonucleotides cannot be non-specific as no other protected regions except region I and III were competed. The requirement for a high concentration of this oligonucleotide for *in vitro* competition could be due to the relative inefficiency of the monomeric sequence to bind to the factor(s). Another possible explanation would involve cooperative interaction between adjoining region I and III, which does appear to occur since protection of all three regions was competed equally by

Figure 3.2 Oligonucleotide competition footprints for the HPV 11 enhancer. Footprints were performed with HeLa cell nuclear extracts. DNA treated with DNaseI in the absence of nuclear extract is in lane F, and in the presence of nuclear extract in lane B. Lane 1 to 5 represents competition experiments where the same amount of protein and probe were incubated with double-stranded oligonucleotides before DNaseI digestion. Oligonucleotides were: the wild type region II sequence in lane 1 (5'TATTGCCCTGCCAA 3'), a mutant region II sequence in lane 2 (5'TATTGCCCTGGGAA 3'), the region IV oligonucleotide in lane 3 (5'TATTGCATGACTAATGT 3'), the region VI oligonucleotide in lane 4 (5'TGGATTGCAGCCAAAGGTT 3') and a control oligonucleotide (5' AAGGGAAGGGATGG 3' of JC virus, Frisque *et al.* 1984) in lane 5. Concentration of oligonucleotides was 250 ng in lane 1, 2 and 5 and 100 ng in lanes 3 and 4.


F B 1 2 3 4 5 F
I
II
III
IV
V
VI
VII

this oligonucleotide. Therefore, the enhancer fragment probe containing all three regions has higher affinity for binding than the monomeric oligonucleotide. The region IV oligonucleotide (100 ng) competed with the homologous region (Fig. 3.2, lane 3). Competition was also observed for region I and II. The DNaseI sensitive regions above and below this region were also eliminated by competition. The region VI oligonucleotide (100 ng) competed for all regions except region IV (Fig. 3.2, lane 4).

3.3.4 Analysis of DNA-protein interactions of region II and VI of HPV 11 by gel retardation

Oligonucleotide competition studies suggested that protection patterns of region II and region VI, although they contain the same GCCAA motif, are differentially affected by homologous and heterologc competitors. To allow more detailed analysis, gel retardation assays were performed. Gel retardations, also called gel mobility shift assays, are useful to evaluate quantitative and qualitative differences in transcription factors among various cell types. The procedure involves incubation of nuclear extracts with end-labelled DNA fragments or oligonucleotides and electrophoresis of this mixture in a native acrylamide gels. In these gels, DNA that is complexed with proteins have retarded mobility compared to free non-complexed DNA. The intensity and the number of DNA-protein complexes obtained with nuclear extracts from different cell

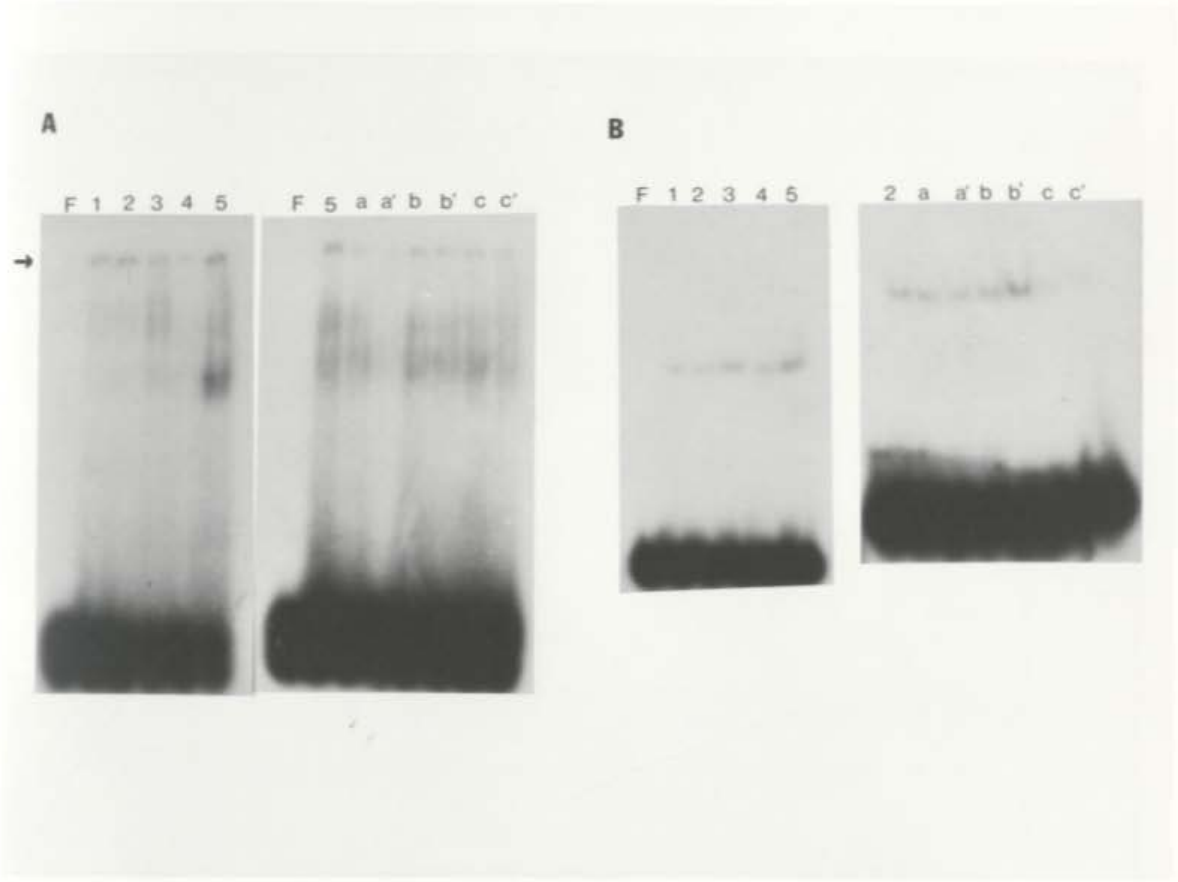
types depends on the quantity and quality of DNA binding proteins present in these cells. Specificity of DNA-protein interaction can be tested by a competition assay similar to DNaseI footprinting-competitions.

As shown in Figure 3.3, there was considerable difference in the DNA-protein complexes formed with region II and region VI oligonucleotides. Two bands with reduced mobility for region II and one for region VI were observed. The retarded bands of region II oligonucleotide was always diffuse even under varying experimental conditions. The region II oligonucleotide result demonstrated an interesting correlation with *in vivo* results (Table 3.1). The faster migrating complex was more abundant in non-expressing fibroblast cells than in the expressing cell lines C33A, SiHa and HeLa. Quantitative differences in factor(s) binding to region VI oligonucleotide were also observed for the various cell types with CaSki nuclear extracts providing the lowest level of interacting factor. This was consisten with DNaseI footprinting of CaSki in which the protections was eliminated with a low concentration of competitor (Fig. 3.1).

Next, I examined the effects of competition in mobility shift experiments (Fig. 3.3). The region II oligonucleotide competed effectively for homologous region II-protein interactions but poorly for those of region VI. A mutation of GCCAA motif of region II resulted in poor competition for both region

Figure 3.3 Gel retardation assays. A) End-labelled region II oligonucleotide was incubated with 4 μg of nuclear extracts from C33A (lanes 1), HeLa (lane 2), SiHa (lane 3), CaSki (lane 4) and 143B cells (lane 5) and 7.5 μg of poly dI.dC and run on 4% polyacrylamide gel (left). Arrow indicates the top of the gel. The right panel shows competition assays performed in 143B cell nuclear extracts. Competitors were region II oligonucleotide (lanes a and a'), mutant region II oligonucleotide (lane b and b') and region VI oligonucleotide (lane c and c'). Lanes a, b and c contained 30 fold-molar excess of competitor DNA, whereas lane a', b' and c' contained 100 fold-molar excess of competitor DNA.

B) Region VI oligonucleotide was used as a probe. Nuclear extracts, competitors and their concentrations are as in A. Competition assays were with HeLa cell extracts.


A
F 1 2 3 4 5
F 5 a a' b b' c c'
B
F 1 2 3 4 5
2 a a' b b' c c'

II and region VI-protein interaction. Competition of mobility shifts of region II oligonucleotide by the region VI oligonucleotide was also very poor. This is in contrast to DNaseI footprint competition results of Figure 3.3 in which the region VI oligonucleotide was a better competitor for region II than region II itself. Again, it would appear that cooperative interactions among adjacent NF-1 sites strongly influence binding to the intact enhancer sequences, a result not appreciated in experiments with isolated regions (see Discussion).

3.3.5 The role of protected region II of HPV 11 in enhancer function

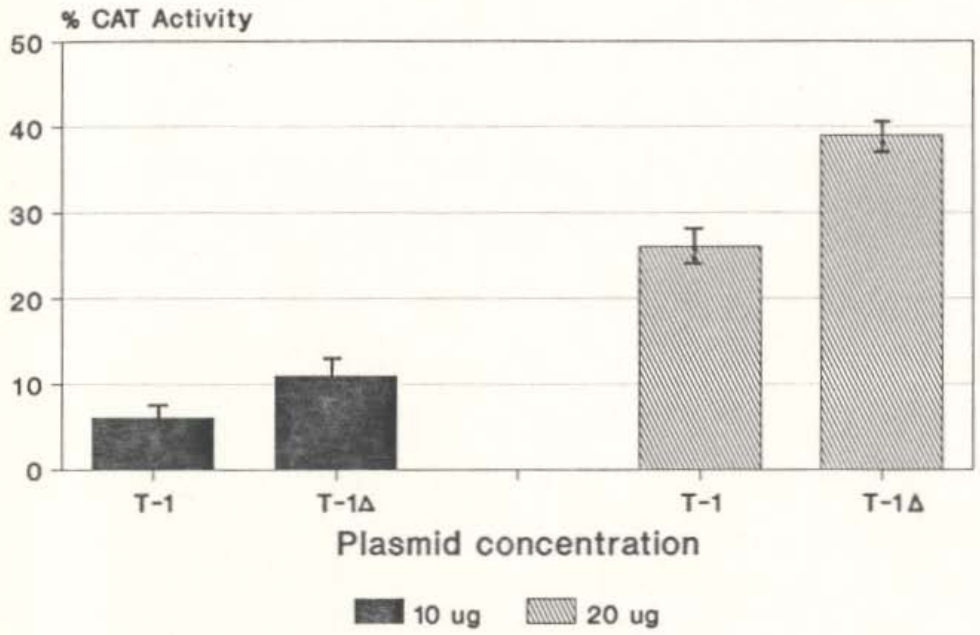
Since gel retardation assays demonstrated cell-type specific differences, particularly for binding to region II sequences, *in vivo* and *in vitro* assays were performed to study the role of this region in HPV 11 expression. *In vivo* assays compared the expression of pT-1 with a mutant (pT-1Δ) lacking regions I, II and III (Fig. 3.4).

The expression of this mutant in HeLa cells is approximately 50% more than pT-1. These results suggested that factors binding to sequences within regions I, II and III negatively regulate HPV 11 expression. I have performed similar experiments in 143B cells. Although pT-1Δ expressed better than pT-1, clear interpretation is difficult due to very low enhancer activity.

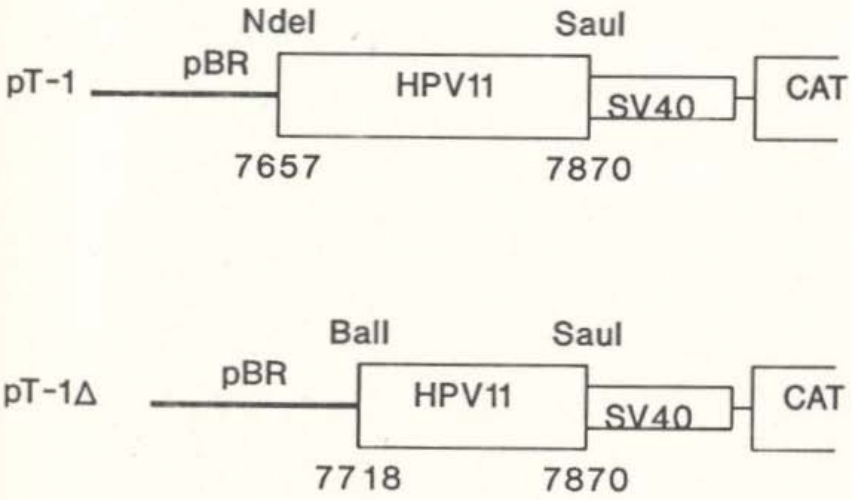
Figure 3.4 Deletion analysis of the HPV 11 constitutive enhancer. A) HeLa cells were transfected with the indicated quantity of plasmid pT-1 or deletion derivative pT-1Δ. Cellular extracts were prepared 48 h post-transfection and CAT assays were performed. The average of four experiments is presented.

B). Schematic representation of plasmids pT-1 and pT-1Δ.

A

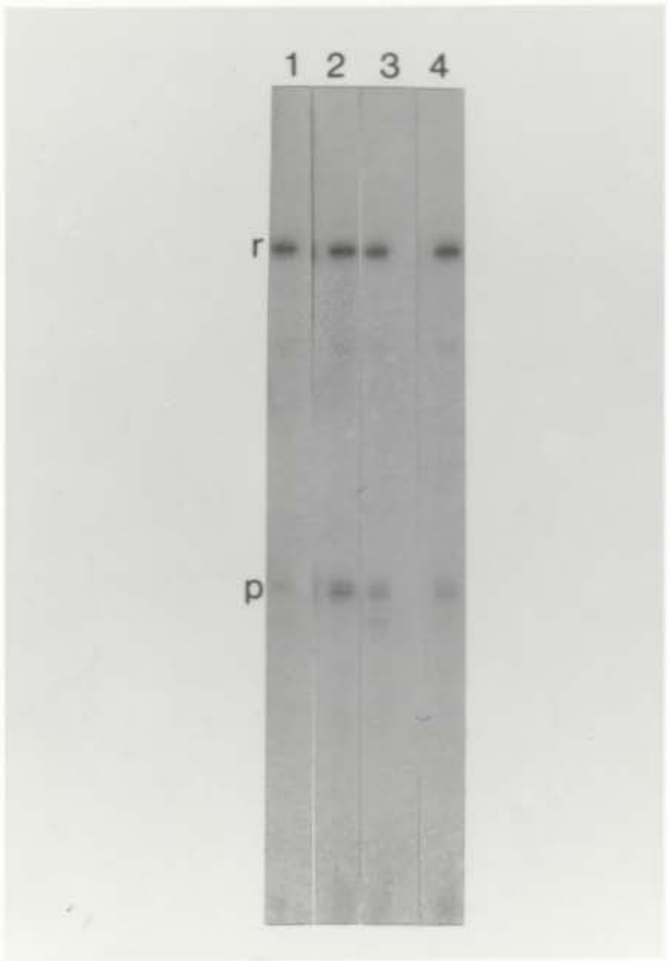

% CAT Activity
50
40
30
20
10
0
T-1
T-1Δ
T-1
T-1Δ
Plasmid concentration
10 ug
20 ug

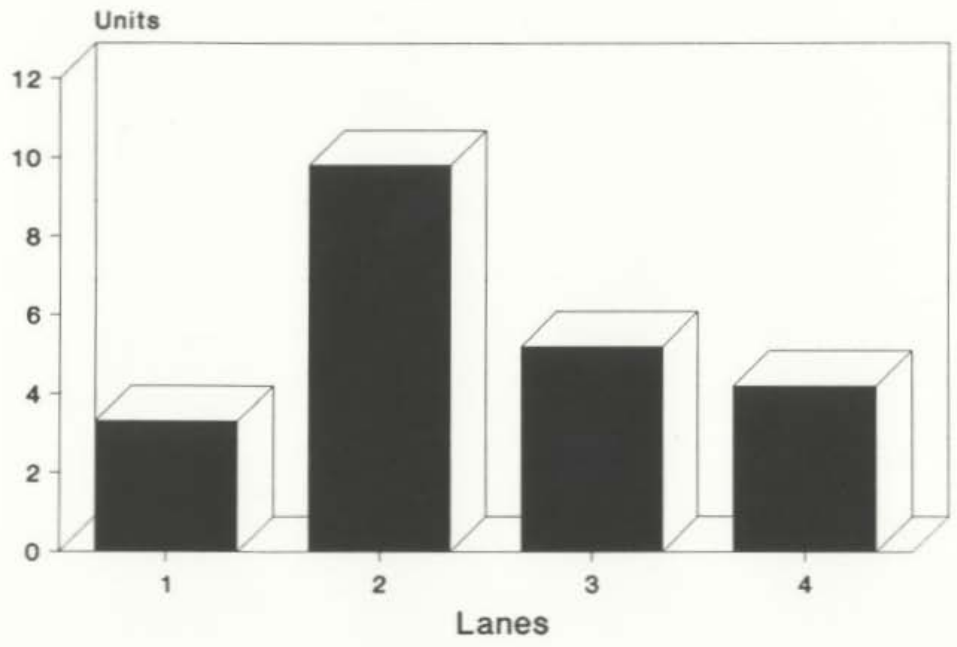
B


NdeI
SauI
pT-1
pBR
HPV11
SV40
CAT
7657
7870
BalI
SauI
pT-1Δ
pBR
HPV11
SV40
CAT
7718
7870

To determine whether the negative regulatory effect is due to factors binding to region II, oligonucleotide competition assays were performed. *In vitro* transcription assays were carried out with HeLa cell nuclear extracts, plasmid pT-1 as template and region II and VI homologous oligonucleotides as competitors. Nuclear extracts used in these assay were the same as those used in the DNaseI footprinting and gel retardation assays. They are presumed to contain protein factors necessary for transcription initiation and elongation. When a template containing an enhancer, promoter and a gene is incubated alongwith nuclear extracts, and all four ribonucleoside triphosphates the gene in the template is transcribed. The quantity of transcript generated is dependent upon the level of positively and negatively acting factors present in the extracts. It is possible to artificially increase the level of factor(s) of interest in the transcription reaction by the addition of purified factor(s). Factor(s) of interest can also be prevented from interacting with the enhancer and promoter of the template by the preincubation of nuclear extracts with synthetic oligonucleotides containing recognition sequences for the factor(s) of interest. Figure 3.5 represents typical results of such experiments in which factors binding to region II and region VI are prevented from interacting with the HPV 11 enhancer. Preincubation of nuclear extracts with region II oligonucleotide increased reproducibly pT-1 expression approximately

Figure 3.5 An oligonucleotide-competition assay for *in vitro* activity of the plasmid pT-1. Top: Transcription of pT-1 (0.15 pmol) was carried out in a HeLa cell nuclear extract (50 μg) with no oligonucleotides (lane 1), region II oligonucleotide (lane 2), mutant region II oligonucleotide (lane 3), and region VI oligonucleotide (lane 4). The molar ratio of T-1 to oligonucleotide is 1:1000. The primer extension products representing transcripts (p) and end-labelled DNA (SV40 enhancer fragment) as an internal control (r) are indicated on the left. Bottom: the level of transcription in each lane of the above autoradiogram was quantitated by densitometry and is presented in bar graph form.


1 2 3 4
r
p


Units
12
10
8
6
4
2
0
1
2
3
4
Lanes

three-fold. Values obtained for mutant region II oligonucleotide and wild type region VI oligonucleotide were not significantly above control. I attempted similar experiments *in vivo* where the dimeric oligonucleotide of region II cloned into pUC-19 was used as a competitor. In each of ten experiments there was an increase in pT-1 expression. Since there was considerable variation among experiments (1.5 fold to 5 fold increase), the results are not presented graphically.

3.3.6 Protein binding motifs within the constitutive enhancer of HPV 16

Figures 3.6A and 3.6B document footprints obtained for the HPV 16 enhancer. Sequences within footprints numbered I to IX are presented in Table 3.3. Most of the protected regions for HeLa extracts were as reported by Gloss *et al.* (1989a); although the protection of region I was not reported, protection of regions II through IX were, and were designated as fp2e, fp3e, fp4e, fp5e, fp6e, fp7e, fp8e, and fp9e, respectively. Regions II, III, V and VII are protected in all five cell

lines tested, although protection of region II was partial in CaSki. Region III protection was extended in C33A (for the coding strand only; Fig 3.6A) and region VII protection was a shorter distance towards the region VI-proximal side in CaSki. Region I is protected completely only by C33A and 143B extracts (for the coding strand, but not the non-coding strand; compare Fig. 3.6A to 3.6B), while region IV is protected completely in 143B and partially in HeLa and CaSki cell

Figure 3.6A DNaseI protection analysis of the HPV 16 constitutive enhancers. The HPV 16 enhancer, radiolabelled on the coding strand, was incubated with nuclear extracts from A) C33A, B) HeLa, C) SiHa, D) CaSki and E) 143B cells and subjected to DNaseI digestion. Chemical cleavage of purines is in lanes A/G. DNA treated with DNaseI in the absence and presence of nuclear extracts are in lanes F and B, respectively. Lanes 1 to 8 represent competition experiments in which nuclear extracts and probe were incubated with cold competitor DNA before DNaseI digestion. Competitors were HPV 11 DNA in lanes 1 and 2, HPV 16 DNA in lanes 3 and 4, HPV 18 DNA in lanes 5 and 6, and pUC-19 DNA in lanes 7 and 8. Odd numbered lanes had 10 ng and even numbered lanes had 50 ng of competitors. Protected regions are indicated by brackets labelled I to IX.

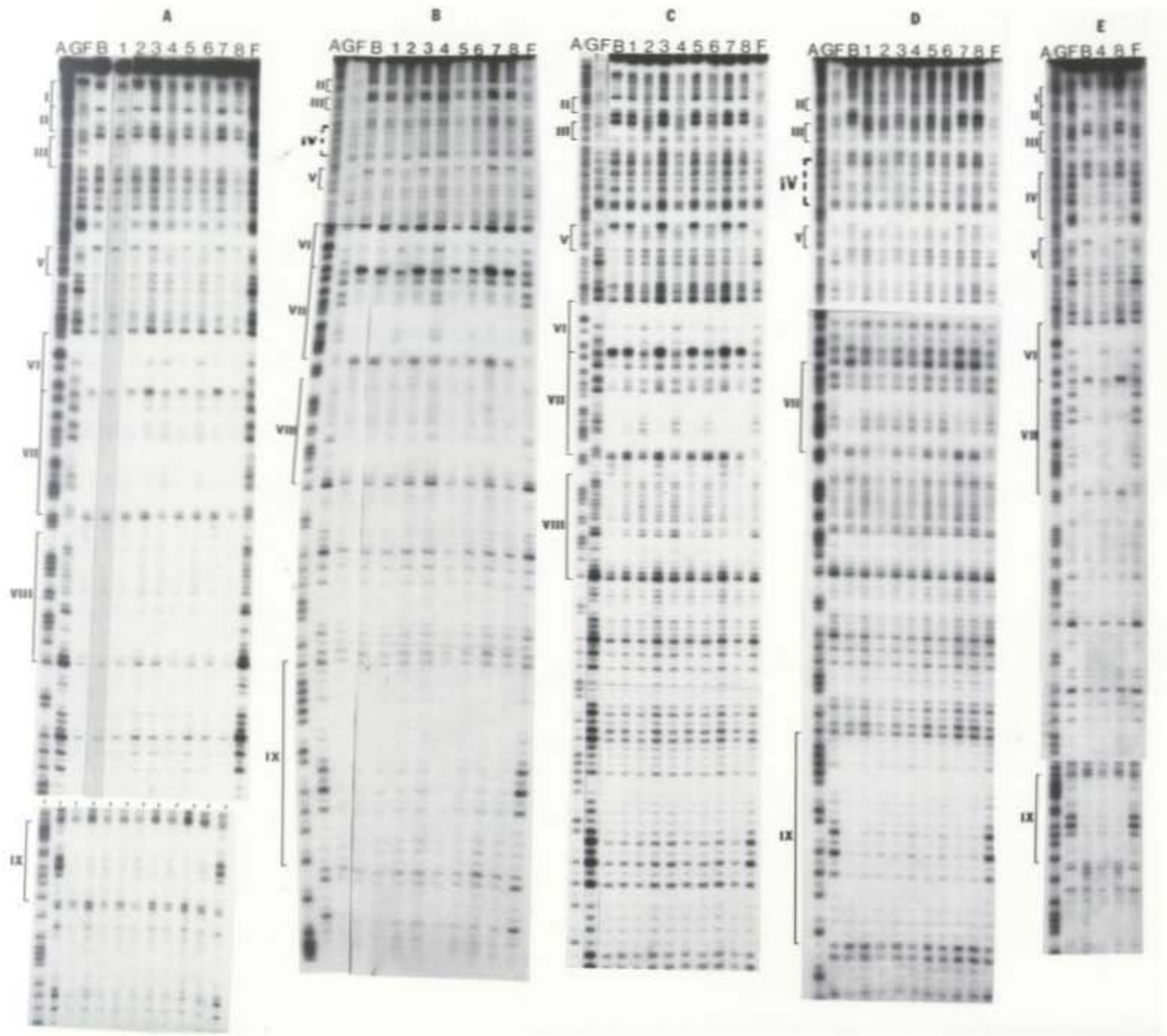

A
AGFB 1 2 3 4 5 6 7 8 F
II
III
V
VI
VII
VIII
IX
B
AGFB 1 2 3 4 5 6 7 8 F
II
III
IV
V
VI
VII
VIII
IX
C
AGFB 1 2 3 4 5 6 7 8 F
II
III
V
VI
VII
VIII
D
AGFB 1 2 3 4 5 6 7 8 F
II
III
IV
V
VII
IX
E
AGFB 4 8 F
II
III
IV
V
VI
VII
IX

Figure 3.6B Non-coding strand footprints of the HPV 16 enhancer. Probes were incubated with nuclear extracts from A) C33A, B) HeLa, C) SiHa, D) CaSki and E) 143B cells. Chemical cleavage of purines are in lanes A/G. DNA treated with DNaseI in the absence and presence of nuclear extracts are in lanes F and B, respectively. Competition with HPV 16 and pUC-19 DNA are in lanes 4 and 8, respectively.

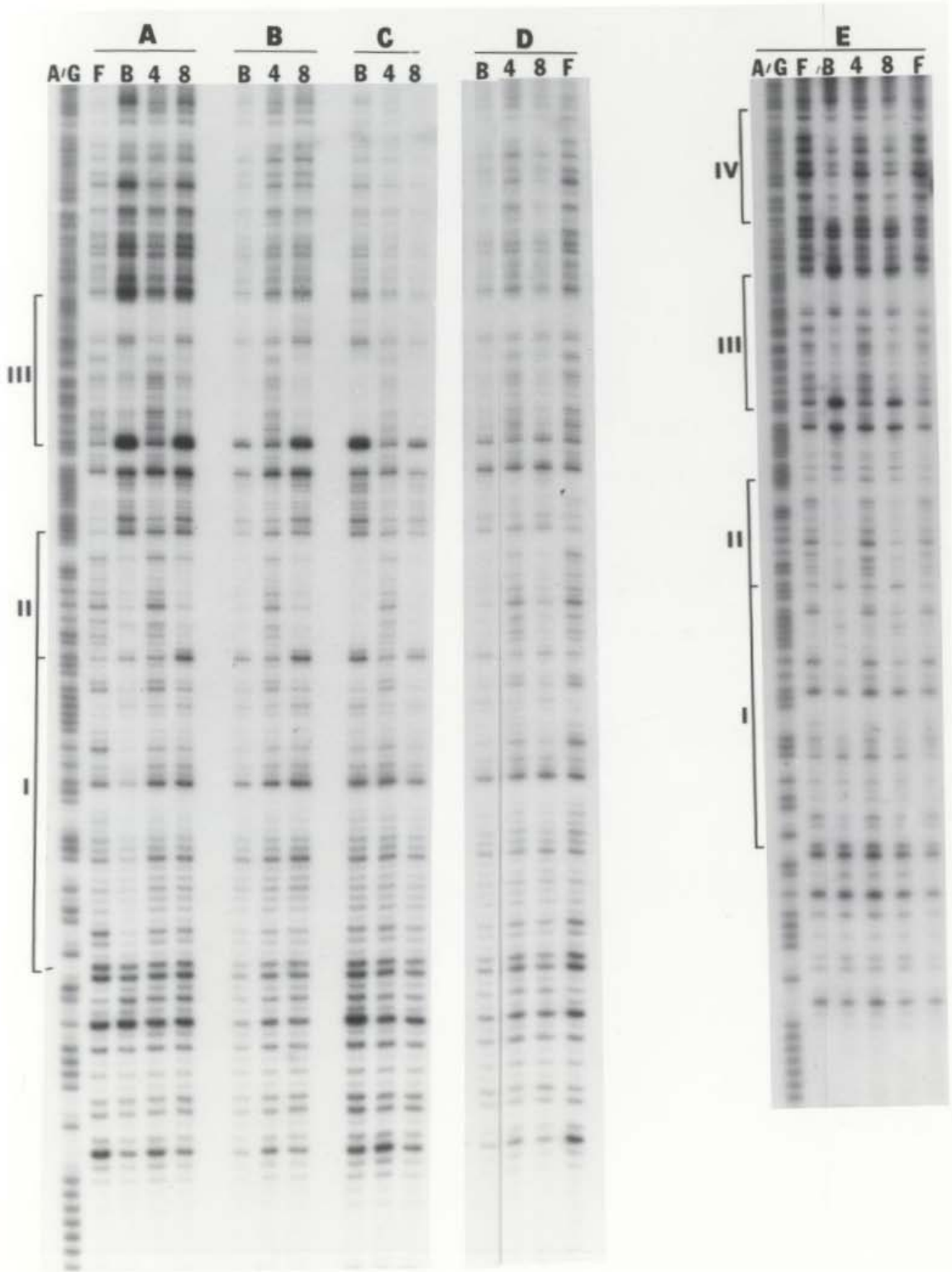

A
B
C
D
E
A/G F B 4 8
B 4 8
B 4 8
B 4 8 F
A/G F B 4 8 F
IV
III
II
I

Table 3.3 Sequences and locations of HPV 16 protected regions and cell lines for which protection is observed

[a] Numerals I to IX are the same as the regions marked in Figure 3.6.

[b] NF-1 GCCAA motifs are underlined. Broken lines represent putative AP1 binding motifs. Repeated sequences TTTTA are indicated by double lines. Sequences shared by region III and HPV 18 regions I and V are boxed. ATTGC sequence shared by HPV 11 and 18 is indicated by wavy lines.

[c] Nucleotide numbers are according to Seedorf *et al.* (1985).

[d] The bracketed sequences were protected for C33A only in the coding strand but for all cell lines in the non-coding strand

[e] The bracketed sequences were not protected for CaSki.

Table 3.3 Sequences and locations of HPV 16 protected regions and cell lines for which protection is observed

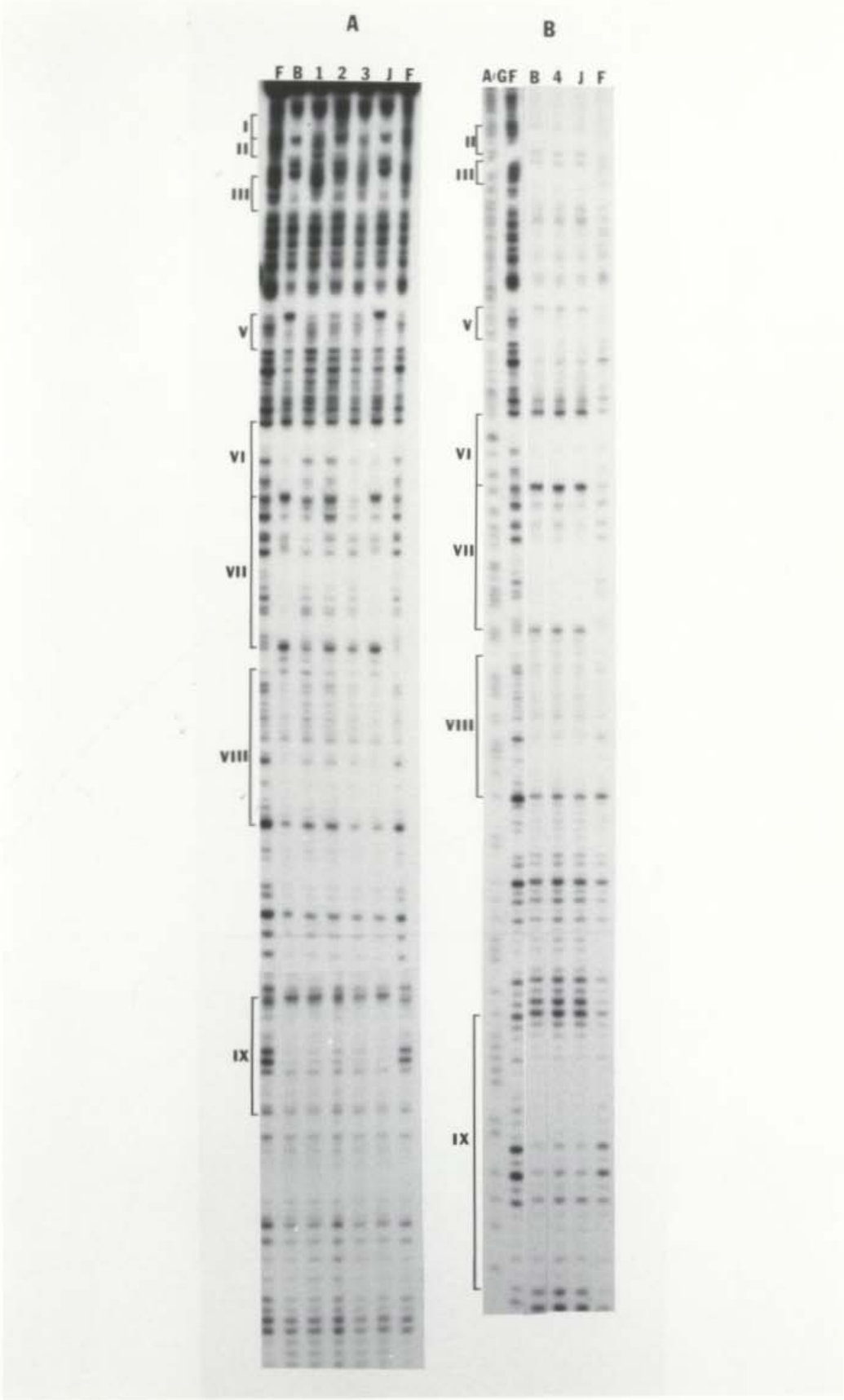
Regions[a]	Sequences[b]	Nucleotide numbers[c]	cell lines C33A	HeLa	SiHa	CaSki	143B
I	TATGGTTTAAACTTGTACGTTTCCTG	7515-7540	+	-	-	-	+
	TGATACCAAATTTGAACATGCAAAGGA	7513-7542	+	-	-	-	+/-
II	CCATGCGTGCCAAATCCCTGT	7545-7565	+	+	+	+	+
	GGTACGCACGGTT	7545-7557	+	+	+	+	+
III	CTGC ACTGCT TGCCAACCATTC(CATTGTTTTTAC)[d]	7575-7595	(+)[d]	+	+	+	+
	G TGACGA ACGGTTGGT	7578-7593	+	+	+	+	+
IV	CTGAATCACTATGTA	7630-7643	-	-	-	+/-	+
V	ACTATGCGCCAAC	7667-7679	+	+	+	+	+
VI	TTGGCTTGTTTTAAC	7710-7724	+	+	+	-	+
VII	(AACC)[e]TAATTGCATATTTGGCATAAGGTTT	7726-7754	+	+	+	(+)[e]	+
VIII	TCTAAGGCCAACTAAA	7760-7775	+	+	+	-	-
IX	GTAAAGGTTAGTCATACA	7803-7820	+	+	-	+	+

extracts. Although the coding strand did not show any protection of region I by HeLa and CaSki nuclear extracts, a partial protection of this region is evident in the non-coding strand (Fig. 3.6B). Closer observation of C33A and 143B footprints in Figure 3.2B reveals that regions I and II are separated by only a single nucleotide in the non-coding strand. Region VI is protected in all cell lines except CaSki, while region VIII is protected more completely in C33A (see also Fig. 3.7 for a darker B lane) and SiHa, poorly in HeLa and not detectably in CaSki and 143B. Region IX showed little or no protection in SiHa. The specificity of binding was confirmed by footprint competition experiments using unlabelled enhancer fragments of HPV 16. As presented in Figure 3.6A and 3.6B, homologous HPV 16 DNA competed for all protected regions, while control pUC-19 DNA failed to eliminate protection in regions II to IX. Protection for region I appears to be competed to a lesser extent by nonspecific competitor pUC-19 DNA. This could be due to homologous or related sequences present in pUC-19 DNA. Since this region shows cell-type variability, I feel that region I represents specific DNA-protein interactions. Enhancer elements of the other two HPVs, types 11 and 18, were also used in competition experiments. HPV 11 competed for all NF-1 motif containing regions, mainly in SiHa and CaSki. HPV 18 competed for all protected regions, although competition was less efficient than with HPV 16.

3.3.7 Oligonucleotide competition studies of DNaseI protected regions in the HPV 16 enhancer

Among the nine footprints for the HPV 16 enhancer, regions II, III, V, VI, VII and VIII contain the NF-1 binding consensus sequence 5' GCCAA 3' (Jones *et al.* 1988; Table 3.3). To study whether the same or related factors bind to these regions, I performed footprint competition experiments with synthetic oligonucleotides corresponding to region III, region VI without the AGCCAA sequence, and region VII. C33A extracts were used, as they gave the strongest DNaseI protection and gave least competition for NF-1-containing regions (Fig. 3.6). All three oligonucleotides competed mainly for regions containing GCCAA motifs that are protected by C33A extracts (regions II, III, V, VI, VII and VIII; Fig. 3.7A, lanes 1, 2 and 3). However, there were marked differences in the competition patterns for these three oligonucleotides. 1) Only the region VI oligonucleotide significantly affected the non NF-1-containing region I footprint. 2) Complete competition of the region II and III footprints by region III oligonucleotide exceeded that by region VII oligonucleotide, which in turn exceeded that by region VI oligonucleotide. 3) Complete competition of the region VII footprint by region III oligonucleotide compared with partial competition by the other two oligonucleotides. Competition experiments were also performed with an oligonucleotide corresponding to region IX (Fig. 3.7B, lane

Figure 3.7 Oligonucleotide competition of footprints for the HPV 16 enhancer. Lanes F are for DNase digestion in the absence of protein or competitors, lanes B are for samples not treated with competitor oligonucleotides and lanes J are for the control JCV oligonucleotide. Competitor oligonucleotides were region III in lane 1 (5'CTGCTTGCCAACCATT 3'), the TGTTTTAA sequence of region VI in lane 2, region VII in lane 3 (5'AACCTAATTGCATATTTGGCATAAGGTTT 3') and region IX in lane 4 (5'TAAAGGTTAGTCATACA 3'). In panels A and B, the cell extracts were C33A and HeLa, respectively. For all oligonucleotides, 100 ng was incubated with extracts before the addition of probes.


A
F B 1 2 3 J F
I
II
III
V
VI
VII
VIII
IX
B
A/GF B 4 J F
II
III
V
VI
VII
VIII
IX

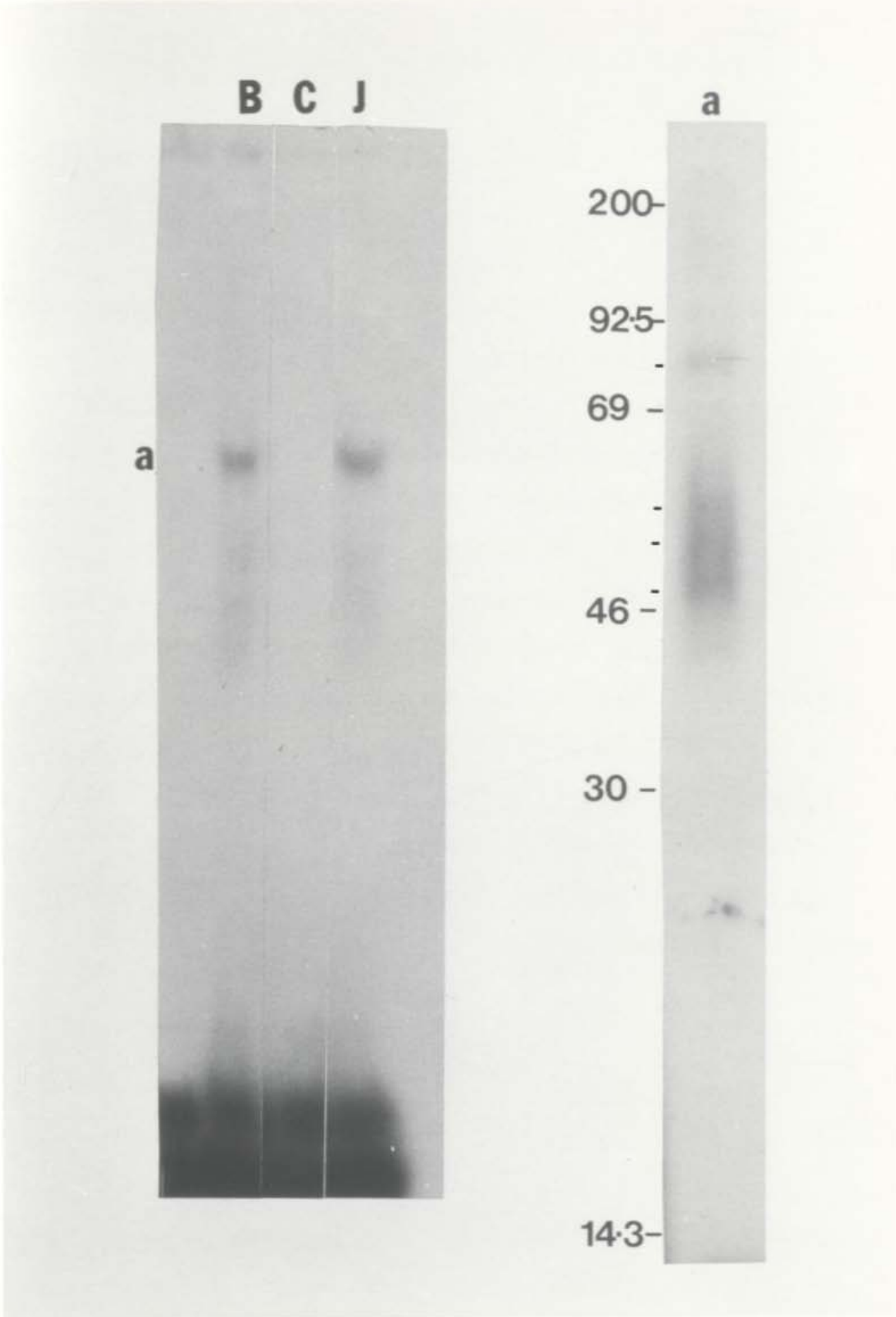
4), which does not contain the GCCAA motif. Weak competition was observed only for the homologous region (Fig. 3.7B, lane 4). In all oligonucleotide competition experiments, the control oligonucleotide, lane J, did not show any effect on footprints.

3.3.8 Characterization of HPV 16 region VI binding factors

To determine whether one or more proteins bind to GCCAA motif-containing regions, UV cross-linking assays were performed. The assay determines the molecular weight of proteins that interact with DNA and involves two steps: gel retardation and SDS-polyacrylamide gel electrophoresis. The probe used for gel retardations are usually nick translated DNA in which the thymidine residues of the DNA are replaced by 5'-bromodeoxyuridine. The probe is incubated with nuclear extracts, cross linked to protein by exposure to UV-light and subjected to electrophoresis on a native polyacrylamide gel. DNA-protein complexes detected in the native gel are excised, eluted and run on SDS polyacrylamide gels. In these gels, the DNA-protein complex runs according to the size of the protein interacting with DNA.

Gel retardation assay with the nick translated sequence TGTTTTAA of HPV 16 region VI motif as a probe yielded one DNA-protein complex with C33A nuclear extracts (Fig. 3.8, left panel). This DNA-protein complex was eliminated upon addition of homologous but not heterologous competitor DNA, thus con-

Figure 3.8 Gel retardation and UV cross-linking of the region VI oligonucleotide of HPV 16. The oligonucleotide (a 28mer containing two direct repeats of the TGTTTTAA sequence) was nick translated in the presence of [^{32}P]dCTP and 5′-bromodeoxyuridine and incubated with C33A extracts, UV-irradiated and then run by electrophoresis on native polyacrylamide gels (left panel). Incubation of probe with 8 μg of C33A extract in lane B, with extract plus unlabelled HPV competitor oligonucleotides in lane C and with extract plus the JCV nonspecific competitor oligonucleotide in lane J. Lane "a" is the gel shift band visualized in lane B but is subjected to SDS-polyacrylamide gel electrophoresis. The size of molecular weight markers is indicated.


B C J
a
a
200
92·5
69
46
30
14·3

firming the specificity of the DNA-protein interaction. The abundance and molecular weights of individual interacting proteins were obtained by electroelution of the DNA-protein complexes and subsequent analysis on an SDS-polyacrylamide gel (Fig. 3.8, lane a). Four major proteins of molecular weights 80, 59, 54 and 50 KD appear to interact with this sequence motif. My efforts to analyze the DNA-protein interaction on the other sequence motifs of HPV 16 by the same technique were unsuccessful. A possible requirement for both T and C residues of the DNA-protein contact point, to obtain both cross links and radioactive label, is one of the limitations of this technique. This might explain why other HPV 16 sequence motifs could not be used successfully.

3.3.9 *In vivo* competition for HPV 16 enhancer motifs

To test the functional relevance of some of the protections observed *in vitro*, I performed *in vivo* competition experiments (Fig. 3.9). The rationale for these experiments is the same as that of *in vitro* transcription-competition assays except for a few procedural differences. The assay involves cotransfection of cells with an enhancer-promoter-reporter gene plasmid and plasmids containing competitor oligonucleotides.

Cotransfection of a plasmid containing the sequences of the HPV 16 DNaseI protected region III decreased expression from the HPV 16 enhancer CAT construct by up to 40%. Region IX plasmid had a similar effect, while a greater effect was

observed for region VI plasmid. Thus, all three protected regions appear to bind transactivating protein factors. The same competitor plasmids had little or no effect on SV40 enhancer dependent pSV2cat expression, suggesting that few transcription factors are shared by the two enhancers.

3.3.10 Protein binding motifs within the constitutive enhancer of HPV 18.

The results of the HPV 18 enhancer footprinting experiments are shown in Figures 3.10A and 3.10B. Nucleotide sequences of the protected regions along with cell lines for which protection is observed are in Table 3.4. Protected regions III and IV had been previously identified by Garcia-Carranca *et al.* (1988) and designated as sites V and VI, respectively. In addition to the five protected regions, three hypersensitive sites were also observed.

Cell-type variability is evident among all protected regions. All five regions are protected by C33A extracts. For HeLa, regions I, II and III showed complete protection, while region IV showed only partial protection. No protection for region V was observed in this cell line. Regions I and II (partial in the non-coding strand) are protected in SiHa, while regions I (in non-coding strand), II (in non-coding strand), III and IV (partial in both strands) are protected in CaSki. Only one of the three hypersensitive sites are observed in CaSki.

Figure 3.9 Competition for HPV 16 enhancer dependent CAT expression in the C33A cell line. In the T-3 graph, 5 μg of HPV 16 CAT plasmid T-3 was cotransfected with 0, 5, 10 and 20 μg of competitor pUC-19 plasmids containing two directly repeated copies of DNaseI protected regions III (Δ), VI (O), and IX (□). One hundred percent is the activity of T-3 in the absence of any competitor. In the pSV2cat graph, 5 μg of pSV2cat was cotransfected with the same competitor plasmids as the control. The total quantity of DNA used in the transfections was adjusted to 25 μg by the addition of pUC-19 DNA. Results are averages of four and two experiments for T-3 and pSV2cat, respectively.

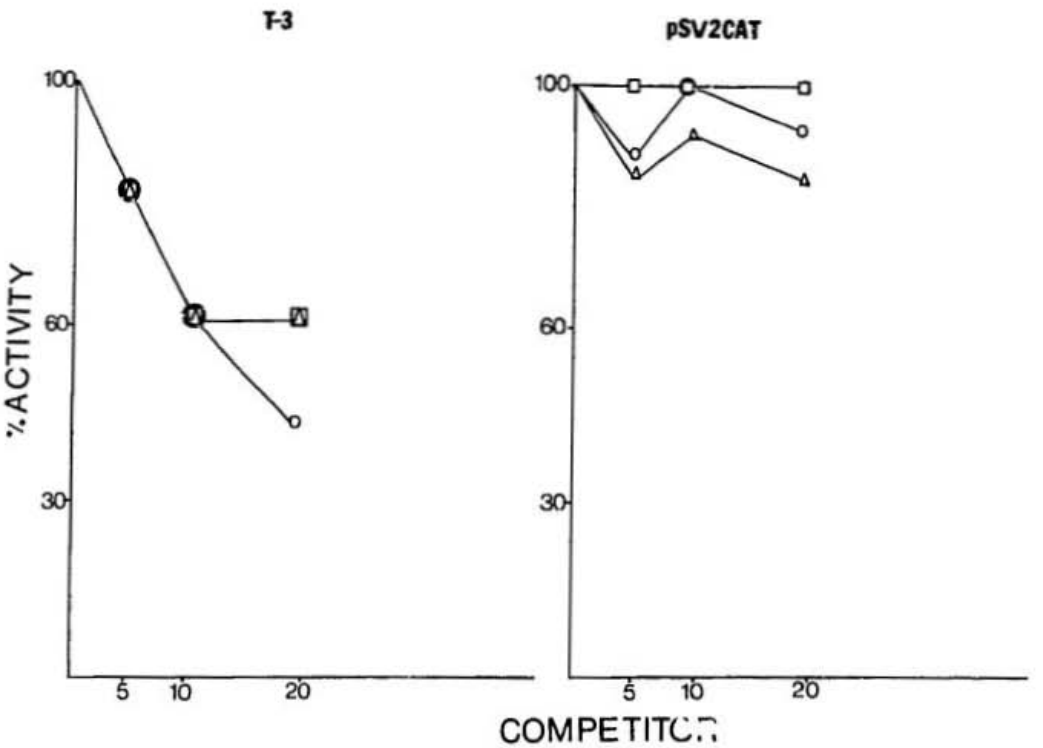

T-3
pSV2CAT
100
60
30
% ACTIVITY
5
10
20
COMPETITOR

Figure 3.10A DNaseI protection analysis of the HPV 18 constitutive enhancer. HPV 18 enhancer radiolabelled on the coding strand was incubated with nuclear extracts from A) C33A, B) HeLa, C) SiHa, D) CaSki and E) 143B cells and subjected to DNaseI digestion. Chemical cleavage of purines is in lane A/G. DNA treated with DNaseI in the absence and presence of nuclear extracts are in lanes F and B, respectively. Lanes 1 to 8 represent competition experiments in which nuclear extracts and probe were incubated with cold competitor DNA before DNaseI digestion. Competitors were HPV 11 DNA in lanes 1 and 2, HPV 16 DNA in lanes 3 and 4, HPV 18 DNA in lanes 5 and 6, and pUC-19 DNA in lanes 7 and 8. Odd numbered lanes had 10 ng and even numbered lanes had 50 ng of competitors. Protected regions are indicated by brackets labelled I to V. Brackets with dotted lines indicate regions of partial protection. Arrows indicate DNaseI hypersensitive sites.

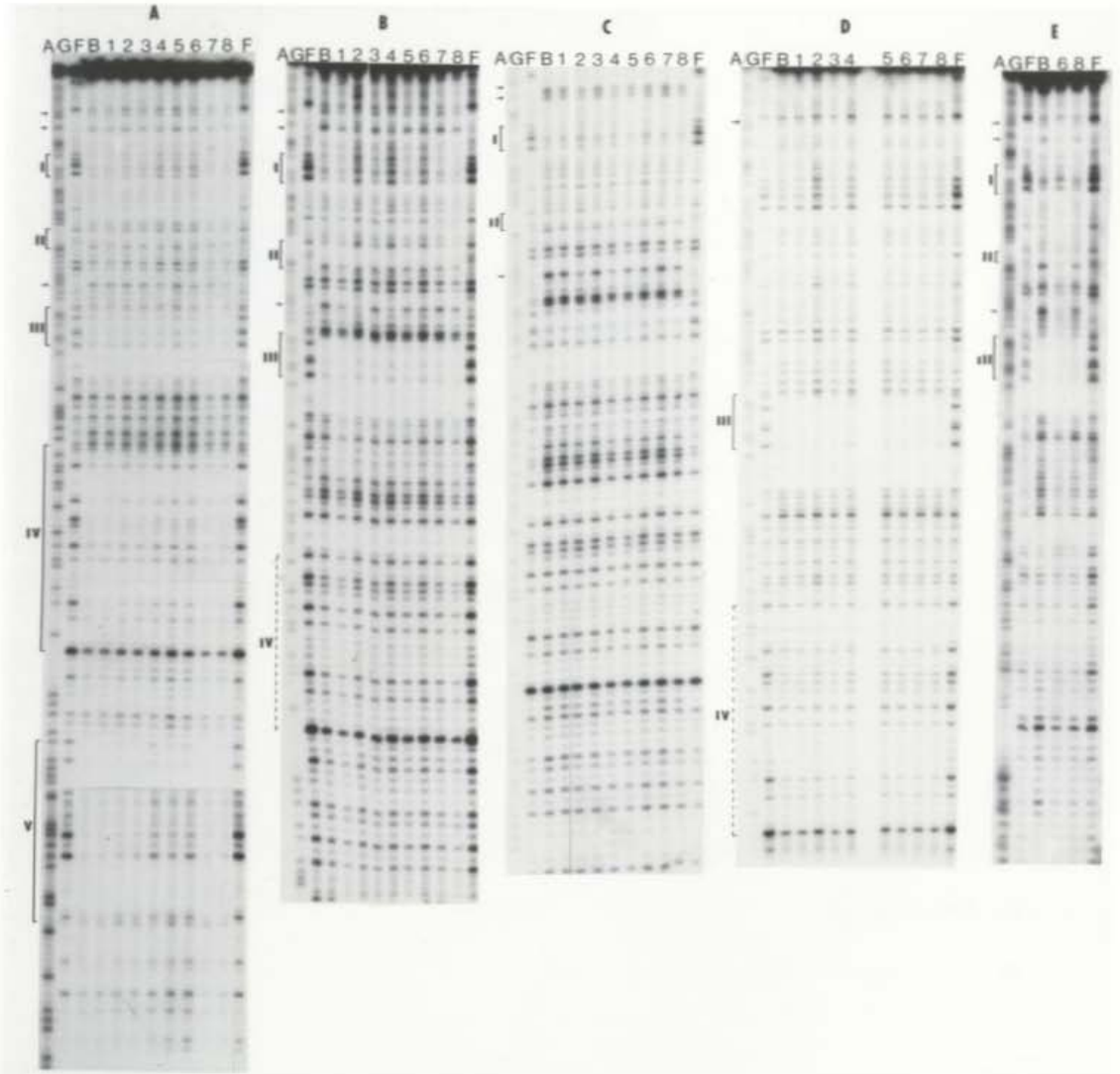

A
AGFB1 2 3 4 5 6 7 8 F
B
AGFB1 2 3 4 5 6 7 8 F
C
AGFB1 2 3 4 5 6 7 8 F
D
AGFB1 2 3 4 5 6 7 8 F
E
AGFB 6 8 F

Figure 3.10B Non-coding strand footprints of the HPV 18 enhancer. Probes were incubated with nuclear extracts from A) C33A, B) HeLa, C) SiHa, D) CaSki and E) 143B cells. Chemical cleavage of purines is in lane A/G. DNA treated with DNaseI in the absence and presence of nuclear extracts are in lanes F and B, respectively. Competition with HPV 18 DNA and pUC-19 DNA are in lane 6 and 8, respectively.

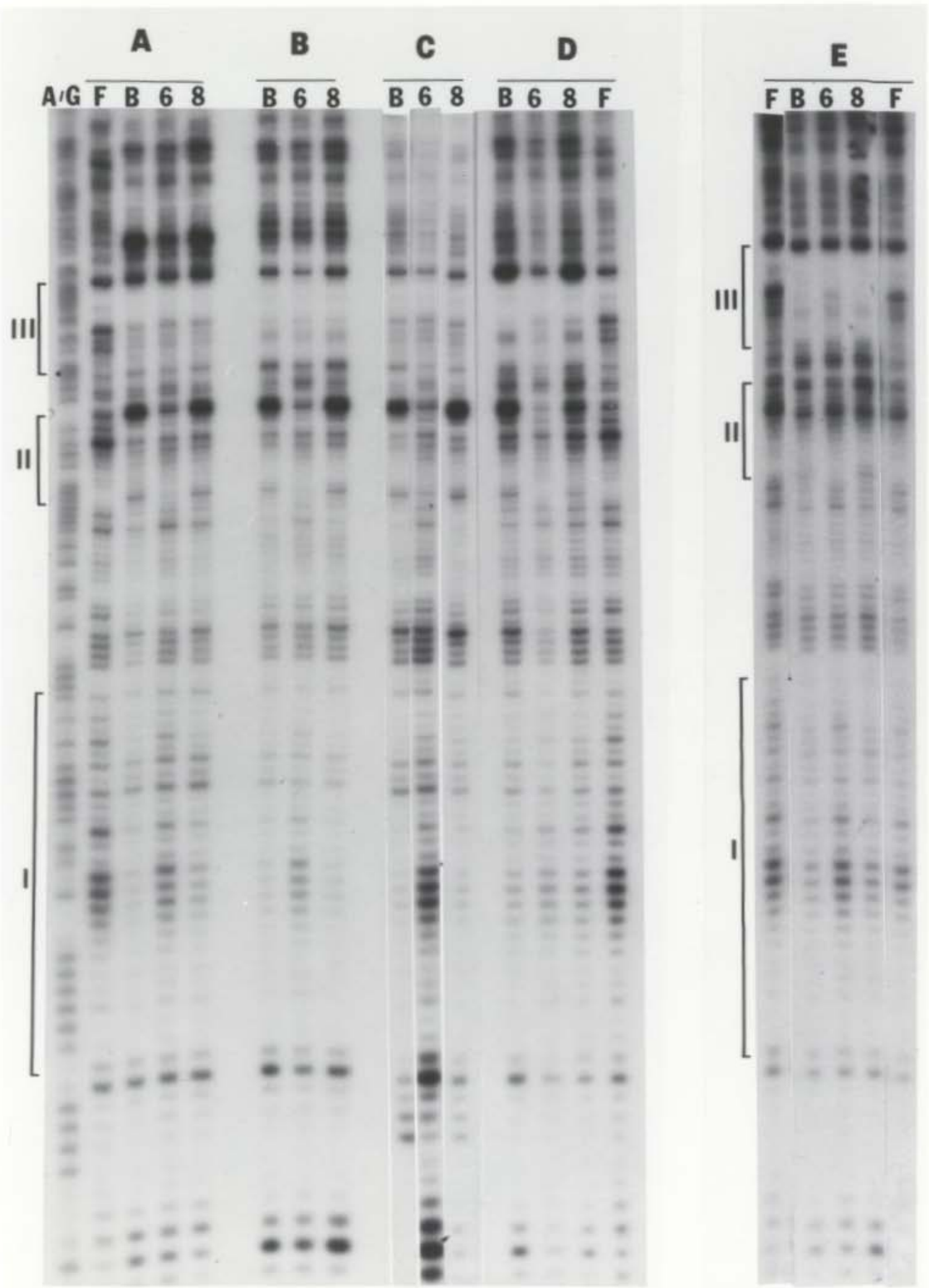

A
B
C
D
E
A/G F B 6 8
B 6 8
B 6 8
B 6 8 F
F B 6 8 F
III
II
I
III
II
I

Table 3.4 Sequences and locations of HPV 18 protected regions and cell lines for which protection is observed

Regions[a]	Sequences[b]	Nucleotide numbers[c]	cell lines C33A	HeLa	SiHa	CaSki	143B
I	CTTTTGGGC ACTGCT CCTACA	7536-7556	+	+	+	-	+
	ACCCG TGACGA GGATGTA	7540-7557	+	+	+	+/-	+
II	CCTCTTTGGC	7577-7586	+	+	+	-	+
	ACCGCGTATA	7583-7592	+	+	+	+	+
III	[TGGTAT]TAGTCATTTTC	7603-7619	+	+	-	+	+
	ACCATA ATCAGTA	7603-7615	+	+	+/-	+	+
IV	AACAATTGCTTGCATAACTATA	7637-7658	+	+	-	+/-	-
V	AATAAA ACTGCT TTTAGGCACATATTTTAGTT	7673-7704	+	-	-	-	-

[a] Numerals I to V are the same as the regions marked in Figure 3.10.

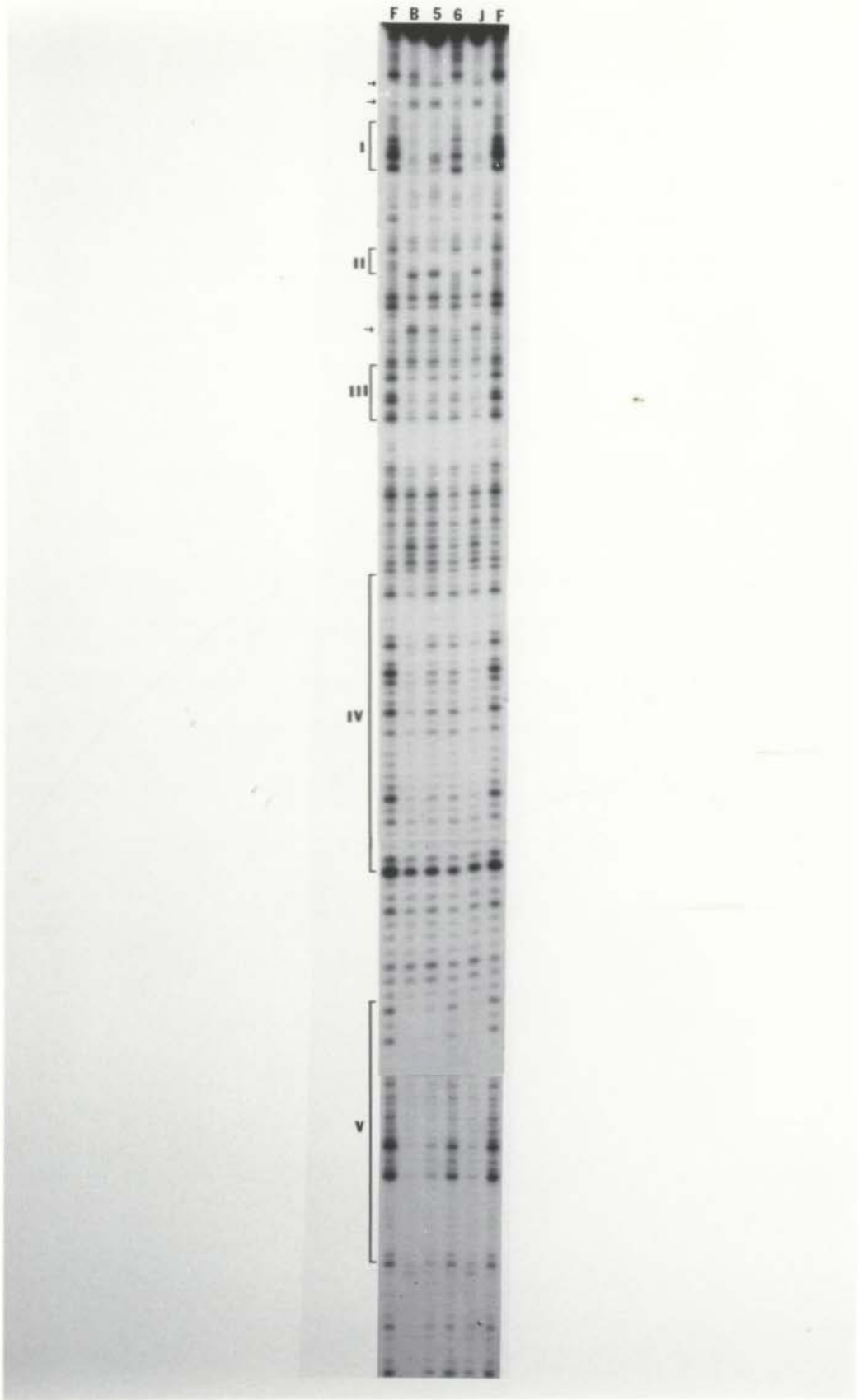
[b] NF1 GCCAA motifs are underlined. Broken lines represent putative AP1 binding motifs. Palindromic sequences are represented by thin arrows. An EF-II motif is overlined. Repeated sequences TTTTA are indicated by double lines. Sequences shared by regions I and V and HPV 16 region III are boxed. Sequences in region III of HPV 18 that are common in adenovirus E1A repressed enhancer elements are shown by squared brackets. Sequence ATTGC shared by HPV 11 and 16 is indicated by a wavy line.

Regions I, II and III are protected in 143B cell lines. The addition of unlabelled HPV 18 DNA to the binding reaction resulted in competition for all of the protected regions, although some of the competition was more evident with non-coding strand footprints, confirming the specificity of binding. Enhancer elements from the other two HPVs, types 11 and 16, were also used in competition experiments. The HPV 11 enhancer competed for region I, mainly in HeLa (Fig. 3.10A). Similarly, protections were eliminated upon addition of unlabelled HPV 16 enhancer into the binding reaction. Hypersensitive sites were also frequently lost or diminished by competition by the three HPVs. No competition was observed for the control pUC-19 DNA, as shown in lane 8, Fig. 3.10A.

3.3.11 Oligonucleotide competition studies of DNaseI protected regions of HPV 18

Oligonucleotide competition experiments of the HPV 18 enhancer are shown in Figure 3.11. Region IV oligonucleotide competed for the homologous region and poorly for region V. Interesting results were obtained with the oligonucleotide for region V. In addition to competing for its homologous region, this oligonucleotide competed effectively for all the protected regions and the three hypersensitive sites. These competitions were stronger than with the total enhancer fragment (Fig. 3.10). This effect of region V oligonucleotide could be due to sharing of some transcription factors by this motif with

Figure 3.11 Oligonucleotide competition of footprints for the HPV 18 enhancer. Lanes F are DNaseI digestion in the absence of protein or competitors, lane B is sample not treated with competitor oligonucleotides and lane J is for the control JCV oligonucleotide. Competitor oligonucleotides were of region IV in lane 5 (5'AACAATTGCTTGCATAACTATA 3') and region V in lane 6 (5'AATAAAACTGCTTTTAGGCACATATTTTAGTT). Competition was done in C33A extracts with 100 ng of indicated competitors.


F B 5 6 J F
I
II
III
IV
V

other regions since I detected the interaction of more than one protein in this region (see below).

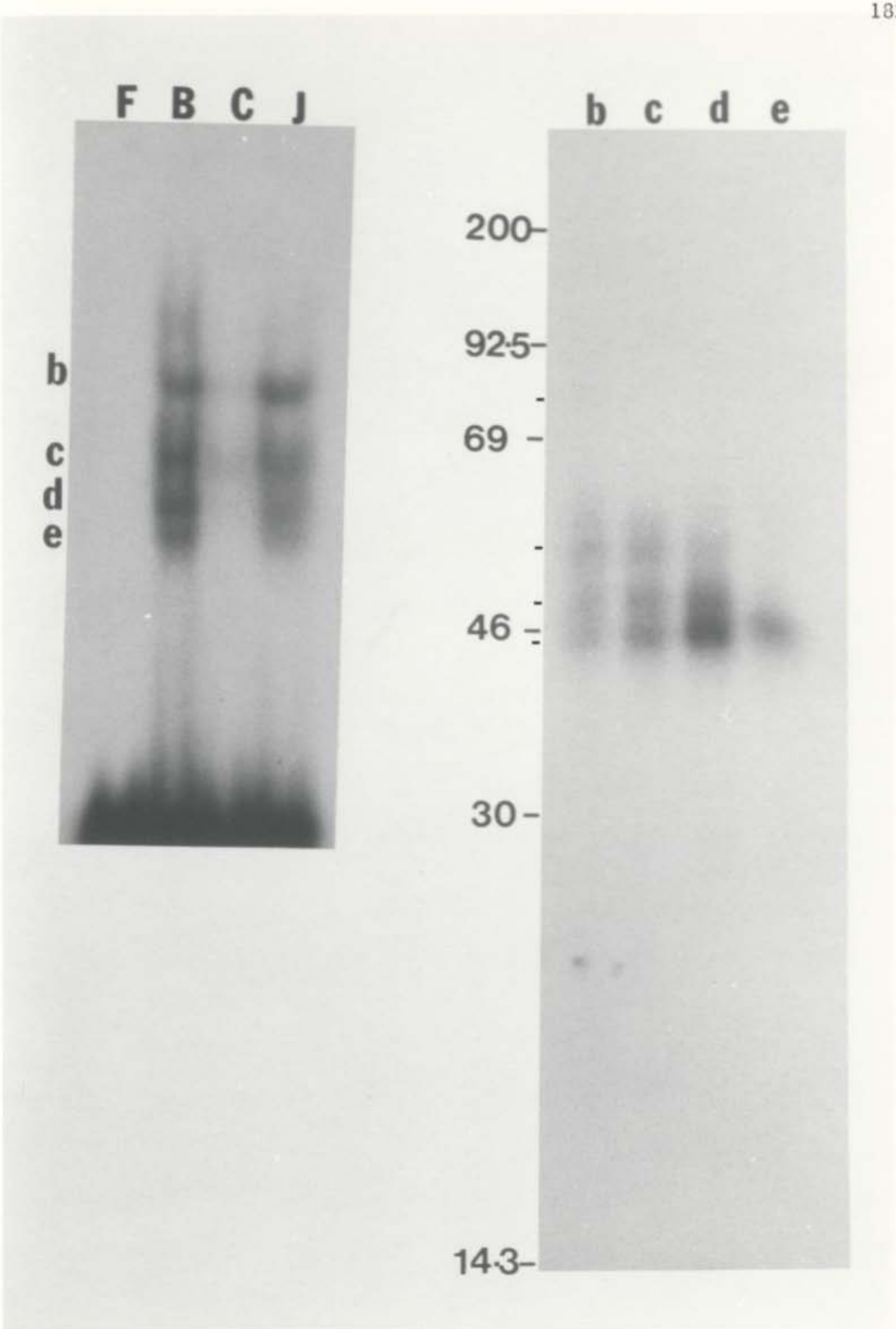
3.3.12 HPV 18 Region V binding factors

The results of UV cross-linking and gel retardation experiments with a monomeric oligonucleotide of region V of HPV 18 are presented in Figure 3.12. A total of four DNA-protein complexes which were competed by homologous competitor DNA shown in the left panel. DNA-protein complexes analyzed on the SDS gel indicate that three major proteins of molecular weights 54, 50 and 46 KD and a minor protein of 80 KD interact with this motif. It appears that for the four bands seen in the gel shift assay on the left panel the slowest migrating species represents a complex of four distinct proteins, the next slowest band is the same complex from which one particular protein has been dissociated, etc. Alternatively, the shift in the mobility could also arise from protein-protein interactions which are not detectable in this technique.

3.4 Discussion

The studies described in this chapter were undertaken to further the understanding of the regulation of HPV 11, 16 and 18 gene expression. The main objective of the study was to correlate activity of constitutive enhancers of these viruses in various cell types with *in vitro* DNA-protein interaction. By DNaseI footprinting, a definite correlation between *in vivo* activity and *in vitro* DNA-protein interaction was not observed

Figure 3.12 Gel retardation and UV cross-linking of region V oligonucleotide of HPV 18. The oligonucleotide was nick translated in the presence of $\alpha[^{32}P]$dCTP and 5′-bromodeoxyuridine and incubated with C33A extracts, UV-irradiated and then run by electrophoresis on native polyacrylamide gels (left panel). Incubation was without extract or competitor oligonucleotides in lane F, with 8 μg of C33A extract in lane B, with extract plus unlabelled HPV competitor oligonucleotides in lane C and with extract plus the JCV nonspecific competitor oligonucleotide in lane J. In lanes b to e, the respective gel shift bands visualized in lane B were subjected to SDS-polyacrylamide gel electrophoresis.


F B C J
b
c
d
e
b c d e
200
92·5
69
46
30
14·3

for any HPVs. However, additional studies involving gel retardation provide evidence for cell-type specific differences in HPV 11 and HPV 16 enhancer binding factors which correlate with enhancer activity.

3.4.1 Tissue specific expression of HPV 11 may, in part, be regulated by a repressor(s)

The constitutive enhancer of HPV 11 used in this study has been shown previously to be highly tissue specific (Chin *et al.* 1989; Marshall *et al.* 1989). My efforts to correlate the tissue specificity with *in vitro* DNA-protein interactions, as judged by DNaseI footprinting, were not successful. As shown in Figure 3.1, the pattern of DNaseI protection in epithelial and fibroblast cells are almost identical except for region IV. Although this region corresponds to a cell-type specific factor binding motif present in the epithelial cell specific enhancer CII of HPV 11 (Chin *et al.* 1989), there is no correlation between protection of this region and enhancer activity. (This region was not analyzed further as initial *in vivo* competition assays provided ambiguous results.) One possible explanation would be that the pattern of interaction seen *in vitro* may not actually occur *in vivo*. For example, Becker *et al.* (1987) have shown that the methylation of CpG dinucleotides in the regulatory elements of the tyrosine aminotransferase gene in non-expressing cells prevents the interaction of factors *in vivo* although the same region interacts with factors *in vitro*. A second possibility is that

the protection observed in fibroblasts is due to the leaky transcription of an epithelial specific gene, as reported recently in a different system (Sarkar *et al.* 1989). A third possibility is that the factors that confer tissue specificity could be piggyback transcription factors which interact with a ubiquitous factor. Interaction of such a piggyback transcription factor, MATα1, with DNA binding protein PRTF is known to activate a set of cell-type specific genes in yeast (Bender & Sprague, 1987). A fourth possibility is a role of the NF-1 GCCAA motif binding proteins in tissue specificity. As mentioned previously, analysis of NF-1 cDNAs indicated that several related proteins with highly conserved N-termini and variable C-termini exist in cells (Gil *et al.* 1988; Paonessa *et al.* 1988; Santoro *et al.* 1988). Lastly, quantitative differences in factors may contribute to the cell-type specificity. The results of gel retardation assays, *in vitro* transcription-competition and deletion analysis provide some evidence for the last possibility (see below).

Gel retardation competition assays performed with protected regions II and VI oligonucleotides as probes indicated i) qualitative difference in factors binding to region II and VI, ii) quantitative differences in factors among cell types. These differences are: 1) the number of DNA-protein complexes were two for region II whereas region VI produced only one complex. 2) Factors binding to region II but not region VI are extremely sensitive to repeated freeze thawing of the

nuclear extract. Since NF-1 has been shown to be resistant to repeated freeze thawing, at least with the procedures of nuclear extract preparation employed in this study (Hennighausen and Lubon, 1987), it is likely that region II binds to additional factors that are sensitive to freeze thawing. 3) As mentioned earlier, CaSki cells contained the least amount of region II and region VI binding proteins.

The differences in gel retarded bands for region II and region V oligonucleotides could be due to divergent sequences within or outside the NF-1 motifs. For example, the region VI oligonucleotide contains sequences that can accommodate two molecules of NF-1/CTF (TGG(N)$_7$CCA; Gil *et al*. 1988). In contrast, region II contains both TGG and TGGCA motifs which can interact with NF-1/CTF and TGGCA-binding factors (Table 3.2). Furthermore, sequences other than GCCAA in region II bear significant homology to the enhancer motifs of Moloney murine leukaemia virus (Mo-MuLV) and Adenovirus E1A (Speck and Baltimore, 1987; Herbst *et al*. 1990). The sequence with homology to Mo-MuLV, 5′ CCTGC 3′, which binds factor LVc and overlaps the GCCAA motif in region II (Table 3.2). The second motif, 5′ CAGGGCAATA 3′ (noncoding strand) has significant homology (underlined region) to the ØAP3 binding motif (5′ TGTGGCAAA 3′) of the adenovirus E1A enhancer. Interestingly, ØAP3 is a transcriptional repressor with three to five fold inhibitory activity and is present at a 10 to 20 fold higher concentration in undifferentiated rodent fetal fibroblasts

than in HeLa cells (Herbst *et al.* 1990). This observation parallels that of my *in vitro* transcription competition, deletion analysis and gel retardation results in that the factors binding to region II possess repressor function and are present at a higher level in the fibroblast cell line, 143B. Based on this, I propose that tissue-specific expression of HPV 11 is regulated, in part, by the balance between positive and negative factor(s). For example, inactivity of HPV 11 in 143B cells could be due to a higher level of negatively acting factors. Similar mechanisms have been proposed for other papovaviruses such as JCV (Tada *et al.* 1989). Inactivity in CaSki cells can be explained by relatively low levels of positive as well as negative factors.

At this stage I can not propose any specific mechanism for negative regulation. Among several models proposed by others, "quenching" (see Chapter 1.1.5) may be ideal for HPV 11 (Levine and Manley, 1989). Further experiments involving site-directed mutagenesis, affinity purification and recombinant *in vitro* transcription are required to elucidate the mechanism of repression.

Oligonucleotide competition of DNaseI footprints and gel retardation provided conflicting results, particularly with the region II oligonucleotide. Competition by the region II oligonucleotide with a mutation of the GCCAA motif for region II protein interactions was equivalent to the wild type oligonucleotide in DNaseI footprints but not in gel

retardations. One possibility is that the mutant oligonucleotide retains the ability to interact with a few of the factors that interact with region II, possibly with lower affinity. This interaction is sufficient to partially disrupt protein-DNA complexes at region II of the HPV 11 enhancer and expose the DNA to DNaseI. It is relevant to point out that the mutant oligonucleotide (TATTGCCCTGGGAA) contains LV-c homology , oAP3 homology and TGG sequences for CTF/NF-1 and differs from the wild type by not possessing the TGGCA-motif. This can also explain the marginal effect of the mutant oligonucleotide in in vitro transcription competition assays. Additional experiments with saturation mutagenesis of sequences of region II in the HPV 11 enhancer are required to study the nature of factors involved.

3.4.2 Ubiquitous and cell-type specific factors interact with the HPV 16 enhancer

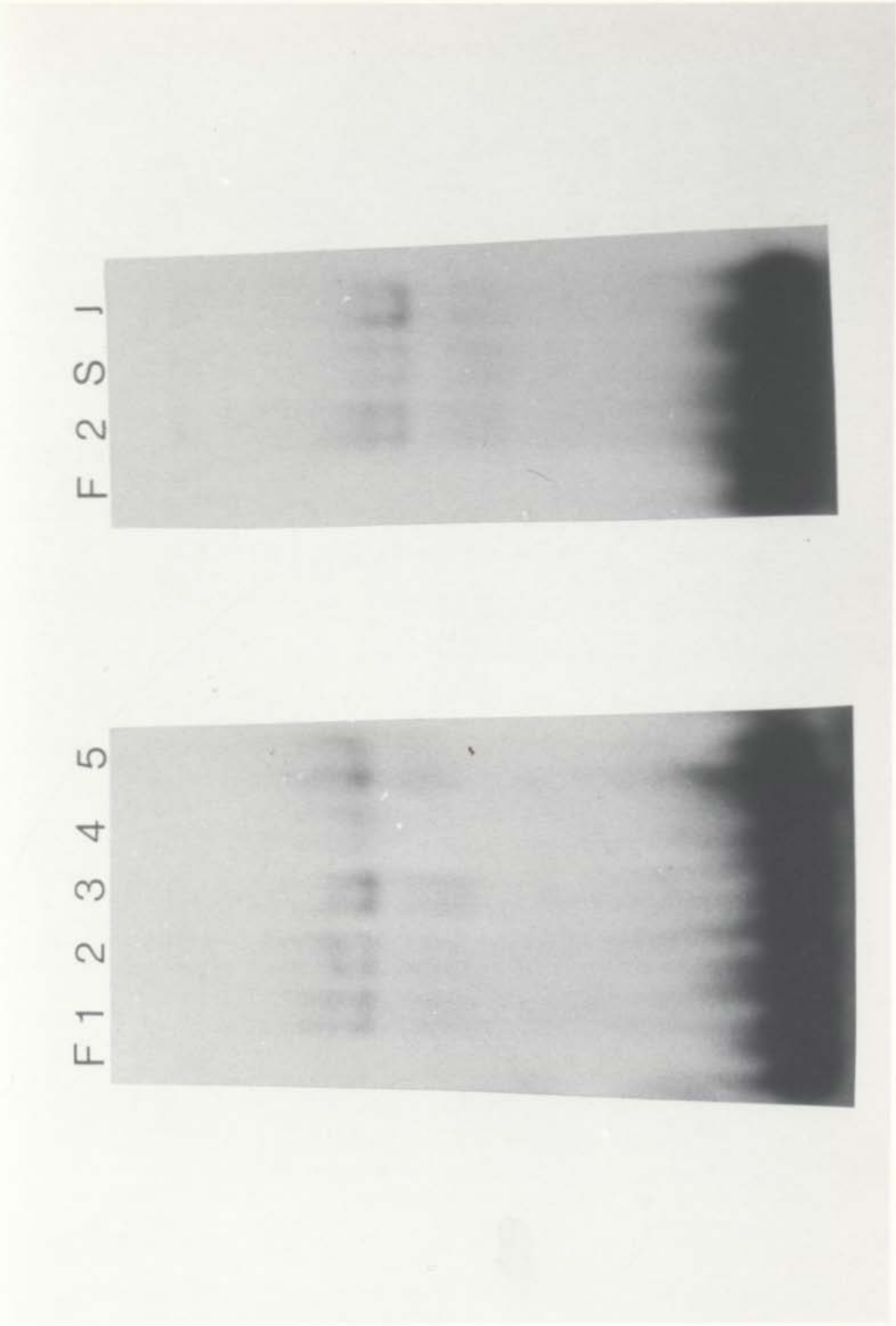
The HPV 16 enhancer fragment contained nine DNaseI protected regions, of which six contained a nuclear NF-1 binding motif. Among them, three regions (II, III and VII) contained TGGCA protein-binding motifs. There are three protected regions which do not contain NF-1 binding sequences. Among them, region I does not show homology to any known transcription factor motif. Region IV, containing the sequence, TGAATCA, which differs from the transcription factor AP1 binding motif, TGAGTCA, by one nucleotide, showed complete protection only for 143B, although others have obtained

protection for this same region with HeLa cell extracts (Gloss et al., 1989a). Region IX, which contains a consensus sequence TTAGTCA for transcription factor AP1, acts as positive modulator of HPV 16 enhancer function in C33A cells (Fig3.9).

Regions II, III and VII, although protected in all cell types tested, show striking differences in the degree and extension of protection. These differences could be attributable to both quantitative and qualitative differences in factors which bind DNA directly and/or influence the binding of other factors. In support of this, gel retardation assays using region VII oligonucleotide show two bands for C33A, HeLa and SiHa, but only one band for CaSki and 143B nuclear extracts (Fig. 3.13). This could explain why CaSki region VII protection is readily competed by low levels of HPV DNA (Fig. 3.6A). Similarly, region III shows extended protection and weakest competition for C33A extracts, suggesting that C33A contains more interacting factor(s) than CaSki or 143B cells. The strength of the protection/competition for region VI of HPV 16 suggests that the levels of factors in cells can be ranked C33A, greater than HeLa, greater than SiHa and 143B greater than CaSki. Some of these differences may contribute to the cell-type specific expression.

An interesting feature of region VI is that it contains two distinct motifs: GCCAA and GTTTTAA. As well as containing the TTTTA sequence repeat of the HPV 16 region III and the HPV 18 region V (Table 3.3 and 3.4), the second motif is identical

Figure 3.13 Gel retardation using region VII oligonucleotide of HPV 16 as a probe. Four μg of nuclear extracts from C33A (lane 1), HeLa (lane 2), SiHa (lane 3), CaSki (lane 4) and 143B (lane 5) were incubated with endlabelled region VII oligonucleotide and run on 4% polyacrylamide gels. Probe without any protein is in lane F. The right panel is a competition experiment in which 100 fold excess of either unlabelled specific oligonucleotide (lane S) or non-specific oligonucleotide (lane J) was incubated along with probe and HeLa cell nuclear extract. Competition by the specific oligonucleotide was not complete possibly due to the presence of excess of factors.


F 2 S J
F 1 2 3 4 5

to a protected region of the cell-type specific enhancer element in the c-*mos* oncogene (van der Hoorn, 1987). UV cross-linking experiments have indicated the interaction of four different proteins with this region (Fig. 3.8) and *in vivo* competition experiments suggest that DNA-protein interactions in this region function as a positive modulator of enhancer function. (Fig. 3.9). This region overlaps the sequence TTTGGCTTT found in enhancer motifs of cytokeratin genes (Blessing *et al.* 1987) which has been thought to have a role in cell type-specific gene expression (Blessing *et al.* 1987; Cripe *et al.* 1987). Although epithelial and fibroblast cell lines produced identical footprints in this region, it is possible that factors binding to this motif are functionally different in different cell types. Since more than one factor binds to this motif (Fig. 3.8), there may be analogy to the octomer motif of SV40 and immunoglobulin enhancers in which two factors can bind the same element, with only one conferring cell-type-specific function (Davidson *et al.* 1986; Gerster *et al.* 1987; Schaffner, 1989; Wirth *et al.* 1987).

Oligonucleotides corresponding to three protected regions of HPV 16 containing NF-1 binding sites were tested for their ability to compete for enhancer binding factors. All three oligonucleotides competed for all six regions containing the GCCAA motif (Fig. 3.7). Similar results have been obtained by others using oligonucleotides corresponding to region III and the adenovirus NF-1 binding site (Gloss *et al.* 1989a,

1989b). However, by using different oligonucleotides, I have been able to find marked differential effects. The competitions observed were quantitatively different for the oligonucleotides tested and regions competed. While oligonucleotides for protected regions III, VI and VII have comparable effects on the protected regions V and VII, regions I, II, III and VI are differentially affected. Only the oligonucleotide for region VI significantly affects region I protection. Competition for regions II and III is greatest by the oligonucleotide for region III, intermediate by the oligonucleotide for region VII, and least by the oligonucleotide for region VI. Competition for region VI was observed only for oligonucleotides for regions III and VI. The differential effects of the oligonucleotides might be explained by one or more of the following mechanisms. 1) Binding of sequence specific factors to sequences adjoining NF-I motif differentially influences binding of common, GCCAA-specific factors. 2) The presence of the TGGCA motifs in only a few of the NF-1 motif containing regions. 3) The factor(s) binding to the GCCAA motif interact cooperatively with factors binding to adjoining regions. The latter mechanism is supported by the observation by Gloss *et al.* (1989b) that for regions IV, V and VI, with increased factor levels, protection on the non-coding strand is fused into one long protected stretch. Also, competition of the GCCAA motif-containing regions by the region VI specific oligonucleotide without its

GCCAA motif could be due to cooperative interaction of some factor(s) other than NF-1, consistent with my UV cross-linking studies which indicate the interaction of at least four proteins.

3.4.3 HPV 18 enhancer binding factors are mostly cell-type specific

The constitutive enhancer of HPV 18 contained five protected regions (Fig. 3.10). Region I, which is protected in four of the five cell types tested, contains the sequence GCACTGCTCC, which is identical to the central portion of the GAL4 binding site present in the GAL10 gene of yeast (Wingender, 1988) and harbours an interesting ACTGCT sequence discussed below. Extension of this protected region to a hypersensitive site (central arrow in Fig. 3.10A) would include a sequence, TGCCCAA, which differs from a GCCAA motif by only one nucleotide. Furthermore, protection of this region is competed by the enhancer elements of HPV 11 and 16 which contain NF-1 binding motifs. Region II contains the NF-1 consensus motif, GCCAA. The sequence motif, TTAGTCA, present in the protected region III of HPV 18, is also a consensus sequence for AP1. The sequence protected in region III also has homology to a consensus sequence, GTGGTATG, present in the enhancer elements repressed by the adenovirus E1A gene product. This could be important, since Swift *et al.* (1987) and Thierry

et al. (1987) have shown that E1A gene products of adenovirus can repress the HPV 18 enhancer function.

Region IV of HPV 18 contains the sequence, TTATGCA, which is homologous to the EF-II binding motif of Rous Sarcoma virus LTR (Sealey and Chalkey, 1987). In addition, this region has homology to the responsive element for the TIF (trans-inducing factor) α gene of herpes virus type I (Kristie and Roizman, 1988). Gius and Lamins (1989) have shown that TIFα can activate HPV 18 enhancer dependent expression in cells of epithelial origin and suggested that the effect could be mediated through sequences overlapping region IV. An interesting stretch of sequences could be ATGCTAAT which is identical to the homolog of octomer element present in TIFα responsive genes. Recent studies have suggested that Oct-1 factor along with TIFα and two other unknown cellular proteins form a multiprotein-DNA complex over these sequences and stimulate transcription (Kristie *et al.* 1989). It is thus essential to determine the importance of these sequences and their interacting proteins in TIFα mediated transactivation of the HPV 18 enhancer. Additionally, a recent report by Garcia-Carranca *et al.* (1988) places HPV 18 region IV within the enhancer region responsible for tissue specific expression. Interestingly, this region is not protected in the fibroblast cell line 143B. Therefore, it will be of interest to mutate these sequences to assess their functional significance.

Attempts to characterize proteins binding to this region by UV cross-linking were unsuccessful.

UV cross-linking studies with oligonucleotides corresponding to region V, as observed for C33A cell extracts, indicate interactions with at least four proteins (Fig. 3.12). This oligonucleotide competed very effectively for all protected regions to a degree greater than that of the complete HPV 18 enhancer fragment (Fig. 3.10 and 3.11). A unique feature of region V is the inversely repeated sequence TTTTA, which is also present within protected regions III and VI of HPV 16. It is possible that some of the proteins interacting with these motifs may be the same. The 50 and 54 KD proteins identified in our UV-cross linking experiments are the likely candidates. The reason for general protection of the TTTTA motif-containing region VI of HPV 16 and C33A-specific protection for HPV 18 region V (while extension of protection is observed for HPV 16 region III) could involve other cell type specific proteins (possibly the 46 KD protein) and/or a higher level of protein found only in C33A. The latter possibility is strengthened by results in one experiment in which I have observed a partial protection of region V in 143B extracts which was competed by a specific as well as a non-specific competitor. Additionally, the TTTTA and the adjacent ACTGCT sequences in HPV 18 region V, show interesting correlations. 1) Neither sequence is present in the enhancer motifs of non-oncogenic HPV 11 (Table 3.2). 2) The region V oligonucleotide of HPV 18,

containing inverse copies of the TTTTA sequence flanking the ACTGCT sequence, competed very effectively with the region I protected site which has homology only to the ACTGCT sequence of region V. An interesting possibility is that the mechanism by which the factor(s) binding HPV 18 region V function(s) is by interacting with all the other sites or the factors bound to them. Since the *in vivo* oligonucleotide competition assays, similar to HPV 16, gave inconsistent results, mutational analyses are required to assess the functional significance of these individual motifs.

3.4.4 NF-1 motifs and ATTGC motifs are common motifs of HPV 11, 16 and 18

Four protected regions of HPV 11, six regions of HPV 16 and one of HPV 18 contained the NF-1 GCCAA motif. Therefore, NF-1 motif binding factors could play an important role in the control of HPV expression. The NF-1 motifs may not only contribute to constitutive expression but also confer inducibility. This is supported by the observation that transforming growth factor β (TGFβ) induced expression of mouse α2 collagen I gene (Rossi *et al*. 1988), and v-src and serum induced expression from Rouse Sarcoma Virus (RSV) long terminal repeat (LTR) (Dutta *et al*. 1990) is mediated through factor(s) binding to the GCCAA motif or related CCAAT motifs.

Several studies have demonstrated that TGFβ_1 inhibits keratinocyte growth and this response is lost in many, but not

all, neoplastically transformed epithelial cell lines (Silberstein and Daniel, 1987; Masui *et al.* 1987). More recently, it was suggested that TGFβ_1 represses transcription of the c-*myc* gene and this reduction of c-*myc* expression plays an integral role in TGFβ_1 mediated inhibition of keratinocyte growth (Pietenpol *et al.* 1990a). Furthermore, it was shown that transforming proteins of DNA tumor viruses, including the E7 protein of HPV 16, somehow block the TGFβ_1 mediated c-*myc* repression thus enforcing continuous cell proliferation (Pietenpol *et al.* 1990b). Since transcription of the E7 gene of HPV 16 is dependent upon the enhancer described above, it is possible that TGFβ_1 may in fact aid in proliferation of transformed cells by inducing E7 proteins through NF-1 motifs.

The other common motif found in all three HPVs is ATTGC. This sequence binds to a rat liver protein CBP more efficiently than the prototype CCAAT box sequence ATTGG (Graves *et al.* 1986). However, for unknown reasons, CBP bound to ATTGC sequences represses transcription from an adjoining promoter, an effect opposite to that of CBP bound to a prototype CCAAT box. Expression of CBP is tissue restricted. In the mouse, CBP is found most abundantly in fat, liver, and lung tissue. Beyond being limited in tissue distribution, CBP is restricted further to terminally differentiated cells (Birkenmeier *et al.* 1989). Additionally, multiple species of CCAAT box binding proteins, some of which are tissue specific, have been identified (Chodosh *et al.* 1988; Cohen *et al.* 1986; Dorn *et*

al. 1987; Johnson *et al*. 1987; Raymondjean *et al*. 1988). There is strong analogy with HPV in this regard which makes speculations very attractive. The analogies are that HPVs are epithelial tissue specific for expression and also demonstrate a remarkable response to differentiation. Certainly experiments testing the effect of ATTGC-containing oligonucleotides on *in vitro* and *in vivo* transcription would be of interest. Whether this factor has a role in tissue specific gene expression of HPVs requires further experiments involving site directed mutagenesis.

3.4.5 Possible role of enhancers in HPV late gene expression

As mentioned previously, one of the major interests in papillomavirus research is the viral life cycle which requires terminally differentiated keratinocytes. However, none of the papillomaviruses have yet been successfully propagated in cell culture to yield virus particles. Although virions derived from planter warts can infect cultured keratinocytes, the viral life cycle ends at the stage of viral DNA replication (Shah and Howely, 1990). Some success has been achieved in identifying HPV region specific transcripts in differentiating epithelial cells of genital warts (Stoler *et al*. 1989). In these warts, abundant E4 and E5 open reading frame messages and a lower level of E6-E7 messages were observed in undifferentiated basal cells of the epithelium. These messages

derived from the early region of the virus were transcribed before the onset of vegetative DNA replication and continued to be expressed in increasing amounts in the maturing epithelium. The late messages encoding capsid proteins were truly late in that they appear concomitant with or after the onset of vegetative viral DNA replication and were present in the superficial strata of the epithelium, which contain the oldest and most differentiated keratinocytes. Even in these cells, early region messages are far more abundant than late region messages (Stoler *et al*. 1989). It is proposed that alternative promoters, splice sites and polyadenylation sites are utilized to achieve this complex pattern of mRNA production.

Six different promoters that are active in transformed cells and one promoter that is active only in productively infected keratinocytes have been identified in the BPV1 genome (Howley, 1990). Similar promoters are believed to exist in HPVs and at least in HPV 11 and 6b, two promoters have been described (Smotkin *et al*. 1989). The activation of these promoters during differentiation of keratinocytes may involve one or more of the following mechanisms. The first is by "transcriptional interference". Control of transcription by a transcriptional interference mechanism involves a switch from one promoter to another, possibly due to activation of a specific set of factors, and subsequent repression of the former promoter by interference in the assembly of the

preinitiation complex (Corbin and Maniatis, 1989). However, this is less likely or restricted to only a few promoters as the progression of the life cycle in HPV 11 does not involve "shut off" of early transcripts and subsequent emergence of late transcripts (Stoler *et al.* 1989). In fact, early transcript levels continues to increase during the progression of the life cycle. The second possibility is that, as differentiation of keratinocytes progresses, there could be an activation of promoter specific factors and/or inactivation of repressors that maintain promoters in an inactive state. The third possibility is that specific enhancer binding factors are induced as the differentiation progresses. Since transcription from all promoters is believed to be activated by a single enhancer (described in this study), it is possible that the increase in the number of active promoters results in a simultaneous increase in the activating ability of the enhancers. This process may be remotely similar to activation of the bacteriophage T4 late promoter described recently (Herendeen *et al.* 1990). The T4 late promoter is activated only in the presence of T4 DNA polymerase accessory proteins bound to a mobile enhancer and a T4 encoded protein which acts as a communicator between RNA polymerase and enhancer bound accessory proteins. In analogy with some stage of the HPV life cycles, both late and early promoters of T4 are transcribed simultaneously. However, only those RNA polymerases that are bound to σ factor, a prokaryotic initiation factor, are compet-

ent to transcribe early and middle genes whereas RNA polymerase bound to the communicator T4 encoded protein can activate late gene transcription. It is possible that during certain stages of the HPV life cycle, transcription factors bound to enhancer motifs described in this study interact with late gene promoters of the virus through a differentiated cell type specific factor and virus encoded communicator protein. NF-1 could be one of the transcription factors involved in this process as NF-1 has already been shown to play a significant role in the adenovirus life cycle (Jones *et al.* 1987). However, this hypothesis is testable only after developing systems that would allow isolation of enough cells at different stages of differentiation. Once that is achieved, *in vitro* transcription assays with late gene promoters as templates is an ideal approach to test the above hypothesis. Also, the same assay system can be supplemented with viral early proteins obtained from either infected cells or *in vitro* transcription-translation to elucidate the role of viral early proteins in late gene transcription.

3.4.6 Possible role of enhancers in determining the oncogenicity of HPVs

All three HPVs used in my study are associated with carcinogenesis. Tumors, which develop after a long latency period, are generally observed in epithelial cells that are permissive for the viral life cycle. Frequently, synergism

between papillomavirus infection and chemical or physical carcinogens have been observed in carcinogenesis. In premalignant lesions, viral DNA is predominantly in an extrachromosomal state and progression to malignancy involves integration of viral DNA, possibly triggered by cofactors, which invariably result in disruption of the E1 or E2 open reading frames (Baker *et al.* 1987; Matsukura *et al.* 1986; Schwarz *et al.* 1985; Pater and Pater, 1985). However, as mentioned previously, the type of lesions induced after infection varies among virus types. HPV 11 is associated with benign proliferative lesions, including condyloma accuminata which rarely progresses into malignant tumors whereas HPV 16 and 18 induce malignant tumors of the genital tract. The site of infection also varies among viruses; HPV 11 has been shown to infect the respiratory tract and the genital tract whereas HPV 16 and 18 can only infect the genital tract. The results of primary epithelial cell transformation assays mimic the clinical observations with regard to the type of lesions induced by different HPV types in that the DNA from HPV 16 and 18 but not HPV 11 can cooperate with an activated *ras* oncogene in transformation (Schlegel *et al.* 1988; Pater *et al.* 1988a; Storey *et al.* 1988; Pecoraro *et al.* 1989). It is proposed that differences in viral regulatory elements as well as coding regions account for the difference in oncogenic potential.

Suggestions for the role of regulatory elements in determining oncogenic potential come from a recent study which

indicated that non-oncogenic HPV 11 and HPV 6b but not oncogenic HPVs 16 and 18 contain a promoter within the E6 open reading frame (Smotkin *et al*. 1989). It was suggested that this additional promoter generates transcripts that encode only the E7 protein at the expense of E6-E7 transcripts. The potential impact of this differential promoter utilization is that rare E6-E7 transcripts may not synthesize the threshold level of E6 protein required for transformation (see below). Also, it is possible that the E7 transcript is translated poorly. Since E6 and E7 proteins are associated with cellular transformation (see below), it is likely that the HPV 11 enhancer, although stronger than the HPV 16 or 18 enhancers (Table 3.1), fails to maintain an appropriate level of E6 and E7 proteins. This appears to be the case as in rare HPV 11 associated cancers, the untranslated regulatory region appears to differ from those of HPV 11 isolated from condylomata and possess stronger enhancer activity than the wild type (Pater *et al*. 1988b; Smotkin *et al*. 1989; Byrne *et al*. 1987).

The E7 proteins of HPVs 16 and 18 have been shown to be required for cellular transformation (Phelps *et al*. 1988; Storey *et al*. 1988; Matlashewski *et al*. 1987; Bedell *et al*. 1987; Pirisi *et al*. 1987; Pater *et al*. 1989). The presence of proteins from the E6 ORF enhance the transformation frequency. One of the important biological activities of E7 is its ability to interact with the products of the retinoblastoma susceptibility gene (Rb), a tumor suppressor

gene (Dyson *et al.* 1989). The exact functions of tumor suppressor gene proteins are not well understood although they may be involved in the negative regulation of cell growth and/or differentiation (reviewed in Klein, 1987). It is believed that interaction of E7 with Rb abrogates the negative regulatory effect of Rb. Comparison of the Rb binding ability of E7 proteins from HPV 16 and 18 with HPV 11 indicated that the E7 proteins of HPV 16 and 18 bind Rb more efficiently than HPV 11 E7 (Munger *et al.* 1989; Gage *et al.* 1990). Furthermore, recently we have observed that a chimeric HPV 11 genome containing a portion of HPV 16 E7 ORF encompassing the Rb binding domain inserted into the E7 gene of HPV 11 transforms baby rat kidney cells; the unaltered HPV 11 genome will not transform baby rat kidney cells (Pater *et al.* unpublished observation). Some differences have also been observed with E6 proteins of different viruses. E6 of HPVs 16 and 18, but not HPV 11, interact with protein p53, the product of tumor suppressor gene p53 (Werness *et al.* 1990). The E6 ORF of HPVs 16 and 18 encode another protein $E6^{*}$ which is derived by an internal splicing (Schneider-Gàdicke and Schwarz, 1986; Smotkin and Wettstein, 1986). The internal splice site required for $E6^{*}$ has not been observed in the E6 open reading frame of HPV 11. It is unresolved whether the $E6^{*}$ gene product has a function <u>per se</u> or whether the mRNA-splicing event is important for efficient expression of E7 protein. Nevertheless, as indicated earlier, the inability of HPV 11 to synthesize enough E6 and E7 proteins

and the low affinity of these proteins for Rb suggests that both regulatory elements and proteins play a significant role in determining oncogenic potential. A way to test this hypothesis is by overexpressing E6 and E7 proteins in cells. The rationale in these experiments is that, in spite of lower affinity of E6 and E7 proteins of HPV 11 for p53 and Rb, overexpression would permit sequestration of all p53 and Rb proteins of the cells. Experiments can also be performed with chimeric plasmids generated by the swapping of enhancer, promoters, E6 and E7 genes of different HPVs.

CHAPTER 4

Future Studies

The studies that have been described thus far in this thesis have involved cell-type specific expression of papovaviral enhancers and the viral DNA sequences that may be important in this process. Methods used have concentrated upon DNaseI protection and competition experiments. Gel retardation assays have been a first step towards an analysis of proteins that may be important. The purpose of this chapter is to outline mechanisms that may be involved and approaches that might be taken to achieve a more thorough understanding of this cell-type specific expression. I feel it would be important to concentrate on the analysis of first and second order protein interactions with specific enhancer sequences and the effect these have on cell-type specific expression and/or differentiation.

The results of the present study involving embryonal carcinoma cells have not only a direct application in furthering our understanding of viral gene expression but also have a broad implications for our understanding of cellular differentiation process. There is a clear indication that cellular differentiation is accompanied by the activation of genes which could be mediated partly through the activation of transcription factors. Activation of transcription factors could involve one or more of the following processes: 1) *De novo* gene transcription. For example, the Egr-1 gene, which

encodes a zinc finger protein with properties of a transcription factor, is induced *de novo* during cardiac and neuronal cell differentiation of the P19 cell line (Sukhotame *et al.* 1988). 2) Unmasking of stored messenger RNAs during differentiation as a prerequisite for their translation. This type of mechanism has been reported in sea urchins where fertilization triggers unmasking of stored maternal mRNA and translation (Grainger and Winkler, 1987). 3) The trapping of transcription factors in the cytoplasm through their association with inhibitory proteins. During differentiation, modification of either the transcription factor(s) or the inhibitory protein may result in the dissociation and nuclear translocation of the transcription factors. An analogous process has been described for NF_{κ}-B, a factor required for immunoglobulin gene expression. NF_{κ}-B is localized in the cytoplasm of nonexpressing cells due to its association with IκB. Agents such as phorbol esters release the NF_{κ}-B from the NF_{κ}-B-IκB complex by phosphorylation of IκB and this allows translocation to the nucleus (Baeuerle and Baltimore, 1988). 4) Post-translational modification such as phosphorylation/dephosphorylation of transcription factors during differentiation. Phosphorylation of the cyclic AMP response element binding protein (CREB) by protein kinase A has been shown to stimulate binding of CREB to DNA and transcriptional activation (Yamamoto *et al.* 1989). In contrast, cyclic AMP can also inactivate transcription factors by phosphorylation as shown

in the case of yeast ADR1 (Cherry *et al.* 1989). Similarly, in a meeting report, Karin and Hunter suggest that the activity of AP1, a heterodimer of c-jun and c-fos, is induced by dephosphorylation (Berk and Schmidt, 1990).

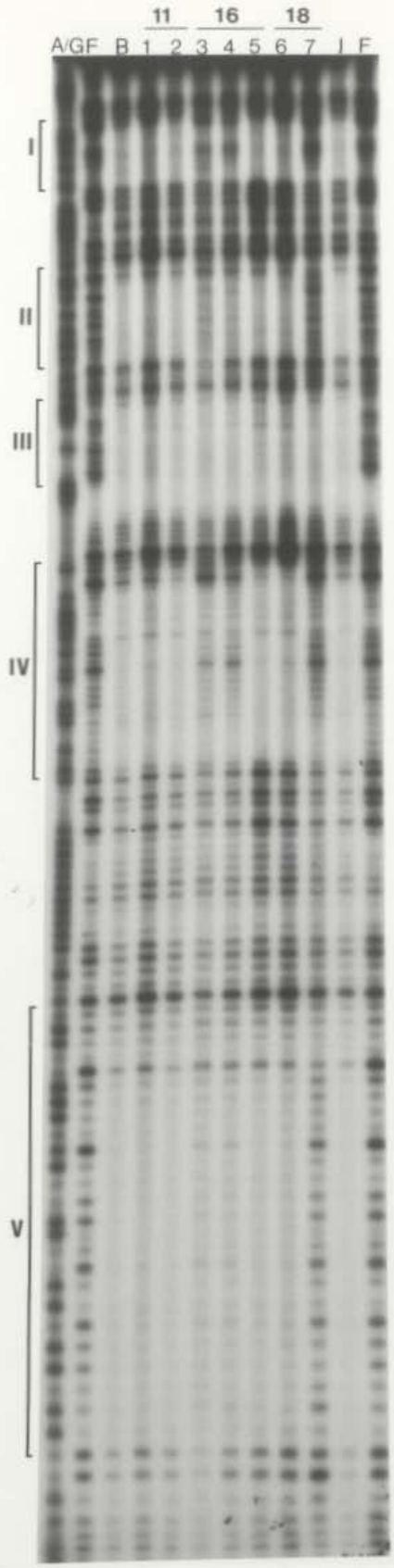
Activation of genes during differentiation could also involve the cessation of transcription of undifferentiated cell-type specific genes that encode transcription repressors. A candidate is the undifferentiated EC cell factor NF-AB which has been shown to bind to an octamer motif of the immunoglobulin heavy chain enhancer and represses transcription (Lenardo *et al.* 1989).

Future experiments should be directed towards identifying which among the above mechanisms are operative during differentiation induced activation of SV40, BK and JC regulatory elements. The methods involving screening of λgt-11 expression library with double stranded DNA as a probe (Singh *et al.* 1988) and subtractive cDNA cloning (Timblin *et al.* 1990; Wieland *et al.* 1990) may be an immediate approach to study induction by *de novo* gene transcription.

An exciting finding from my studies is that the NF-1 motif is a common motif found in enhancers and plays a significant role in gene transcription of viruses with distinct and divergent tissue specificities. Five out of six viral enhancers contained this motif. Competition studies have indicated that NF-1 motifs of all three HPVs interact with similar factors in a cell type. Likewise, BK and JC enhancers

interact with similar factors in retinoic acid differentiated P19 cells. This observation suggests that tissue specific function of NF-1 motif containing enhancers is achieved by more than one mechanism (see below). It is possible that within any one cell type NF-1 motifs, irrespective of the type of enhancer they are in, interact with similar factors. Results from a preliminary experiment support this possibility. As shown in Figure 4.1, DNaseI footprints of the BKV enhancer in C33A extracts were competed by oligonucleotides corresponding to various HPV protected regions. Protected regions I, II, IV and V of BKV contain NF-1 motifs and all of them were competed by HPV 16 region III and VII oligonucleotides (lanes 3, 4). The HPV 11 region II oligonucleotide (lane 1), which has been a poor competitor in previous footprints (Chapter 3), again showed no or poor competition. Among the non-NF-1 motif containing oligonucleotides (region IV of HPV 11 in lane 2, region IX of HPV 16 in lane 5 and region IV and V of HPV 18 in lane 6 and 7, respectively) only the region V oligonucleotide of HPV 18 competed for the BKV NF-1 motif containing regions. It appears that the HPV 18 region V oligonucleotide is unique in that, although it is devoid of a consensus NF-1 motif, it competes effectively for an NF-1 motif containing region of HPV 18 (Fig. 3.11) as well as of BK. One possibility is that this region contains an undefined NF-1 motif. Competition by all of the oligonucleotides appears to be specific for BKV NF-1 motifs as none of the oligonucleotides

Figure 4.1 Oligonucleotide competition of footprints for the BKV enhancer in C33A extracts. Lanes F is for DNase digestion in the absence of proteins or competitors, lane B is a sample not treated with competitor oligonucleotides. Competitors were a region II oligonucleotide of HPV 11 in lane 1; region IV of HPV 11 in lane 2; region III of HPV 16 in lane 3; region VII of HPV 16 in lane 4; region IX of HPV 16 in lane 5; region IV of HPV 18 in lane 6; region V of HPV 18 in lane 7 and a control oligonucleotide (5′ AAGGGAAGGGATGG 3′) in lane J. For all oligonucleotide competitons, 100 ng of oligonucleotide was incubated with the extracts before the addition of probes. Protected regions of BKV are numbered I-V. Oligonucleotides in lanes 1, 3 and 4 and BKV protected regions I, II, IV and V contain NF-1 motifs.


11
16
18
A/GF B 1 2 3 4 5 6 7 J F
I
II
III
IV
V

competed for the GC rich region III of BK. Also, the control oligonucleotide (lane J) did not effect any of the footprints.

Overall, from the results presented above, it appears that the cell-type specific function of NF-1 motif containing enhancers could involve one or more of the following mechanisms: 1) Interaction with cell-type specific species of NF-1 proteins. In this case, the cell-type specific function of the enhancer is due to the ability of a particular species of NF-1 to interact with proteins bound to the remainder of the enhancers and promoter. Although NF-1 proteins of non-expressing cell types can interact with NF-1 motifs of a cell type specific enhancer, they somehow fail to interact productively with the rest of the transcription machinery. 2) A ubiquitous NF-1 may aid in the interaction of a cell-type specific protein to the adjoining motif by cooperative interaction. 3) NF-1 proteins may communicate with the general transcriptional machinery through a cell-type specific "piggyback" transcription factor.

The possibility of cell-type specific differences in NF-1 motif binding factors should be testable with the JCV enhancer in differentiating EC cells. JCV enhancer activity is tissue restricted and depends mostly on NF-1 motif containing regions (Table 2.1). Also, previous work has suggested that factors from HeLa and glial cells that interact with JCV

NF-1 motifs have different mobilities in SDS-polyacrylamide gels (Khalili *et al.* 1988).

Chang *et al.* (1990) reported that fragments of HPV 16 enhancers containing only multimeric NF-1 motifs do not possess any enhancer activity. They suggest that papillomavirus enhancer activity requires the interaction of PVF (papillomavirus factor) in addition to NF-1. Thus, epitheliotropism of HPVs could be regulated, partly, by cooperative interaction of NF-1 with PVF. In addition, my studies using oligonucleotide competition, particularly with the region VI oligonucleotide of HPV 16 (Fig. 3.7) and the region V oligonucleotide of HPV 18 (Fig. 3.11), provide indirect evidence for cooperative interaction of NF-1 motif bound proteins with non-NF-1 motif bound proteins. Direct evidence can be obtained with point mutational analyses of the enhancer. Ideally, if there is cooperative interaction, a point mutation of the NF-1 motifs should render adjoining motifs more susceptible to DNaseI digestion due to weak protein interaction. Further analysis could be done by oligonucleotide competition of DNaseI footprint assay. In these assays, weakly protected regions adjoining a mutated NF-1 motif should be more susceptible to competition by a specific competitor in the NF-1 mutant enhancer than in the wild type enhancer.

Gel retardations with an NF-1 motif containing region VII oligonucleotide of HPV 16 demonstrated two DNA-protein complexes for C33A, HeLa and SiHa cells and one complex for

CaSki and 143B cells (Fig. 3.13). The slower migrating C33A, HeLa and SiHa specific complex may have arisen from the direct interaction of C33A, HeLa and SiHa specific protein with DNA. The other possibility is that the slower migrating complex represents a DNA-protein complex similar to the faster migrating complex except that it contains an additional piggyback transcription factor(s). Alternatively, C33A, HeLa and SiHa contain a specific kinase which phosphorylates the protein interacting with region VII of HPV 16. In this case, the slower migrating complex may contain a phosphorylated version of the protein. An ideal approach to test the first two possibilities is by chromatographic and gradient fractionation of nuclear extracts and subsequent gel retardation assays with individual fractions or combination of fractions. If the slower migrating complex represents an additional or a unique DNA-protein interaction, proteins which form this complex should fractionate differently from the proteins of the faster migrating complex. On the other hand, if a piggyback transcription factor is involved, fractions containing only piggyback factors would not form any specific complex in a gel retardation assay. Gel retarded complexes similar to those with crude nuclear extracts should be obtained only after addition of DNA binding protein containing fraction to the piggyback factor containing fractions. Further confirmation could be persued by UV-crosslinking and by glutaraldehyde crosslinking in combination with UV-crosslinking.

UV-crosslinking would identify only DNA-protein complexes but not DNA-protein-protein complexes. However, the combination of UV-crosslinking and glutaraldehyde crosslinking would identify both DNA-protein complexes and DNA-protein-protein complexes. The functional significance of any of these factors can be tested by affinity purification followed by reconstituted *in vitro* transcription. Affinity purification of piggyback transcription factors may require covalently attached DNA-protein complexes in the affinity column. Cloning of genes coding for piggyback transcription factor(s) can be achieved either by conventional methods involving affinity purification, protein sequencing, and screening of cDNA library with DNA probes deduced from protein sequence or by a recently developed direct cloning technique involving screening of λgt11 fusion protein library using biotinylated DNA binding protein as a probe (Nelbock *et al.* 1990). Involvement of protein phosphorylation in generating a differential gel retardation pattern for region VII of HPV 16 with C33A, HeLa, SiHa, CaSki and 143B cell extracts can be tested by treating nuclear extracts with phosphatases, such as calf intestinal phosphatase, prior to gel retardation. If the slower migrating complex observed with C33A, SiHa and HeLa nuclear extracts contained the phosphorylated version of the protein present in the faster migrating complex, phosphatase treatment of nuclear extracts should produce only one complex corresponding to the faster migrating complex.

Pugh and Tjian (1990) have recently described a new class of transcription factors called "Coactivators". Coactivators act as communicators between upstream regulatory element binding factors and TFIID. Transcription factors Sp1 and CTF/NF-1 require these coactivators for transcriptional activation (Pugh and Tjian, 1990). Coactivators appear to be species specific as coactivators of CTF/NF-1 of Drosophila and human are not functionally interchangeable. Coactivators are also promoter specific since coactivators required for Sp1 mediated transcriptional activation of promoters with and without TATA boxes are different (Pugh and Tjian, 1990). It is also possible that coactivators are cell-type specific. For example, the CTF/NF-1 motif containing HPV enhancer's epitheliotropism and the JCV enhancer's neurotropism could be due to cell-type specific coactivator(s).

Unlike piggyback transcription factors, it has not been possible to separately purify coactivator(s) from general transcription factors (Pugh and Tjian, 1990). One possibility is that coactivators interact with upstream regulatory element binding factors and TFIID only when the latter two factors are in simultaneous contact with DNA. In that case, purification of coactivators would be possible only when enhancer molecules containing both covalently bound enhancer binding factors and TFIID are used in an affinity chromatography column. This may be a possible approach to test for and purify coactivators that may be involved in HPV and JCV cell specific gene expression.

SECTION 2

THE FUNCTIONAL ROLE OF BKV TUMOR ANTIGENS IN TRANSFORMATION

CHAPTER 5

The essential function of the regulatory elements described in the previous section is to initiate transcripts that encode proteins required for the virus life cycle. As stated previously (Section 1), cell types that are permissive for the virus life cycle should have the ability to support virus entry, gene transcription, DNA replication, assembly and release. Infection of cell types that can not support all these requirements of the life cycle but allow virus entry and gene transcription would result in premature termination of the virus life cycle. Occasionally, these infected cells undergo morphological transformation as a consequence of having biologically active viral proteins. Permissive cells can also undergo such morphological transformation when infected with viruses with a defective genome or wild type viruses in the presence of co-factors. All the viruses described in the previous section have been shown to induce morphological transformation of certain cell types. Since viral regulatory elements have a significant contribution in determining the cell types that can be transformed, based on the previous study, it appears that JCV should preferentially transform cells of neuronal origin, HPVs should preferentially transform cells of epithelial origin and BK and SV40 should transform a variety of cell types. Evidence for this is abundant in the literature (Tooze, 1981; Fields, 1990).

Analyses of cells derived from tumors and *in vitro* transformation assays have allowed the identification of viral proteins involved in cellular transformation. They have been termed viral oncoproteins. Genes encoding large T- and small t-antigens of SV40, BKV and JCV, and E6 and E7 of HPVs 16 and 18 are capable of transforming cells and a few of these proteins have also been detected in tumor cells (Section I; Tooze, 1981).

I have examined the role of BKV tumor antigens in cellular transformation. These studies are presented and discussed in this section. In the introduction, a brief description of cellular transformation, oncogenes and papovaviral tumor antigens is presented.

5.1 Introduction

5.1.1 Oncogenes and cellular transformation

One of the most rigorously studied aspects of molecular biology is cellular transformation. The last two decades have witnessed exciting breakthroughs in this field. Sufficient evidence has accumulated to support J.M. Bishop's postulation that "Cancer may be malady of genes, arising from genetic change of diverse sorts - recessive and dominant mutations, large rearrangements of DNA and point mutations, all leading to distortions of either the expression or biochemical function of genes" (Bishop, 1987). In normal cells, homeostasis is maintained by the interaction of positive and negative signals

of cell growth. These interacting signals afford chromosomal stability, the capacity to undergo terminal differentiation, and programmed cellular proliferation. Positive signals are derived by a cascade of 'proto-oncogenes' while negative signals are generated by 'anti-oncogenes' (for reviews see Herrlich and Ponta, 1989; Sagar, 1989). Distortion in any of these signaling pathways could lead to cellular transformation.

5.1.1.1 Most oncogenes are involved in generating mitotic signals

The oncogene theory began 80 years ago when Peyton Rous reproduced a tumor by using cell free filtrates of a tumor from a Plymouth Rock hen (reviewed in Varmus, 1985). Rous sarcoma virus, an RNA tumor virus, isolated in these experiments served as an effective tool in the discovery of oncogenes sixty years later. Subsequent studies showed that the src oncogene of this retrovirus is a nearly exact copy of a gene found in all chicken cells (Stehelin *et al.* 1976). The majority of retroviruses have incorporated a normal 'proto-oncogene' into their genome in the course of infection, and in the process convert them to oncogenes (Bishop, 1983). This oncogenic conversion of a proto-oncogene could occur by two mechanisms: 1) Quantitative: excessive expression of an otherwise normal gene might convert cells to the neoplastic phenotype. Transduction by a retrovirus places genes under powerful signals in the viral genome and promote their expression in great abundance. 2)

Qualitative: mutations in the coding domain of a proto-oncogene during or subsequent to transduction could alter the biological activity of the proto-oncogene products. Apart from retroviruses, carcinogens can also convert proto-oncogenes to oncogenes by damaging DNA.

It is now known that cells carry several different proto-oncogenes whose products are distributed in specific compartments of the cell (e.g. in the plasma membrane, cytoplasm and nucleus) and are functionally interlinked (for review see Adamson, 1987; Herrlich and Ponta, 1989). Proto-oncogenes are grouped into five classes based on their biological functions: 1) those encoding secretory growth factors, e.g. c-*sis*; 2) growth factor receptors, e.g. c-*erb*-B, c-*neu* and c-*fms*; 3) tyrosine kinases, e.g. c-*src*, c-*fes* and c-*fps*; 4) GTPases. e.g. the *ras* family; 5) nuclear proteins. e.g. the c-*myc* family, the c-*jun* family, the c-*fos* family and c-*erb*-A. (Adamson, 1987). One of the best studied oncogenes which is frequently activated in tumors (15% of tumors) is the *ras* oncogene (for review see Barbacid, 1987). Since the *ras* oncogene is used in some of the experiments presented in this thesis, the biochemical properties of *ras* gene products are discussed below.

To date, three *ras* genes encoding closely related proteins of molecular weight 21 KD (p21) have been identified in the mammalian genome. They have been designated as H-*ras*-1, K-*ras*-2 and N-*ras* (Barbacid, 1987). Two of them, K-*ras*-2 and

H-*ras*-1, were first identified as transforming genes of the Kirsten and Harvey strains of rat sarcoma viruses, two acutely transforming retroviruses generated by transduction of the rat K-*ras*-2 and H-*ras*-1 cellular genes, respectively (DeFeo *et al.* 1981; Ellis *et al.* 1981). N-*ras* was identified by DNA-mediated transfection from neuroblastoma cells (Taparowsky *et al.* 1983). The EJ-*ras* oncogene used in my studies is derived from the EJ bladder carcinoma cell line and carries a mutation at the 12th codon which encodes valine instead of glycine (Tabin *et al.* 1982).

Tumorigenic conversion of cells in culture has been a convenient experimental system to study oncogenesis. Echoing the general belief that cancer is a multistep process, two steps of transformation can be distinguished in this system: a first step in which cells are immortalized (rescue from senescence), and a second step in which cells acquire a malignant phenotype (completing step). Such studies with the *ras* oncogene provided conflicting results. Most results suggested that *ras* alone can transform established cell lines, while transformation of primary cells required cotransfection of oncogenes such as c-*myc*, the adenovirus E1A gene, N-*myc*, a mutant p53 oncogene, the polyoma virus middle T gene, or SV40 or BK virus T-antigen gene along with ras (Barbacid, 1987). These results imply that *ras* oncogenes are competent for the second step but cannot rescue cells from senescence. However, this interpretation has been challenged by Spandidos and Wilkie

(1982) who show, in a different assay system, that the over-expression of *ras* alone can transform primary cells. In spite of the discrepancy in results, which may be due to an assay system incapable of identifying additional mutations that occur during experimental manipulation, there is evidence that *ras* is a major factor in carcinogenesis.

As the importance of *ras* in cellular transformation became clear, the search for its biological activity was intensified. It was determined that *ras* proteins are localized in the inner side of the plasma membrane, share significant homology to G proteins, bind guanine nucleotides and possess intrinsic GTPase activity (Barbacid, 1987).

The location of ras on the inner surface of the plasma membrane along with their similarity to G proteins has raised the possibility that ras participate in signal transduction. Signal transduction is a multistep process involving the net-like interconnection of oncogenes, growth factors, and growth factor receptors. It has been demonstrated that signal transduction by growth factors such as PDGF, and possibly EGF, growth factor receptor-like oncogenes-fes, fms and tyrosine kinase src can be blocked by antibodies against ras (Herrlich and Ponta, 1989; Marhsall, 1987). This suggests that PDGF, EGF, fes, fms and src are the upstream receptor systems of signal transduction that interact with ras proteins.

Microinjection of ras and antisense RNA have identified protein kinase C, raf and fos as downstream effectors of ras-

dependent signal transduction (Gauthier-Rouvière *et al*. 1990). *Ras* can induce the transcription of the *fos* gene and antisense RNA to *fos* can block *ras*-mediated mitotic induction (Herrlich and Ponta, 1989; Gauthier-Rouivère *et al*. 1990). Fos is a transcription factor and is involved in both positive and negative regulation of gene transcription (Distal *et al*. 1987; Mitchel and Tjian, 1989). One gene which is positively regulated by fos is the growth factor TGF alpha gene (Derynck, 1988). Secreted growth factors may interact with the epidermal growth factor receptor of the producer cells to amplify mitotic signals (Barbacid, 1987). These results suggest that *ras* plays a crucial role in a network of cellular oncoproteins and growth factors that link extracelluar stimuli to the transcriptional machinery to generate positive signals required for cell growth.

As with any enzyme in a biological system, the activity of *ras* protein is tightly regulated. Control appears to be at the level of nucleotide binding. The GTP-bound form but not the GDP-bound form of *ras* protein is the active component in the signal transduction pathway (Trahey and McCormick, 1987). Normal *ras* proteins remain mostly in a GDP-bound form. This is achieved by the association of *ras* protein with a cytoplasmic protein called GAP which stimulates the GTPase activity of *ras* proteins (McCormick, 1989). Most mutations that generate oncogenic *ras* proteins, prevent GAP from stimulating GTPase activity, thereby allowing these mutant

proteins to remain in the active GTP-bound form (Trahey and McCormick, 1987; Barbacid, 1987). Over-expression of *ras* due to pro-viral insertion can also convert cells to an oncogenic phenotype, possibly due to saturation of the GAP protein (George *et al.* 1987). Certain growth factors which use ras in their signal transduction pathway may block GTPase activity of ras by inactivating the GAP protein through phosphorylation (Hall, 1990). As a consequence of ras-mediated signal transduction several genes could be induced. The newly synthesized protein may stimulate cell growth either as a component of the cell cycle or as a enzyme that abrogates the function of negative regulators of cell growth which are described below.

5.1.1.2 'Anti-oncogenes' negatively regulate cell growth

Studies involving normal and tumor cell hybrids established that normal cells carry some functions which can effectively suppress the oncogenic phenotype. These studies, along with the realization that some heritable tumors involve loss-of-function mutations, provided firm evidence for the class of genes called "anti-oncogenes" (Knudson, 1985). The best studied example is the retinoblastoma susceptibility gene (Rb). Although the biological role of the Rb protein is not known, emerging evidence indicates that it has a role in regulating the cell cycle (Cooper and Whyte, 1989). It is postulated that the dephosphorylated form of Rb seen in the G0/G1 phase of the

cell cycle is the biologically active form and its activity prevents cells from entering S phase. Phosphorylation-mediated inactivation of this protein commits the cells to enter S and G2 (DeCaprio *et al*. 1989). Another anti-oncogene that has received considerable attention is p53 (for review see Levine, 1990). Structurally, p53 resembles transcription factors as it is a nuclear phosphoprotein with an acidic N-terminal domain and a basic C-terminal DNA binding domain. Two functions have been proposed for p53. First, it may negatively regulate growth via the regulation of transcription of a specific set of cell cycle dependent genes (Levine, 1990). Second, p53 prevents cellular DNA replication by inhibiting the activity of ATPase-helicase involved in unwinding of DNA (Lane and Benchimol, 1990).

Although the involvement of oncogenes in cancer has been questioned (Duesberg, 1988), the emerging theme is that cancer is a multistep process involving altered function of both oncogenes and anti-oncogenes (Land *et al*. 1983; Mark, 1989; Green 1989). For example, in colon cancer, anti-oncogenes are apparently lost from chromosomes 5, 18, and 17 and the *ras* oncogene is activated (Mark, 1989).

5.1.2 SV40 and BK virus tumor antigens

Apart from the cellular oncogenes, genes of viruses (viral oncogenes) have transforming potential. In fact, most of our knowledge about oncogenes were originally derived from studies

on viruses. Both RNA and DNA tumor viruses have been shown to contain these genes and are implicated in human cancer. Among RNA tumor viruses, retroviruses were the first viruses to be recognized as tumorigenic. They can infect and induce tumors in wide range of animals (Weis *et al*. 1985). These viruses can induce tumors either by introducing a transforming gene or by integration of viral sequences to perturb cellular genes. DNA tumor viruses that are implicated in cancer include human papillomaviruses (particularly HPV 16 and 18 which were discussed and studied in Chapter 3), hepatitis-B virus and Epstein-Barr virus. Other members of this group, although rarely detected in tumors, serve as effective tools for functional dissection of cellular transformation (Tooze, 1981). Although the mechanism of cellular transformation by viral proteins is a subject of intense speculation, an attractive model predicts that viral proteins possess activities that generate positive signals for cell growth and/or inactivate negative signals for growth.

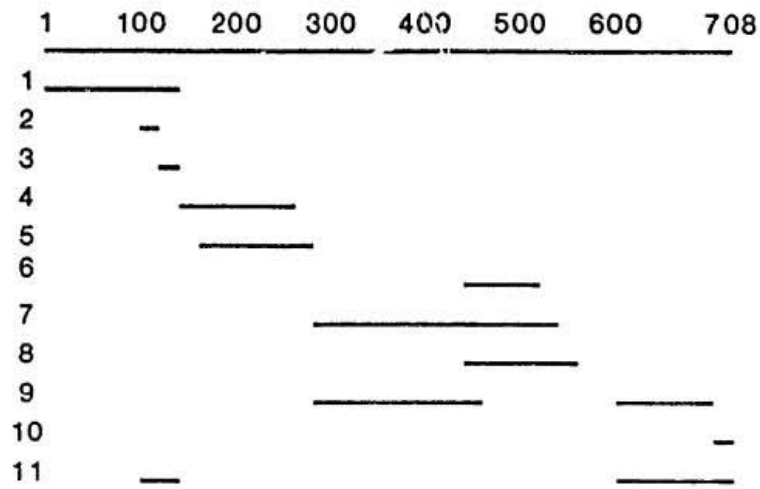
The DNA tumor viruses occur in at least four distinct families: Hepadna-, Papova-, Adeno- and Herpes viruses. Since the transformation studies in this thesis involve only papovaviruses, transforming proteins of two papovaviruses, SV40 and BK, are described below.

In addition to its normal role in viral DNA replication and late gene expression, the SV40 T-antigen can convert rodent cells to tumorigenic cell lines (Livingston and Bradley, 1987).

Deletion and point mutation studies have been employed to correlate transforming function with various biochemical functions of the T-antigen. These studies indicate that SV40 T-antigen is a pleiotropic oncogene performing a number of functions, each of which is usually confined to a different cellular gene product.

The functional domains of the SV40 T-antigen are presented diagrammatically in Figure 5.1. Although the protein is localized predominently in the nucleus, 2% of the T-antigen is embedded in the plasma membrane via a covalently linked palmitic acid (Livingston and Bradley, 1987). Nuclear localization of T-antigen is greatly facilitated by the nuclear localization domain located between amino acid residues 126 and 132. When residues 126 to 132 are appended to the sequences of certain unrelated non-nuclear proteins, they lead to the efficient nuclear localization of these molecules (Kulderon *et al.* 1984). This nuclear localization function is not required for the transformation of established cell lines, although it may be required for immortalization and transformation of at least some primary cell types (Livingston and Bradley 1987). However, transformation of primary cells by nuclear transport defective mutants can be achieved by cotransfection of either an activated ras oncogene or polyoma middle T-antigen gene (Michalovitz *et al.* 1987; Vass-Marengo *et al.* 1986).

Figure 5.1 The domains of the SV40 T-antigen and their homology to the BK T-antigen. The full length SV40 T-antigen is represented by a line with periodical amino acid numbers given above. The domains of the T-antigen are indicated by short lines which are numbered 1-11. The name and amino acid residue numbers of each domain are also indicated.



	Domains	SV40 T-antigen amino acids	Homology (%) to BK T-antigen
1	Immortalization	1-137	82
2	Retinoblastoma & p120 binding site	105-120	62
3	Nuclear localization	126-132	100
4	DNA binding	131-259	85
5	Cellular DNA synthesis	160-272	83
6	rRNA synthesis	420-509	87
7	p53 binding	272-517	80
8	ATP binding	413-528	84
9	ATPase	271-449, 570-669	76, 47
10	Adenovirus helper/ Host range effect	674-708	25
11	Oligomerization	114-152, 591-708	87, 37

A DNA binding domain, localized to amino acid residues 131 to 259 and forming an alpha helical structure, is required for the interaction of the T-antigen with the origin of replication (Clark *et al.* 1983; Fanning *et al.* 1989; Paucha *et al.* 1986). This function of the protein is not required for transformation as two point mutants, one at amino acid 153 and the other at 214, that fail to bind DNA can transform rat cells (Stringer, 1982; Prives *et al.* 1983). Amino acid residues between 302 and 320 can form zinc fingers; mutants at this region can still bind DNA but have varying effects on replication and transformation. The functional significance of this region may be to maintain the overall structure of the protein (Loeber *et al.* 1989).

Some of the biological activity is regulated by phosphorylation of the serine and threonine residues clustered at the amino terminus and carboxy terminus of the protein (Fanning *et al.* 1989). Much of the protein in infected and transformed cells occurs in phosphorylated oligomers while a smaller fraction of newly synthesized protein is unphosphorylated and monomeric (Fanning *et al.* 1981). Phosphorylation of the T-antigen is essential for its role in viral DNA replication (Prives, 1990). Phosphorylation at serine residues, particularly at the amino terminus, inhibits the replication function of the T-antigen whereas phosphorylation at the threonine residues enhances the replication function. Two host cell proteins, protein kinase cdc2 and phosphatase PP2A,

determine the phosphorylation status of T-antigens (McVey *et al.* 1989; Prives, 1990). Viral DNA replication requires sequence specific binding of the T-antigen to the origin of replication and unwinding of the DNA duplex by ATPase/helicase activity of the T-antigen. Which among these functions of the T-antigen is influenced by phosphorylation is not known.

Amino acid residues between 160 and 272 are essential for the induction of host cell DNA synthesis. Similarly, the amino acids of the T-antigen required for host rRNA synthesis extend from residues 420 to 509 (Soprano *et al.* 1983). By comparing the requirements for transformation, stimulation of cellular DNA synthesis and activation of rRNA genes, Soprano *et al.* (1983) concluded that the latter two activities are insufficient for transformation.

Oligomeric, but not monomeric, forms of T-antigen possess ATPase activity (Bradley *et al.* 1982). Oligomerization requires amino acid residues 114 to 152 and 591 to 708 (Montenarh *et al.* 1986). Amino acids between residues 271 and 416, and 570 and 669 are required for ATPase activity (Clark *et al.* 1983). The point mutants with reduced ATPase activity are transformation competent whereas deletion of ATPase domains renders the T-antigen transformation defective (Manos and Gluzman, 1984; Clark *et al.* 1983). The T-antigen can undergo ADP ribosylation which requires the amino acids between 412 and 528 (Clertant *et al.* 1984). The two activities, ATPase and ADP ribosylation, appear to be linked as the T-antigens

with enhanced ATPase activity are preferentially ribosylated (Tack *et al.* 1988).

Two additional functions, the ability to help human adenovirus to grow in monkey cells and the control of viral capsid synthesis in CV-1P cells, can be autonomously performed by the C-terminal 32 amino acid fragment of the protein (Tornow *et al.* 1985). These two functions are dispensable for transformation of rat cells (Soprano *et al.* 1983). However, these sequences are the principle determinant of host range of the virus and, as expected, show least homology to the closely related viruses BK and JC (Seif *et al.* 1979; Frisque *et al.* 1984).

In addition to the biochemical properties described above, the T-antigen forms a non-covalent complex with several cellular proteins. These cellular proteins generally coprecipitate with the T-antigen during the immunoprecipitation with antibodies against T-antigen. The main focus has been on three proteins of molecular weight 120 KD, 105 KD, and 53 KD (Ewen *et al.* 1989; Lane and Crawford, 1979; Linzer and Levine, 1979; Oren *et al.* 1981). The 105 KD and 53 KD proteins are the products of the retinoblastoma susceptibility gene (Rb) and the p53 anti-oncogene, respectively (DeCaprio *et al.* 1988; Finlay *et al.* 1989). The binding of Rb and p120 requires amino acid residues at positions 105 to 120, a region indispensable for cellular transformation (Ewen *et al.* 1989). Recent studies have indicated that oligomers of T-antigen bind preferentially

to the unphosphorylated form of Rb; such complexes exist in G1 phase and dissociate at or near the G1/S boundary of the cell cycle which coincides with phosphorylation of Rb (Ludlow *et al.* 1989, 1990). Given the high efficiency of T-Rb binding and release, the fact that other unrelated viral oncoproteins bearing a common consensus binding sequence also interact with Rb, and the finding that certain spontaneous Rb mutants lacking normal growth suppression function are also defective in T-antigen binding, one could speculate that the free unphosphorylated form of Rb blocks the entry of cells into the S phase and the T-antigen renders Rb inactive thus allowing the cells to enter the S phase (Ludlow *et al.* 1990; Dyson *et al.* 1989; Moran, 1988; Horowitz *et al.* 1989).

Like the Rb protein, p53 is also involved in growth suppression (Levine, 1990). Studies with deletion and point mutants of the T-antigen have indicated that the region within and flanking the ATPase domain are required for p53 interaction (Tack *et al.* 1988). This domain is required for the transformation of primary rat cells but not established rat cell lines (Clayton *et al.* 1982; Clark *et al.* 1983; Sompayrac and Danna, 1984).

The early region of SV40 also encodes a protein of 174 amino acids called the small t-antigen (Tooze, 1981). Functional analysis of this protein indicated that it is required along with large T-antigen for the maintenance of transformation (Bikel *et al.* 1986), to enhance the transforming activity

of limiting concentrations of large T-antigen (Bikel *et al.* 1987), for loss of actin cable in rat cells (Graessmann *et al.* 1980), and for the transactivation of RNA polymerase class II and III genes (Loeken *et al.* 1988). Based on the structural homology, Friedmann *et al.*(1978) suggested that there may be some functional similarity between t-antigen and some members of the growth-promoting glycoprotein hormone family. Two cellular proteins of 36KD and 63KD, form a complex with the small t-antigen (Yang *et al.* 1979).

Recent studies have indicated that the 36KD and 63KD proteins are catalytic and regulatory subunits of protein phosphatase 2A (PP2A), respectively, (Pallas *et al.* 1990). How interaction of PP2A with small t-antigen affects its function is not known. Two models have been proposed. The first model predicts that small t-associated PP2A is incapable of dephosphorylating kinases such as S6 kinase. S6 kinase is a cell cycle regulated protein which is activated at the G0/G1 transition by serine and threonine phosphorylation (Ballou *et al.* 1989). It is believed that this phosphorylated S6 kinase phosphorylates ribosomal proteins which triggers the increased rate of protein synthesis required for cell division. In SV40 transformed cells, small t-antigen prevents dephosphorylation of S6 kinase and thus stimulates the cell cycle.

The second model predicts the opposite. According to this model, small t-associated PP2A actively dephosphorylates serine residues of the large T-antigen and stimulates its DNA repli-

cation function (see above, Prives, 1990). Indeed, it has been demonstrated that purified PP2A can specifically dephosphorylate serine residues of the large T-antigen (Virshup *et al.* 1989).

So far, it has not been possible to assign a specific biochemical activity of the SV40 tumor antigens for their transformation function. It appears that transformation requires several combinations of these biochemical functions which again is subject to variation depending on cell type. It is possible that, in transformed cells, activities such as ATPase and stimulation of host cell DNA synthesis generate positive signals for cell growth whereas interaction with Rb and p53 inactivate negative signals of cell growth.

Unlike SV40, tumor antigens of BKV have received little attention. Few studies have been conducted using transgenic mice, *in vitro* infection and transformation assays. In transgenic mice, the BKV early region can induce primary renal and hepatocellular carcinomas and thymoproliferative disorders (Small *et al.* 1986; Dalrymphle and Beemon, 1990). Syrian golden hamsters infected with BKV develop insulinomas (Uchida *et al.* 1979). Under laboratory conditions, BKV can productively infect human foetal brain cells, foetal pancreatic cells and embryonic kidney (HEK) cells (Takemoto *et al.* 1979b; van der Noordaa *et al.* 1986; Purchio and Fareed, 1979). Productive infection generated several variants with altered DNA sequences, predominantly in the regulatory region of the virus,

with altered infection and transformation capacity (Watanabe and Yoshiike, 1982; Pater *et al*. 1983).

Several cell types such as HEK, human fibroblast cells with deletions in the short arm of chromosome 11, human foetal brain and pancreas cells, hamster and rat cells are susceptible to BKV DNA dependent transformation (Purchio and Fareed, 1979; Pater and Pater, 1986; Ronde *et al*. 1988; Takemoto *et al*. 1979b; van der Noordaa *et al*. 1986; Beth *et al*. 1981; Watanabe *et al*. 1984). HEK cells transformed by BKV DNA selected in focus assays show partial transformation as these cells are incapable of growth in soft agar. However, HEK cells transformed after their transfection with a combination of BKV DNA and an activated H-*ras*-oncogene (EJ-H-*ras*-1) attain the ability to grow in soft agar (Pater and Pater, 1986). All transformed cells retain an intact early region of BKV DNA and expressed the large T- and small t-antigen of this virus. Variants of BKV with a deletion in the T-antigen coding early region cannot transform whereas variants with only the early region can transform (Pater and Pater, 1986). These observations indicate the essential function of T-antigen as a transforming protein.

The functional domains of the BKV tumor antigen are not fully known. Their structural similarities with SV40 tumor antigens are the only criteria which have been used to assign a given function to a region of the protein. The results of studies that indicate similarities and differences in the

properties of tumor antigens of the two viruses are listed below:

1) Methionine tryptic peptides of SV40- and BKV T-antigens have shown that a total of 20 and 21 tryptic peptides are generated for SV40 and BKV, respectively, and only seven among them are similar for both proteins (Simmons *et al.* 1977).

2) The large T-antigen of BKV is less stable than the SV40 large T-antigen in transformed Rat 2 cells (Bollog *et al.* 1989).

3) SV40 transformed HEK cells readily show anchorage independent growth (Major and Matsumura, 1984) whereas HEK cells transformed by a combination of BKV DNA and the activated *ras* oncogene but not those transformed by BKV alone exhibit such a property (Pater and Pater, 1986).

4) Both SV40 and BKV require T-antigen for viral DNA replication and late gene transcription. Major and Matsumura (1984) have reported that HEK cells producing T-antigen of SV40 can support the replication of T-antigen defective BKV by binding to the BKV origin of replication. However, in the same cells, no virion or capsid proteins of BKV are synthesized indicating that SV40 T-antigen is not competent for activating BKV late genes.

5) The BKV small t-antigen, like the SV40 t-antigen, interacts with two cellular proteins of 36 and 63 KD (Rundell *et al.* 1981).

5.1.3 Objectives of the Study

Some of the differences in the activities of the SV40 and BKV T-antigens noted above could be due to variations in some of the biochemical properties of the two proteins. Prior to understanding these differences, it is essential to study the biology of the BKV T-antigen. For this purpose, my studies were aimed at: 1) identifying the role of the BKV tumor antigens in the maintenance of the transformed phenotype, and 2) identifying domains of the tumor antigens which are required for transformation.

The role of the SV40 T-antigen in the maintenance of transformation has been examined using temperature sensitive mutants (ts mutants) (Brugge and Butel, 1975; Kimura and Dulbecco, 1973). However, a similar approach is not possible with BKV due to the nonavailability of mutants as well as the restricted growth of BKV to HEK cells. As an alternative, I have utilized a technique involving controlled expression of antisense RNA (Izant and Weintraub, 1984).

Plasmids carrying sequences for BKV T-antigen expressed in an antisense orientation under the control of the human metallothionein promoter were introduced into BK-transformed cells. Cells which express the antisense RNA failed to grow in agar, indicating that the BKV T-antigen is required for the continued expression of the transformed phenotype. Additionally, several deletion and translation termination linker mutants

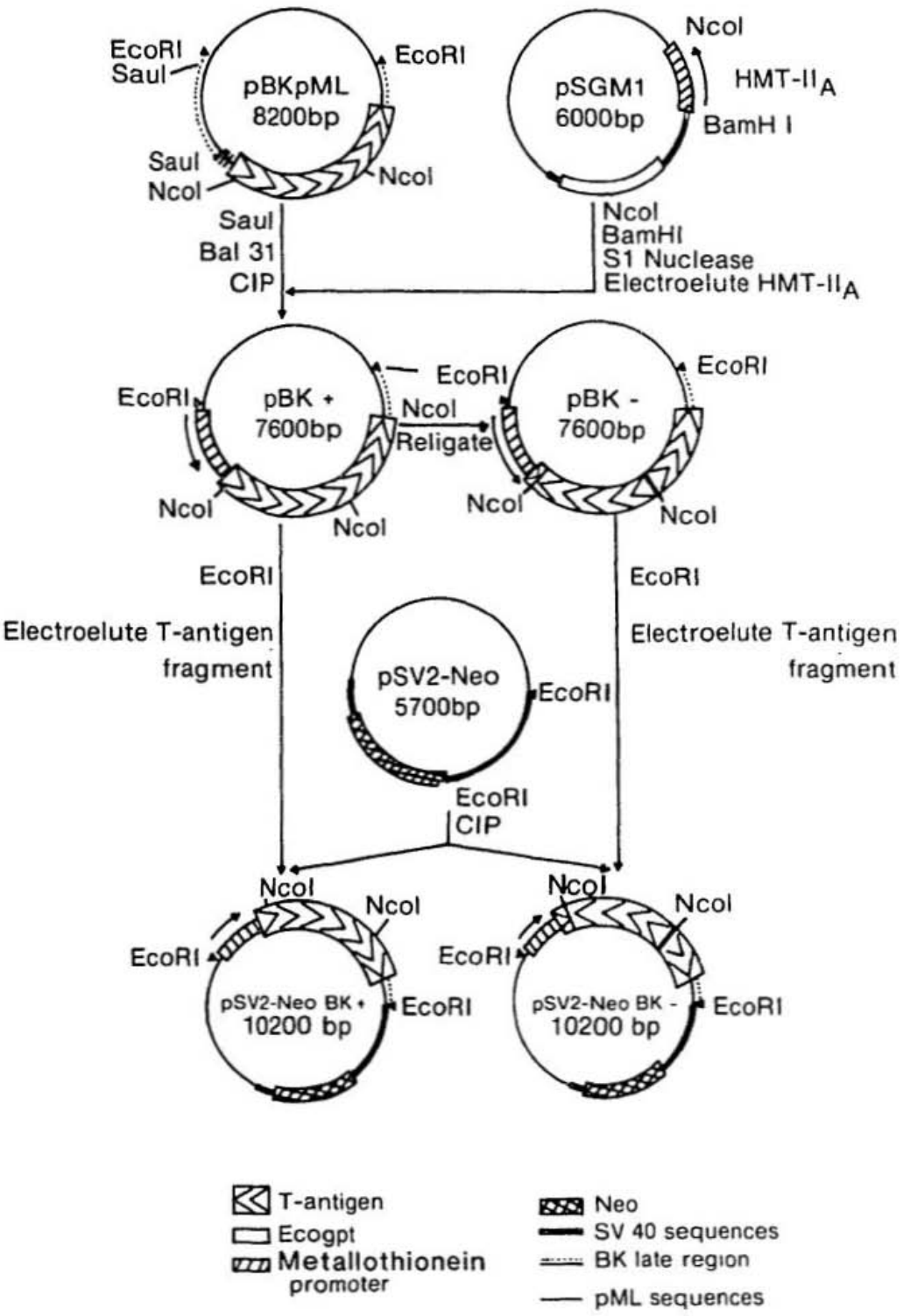
of T-antigen were tested for their ability to transform primary as well as established cell lines.

5.2 Materials and Methods

5.2.1 Construction of sense and antisense BKV T-antigen expressing plasmids

Plasmid pBK-pML, with BKV DNA inserted into the bacterial plasmid pML at its unique EcoRI site, was first digested with SauI (to remove nucleotides 3499-5100),treated with nuclease Bal 31 to delete approximately 25 base pair (bp) from the remaining BKV 5' regulatory region and part of the late coding sequences and then treated with calf intestinal phosphatase (Figure 5.2). The human metallothionein (HMT-IIA) promoter fragment was then ligated into this DNA in the correct orientation to direct transcription of the BKV T-antigen. The HMT-IIA promoter sequence was isolated from the plasmid $pSGM_1$ (a gift from M.L. Breitman) after digestion with restriction endonuclease NcoI, S1 nuclease, and BamHI and subsequent filling in with reverse transcriptase (Figure 5.2). The resulting plasmid expressing BKV T-antigen coding sequences (pBK+) was digested with restriction endonuclease NcoI and religated to yield the plasmid which has BKV T-antigen coding sequences from nucleotides 3279 to 1600 in an antisense orientation (pBK-). The EcoRI T-antigen fragments from plasmids pBK+ and pBK- were then ligated into the EcoRI site of the plasmid pSV_2-neo (Southern and Berg, 1982) to generate pSV_2-neo BK+ and pSV_2-neo BK- plasmids

Figure 5.2 Flow diagram for the construction of sense (pSV_2-neoBK+) and antisense (pSV_2neoBK-) BKV T-antigen expressing plasmids (see Materials and Methods for details). CIP: calf intestinal phosphatase. Ecogpt: E. coli xanthine-guanine phosphoribosyltransferase gene. HMT-IIA: Human metallothionein promoter.


EcoRI
SauI
pBKpML
8200bp
EcoRI
SauI
NcoI
NcoI
pSGM1
6000bp
NcoI
HMT-II$_A$
BamH I
SauI
Bal 31
CIP
NcoI
BamHI
S1 Nuclease
Electroelute HMT-II$_A$
pBK +
7600bp
EcoRI
NcoI
Religate
pBK -
7600bp
EcoRI
EcoRI
NcoI
NcoI
NcoI
NcoI
EcoRI
Electroelute T-antigen fragment
EcoRI
Electroelute T-antigen fragment
pSV2-Neo
5700bp
EcoRI
EcoRI
CIP
NcoI
NcoI
EcoRI
pSV2-Neo BK +
10200 bp
EcoRI
NcoI
NcoI
EcoRI
pSV2-Neo BK -
10200 bp
EcoRI
T-antigen
Ecogpt
Metallothionein promoter
Neo
SV 40 sequences
BK late region
pML sequences

5.2.2 Construction of translation termination linker mutants and in frame deletion mutants.

The plasmid designation, site of insertion of linker and region of deletions are shown in Figure 5.3. The total number of amino acids of the truncated proteins derived from linker insertion mutation are given in Table 5.1. For deletion mutants deleted amino acid numbers are given. Plasmid H3751 has the wild type BKV T-antigen coding sequences and its 5′ and 3′ noncoding regulatory sequences. This plasmid was constructed by inserting the PvuII fragment (nt 3751-510) of BKV into the HindIII site of plasmid pML which was filled in with reverse transcriptase. All the deletion and nonsense mutants were derived from this plasmid. Plasmid H3475 and H3279 were obtained by partial digestion with restriction endonuclease NcoI, filling in with reverse transcriptase and inserting XbaI linkers (5′ CTAGTCTAGACTAG 3′) having termination codons in all three reading frames. Insertion of the same linker into the RsaI sites of H3751 (partialy cleaved) gave rise to plasmids H2313, H1813, H1559 and H827. Plasmid H1813 has a deletion of 99 nucleotides from nt 1813 to 1714 and the linker is within this deletion. Partial cleavage with restriction endonucleases StuI and BspMI and insertion of linkers resulted in plasmids H3260 and H2169, respectively. To derive

Figure 5.3 Diagrams of nonsense and deletion mutants of the BKV tumor antigen. Nonsense mutants were constructed by the insertion of translational termination linker codon (TAG) into the wild type plasmid, H3751. Numbers below TAG in each plasmid indicates the site of insertion of the termination codon. Deletion mutants were constructed from the same parental plasmid using suitable restriction endonucleases as described in **Materials and Methods**. Empty boxes in H1813, H Δ69, H Δ363 and H Δ183 represents deletions. Initiation codon (ATG) and termination codon (TAA) for the wild type plasmid are indicated. Broken lines (- - -) represent intron sequences.

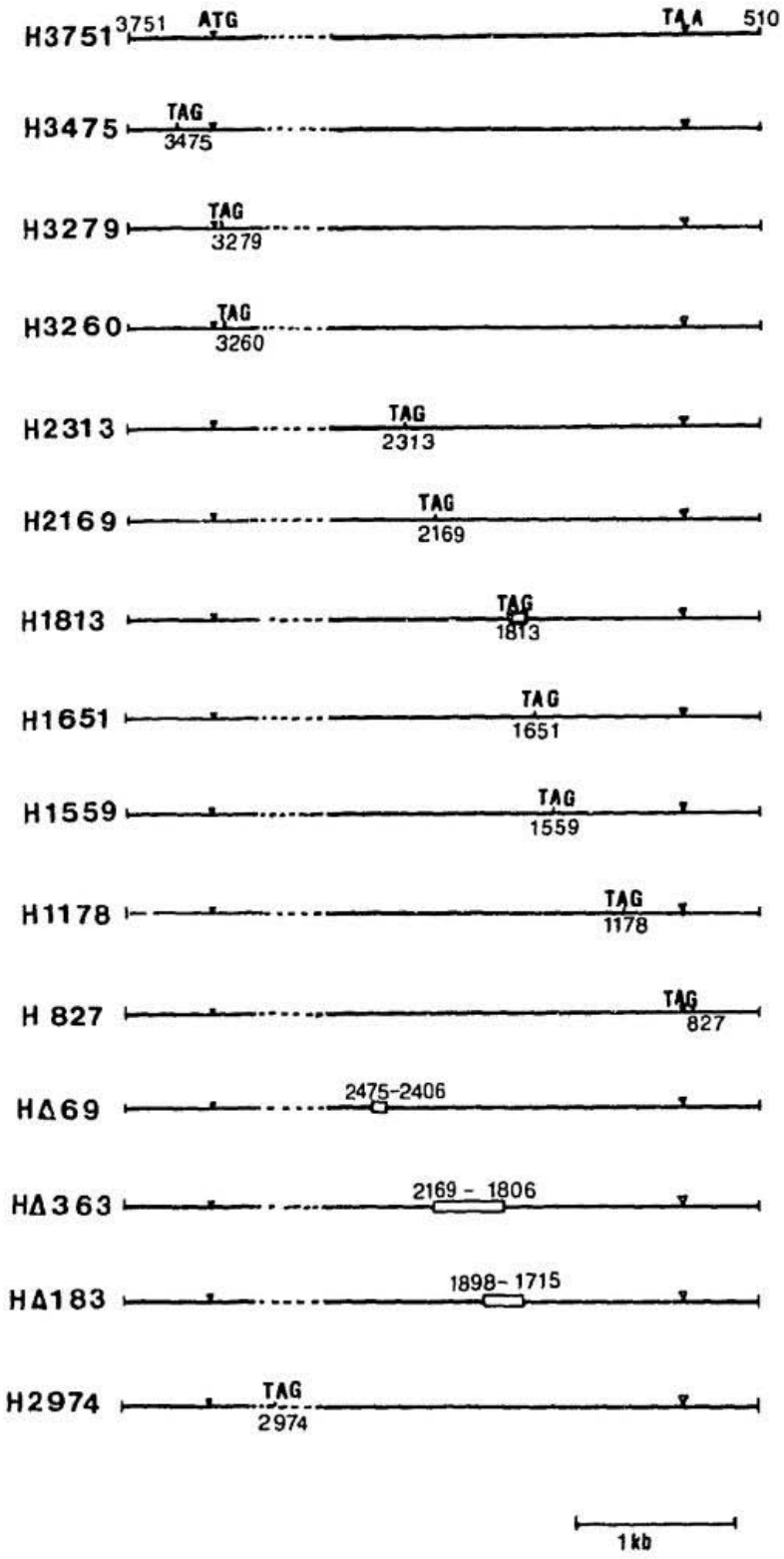

H3751
3751
ATG
TAA
510
H3475
TAG
3475
H3279
TAG
3279
H3260
TAG
3260
H2313
TAG
2313
H2169
TAG
2169
H1813
TAG
1813
H1651
TAG
1651
H1559
TAG
1559
H1178
TAG
1178
H 827
TAG
827
HΔ69
2475-2406
HΔ363
2169 - 1806
HΔ183
1898- 1715
H2974
TAG
2974
1kb

Table 5.1: Transformation of BRK and BHK cells with plasmids expressing BK T-antigen and its mutants.

Plasmids[a]	BRK cells[b]	BHK cells[c]	Amino acid[d]
H3751 (wild type)	80	890	695
H3475	35	356	695
H3279	0	0	10
H3260	0	0	17
H2313	0	0	217
H2169	0	0	266
H1813	6	70	384
H1651	7	110	438
H1559	2	185	470
H1178	10	190	596
H827	15	700	695e
H Δ69	10	156	164-187
H Δ363	0	0	265-386
H Δ183	0	0	356-417
H2974	29	84	112[f]

[a]The plasmid numbers indicate the nucleotide number at which an XbaI termination linker was inserted or the size of the deletion (Δ) in nucleotides.

[b]Kidney cells (10^7 cells/plate) from 8 day old rats were plated in two 100 mm^2 plates. After 48 hours, the cells were cotransfected with the indicated plasmids (10 μg/ plate) and the same amount of EJ-H-*ras*-1-pML. Tabulated are the average number of foci detected after 20 days per two plates or $2x10^7$ cells in two experiments. No foci were formed in plates transfected with EJ-H-ras-1-pML alone.

[c]Baby hamster kidney cells (two 60 mm^2 plates, 10^5 cells/plate) were transfected with 10 μg of the indicated plasmids. At 48 hours after transfection the cells were expanded into two 100 mm^2 plates. After another 48 hours the cells were plated in ten 60 mm^2 dishes containing 0.28% soft agar. Colonies were counted after 20 days. Numbers represent the average of duplicate experiments in which two plates or 2×10^5 cells were transfected.

[d]For termination insertion mutants the number represents the terminal amino acid; for deletion mutants the number represents deleted amino acids. Plasmid H3475 had an insertion into a noncoding region.

[e]Insertion after the polyadenylation signal.

[f]Truncated small t-antigen.

the plasmid pH1651, plasmid H3751 was partially digested with restriction enzyme Pst1, S1 nuclease treated and religated to bring an out of frame nonsense codon in frame. Likewise, treatment with restriction enzyme XbaI, reverse transcription and religation gave rise to nonsense mutant H1178.

Plasmid H2974 has a termination linker inserted at its SphI site resulting in truncated small t-antigen coding sequences without effecting the large T-antigen coding sequences. All the insertions have been verified by restriction mapping. Plasmid H1813 has been verified by sequencing.

To generate deletion mutant H Δ69, plasmid H3751 was digested with restriction enzyme AccI to delete the 71 nucleotides from 2475-2404. Subsequent filling in with reverse transcriptase and religation yielded this in frame deletion mutant. Likewise, digestion with restriction endonuclease BspMI, filling in and religation gave rise to plasmid H Δ363 with the deletion of nucleotides 2169 to 1806. Deletion of 183 nucleotides in region 1898 to 1715 was achieved by digesting with NdeI, S_1 nuclease, and partial digestion with RsaI and subsequent circularization to yield the plasmid H Δ183. The extent of deletions have been analysed by acrylamide gel electrophoresis. Plasmid H Δ183 has been verified by DNA sequencing.

5.2.3 Cell culture and transfection

NIH 3T3 cells were grown in Dulbecco's modified Eagles (DME) medium supplemented with 10% foetal calf serum (FCS). BHK-21 cells were grown in the same medium with the addition of 10% tryptose phosphate broth. Neomycin (neo) resistant colonies were selected by their growth in the presence of 500 μg/ml and 400 μg/ml G418 for BHK-21 and NIH 3T3 cells, respectively. Inducible expression from the metallothionein promoter was with 50 μM $ZnSO_4$ (Trimble *et al*. 1986). Baby rat kidney (BRK) cells were prepared from 8 day old Fischer rats and maintained in culture in DME + 10% FCS and changed to 2% fetal calf serum containing medium 48 hours after transfection. All the transfections were carried out by the calcium phosphate DNA coprecipitation procedure as described in Chapter 2.

5.2.4 Isolation of high molecular weight DNA and Southern blotting

Cells grown in tissue culture plates were washed with PBS, collected by a rubber policeman and stored at -20°C. The stored cells were thawed on ice and resuspended in 3 ml of 1x TE (10 mM Tris HCl, 1 mM EDTA). SDS (0.25 ml of 10%) and pronase (0.25 ml of 10 mg/ml) were added to thawed cells and incubated at 37° for 3 hours. After incubation, samples were phenol extracted and dialyzed successively in 0.01 M Tris (pH 8.4), 1 M NaCl for 5 hours and in 0.01 M Tris (pH 7.9) overnight. On the next day, dialyzed sampels were further digested with 0.25 ml of RNase (4 mg/ml) at 37° for 2 hours followed by

0.1 ml pronase (10 mg/ml) at 37° for another 3 hours. As before, samples were phenol extracted and dialyzed for 24 hours in 0.01 M Tris (pH 8.4), 1 M NaCl and for another 24 hours in 0.01 M Tris (pH 7.9). DNA was collected by ethanol precipitation.

Southern blotting was done according to Whal *et al.* (1979) as described by Pater *et al.* (1980). High molecular weight DNAs were digested with restriction enzymes and run by electophoresis through 1% agarose gels. The DNA was partially hydrolysed by soaking the gel in 0.25 M HCl for 15 minutes and then denatured by soaking the gels in denaturing buffer (1.0 M NaCl, 0.5 M NaOH) for 30 minutes. This was followed by treatment with neutralizing buffer (180 mM sodium acetate pH 4.0). DNA from the neutralized gel was transferred onto diazobenzyloxymethyl (DBM) paper which had been pretreated with 1.2 M HCl containing 0.06 gm/100 ml of sodium nitrite. The procedure for transfer was according to Whal *et al.* (1979).

Prehybridization and hybridization were done according to Whal *et al.* (1979) with minor modifications. Prehybridization was at 42° for 4 hours in a solution containing 5X SSC (20 X SSC is 3 M NaCl, 0.3 M sodium citrate), 50% formamide, 50 mM sodium phosphate (pH 6.5), 0.1% glycine, 100 μg/ml sonicated, denatured, salmon sperm DNA and 4.75X hybridization buffer (1X is 0.02% Ficoll, 0.02% Bovine serum albumin and 0.02% polyvinyl pyrrolidone in 6 X SSC). For hybridization, denatured radioactive probe was added to 4.4

X SSC, 0.9% hybridization buffer, 18 mM sodium phosphate (pH 6.5), 45% formamide, 100 µg/ml denatured, sonicated salmon sperm DNA. After hybridization at 42° for 20 hours, the filters were washed in 2 X SSC, 0.1% SDS for 15 minutes at room temperature and in 0.1 X SSC, 0.1% SDS for 2 hours at 50°. The filters were dried and exposed for autoradiography at -70°.

5.2.5 Detection of double stranded RNA and analysis of poly A^+ RNA

The nuclear and cytoplasmic RNAs were prepared for ds RNA analysis as described (Favaloro *et al.* 1980). The cells were washed in ice-cold PBS and collected into a corex tube with the aid of a rubber policeman. The cells were washed once again with PBS and resuspended in ice-cold lysis buffer [140 mM NaCl, 1.5 mM $MgCl_2$, 10 mM Tris HCl, pH 8.6, 0.5% NP40, 0.015% (w/v) macloid]. After vortexing, the cellular suspension was underlayered with an equal volume of 24% (w/v) sucrose, 1% NP40, in lysis buffer. After 5 minutes at 0°, the lysate was centrifuged at 8000 rpm, 4°, for 20 min. The cytoplasmic supernatant was removed to a separate tube and an equal volume of 2 X pronase buffer (200 mM Tris HCl, pH 7.5, 2.5 mM EDTA, 300 mM NaCl, 2% w/v SDS, 400 µg/ml pronase) was added. To the nuclear pellet, lysis buffer containing 2% NP40 was added and the nuclei were disrupted by repeated shearing through a 19 gauge syringe needle. An equal volume of 2 X pronase buffer was added to the disrupted nuclei. Both nuclear

and cytoplasmic fractions were incubated at 37° for 30 minutes. This was followed by phenol extraction and ethanol precipitation. The collected RNA was digested with 250 μg of RNase-free DNaseI (Worthington Diagnostics) per ml in 50 mM Tris (pH 7.4), 10 mM $MgCl_2$ for 1 h at 37° (Kim and Wold, 1985). RNA samples were extracted by chloroform-isoamyl alcohol (1:1) and precipitated with 0.5 M sodium acetate and 2 volumes of 95% ethanol. RNA samples in 10 mM Tris HCl, pH 7.8, 5 mM EDTA, 30 mM NaCl were subjected to RNase A and T_1 (40 μg/ml and 2 μg/ml, respectively) digestion at 30° for 30 minutes (Kim and Wold, 1985). The RNases were inactivated by the addition of 10 μg of 20% SDS, 50 μg of pronase and incubation at 37° for 15 minutes. The RNA was then extracted and precipitated as before. The samples were analyzed by northern blotting (Maniatis *et al*. 1982). For this purpose, samples were run on formaldehyde-agarose gels [1% agarose, 2.2 M formaldehyde, 0.2 M morpholinopro-panesulfonic acid (MOPS), pH 7.0, 50 mM sodium acetate, 1 mM EDTA, pH 8.0]. After electrophoresis, the gels were soaked for 5 minutes in several changes of water. This step was followed by partial alkaline hydrolysis for 15 minutes in 50 mM NaOH and 10 mM NaCl. The gel was neutralized in 0.1 M Tris-HCl (pH 7.5) for 45 minutes followed by 20 X SSC treatment for 1 hour. Transfer of RNA from the gel onto a nitrocellulose filter was according to Maniatis *et al*. (1982). Following transfer, the filter was dried and baked for 3-4 hours at 80° under vacuum. Prehybridization and hybridization

solutions are similar to the Southern blot procedure except that solutions contained 0.2% SDS.

Poly A^+ RNA for Northern blotting was prepared as per Maniatis *et al.* (1982). The cells grown to semi-confluence were washed with ice-cold balanced salt solution (0.17 M NaCl, 0.035 KCl, 0.01 M sodium phosphate dibasic and 0.01 M potassium phosphate monobasic). Washed cells were treated with isotonic buffer (0.15 M NaCl, 0.01 M Tris, 0.0015 M $MgCl_2$) containing 0.65% NP40 and 20 U/ml of RNasin. The cell lysate was collected by a rubber policeman and centrifuged at 15,000 rpm for 15 min at 4°. The supernatant containing polysomes was centrifuged again at 37,500 rpm in the 60 Ti rotor (Beckman) for 1 h. The polysomal pellet was resuspended in binding buffer (25 mM Tris HCl, 0.5 M NaCl, pH 7.5, 0.5% SDS) and passed through an oligo-dT column. Poly A^+ RNA was eluted from the column and characterized by Northern blotting as described above.

5.2.6 Synthesis of RNA probes with SP6 polymerase

The 1.9 kb HindIII B fragment (nucleotides 2797 to 872) of T-antigen coding sequences of BKV (Seif *et al.* 1979) was inserted in either orientation into the HindIII site of plasmid vector pGEM4 (Promega). *In vitro* RNA synthesis by SP6 polymerase was carried out in 40 mM Tris, pH 7.4, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM DTT, RNAsine (1 unit/μl), 100 μg/ml BSA, 500 μm of each rNTP and 100 μg/ml of template plasmid DNA linear-

ized at the PvuII site of the vector (Melton *et al*. 1984). [α^{32}P]UTP was used for labelling the transcripts. Sense and antisense probes prepared by this protocol were used to probe the RNA blot.

5.2.7 Immunoprecipation

The cells were preincubated for 1 h in methionine-free MEM medium and then incubated for 3 h in the presence of medium with 100 μCi/ml [^{35}S]methionine (Pater and Pater, 1986). Extraction was in PBSTDS (10 mM dibasic sodium phosphate, pH 7.2, 0.9% sodium chloride, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 0.2% sodium azide, 0.004% sodium fluoride, pH adjusted to 7.25 with 1.0 M monobasic sodium phosphate) as described by the supplier (Oncogene Science). The extracts were then incubated for 18 h at 4° in the presence of protein-A agarose (Pharmacia) and either a monoclonal antibody against the SV40 T-antigen in culture medium (Gurney *et al*. 1986) or normal medium. The immune complexes were pelleted, washed four times in PBSTDS and subjected to 12% SDS-polyacrylamide gel electrophoresis [12% acrylamide-bisacrylamide (37:1), 3.75 M Tris HCl, pH 8.8, 0.1% SDS for lower gel and 0.03% acrylamide-bisacrylamide, 125 mM Tris HCl, pH 6.8, 0.1% SDS for upper gel]. The sample loading buffer and electrophoresis buffer contained 2% SDS, 0.05 M Tris, pH 6.8,0.1 M DTT, 10% glycerol, 0.005% bromophenol blue and 50 mM Tris HCl, 0.1% SDS, 384 mM glycine, respectively. Samples were resuspended in loading buffer, incubated at 85° for 10 minutes and subjected to elec-

trophoresis. After electrophoresis, the gel was treated with enhance (NEN), dried and subjected to fluorography (Pater and Pater, 1986).

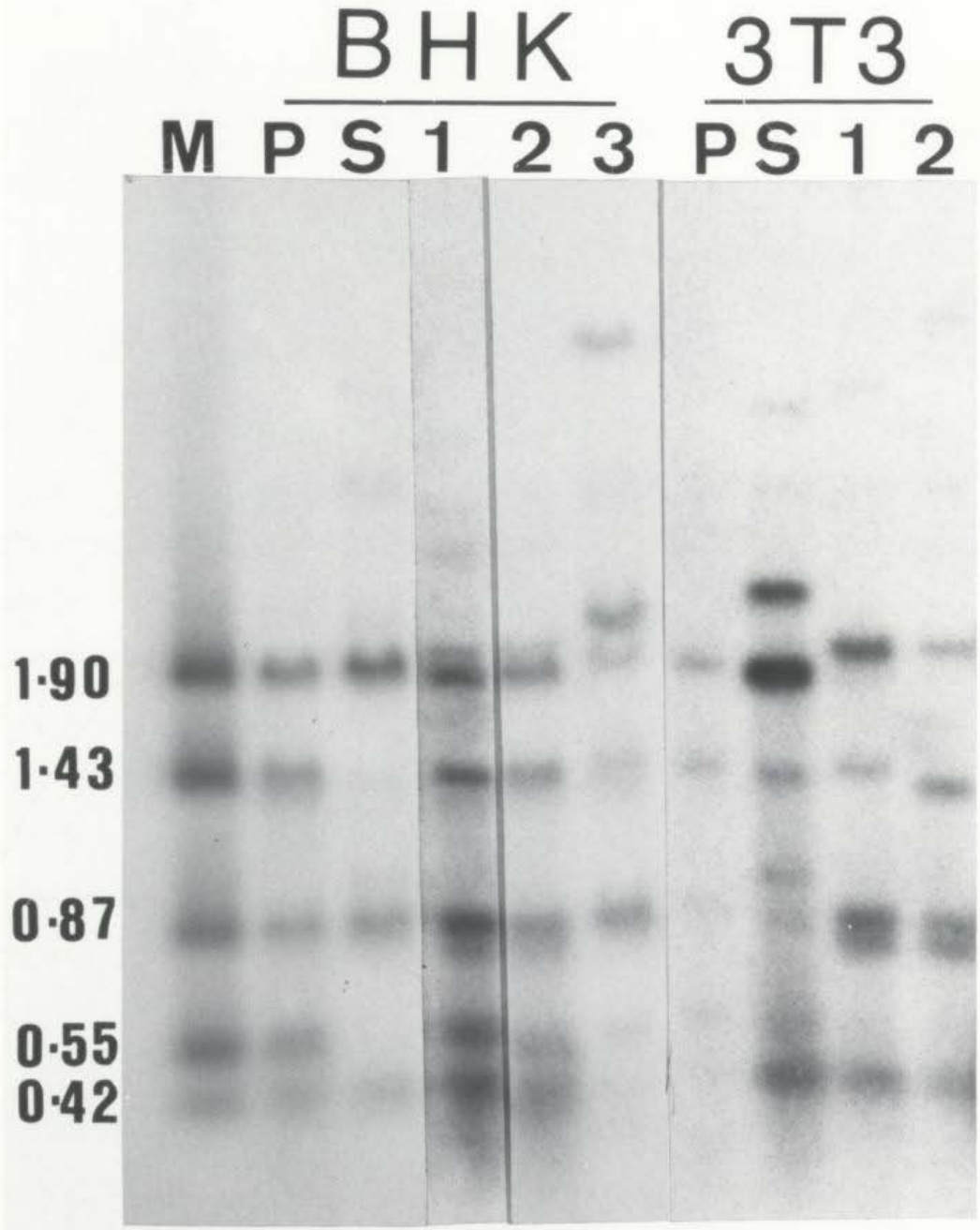
5.3 Results

5.3.1 The role of large T-antigen of BKV in the maintenance of transformation

BK virus transformation of BHK-21 and NIH 3T3 cells to anchorage independence has been shown previously and expression of T-antigen of this virus was shown to be essential for transformation (Tooze, 1981). To determine whether the continued expression of T-antigen is required, I have used antisense RNA to block the synthesis of T-antigen in transformed cells. bHK-21 and NIH 3T3 cells were transfected with a plasmid containing the intact early region of this virus (pBK-pML, Figure 5.2) and plated in 0.28% soft agar for colony formation. Individual colonies were selected and propagated as monolayers. The presence of integrated copies of the intact early region of BKV DNA in the transformed cells was confirmed by Southern blot hybridization of cellular DNA (Figure 5.4 lane P). For both BHK 21 and NIH 3T3 cells, one clone containing one or two copies of viral DNA was selected to be the recipient of the antisense RNA expressing plasmid.

The construction of both antisense and control sense plasmids are as detailed in Figure 5.2. In these plasmids, transcription of the BKV early region (either in a sense or

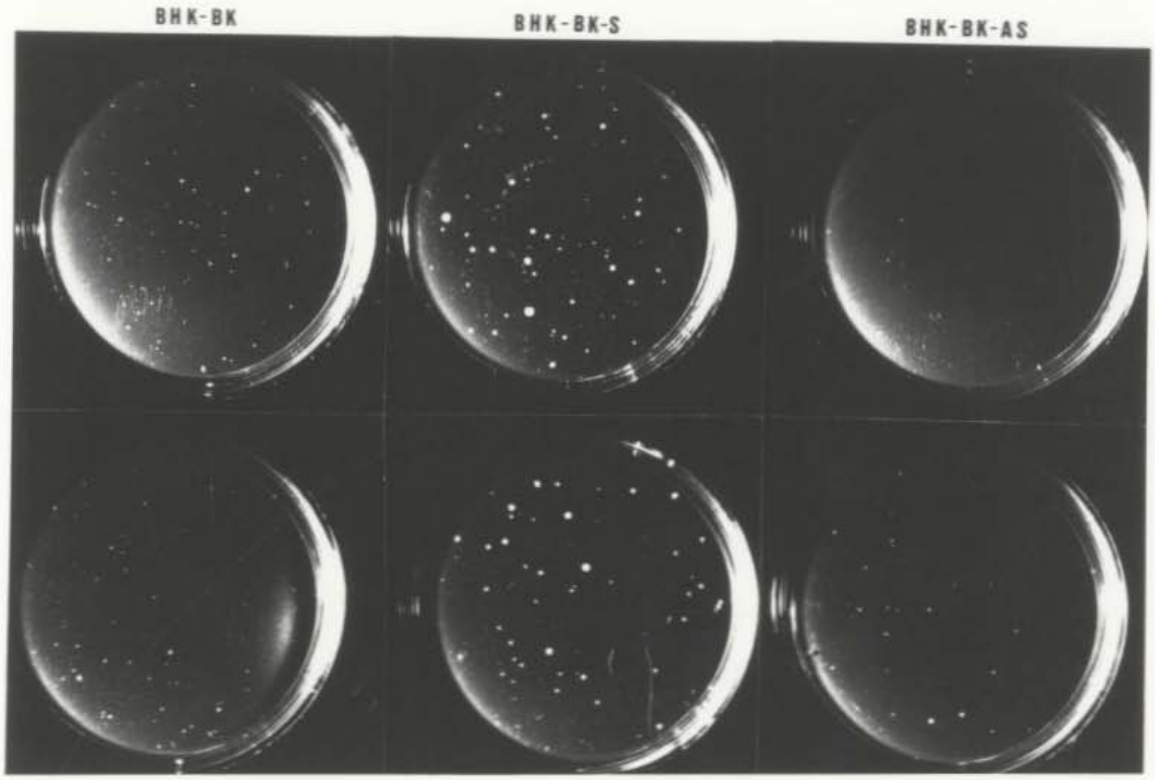
Figure 5.4 Southern blot-hybridization of BKV DNA in transformed, sense and antisense BHK and 3T3 clones. Total cellular DNA (10 μg) digested with a combination of the restriction endonucleases HindIII and EcoRI and using BKV DNA as a probe (Pater *et al*. 1982). **M**: 100 picograms of BKV DNA digested with HindIII and EcoRI. **P**: parental BK transformed clone. **S**: A clone isolated after transfection of the BK-transformed parental clone with pSV_2-neoBK+ followed by selection for neo. Lanes marked 1, 2, and 3 represent three different antisense clones isolated after transfection of the BK transformed parental clone with pSV_2-neo BK- and selection for neo. The position and size of the molecular weights in kilobases are indicated. Cleavage of antisense plasmids with HindIII generates a band of 0.85 kb in addition to the bands detected for the parental clone.


BHK
3T3
M P S 1 2 3 P S 1 2
1·90
1·43
0·87
0·55
0·42

antisense orientation) is under the control of the human metallothionein promoter which is active in rodent cells(Mayo *et al.* 1982). These plasmids also contain pSV_2-neo sequences which confer resistance to the drug G418 (Southern and Berg, 1982). These plasmids were transfected into BK-transformed BHK-21 and NIH 3T3 cells. Twenty BHK-BK antisense clones, fifteen 3T3-BK antisense clones and four sense clones for both cell lines were isolated by selection for G418 resistance.

To study the effect of antisense RNA on the expression of the transformed phenotype, I plated cells from both sense and antisense clones in 0.28% soft agar either in the presence or absence of 50 μM $ZnSO_4$. All sense clones grew in soft agar and formed colonies that were larger than parental clones. Five of twenty BHK-BK and 14 of 15 3T3-BK antisense clones failed to grow in soft agar when incubated in the presence of $ZnSO_4$. Cells from three of the five BHK-BK antisense clones that failed to grow in soft agar were plated in the absence of $ZnSO_4$. All three clones could grow to form colonies in soft agar, but at reduced frequency. This is probably due to inhibition of T-antigen synthesis by the basal level expression of antisense RNA in at least some of the transformed cells. Growth in agar of the parental, one sense and one antisense clone in the presence and absence of $ZnSO_4$ are shown in Figure 5.5 as examples. In a control experiment, clones selected after transfection with pSV_2-neo alone showed the same pattern

Figure 5.5 Soft agar colonies formed by BHK clones. BK-transformed BHK clone (BHK-BK, designated P in Figure 5.4), a clone containing the sense plasmid (BHK-BK-S, designated S in Figure 5.4) and a clone containing the antisense plasmid (BHK-BK-AS, designated 1 in Figure 5.4) were plated in 0.28% soft agar (10^3 cells) in the presence of 50 μM $ZnSO_4$ (top row) and in the absence of $ZnSO_4$ (bottom row). The plates were then incubated at 38.5° for 3 weeks. The number colonies obtained with parental clones and antisense clones in the absence of zinc was ~100 and ~ 20, respectively


BHK-BK
BHK-BK-S
BHK-BK-AS

of growth in agar as the parental and sense clones, indicating that failure of antisense clones to grow in agar is not due to the integration of pSV_2-neo.

To study the status of antisense RNA coding DNA sequences in the antisense clones, total cellular DNA from parental cells, one sense, two antisense clones that failed to grow in soft agar and one antisense clone of BHK-BK that grew in soft agar was analyzed by Southern blotting (Figure 5.4). The change in the orientation of T-antigen coding sequences in antisense plasmids generates a fragment of 0.85 kb when digested with HindIII. Clones that failed to grow in soft agar have this fragment intact (Figure 5.4, for BHK lanes 1 and 2, for 3T3 lanes 1 and 2) in addition to the restriction fragments from parental DNA. Note, however, that the antisense clone that retained its ability to grow in agar did not contain this fragment (Figure 5.4, BHK lane 3). The number of copies of parental BKV DNA in the sense clone appear to have been reduced as the 1.43 and 0.55 kb bands are less intense than in the parental clone (Figure 5.4, BKV compare lanes P with S). Also, some rearrangement was observed in 3T3 antisense clones (Figure 5.4). However, this rearrangement is in a region other than T-antigen coding and BK regulatory region sequences. The 1.9, 0.87 and 0.42 kb fragments representing T-antigen coding and regulatory region sequences remain intact in these cells (Figure 5.4).

To examine the expression and distribution of antisense RNA, nuclear and cytoplasmic RNA from cells that were induced or not induced by Zn^{++} were isolated. To identify the RNA:RNA hybrids, both cytoplasmic and nuclear RNAs were digested with a combination of RNase A and RNase T1 using conditions which selectively cleaves only single stranded RNA. RNase resistant RNA hybrids were then identified by electrophoresis through formaldehyde gels followed by transfer onto nitrocellulose paper and probing with *in vitro* labeled sense and antisense RNA probes. The results of blot hybridization with an antisense transcript probe is shown in Figure 5.6A. An RNase protected fragment of 1.7 kb was observed in antisense clones but not in the sense clones. The double stranded RNA was more predominant in the nucleus than in cytoplasm and was about 15- to 20-fold more abundant in induced cells as compared to non-induced cells. Based on the size of the RNase protected region, it appears that double stranded RNA formation prevented splicing. Further, the amount of antisense RNA transcribed in these cells is sufficient to form hybrids with all of the sense RNA as judged by the absence of sense RNA in polysomes of antisense clones (Figure 5.6B).

I performed immunoprecipitation of ^{35}S-methionine labeled cellular extracts with the monoclonal antibody PAb 108 directed against the SV40 T-antigen which also cross reacts with the BK T-antigen (Gurney *et al.* 1986, Figure 5.7). The amounts

Figure 5.6 (A) Northern blot analysis of double stranded RNA from BHK clones. Nuclear (N) and cytoplasmic (C) RNA from sense (S, Figure 5.4) and antisense (AS, 1 and 2 of Figure 5.4) clones, grown in the presence of Zn^{++}(+) or absence of Zn^{++}(-), were isolated and then digested with RNase A and T1 under a condition that degrades single stranded RNA but spares double-stranded RNA. RNase was eliminated by pronase 120 KBdigestion. The nuclease resistant RNA was subjected to northern blot analysis and probed with [^{32}P]UTP-labeled *in vitro* synthesized antisense transcripts. The probe was prepared using SP6 polymerase and includes the 1.9 kb HindIII B fragment of BKV DNA in an antisense orientation. BKV DNA digested with HindIII was used as a marker (M). **(B) Analysis of poly A$^+$ RNA from the polysomes of BHK clones.** Poly A$^+$ RNA from parental (P), sense (S) and antisense (AS1) clones grown in the presence of Zn^{++} were analysed by northern blotting using nick translated BKV DNA as probe. Hind III digested BK-pML DNA was used as a marker (M).

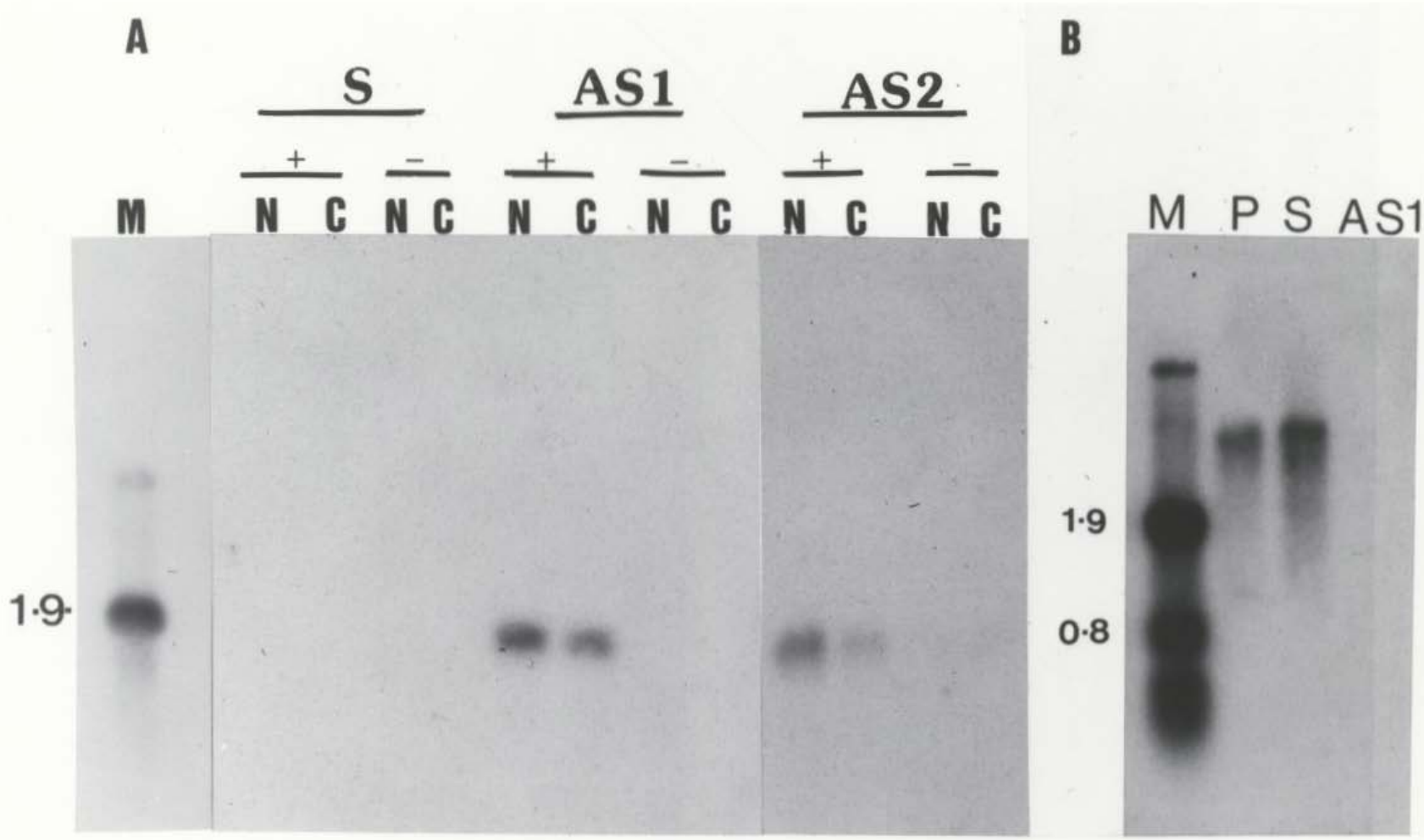
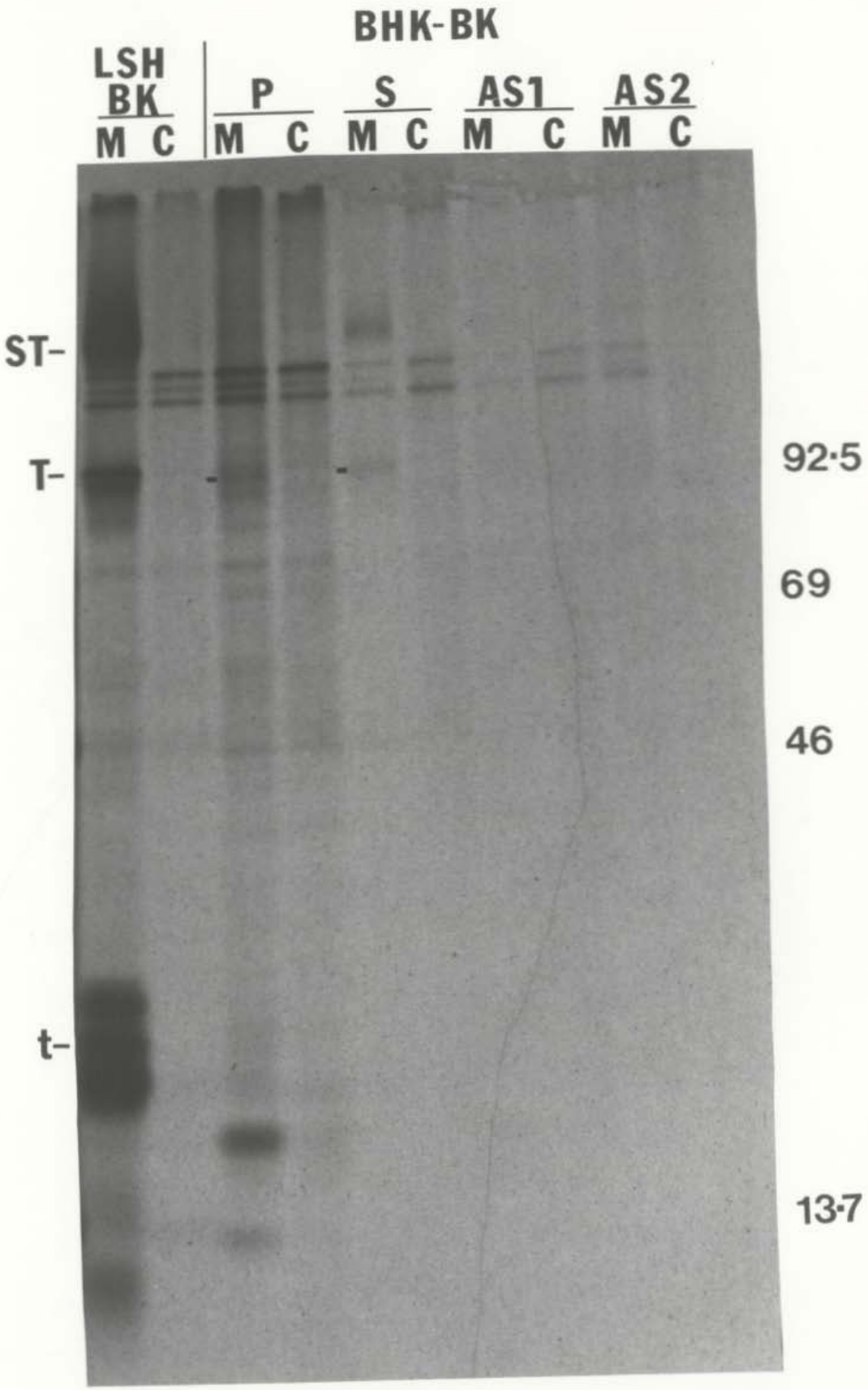

A
S
AS1
AS2
+
−
+
−
+
−
M
N C N C N C N C N C N C
1·9
B
M P S AS1
1·9
0·8

Figure 5.7 Immunoprecipitation of BKV T-antigen from BHK clones. ^{35}S-methionine labeled cellular extracts from the BK-transformed clone (P), the sense (S, Figure 5.4) and the antisense clone (AS1 and AS2, corresponding to clones 1 and 2 of Figure 5.4) were incubated in the presence of normal medium (C) or in the presence of monoclonal antibody PAb 108 (M) containing medium. Immunoprecipitates were run by electrophoresis on SDS-polyacrylamide gels followed by autoradiography. BK-transformed LSH (LSH-BK) was used as a control (Pater *et al.* 1982). The numbers indicate molecular weights of ^{14}C-labeled marker proteins in kilodaltons. **T**, large T-antigen, **t**, small t-antigen. **ST**, super T-antigen


BHK-BK
LSH BK
P
S
AS1
AS2
M C M C M C M C M C
ST-
T-
t-
92·5
69
46
13·7

of T-antigen detected were small, which I find consistently for cell lines as opposed to primary cells. The results do indicate that for BHK transformed cells both the parental and sense clones produced T-antigen. No or very little T-antigen was immunoprecipitated from the antisense clones (Figure 5.7). These results indicate that the failure of cells expressing antisense T-antigen RNA to grow in agar is due to the insufficient synthesis of T-antigen in these cells.

5.3.2 Domains of the BKV T-antigen essential for transformation

To identify the region(s) of BKV T-antigen essential for transformation, I constructed 11 translation termination linker insertion (nonsense) mutants and three in frame deletion mutants. The sites of insertion of the linker and the regions of deletion in T-antigen coding sequences are shown in Figure 5.3. All the nonsense and deletion mutants were transfected into BRK and BHK-21 cells to determine the transformation competence of these plasmids. The results of these experiments are presented in Table 5.1. The mutants with linker insertion in either 5′ or 3′ non-coding regions (H3475 and H827 respectively) could transform both cell types but with reduced efficiency. Plasmids expressing as few as the first 384 amino acids (H1178, H1559, H1651, H1813) transformed both cell lines. Plasmids expressing only the first 266 or fewer amino acids

(H2169, H2313, H3260, H3279) failed to transform. The efficiency of transformation of BRK cells by a mutant with an intact large T-antigen coding sequence but a truncated small t-antigen coding sequence (H2974) was reduced to 35% whereas the efficiency was only 10% with BHK-21 cells. BRK cells transformed by this plasmid showed poor attachment to the tissue culture plates when compared to cells transformed with the wild type plasmid. This indicates that the small t-antigen has some additive function in transformation. Also, BRK cells transformed by plasmid H1813 (expressing the first 384 amino acids) showed an anomolous phenotype. Among the ten foci selected only two could be propagated as monolayer cultures and the cells from these two foci were flat and did not adhere to each other. Therefore, the amino acid sequences essential for T-antigen transformation function in BRK cells are probably localized close to or within this region.

Deletion of the 23 amino acids from position 164 to 187 (H Δ69) did not abolish transformation whereas deletion of the 61 amino acids from position 356 to 417 (H Δ183) and 121 amino acids spanning the region 265-386 (H Δ363) abolished transformation (Table 5.1) These results, together with the results of linker insertion mutagenesis, demonstrate that the amino acid residues in region 356 to 384 (corresponding to nucleotides 1898 to 1813) are essential for the transformation function of BKV T-antigen.

Southern blot analysis of total cellular DNA from one BRK transformed clone for each of the transforming mutants showed the presence of integrated unrearranged copies of BKV T-antigen gene coding DNA (Figure 5.8).

Next, I examined the properties of T-antigen using immunoprecipitation of [^{35}S]methionine labeled cellular extract (Figure 5.9). All the transformed cells tested produced T-antigen. However, to my surprise, truncated T-antigen species showed only slightly greater migration. It is interesting to note that Clark *et al.* (1983) have obtained similar differences in predicted size and mobility of mutant SV40 T-antigens in SDS gels. This could be due to poor resolving capacity of the gel system used in this study. Additionaly, post-translational modifications of T-antigen which influences the mobility of proteins in SDS-polyacrylamide gels may not have been affected by the mutations. The other possibility is that the BRK cells contains suppressor tRNA for TAG stop codon which allows wild type T-antigen synthesis from translation termination linker mutants. However, this is less likely as few of the translation termination linker mutants are non-transforming. Further analysis is possible by immunoprecipitation using monoclonal antibody directed against C-terminus of T-antigen. Note also that a protein of more than 150 KD was immunoprecipitated from transformed cell extracts by the same monoclonal antibody (Figure 5.9, ST). The presence of such

Figure 5.8 Southern blot analysis of viral DNA in transformed BRK cells. Total cellular DNA (5 μg) from transformed clones was digested with restriction endonuclease AvaII and Southern blot-hybridized using BKV DNA as the probe. BKV DNA digested with AvaII was used as a marker. The numbers on the top designate the plasmids used to derive the clones. Numbers on the left indicate molecular weights in kilobases. The fragments with sizes 1.2, 0.9 and 0.4 kb represent BKV T-antigen coding sequences whereas the 1.3 kb fragment is the adjoining BKV and pML sequence which is not present in the marker lane since Pvu II cleaved and electroeluted BKV DNA was used as a marker DNA. The 1.3 kb fragment appears modified in H3751 and H3475 clones.

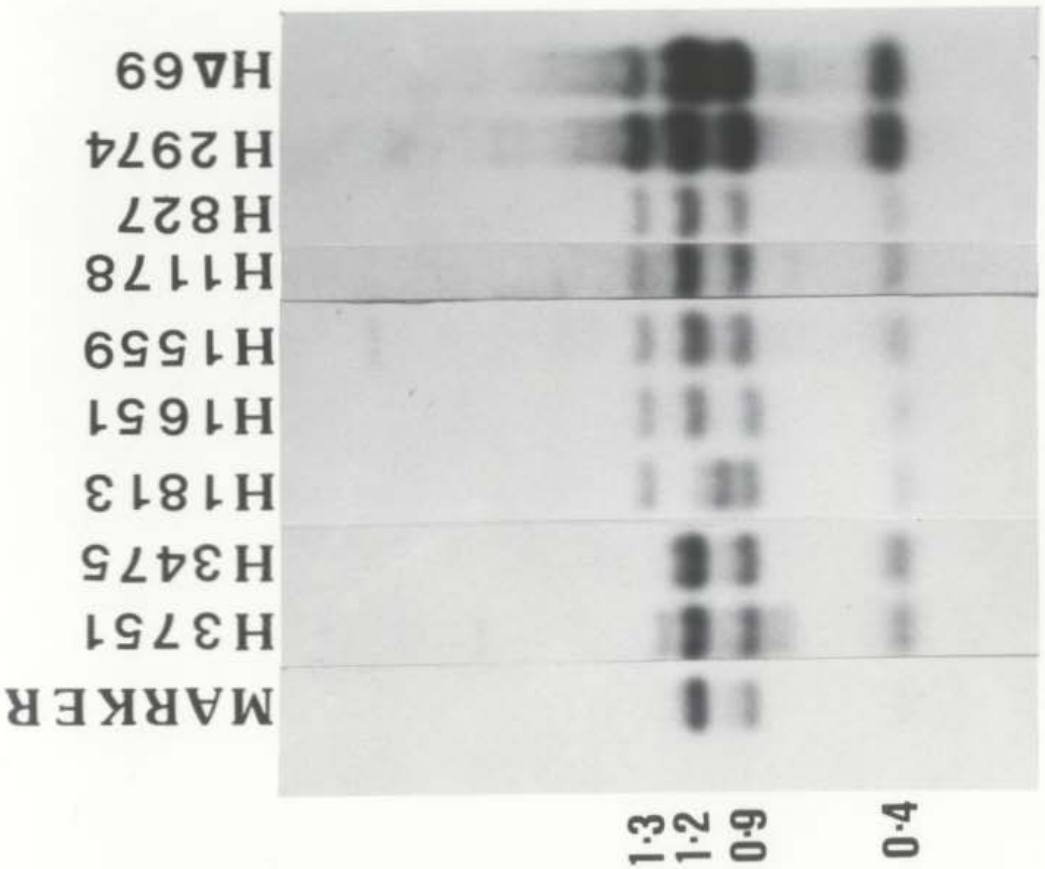
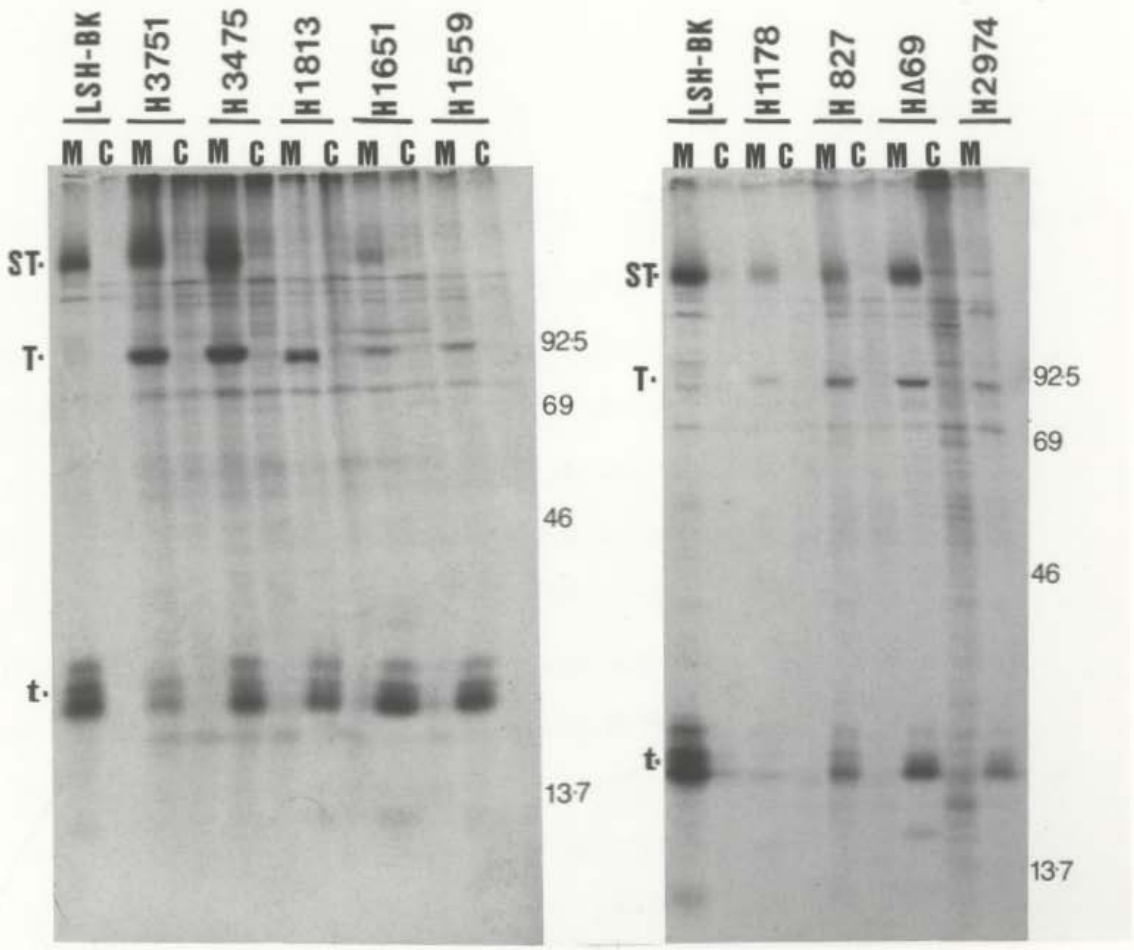

MARKER
H3751
H3475
H1813
H1651
H1559
H1178
H827
H2974
HΔ69
1·3
1·2
0·9
0·4

Figure 5.9 Immunoprecipitation of BKV T-antigen from cells transformed by T-antigen mutants. ^{35}S-methionine labelled cellular extracts from BK transformed BRK cells were incubated in the presence of normal medium (C) or in the presence of the monoclonal antibody PAb 108 (M) containing medium. Immunoprecipitates were run by electrophoresis on 12% SDS-polyacrylamide gels followed by autoradiography. BK-transformed LSH (LSH-BK) was used as a control (Pater *et al*. 1982). The numbers on the top designate the plasmids used to derive the clones. The numbers on the right indicate molecular weights of ^{14}C-labelled marker proteins in kilodaltons. T, large T-antigen; t, small t-antigen; ST, super T-antigen.


LSH-BK
H3751
H3475
H1813
H1651
H1559
M C M C M C M C M C M C
ST·
T·
t·
92·5
69
46
13·7
LSH-BK
H1178
H827
HΔ69
H2974
M C M C M C M C M
ST·
T·
t·
92·5
69
46
13·7

a such a protein does not appear to be unique for BK as a similar protein immunoprecipitated in SV40 transformed BRK cells (Figure 5.10). This protein could be a cellular protein that coprecipitates with the T-antigen or the tetrameric form of the tumor antigens (not dissociated under the denaturing conditions used in these experiments).

The results of Figure 5.9 also suggest that the T-antigen coded by the mutants are stable as there appeared to be no detectable degradation products. Even cells transformed by mutant H1813 with an abnormal phenotype did not have greatly reduced T-antigen protein compared to wild type or other transforming mutants. To further test the stability of mutant T-antigens, pulse chase experiments were used. These revealed that all the tested mutant T-antigens are as stable as wild type T-antigens (Figure 5.11). Since few non-specific proteins immunoprecipitated along with T-antigen, exact quantitation of T-antigen during different puse chase period requires microdensitometric analysis. Note that the turnover rate of small t-antigen is faster than the large T-antigen (Figure 5.11).

5.4 Discussion

In this study, I have presented evidence for the role of BKV T-antigen in the maintenance of the transformed phenotype and have localized the domain of BKV T-antigen essential for transformation. As mentioned earlier, BKV T-antigen shares

Figure 5.10 Immunoprecipitation of T-antigens from SV40 transformed BRK cells. ^{35}S-methionine labelled cellular extracts from SV40 transformed BRK cells were incubated in the presence of normal medium (C) or in the presence of the monoclonal antibody PAb 108 (M) containing medium. Immunoprecipitates were run by electrophoresis on SDS-polyacrylamide gels followed by autoradiography. The numbers on the right indicate molecular weight of ^{14}C-labelled marker proteins in kilodaltons. T, large T-antigens; t, small t-antigen; ST super T-antigen.

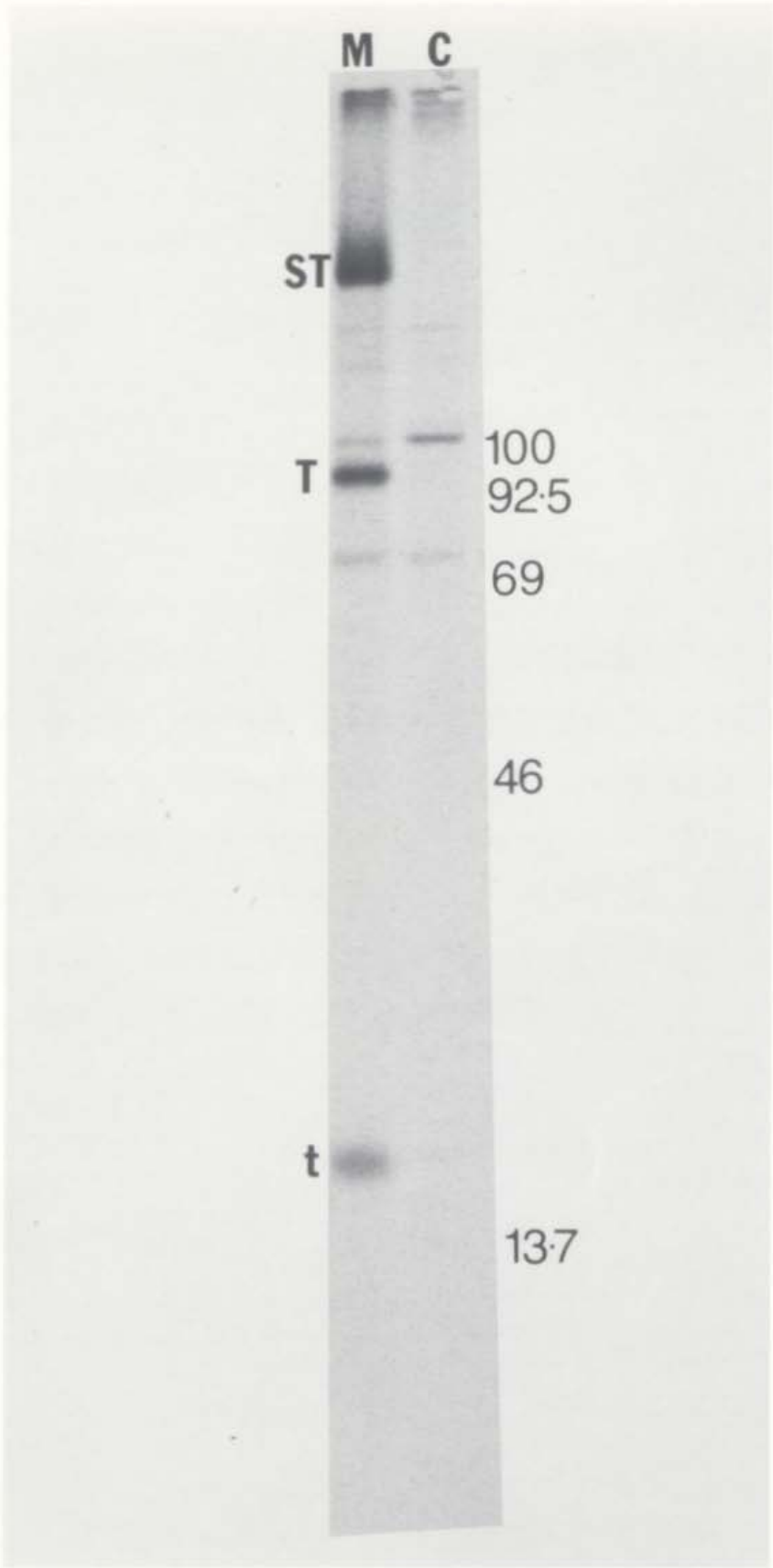
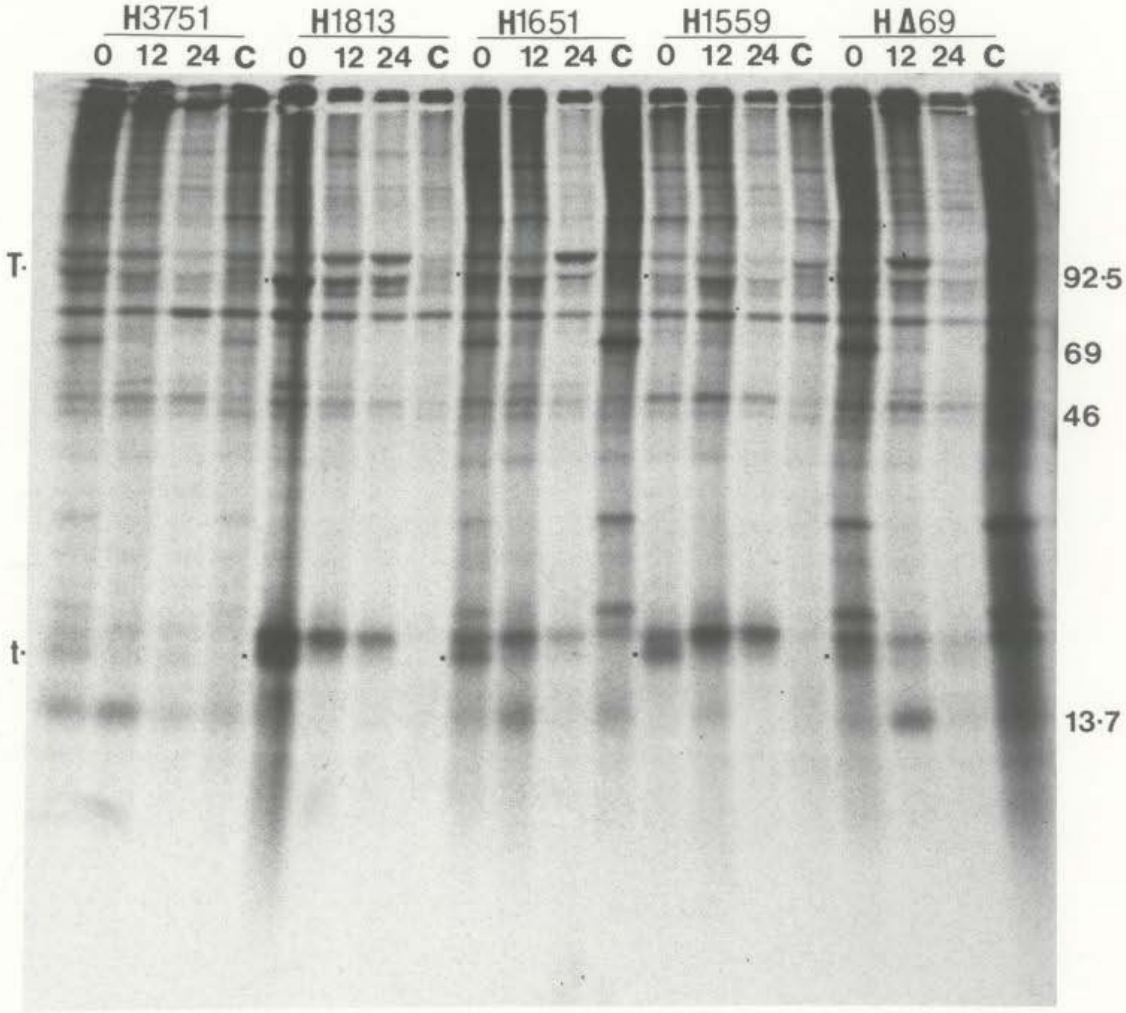

M
C
ST
T
t
100
92·5
69
46
13·7

Figure 5.11 Stability of wild type and mutant BKV T-antigens.

Cells were pulse labeled with ^{35}S methionine containing medium and chased with normal medium for 0, 12 and 24 hours. Immunoprecipitation was performed as described in Figure 5.7. Lane C represents a control in which cellular extracts (0 hours) were immunopreciptated using normal medium.


H3751
0 12 24 C
H1813
0 12 24 C
H1651
0 12 24 C
H1559
0 12 24 C
H Δ69
0 12 24 C
T·
t·
92·5
69
46
13·7

extensive amino acid sequence homology with the transforming protein of another papovavirus, SV40 (529 out of 695 amino acid residues are identical). A large body of useful information on the structure and biological effect of SV40 T-antigen is available (see introduction). The studies involving the role of SV40 T-antigen in the maintenance of the transformed phenotype were greatly aided by the availability of temperature sensitive (tsA) mutants (Brugge and Butel, 1975). However, the results of such experiments are likely to be affected by the cellular response to heat shock. To avoid such ambiguities and also due to the lack of a satisfactory tissue culture system for the isolation of *ts* mutants of BKV, I adopted the antisense RNA approach for such studies (Izant and Weintraub, 1984).

The inducible metallothionein promoter was used to allow a clear demonstration of the role of antisense RNA in the inhibition of transformation as detected by growth in soft agar. Most clones which received antisense plasmids (as determined by independent neo selection) were able to grow in soft agar only when not induced with Zn^{++} (Figure 5.5). The antisense clone which could support soft agar growth was found to contain a deletion of the antisense fragment. In confirmation of the role of antisense RNA, the formation of double stranded RNA (Figure 5.5A) specifically with Zn^{++} induction of antisense RNA was observed. Double stranded RNA

was in an apparently unspliced form and most was in the nucleus. The cytoplasmic double stranded RNA may have been due to some leakage of RNA from the nuclei *in vivo* or during isolation of nuclei. Interference of antisense RNA with T-antigen mRNA function was also clearly demonstrated by the absence of detectable RNA in preparations from polysomes of antisense clones (Figure 5.6B). Finally, the absence of detectable T-antigen after Zn^{++} induction supports the role of BKV T-antigen in the maintenance of transformation. Thus, while SV40 and BKV transformed cells do not share all the same transformation phenotypes, they do share the requirement for continuous synthesis of T-antigen for the maintenance of transformation.

The similarity in function of the tumor antigens of BK and SV40 does not extend to the overall role of the two antigens in transformation. For example, while SV40-transformed HEK cells show anchorage independent growth (Major and Matsumara, 1984), anchorage independent growth of HEK cells requires a combination of BKV DNA and the activated *ras* oncogene (Pater and Pater, 1986). To delineate the differences in transformation by these two proteins, the present results of BKV T-antigen mutants and published results for the SV40 T-antigen can be compared. It appears that amino acids 356 to 384 of both proteins have a significant role in transformation: 1) The majority of tsA mutants of the SV40 T-antigen are located within the sequences for amino acid

residues 272-447 (Lai and Nathans, 1975). The tsA 28, 58, 207, 209, 239, 241, 255 and 276 mutants are located in the coding region for amino acid residues 328-447 whereas tsA 30, 40, 47 and 57 are located in the coding region for amino acid residues 272-328. 2) Clayton *et al.* (1982) have reported an SV40 transformed BALBc/3T3 cell clone in which integration of SV40 DNA into the genomic DNA resulted in the synthesis of a chimeric protein with the first 368 amino acids from the SV40 T-antigen and the rest coded by mouse DNA. 3) Extensive deletion mutational analysis by Pipas *et al.* (1983) demonstrated that amino acids 272-447 appear to be important for the transformation by the SV40 T-antigen. 4) The domain of BKV T-antigen essential for transformation spans from amino acid 356 to 384 (Figure 5.3 and Table 5.1). However, this transformation domain of BKV T-antigen has only 54% amino acid homology to that of SV40 (Seif *et al.* 1979). Since the majority of tsA mutants have mutations in this region of the SV40 T-antigen, it is possible that these amino acids are important for overall stability of the protein. It is arguable that differences in amino acids in this region of BKV T-antigen impart lower stability for the protein. Interestingly, Bollog *et al.* (1989) have observed lower stability for BKV T-antigen than SV40 T-antigen in transformed Rat 2 cells.

The domain of SV40 T-antigen essential for DNA binding (amino acid residues 131 to 259) shows 85% homology to that of BKV. Similarly, the nuclear localization signals of the

two proteins share 100% homology. Both of these properties are essential but not sufficient for transformation (Clark *et al.* 1983; Montenarh *et al.* 1986). This suggests that the limited amino acid divergence in these regions cannot account for the difference in the biological properties of the two antigens in transformation. It should also be noted that mutant H Δ69 contains a deletion of part of the DNA binding domain while retaining transformation competence (Table 5.1), indicating that not all of the domain is required for transformation by BKV DNA.

The SV40 T-antigen complexes with the host anti-oncogene protein p53 (Lane and Crawford, 1979), a property shared by several other DNA tumor viruses such as Adenovirus (Sarnow *et al.* 1982), human papillomavirus 16 (Werness *et al.* 1990) and human JC virus (Haggerty *et al.* 1989). The p53 has a tumor suppression property and it is suggested that tumor antigens of DNA tumor viruses eliminate the tumor suppressing property of p53 by forming an inactive complex (Finely *et al.* 1989; Levine, 1990). Recently, Bollag *et al.* (1989) reported that both SV40 and BK T-antigen form stable complexes with p53 with equal efficiency. However, in my immunoprecipitation with the monoclonal antibody PAb 108 only the BKV T-antigen was precipitated (Figure 5.9). This could be due to the use of different monoclonal antibodies to immunoprecipitate BKV T-antigen or due to a very short life of p53 which results in early degradation particularly in the experimental conditions used in this

study (Lane and Benchimol, 1990). The finding that both BKV and SV40 T-antigens complex with p53 is not surprising, as the p53 binding domain of SV40 T-antigen and the corresponding region of BKV share 80% homology (Figure 5.1). Thus, from previously published results, it can be concluded that the difference in the biological properties of the two T-antigens is not due to differential complexing with p53.

It has been shown recently that the SV40 T-antigen can bind to two other host cell proteins of 105 KD and 120 KD, one of which (105 KD) is the product of the retinoblastoma susceptibility gene (see Introduction). Mutational analysis of the T-antigen revealed that amino acid residues 105 and 120, particularly DLXCXE amino acid residues conserved among DNA tumor viruses antigens, are important for binding of both proteins (Ewen *et al.* 1989; Moran, 1988; Whyte *et al.* 1988). This region of the BKV T-antigen shows 69% homology with SV40 and contains a DLXCXE sequence. Thus it is likely that BKV T-antigen interacts with both Rb and 120 KD protein. However, except for the protein of greater than 150 KD (ST), no protein of 105 KD or 120 KD is observed in any of my immunoprecipitates. This could be due to lack of sensitivity of the assay or ST itself represents either the Rb or 120 KD protein. The latter possibility is likely as: 1) both SV40 and BK transformed cells contain this protein which coprecipitated with the respective T-antigens (Figure 5.10), 2) recent studies have indicated that the Rb protein has greater affinity for the

oligomeric than the monomeric form of SV40 T-antigen (Ludlow *et al.* 1990). Mutants of the SV40 T-antigen lacking the oligomerization domain (amino acid residues 114 to 152 and 591 to 708, Figure 5.1) coprecipitate less Rb protein. The protein ST immunoprecipitated with BKV T-antigen shows a similar pattern, i.e. the amount of protein immunoprecipitated by mutant T-antigens is lower than that immunoprecipitated by the wild type T-antigen (Figure 5.9). This could also provide a possible explanation for the lower transforming ability of mutant T-antigens. The discrepancy in molecular weights of ST and Rb and 120 KD proteins is most likely due to differences in electrophoresis conditions used in this study. Further confirmation of this prediction requires immunoprecipitation with antibodies against the Rb protien. If ST is the rat equivalent of Rb, then the differences in the biological properties of the two papovaviral T-antigens may not be due to a difference in the ability of their antigens to bind the Rb protein. However, I cannot exclude the possibility that the difference in affinity of the two T-antigens for the Rb protein, if any, contributes to the difference in the biological activity of the T-antigen.

The small t-antigens of BK and SV40 show 73% overall homology and 59% homology in their unique carboxy terminal halves (Seif *et al.* 1979). The SV40 small t-antigen has been shown to have some effect on transformation by SV40 either directly or indirectly (Bikel *et al.* 1981). The data presented

here (Table 5.1) also demonstrates that the mutants with a truncated BKV small t-antigen have reduced efficiency in transformation. Tumors induced by small t-antigen mutants of SV40 have been shown to have both enhanced metastatic potential and altered tissue specificity (Dixon *et al.* 1982; Mathews *et al.* 1987). Our preliminary data (not shown) also suggest a similar property for the BKV small t-antigen. Thus, the small t-antigens of both viruses appear to have quite similar functions.

The cellular DNA synthesis (amino acid residues 160-272) and rRNA synthesis (amino acid residues 420-509) domains of SV40 T-antigen have been shown to have some influence on, but are not sufficient for, cellular transformation (Figure 5.1; Soprano *et al.* 1983; Clark *et al.* 1983). Substantial homology has been observed for corresponding regions in the BKV T-antigen. Similar to the SV40 results, from a comparison of the regions required for transformation and the regions required for cellular DNA and rRNA synthesis it is clear that the latter activities are not absolutely necessary for cellular transformation. For example, mutant H1813 which lacks the rRNA synthesis domain and mutant Δ69 which has a deletion in a portion of the cellular DNA synthesis domain are transformation competent.

As mentioned previously, the point mutants of the SV40 T-antigen with reduced ATPase activity are transformation competent whereas deletion of the ATPase domain renders the

T-antigen transformation defective (see Introduction). This domain spans amino acid residues 271 to 449 and 570 to 669 where the carboxy terminal part (570 to 669) has a greater influence on activity (Clark *et al.* 1983). In the BKV T-antigen, a substantial portion of the region corresponding to the carboxy terminal ATPase domain of SV40 (amino acid residues 640 to 661 of SV40 T-antigen) is not present and the remaining portion of this region shows only 46% homology. Although the present data does not directly evaluate the ATPase activity of the BKV T-antigen, the comparative analysis presented above suggests low or no ATPase activity for the BKV T-antigen. In keeping with this argument, I suggest that differences in ATPase activity along with the differences in amino acids in the region of the identified transformation domain (amino acid residues 356 to 384) may account for the variation in the transformation potential of the two proteins. Further experiments with point mutations in the identified domain of the BKV T-antigen and interchanging of carboxy terminal halves of the two proteins are needed to clarify these possibilities.

Previous studies have indicated that cellular homeostasis is maintained by the interaction of positive and negative growth signals (see Introduction). In principle, cellular transformation involves abberation of both of these signals. Since the BKV T-antigen requires cotransfection of the *ras* oncogene for transformation, it is possible that the BKV T-antigen may not contain all activities required to abrogate

both types of signals. A possible model for BKV T-antigen plus *ras* transformed cells would involve the BKV T-antigen functioning to inactivate negative signals of cell growth (by binding to p53 and possibly Rb) while the *ras* oncogene elevates positive signals of cell growth (see Introduction). In contrast, SV40 T-antigen alone can perform both functions as it is capable of interacting with p53 and Rb proteins and inducing *ras* gene expression (Segawa and Yamaguchi, 1987). This model is further supported by the observation that transformation defective deletion mutants of the SV40 T-antigen, most of them with deletions in the nuclear localization signals and the ATPase domain, can be made transformation competent by cotransfection with an activated H-*ras* oncogene (Michalovitz *et al.* 1987). As a consequence of *ras* induction, SV40 T-antigen can elevate positive signals of cell growth by inducing other cellular genes. For example, recent studies have indicated that the gene coding for the transcription factor Sp1 is induced by the SV40 T-antigen (Saffer *et al.* 1990). Additionally, immediate response genes such as *fos* may also be induced by the SV40 T-antigen. Whether similar changes occur in BKV plus *ras* transformed cells is not known. Since BKV alone can induce some degree of transformation (to foci in tissue culture plates), comparative analysis of cellular changes in BK and BK plus *ras* transformed cells may provide some insight into various steps involved in cellular transformation and further evidence for the model stated above.

REFERENCES

Abel, T. and Maniatis, T. (1989) Action of leucine zippers. Nature (London) 341, 24-25.

Adamson, E.D. (1987) Oncogenes in development. Development 99, 449-471.

Akerblom, I.W., Slater, E.P., Beato, M., Baxter, J.D. and Mellon, P.L. (1988) Negative regulation by glucocorticoids through interference with cAMP responsive enhancer. Science 241, 350-353.

Amati, P. (1985) Polyoma regulatory region: A potential probe for mouse cell differentiation. Cell 43, 561-562.

Amemiya, K., Traub, R., Durham, L. and Major, E.O. (1989) Interaction of a nuclear factor-1 like protein with the regulatory region of the human polyomavirus JC virus. The Journal of Biological Chemistry 264, 7025-7032.

Amirhaeri, S., Wohlrab, F., Major, E.O. and Wells, R.D. (1988) Unusual DNA structure in the regulatory region of the human papovavirus JC virus. Journal of Virology 62, 922-931.

Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R.J., Bahnsdorf, H.J., Jonat, C., Heirlich, P. & Karin, M. (1987) Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. Cell 49, 729-739.

Auborn, K.J., Galli, R.L., Dilorenzo, T.P. and Steinberg, B.M. (1989) Identification of DNA-protein interactions and enhancer activity at the 5′ end of the upstream regulatory region in Human papillomavirus type 11. Virology 170, 123-130.

Ayer, D.E. and Dynan, W.S. (1988) Simian virus major late promoter: a novel tripartite structure that includes intragenic sequences. Molecular and Cellular Biology 8, 2021-2033.

Baeuerle, P.A. and Baltimore, D. (1988) Activaton of DNA-binding activity in an apparently cytoplasmic precursor of the NFkB transcription factor. Cell 53, 211-217.

Baker, C.C., Phelps, W.C., Lindgren, V., Braun, M.J., Gonda, M.A. and Howley, P.M. (1987) Structural and

transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. Journal of Virology 61, 1962-971.

Ballou, L.M., Jeno, P. and Thomas, G. (1988) Protein phosphatase 2A activates the mitogen-stimulated S6 kinase from Swiss mouse 3T3 cells. Journal of Biological Chemistry 263, 1188-1194.

Banerji, J., Rusconi, S. and Schaffner, W. (1981) Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell 27, 299-308.

Barbacid, M. (1987) ras Genes. Annual Review of Biochemistry 56, 779-827.

Barrera-Saldana, H., Takahaski, K., Vigneron, M., Wildeman, A., Davidson, I. and Chambon, P. (1985) All six GC-motifs of the SV40 early upstream element contribute to promoter activity *in vivo* and *in vitro*. The EMBO Journal 4, 3839-3849.

Beckemann, A.M., Shah, K.V. and Padgett, B.L. (1982) Propagation and primary isolation of papovavirus JC in epithelial cells derived from human urine. Infection and Immunity 38, 774-777.

Becker, P.B., Ruppert, S. & Schutz, G. (1987) Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors. Cell 51, 435-443.

Bedell, M.A., Jones, K.H. and Laimins, L.A. (1987) The E6 and E7 open reading frames of human papillomavirus type 18 are sufficient for transformation of NIH 3T3 and Rat-1 cells. Journal of Virology 61, 3635-3640.

Beggs, A.H., Frisque, R.J. and Scangos, G.A. (1988) Extinction of JC virus tumour-antigen expression in glial cell-fibroblast hybrids. Proceedings of the National Academy of Sciences USA 85, 7632-3636.

Bender, A. and Spraque, G.F. (1987) Mat αL protein, a yeast transcription activator, binds synergistically with a second protein to a set of cell-type-specific genes. Cell 50, 681-691.

Benezra, R., Davis,R.L., Lockshon, D., Turner, D.L. and Weintraub, H. (1990) The protein Id: A negative regulation of Helix-loop-Helix DNA binding proteins. Cell 61, 49-59.

Berg, D.T., Wells, J.D. and Grinnell, B.W. (1988) A variant enhancer/regulatory region from a cloned human prototype BK virus genome. Nucleic Acids Research 16, 9057.

Berg, J.M. (1989) DNA binding specificity of steroid receptors. Cell 57, 1065-1068.

Berk, A.J. and Schmidt, M.C. (1990) How do transcription factors work? Genes & Development 4, 151-155.

Bernard, B.A., Bailly, C., Lenoir, M-C., Darmon, M., Thierry, F. and Yaniv, M. (1989) The human papillomavirus type 18 (HPV 18) E2 gene product is a repressor of the HPV 18 regulatory region in human keratinocytes. Journal of Virology 63, 4317-4324.

Beth, E., Giraldo, G., Schmidt-Ullrich, R., Pater, M.M., Pater, A. and diMayorca, G. (1981) BK virus-transformed inbred hamster brain cells. 1. Status of the viral DNA and the association of BK virus early antigens with purified plasma membranes. Journal of Virology 40, 276-284.

Bikel, I., Montano, X., Agha, M.E., Brown, M., McCormack, M., Boltax J. and Livingston, D.M. (1987) SV40 small t-antigen enhances the transformation activity of limiting concentration of SV40 large T-antigen. Cell 48, 321-330.

Birkenmeier, E.H., Gwynn, B., Howard, S., Jerry, J., Gordon, J.I., Landschultz, W.H. and McKnight, S.L. (1989) Tissue-specific expression, developmental regulation and genetic mapping of the gene encoding CCAAT/enhancer binding protein. Genes & Development 3, 1146-1156.

Bishop, J.M. (1987) The molecular genetics of cancer. Science 235, 305-311.

Bishop, J.M. (1983) Cellular oncogenes and retroviruses. Annual Review of Biochemistry 52, 301-354.

Blessing, M., Zentgraf, M. and Jorcano, J.L. (1987) Differentially expressed bovine cytokeratin genes. Analysis of gene linkage and evolutionary conservation of 5′ upstream sequences. The EMBO Journal 6, 567-575.

Bodner, M., Castrillo, J-L., Theill, L.E., Deerinck, T., Ellisman, M. and Karin, M. (1988) The pituitary-specific transcription factor GHF-1 is a homeobox-containing protein. Cell 50, 505-518.

Bodner, M. and Karin, M. (1987) A pituitary specific transcription factor can stimulate transcription from the growth hormone promoter in extracts of nonexpressing cells. Cell 50, 267-275.

Bohmann, D., Keller, W., Dale, T., Schöler, H.R., Tebb, G. and Mattaj, I.W. (1987) A transcription factor which binds to the enhancers of SV40, immunoglobulin heavy chain and U2 snRNA genes. Nature (London) 325, 268-272.

Bollag, B., Chuke, W-F. and Frisque, R.J. (1989) Hybrid genomes of the polyomaviruses JC virus, BK virus and Simian virus 40: Identification of sequences important for efficient transformation. Journal of Virology 63, 863-872.

Boshart, M., Gissmann, L., Ikenberg, H., Kleinheinz, A., Scheurlen, W. and zur Hausen, H. (1984) A new type of papillomaviruses DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer. The EMBO Journal 3, 1151-1157.

Boshart, M., Weih, F., Schmidt, A., Fournier, R.E.K. and Schütz. (1990) A cyclic AMP response element mediates repression of tyrosine aminotransferase gene transcription by the tissue-specific extinguisher locus Tsc-1. Cell 61, 905-916.

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

Bradley, M.K., Griffin, J.D. and Livingston, D.M. (1982) Relationship of oligomerization to enzymatic DNA-binding properties of the SV40 large T-antigen. Cell 28, 125-134.

Brady, J., Radonovich, M., Vodkin, M., Natarajan, V., Thoren, M., Das, G., Janik, J. and Salzman, N.P. (1982) Site-specific base substitution and deletion mutations that enhance or suppress transcription of the SV40 major late RNA. Cell 31, 625-633.

Brandl, C.J. and Struhl, K. (1989) Yeast GCN4 transcriptional activator protein interacts with RNA polymerase II *in vitro*. Proceedings of the National Academy of Sciences USA 86, 2652-2656.

Breathnach, R. and Chambon, P. (1981) Organization and expression of eucaryotic split genes coding for proteins. Annual Review of Biochemistry 50, 349-383.

Britten, R.J. and Davidson, E.H. (1969) Gene regulation for higher cells: A theory. Science 165, 349-357.

Brugge, J.S. and Butel, J.S. (1975) Role of simian virus 40 gene A function in maintenance of transformation. Journal of Virology 15, 619-635.

Buratowski, S., Hahn, S., Guarente, L. and Sharp, P.A. (1989) Five intermediate complexes in transcription initiation by RNA polymerase II. Cell 56, 549-561.

Byrne, J.C., Tsao, M-S., Fraser, R.S. and Howley, P.M. (1987) Human Papillomavirus-11 DNA in a patient with chronic laryngotracheobronchial papillomatosis and metastatic squamous-cell carcinoma of the lung. The New England Journal of Medicine 317, 873-878.

Caputo. A., Coralini, A., Grossi, M.P., Balboni, P.G., Negiri, M., Milanesi, G., Federspil, G., Barbanti-Brodano, G. (1983) Episomal DNA of a BK virus variant in a human insulinoma. Journal of Medical Virology 12, 37-49.

Cassill, J.A. and Subramani, S. (1989a) A naturally occuring deletion in the enhancer repeats of the human papovavirus BK optimizes early enhancer function at the expense of late promoter activity. Virology 170, 296-298.

Cassill, J.A., Deyerle, K.L. and Subramani, S. (1989b) Unidirectional and linker scan analysis of the late promoter of the human papovavirus BK. Virology 169, 172-181.

Cedar, H. (1988) DNA methylation and gene activity. Cell 53, 3-4.

Chan, W-K., Klock, G. and Bernard, H-U. (1989) Progesterone and glucocorticoid response elements occur in the long control regions of several human papillomaviruses involved in Anogenital neoplasia. Journal of Virology 63, 3261-3269.

Chan, W-K., Chong, T., Gloss, B., Bernard, H-U. and Klock, G. (1990) Glucocorticoid and progesterone response elements in the enhancer regions of human papillomaviruses involved in anogenital neoplasia. Papillomaviruses (UCLA Symphosia on Molecular and Cellular Biology) 124, 169-178.

Chatis, P.A., Holland, C.A., Silver, J.E., Frederickson, T.N., Hopkins, N. and Hartley, J.W. (1984) A 3' end fragment encompassing the transcriptional enhancers of nondefective friend virus confers erythroleukemogenicity on moloney leukemia virus. Journal of Virology 52, 248-254.

Cherry, J.R., Johnson, T.R., Dollard, C., Shuster, J.R. and Denis, C.L. (1989) Cyclic AMP-dependent protein kinase phosphorylates and inactivates the yeast transcriptional activator ADR1. Cell 56, 409-419.

Chesters, P.M., Heritage, J. and McCance, D.J. (1983) Persistance of DNA sequences of BK virus and JC virus in normal human tissues and in diseased tissues. Journal of Infectious Diseases 147, 676-684.

Chin, M.T., Hirochika, R., Hirochika, H., Broker, T.R. and Chow, L.T. (1988) Regulation of human papillomavirus type 11 enhancer and E6 promoter by activating and repressing proteins from the E2 open reading frame: functional and biochemical studies. Journal of Virology 62, 2994-3002.

Chin, M.T., Broker, T.R. and Chow, L.T. (1989) Identification of a novel constitutive enhancer element and an associated binding protein: Implications for human papillomavirus type 11 enhancer function. Journal of Virology 63, 2967-2976.

Chiu, R., Imagawa, M., Imbra, R.J., Bockoven, J.R. and Karin, M. (1987) Multiple cis-and trans-acting elements mediate the transcriptional response to phorbol esters. Nature (London) 329, 648-651.

Chodosh, L.A., Baldwin, A.S., Carthew, R.W. and Sharp, P.A., (1988) Human CCAAT-binding proteins have heterologous subunits. Cell 33, 11-24.

Chong, T., Chan, W-K. and Bernard, H-U. (1990) Transcription activation of human papillomavirus 16 by nuclear factor 1, AP1, steroid receptors and a possibly a novel transcription factor, PVF: a model for the composition of genital papillomavirus enhancers. Nucleic Acids Research 18, 465-470.

Clark, R., Peden, K., Pipas, J.M., Nathans, D. and Tjian, R. (1983) Biochemical activities of T-antigen proteins encoded by simian virus 40 A gene deletion mutants. Molecular and Cellular Biology. 3, 220-228.

Clayton, C.E., Murphy, D., Lovett, M. and Rigby, P.W.J. (1982) A fragment of the SV40 large T-antigen gene transforms. Nature (London) 299, 59-61.

Clertant, P., Gaudray, P., May, E. and Cuzin, F. (1984) The nucleotide binding site detected by affinity labelling of large T proteins of polyoma and SV40 viruses is distinct from their ATPase catalytic site. Journal of Biological Chemistry 259, 15196-15203.

Cockerill, P.N. and Garrard, W.T. (1986) Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. Cell 44, 273-282.

Cohen, R.B., Shiffery, M. and Kim, C.G. (1986) Partial purification of a nuclear protein that binds to the CCAAT box of the mouse globin gene. Molecular Cellular Biology. 6, 821-832.

Cole, S.T. and Danos, O. (1987) Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome. Phylogeny of papillomaviruses and repeated structure of the E6 and E7 gene products. Journal of Molecular Biology 193, 599-608.

Cooper, J.A. and Whyte, P. (1989) RB and the cell cycle: Entrance or Exit? Cell 58, 1009-1011.

Corbin, V. and Maniatis, T. (1989) Role of transcriptional interference in the Drosophila melanogaster Adh promoter switch. Nature (London) 337, 279-282.

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. and Chambon, P. (1980) Promoter sequences of eukaryotic protein-coding genes. Science 209, 1406-1414.

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

Courey, A.J., Holtzman, D.A., Jackson, S.P. and Tjian, R. (1989) Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. Cell 59, 827-836.

Cripe, T.P., Haugen, T.H. Turk, J.P., Tabatabai, F., Schmid, P. G., Durst, M., Gissmann, L., Roman, A. and Turek, L.P. (1987) Transcriptional regulation of the human papillomavirus-16 E6-E7 promoter by a keratinocyte-dependent enhancer and by a viral E2 trans-activator and

repressor genes products: implications for cervical carcinogenesis. The EMBO Journal 6, 3745-3753.

Dalrymphle, S.A. and Beemon, K.L. (1990) BK virus T-antigens induce kidney carcinomas and thymoproliferative disorders in transgenic mice. Journal of Virology 64, 1182-1191.

Dartmann, K., Schwarz, E., Gissmann, L. and zur Hause, H. (1986) The nucleotide sequence and genome organization of human papillomavirus type 11. Virology 151, 124-130.

Davidson, B.L., Egly, J-M., Mulvihill, E.R. and Chambon, P. (1983) Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. Nature (London) 301, 680-686.

Davidson, I., Fromenthal, C., Augerau, P., Wildeman, A., Zenke, M. and Chambon, P. (1986) Cell-type specific protein binding to the enhancer of simian virus 40 in nuclear extracts. Nature (London) 323, 544-548.

Davidson, I., Xiao, J.H., Rosales, R., Staub, A., & Chambon, P. (1988) The HeLa cell protein TEF-1 bind specifically and cooperatively to two SV40 enhancer motifs of correlated sequence. Cell 54, 931-942.

DeCaprio, J.A., Ludlow, J.W., Figge, J., Shew, J-Y, Huang, C-M, Lee, W-H., Marsilio, E., Paucha, E. and Livingston, D.M. (1988) SV40 large T-antigen forms a specific complex with the product of the retinoblastoma susceptibility gene. Cell 54, 275-283.

DeCaprio, J.A., Ludlow, J.W., Lynch, D., Furukawa, J., Griffin, J., Piwnica-Worms, H., Huang, C-H. and Livingston, D.M. (1989) The product of the retinoblastoma susceptibility gene has properties of a cell cycle regulatory element. Cell 58, 1085-1095.

DeFeo, D., Gonda, M.A., Young, H.A., Chang, E.H., Lowy, D.R., Scolnick, E.M. and Ellis, R.A. (1981) Analysis of two divergent rat genomic clones homologous to the transforming gene of Harvey murine sarcoma virus. Proceedings of the National Academy of Sciences USA 78, 3328-3332.

deGroot, R.P., Schoorlemmer, J., vanGenesen, S.T. and Kruijer, W. (1990) Differential expression of jun and fos genes during differentiation of mouse P19 embryonal carcinoma cells. Nucleic Acids Research 18, 3195-3202.

Derynck, R. (1988) Transforming growth factor α. Cell 54, 593-595.

deVillers, E.M. (1989) Heterogeneity of the human papillomavirus group. Journal of Virology 63, 4898-4903.

Deyerle, K.L., Cassill, J.A. and Subramani, S. (1987) Analysis of the early regulatory region of the human papovavirus BK. Virology 158, 181-193.

Deyerle, K.L., & Subramani, S. (1988) Linker scan analysis of the early regulatory region of the human papovavirus BK. Journal of Virology 62, 3378-3387.

Deyerle, K.L. and Subramani, S. (1989) Human papovavirus BK early gene regulation in nonpermissive cells. Virology 169, 385-396.

DiMaio, D. and Nathans, D. (1982) Regulatory mutants for Simian virus 40: effect of mutations at a T-antigen binding site on DNA replication and expression of viral genes. Journal of Molecular Biology 156, 531-548.

Distal, R.J., Ro, H.S., Rosen, B.S., Groves, D.L. and Spiegelman, B.M. (1987) Nucleoprotein complexes that regulate gene expression in adipocyte differentiation: direct participation of c-fos. Cell 49, 835-844.

Dixon, K., Ryder, B.-J. and Burch-Jaffe, E. (1982) Enhanced metastasis of tumors induced by a SV40 small T deletion mutant. Nature (London) 296, 672-675.

Dorn, A., Bollekins, J., Staub, A., Benoist, C. and Mathis, D. (1987) A multiplicity of CCAAT box-binding proteins. Cell 50, 863-872.

Dorries, K., Johnson, R.T. and ter Meulen, V. (1979) Detection of polyomavirus DNA in PML brain tissue by (*in situ*) hybridization. Journal of General Virology 42, 49-57.

Duesberg, P. (1987) Cancer genes: rare recombinants instead of activated oncogenes (a review) Proceedings of the National Academy of Sciences USA 84, 2117-2124.

Dutta, A., Stocekle, M.Y. and Hanafusa, H. (1990) Serum and V-src increase the level of a CCAAT-binding factor required for transcription from a retroviral long terminal repeat. Genes & Development 4, 243-254.

Dynan, W.S. and Tjian, R. (1985) Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature (London) 316, 863-872.

Dynan, W.S. (1989) Modularity in promoters and enhancers. Cell 58, 1-4.

Dyson, N., Howley, P.M., Münger, E. and Harlow, E. (1989) The human papilloma virus 16 E7 oncoprotein is able to bind to the retinoblastoma gene product. Science 243, 934-936.

Ellis, R.W., DeFeo, D., Shih, T.Y., Gonda, M.A., Young, H.A., Tsuchida, M., Lowy, D.R. and Scolnick, E.M. (1981) The p21 src genes of Harvey and Kirsten sarcoma viruses originate from divergent members of a family of normal vertebrate genes. Nature (London) 292-506-511.

Elsholtz, H.P., Albert, V.R., Treacy, M.N. and Rosenfeld, M.G. (1990) A two-base change in a POU factor-binding site switches pituitary specific to lymphoid-specific gene expression. Genes & Development 4, 43-51.

Epstein, W. and Beckwith, J.R. (1968) Regulation of gene expression. Annual Review of Biochemistry 37, 411-436.

Evans, R.M. and Hollenberg, S.M. (1988) Zinc fingers: gilt by association. Cell 52, 1-3.

Ewen, M.E., Ludlow, J.W., Marsilio, E., DeCaprio, J.A., Millikan, R.C., Cheng, S.H., Paucha, E. and Livingston, D.M. (1989) An N-terminal transformation-governing sequence of SV40 large T-antigen contributes to the binding of both p110Rb and a second cellular protein, p120. Cell 58, 257-267.

Fanning, E., Nowak, B. and Burger, C. (1981) Detection and characterization of multiple forms of simian virus 40 large T antigen. Journal of Virology 37, 92-102.

Fanning, E., Schneider, J., Arthur, A., Höss, A., Moarefi, I. and Modrow, S. (1989) Structure and function of SV40 large T-antigen: communication between functional domains. Current Topics in Microbiology and Immunology 144, 9-22.

Favaloro, J., Treisman, R. and Kamen, R. (1980) Transcription maps of polyoma virus-specific DNA: Analysis by two-dimensional nuclease S1 gel mapping. Methods in Enzymology 65, 718-749.

Feigenbaum, L., Khalili, K., Major, E. and Khoury, G. (1987) Regulation of the host range of human papovavirus JCV. Proceedings of the National Academy of Sciences USA 84, 3695-3698.

Fields (ed.) (1990) Virology (second edition) Raven Press, New York.

Finlay, C.A., Hinds, P.W. and Levine, A.J. (1989) The p53 proto-oncogene can act as a suppressor of transformation. Cell 57, 1083-1093.

Flavell, R.A. (1980) The transcription of eukaryotic genes. Nature (London) 285, 356-357.

Friedmann, T., Doolittle, R.F. and Walter, G. (1978) Amino acid sequence homology between polyoma and SV40 tumour antigens deduced from nucleotide sequences. Nature (London) 274-291-393.

Frisque, R.J., Bream, G.L., & Cannella, M.T. (1984) Human polyomavirus JC virus genome. Journal of Virology 51, 458-469.

Fromental, C., Kanno, M., Nomiyama, H., & Chambon, P. (1988) Cooperativity and hierarchial levels of functional organization in the SV40 enhancer. Cell 54, 943-953.

Gage, J.A., Meyers, C. and Wettstein, F.O. (1990) The E7 proteins of the nononcogenic human papillomavirus type 6b (HPV-6b) and of the oncogenic HPV 16 differ in retinoblastoma protein binding and other properties. 64, 723-730.

Gannon, F., O'Hare, K., Perrin, F., Lepennec, J.P., Benoist, C., Cochet, M., Breathnach, R., Royal, R., Garapin, A., Cami, B. and Chambon, P. (1979) Organisation and sequences at the 5' end of a cloned complete ovalbumin gene. Nature (London) 278, 429-434.

Garcia-Carranca, A., Thierry, F. and Yaniv, M. (1988) Interplay of viral and cellular proteins along the long control region of human papillomavirus type 18. Journal of Virology 62, 4321-4330.

Gauthier-Rouvière, C., Fernandez, A. and Lamb, N.J.C. (1990) ras-induced c-fos expression and proliferation in living rat fibroblasts involves C-kinase activation and the serum response element pathway. The EMBO Journal 9, 171-180.

Gentz, R., Rauscher, J.F. III, Abate, C. and Curran, T. (1989) Parallel association of Fos and Jun leucine zippers juxtaposes DNA binding domains. Science 243, 1695-1699.

George, D.L., Glick, D., Trusko,S. and Freeman, N. (1986) Enhanced c-ki-ras expression associated with Friend virus

integration in a bone marrow-derived mouse cell line. Proceedings of the National Academy of Sciences USA 83, 1651-1655.

Gerster, T., Matthias, D., Thali, M., Jiricny, J. and Schaffner, W. (1987) Cell type specificity elements of the immunoglobulin heavy chain gene enhancer. The EMBO Journal 6, 1323-1330.

Gil, G., Smith, J.R., Goldstein, J.L., Slaughter, C.A., Orth, K., Brown, M.S. and Osborne, T.F. (1988) Multiple genes encode nuclear factor 1-like proteins that bind to the promoter for 3-hydroxy-3-methylglutaryl-coenzyme A reductase. Proceedings of the National Academy of Sciences USA. 85, 8963-8967.

Gill, G. and Ptashne, M. (1988) Negative effect of the transcriptional activator GAL4. Nature (London) 334, 721-724.

Giri, I. and Danos, O. (1986) Papillomavirus genomes: from sequence data to biological properties. Trends in Genetics 2, 227-232.

Gissmann, L., Diehl, V., Schultz-Coulon, H.J. and zur Hausen, H. (1982) Molecular cloning and characterization of human papillomavirus DNA derived from a laryngeal papilloma. Journal of Virology 44, 393-400.

Gissmann, L., Wolnik, L., Ikenberg, H., Koldovsky, U., Schniirch, G.H. and zur Hausen, H. (1983) Human papillomavirus types 6 and 11 DNA sequence in genital and laryngeal papillomas and in some cervical cancers. Proceedings of the National Academy of Sciences USA 80, 560-563.

Gius, D., Grossman, S., Bedell, M.A. and Laimins, L.A. (1988) Inducible and constitutive enhancer domains in the non coding region of human papillomavirus type 18. Journal of Virology 62, 665-672.

Gius, D. and Laimins, L.A. (1989) Activation of human papillomavirus type 18 gene expression by herpes simplex virus type 1 viral transactivators and a Phorbol ester. Journal of Virology 63, 555-563.

Gloss, B., Bernard, H-U., Seedorf, K. and Klock, G. (1987) The upstream regulatory region of the human papilloma virus 16 contains an E2 protein-independent enhancer which is specific for cervical carcinoma cells and regulated by glucocorticoid hormones. The EMBO Journal 6, 3735-3743.

Gloss, B., Chong, T. and Bernard, H.U. (1989a) Numerous nuclear proteins bind the long control region of human papillomavirus type 16: A subset of 6 of 23 DNaseI-protected segments coincides with the location of the cell-type-specific enhancer. Journal of Virology 63, 1142-1152.

Gloss, B., Yeo-Gloss, M., Meisterernst, M., Rogge, L., Winnacker, E. L. and Bernard, H-U. (1989b) Clusters of nuclear factor I binding sites identify enhancers of several papillomaviruses but alone are not sufficient for enhancer function. Nucleic Acid Research 17, 3519-3533.

Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) Recombinant genomes which express Chloramphenicol acetyltransferase in mammalian cells. Molecular Cellular Biology 2, 1044-1051.

Gorman, C.M., Rigby, P.W.J. and Lane, D.P. (1985) Negative regulation of viral enhancers in undifferentiated embryonic stem cells. Cell 42, 579-526.

Goudsmit, J., Wertheim-vanDillen, P., van Strien, A. and van der Noordaa, J. (1982) The role of BK virus in acute respiratory tract disease and the presence of BKV DNA in tonsils. Journal of Medical Virology 10, 91-99.

Gounari, F., deFrancesco, R., Schmitt, J., van der Vliet, P.C., Cortese, R. and Stunnenberg, H. (1990) Amino-terminal domain of NF-1 binds to DNA as a dimer and activates adenovirus replication. The EMBO Journal 9, 559-566.

Graessmann, A., Graessmann, M., Tjian, R. and Topp, W.C. (1980) Simian virus 40 small t protein is required for loss of actin cable network in rat cells. Journal of Virology 33, 1182-1191.

Grainger, J.L. and Winkler, M.M. (1987) Fertilization triggers unmasking of maternal mRNAs in sea urchin eggs. Molecular and Cellular Biology 7, 3947-3954.

Graves, B.J., Johnson, P.F. and McKnight, S.L. (1986) Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. Cell 44, 565-576.

Green, S. and Chambon, P. (1988) Nuclear receptors enhance our understanding of transcription regulation. Trends in Genetics 4, 309-314.

Green, M.R. (1989) When the products of oncogenes and antioncogenes meet. Cell 56, 1-3.

Grinnell, B.W., Berg, D.T., & Walls, J.D. (1988) Negative regulation of the human polyomavirus BK enhancer involves cell-specific interaction with a nuclear repressor. Molecular and Cellular Biology 8, 3448-3457.

Gross, D.S. and Garrard, W.T. (1987) Poising chromatin for transcription. Trends in Biochemical Sciences 12, 293-297.

Gurney, E.G., Tomowski, S. and Deppert, W. (1986) Antigenic binding sites of monoclonal antibodies specific for simian virus 40 large T-antigen. Journal of Virology 57, 1168-1172.

Haggerty, S., Walker, D.L. and Frisque, R.J. (1989) JC virus-simian virus 40 genomes containing heterologous regulatory signals and chimeric early regions: identification of regions restricting transformation by JC virus. Journal of Virology 63, 2180-2190.

Hai, T., Horikoshi, M., Roeder, R.G. and Green, M.R. (1988) Analysis of the role of transcription factor ATF in the assembly of a functional preinitiation complex. Cell 54, 1043-1051.

Hall, A. (1990) ras and GAP - who's controlling whom? Cell 61, 921-923.

Han, S., Levine, M.S. and Manley, J.L. (1989) Synergistic activation and repression of transcription by Drosphila homeobox proteins. Cell 56, 573-583.

Hennighausen, L. and Lubon, H. (1987) Interaction of protein with DNA *in vitro*. Methods in Enzymology 152, 721-735.

Herbst, R.S., Pelletier, M., Boezko, E-M. and Babiss, L.E. (1990) The State of cellular differentiation determines the activity of the adenovirus E1A enhancer element: Evidence for negative regulation of enhancer function. Journal of Virology 64, 161-172.

Herendeen, D.R., Williams, K.P., Kassavetis, G.A. and Geiduschek, E.P. (1990) An RNA polymerase-binding protein that is required for communication between an enhancer and a promoter. Science 248, 573-578.

Heritage, L., Chesters, P.M. and McCance, D.J. (1981) The persistance of papovavirus BK DNA sequences in normal

human renal tissue. Journal of Medical Virology 8, 143-150.

Herr, W. and Clarke, J. (1986) The SV40 enhancer is composed of multiple functional elements that can be compensated for one another. Cell 45, 461-470.

Herr, W., Sturn, R.A., Clerc, R.G., Corcoran, L.M., Baltimore, D., Sharp, P.A., Ingraham, H.A., Rosenfeld, M.G., Finney, M., Ruvkun, G. and Horvitz, H.R. (1988) The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, caenorhabditis elegans unc-86 gene products. Genes & Development 2, 1513-1516.

Herrlich, P. and Ponta, H. (1989) 'Nuclear' oncogenes convert extracellular stimuli into changes in the genetic program. Trends in Genetics 5, 112-116.

Hilberg, F., Stocking, C., Ostertag, W. and Grez, M. (1987) Functional analysis of a retroviral host-range mutant: Altered long terminal repeat sequences allow expression in embryonal carcinoma cells. Proceedings of the National Academy of Sciences USA 84, 5232-5236.

Hirochika, H., Broker, T.R. and Chow, L.T. (1987) Enhancers and trans-acting E2 transcription factors of papillomaviruses. Journal of Virology 61, 2599-2606.

Hirochika, H., Hirochika, R., Broker, T.R. and Chow, L.T. (1988) Functional mapping of the human papillomavirus type 11 transcriptional enhancer and its interaction with the trans-acting E2 proteins. Genes & Development 2, 54-67.

Hope, I.A., Mahadevan, S. and Struhl, K. (1988) Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. Nature (London) 336, 635-640.

Horikoshi, M., Hai, T., Lin, Y.S., Green M.R. and Roeder, R.G. (1988) Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. Cell 54, 1033-1042.

Horikoshi, M., Wang, C.K., Fujii, H., Cromlish, J.A., Weil, P.A. and Roeder, R.G. (1989) Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. Nature (London) 341, 299-303.

Horowitz, J., Yandell, D.W., Park, S-H, Canning, S., Whyte, P., Buchkovich, K., Harlow, E., Weinberg, R.A.and Dryja,

T.P. (1989) Point mutational inactivation of the retinoblastoma antioncogene. Science 243 937-940.

Houff, S.A., Major, E.O., Katz, D.A., Kufta, C.V., Sever, J.L., Pittaluga, S., Roberts, J.R., Gitt, J., Saini, N. and Lux, W. (1988) Involvement of JC virus-infected mononuclear cells from the bone marrow and spleen in the pathogenesis of progressive multifocal leukoencephalopathy. The New England Journal of Medicine 318, 301-305.

Howley, P.M. (1990) Papillomavirinae and their replication. In: Virology (second edition), edited by Fields, B.N. Raven Press, New York, p. 1625-1650.

Huang, H.C., Sundseth, R. and Hansen, U. (1990) Transcription factor LSF binds two variant bipartite sites within the SV40 late promoter. Genes & Development 4, 287-298.

Imler, J.L., schatz, C., Wasylyk, C., Chatton, B. and Wasylyk, B. (1988) A harvey-ras responsive transcription element is also responsive to a tumour-promoter and to serum. Nature (London) 337- 275-278.

Izant, J G. and Weintraub, H. (1984) Inhibition of thymidine kinase gene expression by antisense RNA: A molecular approach to genetic analysis. Cell 36, 1007-1015.

Johnson, P.F., Landschulz, W.H., Gravis, B.J. & McKnight, S.L. (1987) Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. Genes & Development 1, 133-146.

Johnson, P.F. and McKnight, S.L. (1989) Eukaryotic transcriptional regulatory proteins. Annual Review of Biochemistry 58, 799-839.

Jones, K.A., Yamamoto, K.R. and Tjian, R. (1985) Two distinct transcription factors bind to the HSV thymidine kinase promoter *in vitro*. Cell 42, 559-572.

Jones, K.A., Kadonaga, J-T., Rosenfeld, M.G., Kelly, T.J. and Tjian, R. (1987) A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell 48, 79-89.

Jones, N.C., Rigby, P.W.J., & Ziff, E.B. (1988) Trans-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses. Genes & Development 2, 267-281.

Jones, N. (1990) Transcriptional regulation by dimerization: two sides to an incestuous relationship. Cell 61, 9-11.

Kadonaga, J.T. and Tjian, R. (1986) Affinity purification of sequence specific DNA binding proteins. Proceedings of the National Academy of Sciences USA 83, 5889-5893.

Kadonaga, J.T., Carner, K.R., Masiarz, F.R. and Tjian, R. (1987) Isolation of cDNA encoding transcription factor Sp1 and functional analysis for the DNA-binding domain. Cell 51, 1079-1090.

Kageyama, R. and Pastan, I. (1989) Molecular cloning and characterization of a human DNA binding factor that represses transcription. Cell 59, 815-825.

Kamholz, J., Toffenetti, J., Lazzarini, R.A. (1988) Organization and expression of the human mylin basic protein gene. Journal of Neuroscience Research 21, 62-70.

Kanno, M., Fromental, C., Staub, A., Ruffenach, F., Davidson, I. and Chambon, P. (1989) The SV40 TC-II (κB) and the related H-2K^{B} enhansons exhibit different cell type specific and inducible proto-enhancer activities, but the SV40 core sequence and the AP-2 binding site have no enhanson activities. The EMBO Journal 8, 4205-4214.

Karin, M., Haslinger, A., Holtgreve, H., Richards, R.I., Krauter, P., Westphal, H.M. and Beato, M. (1984) Characterization of DNA sequences through which cadmium and glucocorticoid hormones induce metallothionein-IIA gene. Nature (London) 308, 513-519.

Keller, J.M. and Alwine, J.C. (1985) Analysis of an activatable promoter: sequences in the simian virus 40 late promoter required for T-antigen-mediated transactivation. Molecular and Cellular Biology 5, 1859-1869.

Kenney, S., Natarajan, V., Strike, P., Khoury, G. and Salzman, N.P. (1984) JC virus enhancer-promoter active in human brain cells. Science 226, 1337-1339.

Khalili, K., Rappaport, J. and Khoury, G. (1988) Nuclear factors in human brain cells bind specifically to the JCV regulatory region. The EMBO Journal 7, 1205-1210.

Kim, S.K. and Wold, B.J. (1985) Stable reduction of thymidine kinase activity in cells expressing high levels of antisense RNA. Cell 42, 129-138.

Kimura, G. and Dulbecco, R. (1973) A temperature-sensitive mutant of simian virus 40 affecting transforming ability. Virology 52, 529-534.

Klein, G. (1987) The approaching era of the tumor suppressor genes. Science 238, 1539-1545.

Knudson, A.G. (1985) Heriditory cancer, oncogenes and antioncogenes. Cancer Research 45, 1437-1443.

Kouzarides, T. and Ziff, E. (1988) The role of the leucine zipper in the fos-jun interaction. Nature (London) 336, 646-651.

Kristie, T.M. and Roizman, B. (1988) Differentiation and DNA contact points of host proteins binding at the cis site for virion mediated induction of αgenes of herpes simplex virus 1. Journal of Virology 62, 1145-1157.

Kristie, T.M., LeBowitz, J.H. and Sharp, P.A. (1989) The octamer-binding proteins form multi-protein-DNA complexes with the HSV αTIF regulatory protein. The EMBO Journal 8, 4229-4238.

Kryszke, M.H., Piette, J. and Yaniv, M. (1987) Induction of a factor that binds to the polyoma virus A enhancer on differentiation of embryonal carcinoma cells. Nature (London) 328, 254-246.

Kulderon, D., Richardson, W., Markham, A. and Smith, A. (1984) Sequence requirements for nuclear localisation of SV40 large-T. Nature (London) 3211, 33-38.

Lai. C.J. and Nathans, D. (1975) A map of temperature-sensitive mutants of simian 40. Virology 66, 70-81.

Laimins, L.A., Khoury, G., Gorman, C., Howard, B. and Gruss, P. (1982) Host-specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. Proceedings of the National Academy of Sciences USA 79, 6453-6457.

Lancaster, W.D. and Olson, C. (1982) Animal papilloma-viruses. Microbiological Reviews 46, 191-207.

Land, H., Parada, L.F. and Weinberg, R.A. (1983) Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. Nature 304, 596-602.

Lane, D.P. and Crawford, L.V. (1979) T-antigen is bound to a host protein in SV40 transformed cells. Nature (London) 278, 261-263.

Lane, D.P. and Benchimol, S. (1990) p53: Oncogene or anti-oncogene? Genes & Development 4, 1-8.

Lanschulz, W.H., Johnson, P.F. and McKnight, S.L. (1988) The leucine zipper: a hypothetical structure common to a new class of DNA-binding proteins. Science 240, 1759-1764.

Lanschulz, W.J., Johnson, P.F. and McKnight, S.L. (1989) The DNA binding domain of the rat liver nuclear protein c/EBP is bipartite. Science 243, 1681-1688.

Lassar, A.B., Buskin, J.N., Lockshon, D., David, R.L., Apone, S., Hauschka, S.D. and Weintraub, H. (1989) MyoD is a sequence-specific DNA binding protein requiring a region of *myc* homology to bind to the muscle creatine kinase enhancer. Cell 58, 823-831.

Lee, W., Haslinger, A., Karin,M. & Tjian, R. (1987) Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. Nature (London) 325, 368-372.

Lenardo, M., Pierce, J.W. and Baltimore, D. (1987) Protein-binding sites in Ig gene enhancers determine transcriptional activity and inducibility. Science 236, 1573-1577.

Lenardo, M.J., Staudt, L., Robbins, P., Kuang, A., Mulligan, R.C. and Baltimore, D. (1989) Repression of the IgH enhancer in teratocarcinoma cells associated with a novel octamer factor. Science 243, 544-546.

Levine, M. and Hoey, T. (1988) Homeobox proteins as sequence-specific transcription factors. Cell 50, 537-540.

Levine, A.J. (1990) Tumour suppressor genes. BioEssays 12, 60-65.

Levine, M. and Manley, J.L. (1989) Transcriptional repression of eukaryotic promoters. Cell. 59, 405-408.

Lichsteiner, S. and Schibler, U. (1989) A glycosylated liver-specific transcription factor stimulates transcription of the albumin gene. Cell 57, 1179-1187.

Lillie, J.W. and Green, M.R. (1989) Activator's target in sight. Nature (London) 341, 279-280.

Linzer, D.I.H. and Levine, A.J. (1979) Characterization of a 54K dalton cellular antigen present in SV40 transformed cells and uninfected embryonal carcinoma cells. Cell 17, 43-52.

Livingston, D.M. and Bradley, M.K. (1987) The simian virus large T-antigen: A lot packed into a little. Molecular Biology Medicine 4, 63-80.

Loeber, G. and Dorries, K. (1988) DNA rearrangements in organ-specific variants of polyomavirus JC strains GS. Journal of virology 62, 1730-1735.

Loeber, G., Parsons, R. and Tegtmeyer, P. (1989) A genetic analysis of the zinc finger of SV40 large T-antigen. Current Topics in Microbiology and Immunology 144, 21-30.

Loeken, M., Bikel, I., Livingston, D.M. and Brady, J. (1988) Trans-activation of RNA polymerase II and III promoters by SV40 small t antigen. Cell 55, 1171-1177.

Ludlow, J.W., DeCaprio, J.A., Huang, C M., Lee, W-H, Paucha, E. and Livingston, D.M. (1989) SV40 large T-antigen binds preferentially to an underphosphorylated member of the retinoblastoma susceptibility gene product family. Cell. 56, 57-65.

Ludlow, J.W., Shon, J., Pipas, J.M., Livingston, D.M. and DeCaprio, J. (1990) The retinoblastoma susceptibility gene product undergoes cell cycle-dependent dephosphorylation and binding to and release from SV40 large T. Cell 60, 387-396.

Lusky, M. and Botchan, M.R. (1984) Characterization of the bovine papillomavirus plasmid maintenance sequences. Cell 36, 391-401.

Ma, J. and Ptashne, M. (1987) A new class of yeast transcriptional activators. cell 51, 113-119.

Macchi, M., Bornert, J.M., Davidson, I., Kanno, M., Rosales, R., Vigneron, M., Xiao, J-H, Fromental, C. and Chambon, P. (1989) The SV40 TC-II (κB) enhanson binds ubiquitous and cell type specifically inducible nuclear proteins from lymphoid and non-lymphoid cell lines. The EMBO Journal 8, 4215-4227.

Major, E.O. and Matsumura, P. (1984) Human embryonic kidney cells: Stable transformation with an origin-defective simian virus 40 DNA and use as hosts for human papovavirus replication. Molecular and Cellular Biology 4, 379-382.

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

Maniatis, T., Goodbourn, S. and Fischer, J.A. (1987) Regulation of inducible and tissue-specific gene expression. Science 236, 1231-1245.

Manley, J.L., Fire, A. Cano, A., Sharp, P.A. and Gefter, M.L. (1980) DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. Proceedings of the National Academy of Sciences USA 77, 3855-3859.

Manos, M.M. and Gluzman, Y. (1985) Genetic and biochemical analysis of transformation-competent, replication defective simian virus 40 large T-antigen mutants. Journal of Virology 53, 120-127.

Mark, J. (1989) Many gene changes found in Cancer. Science 246, 1386-1388.

Markowitz, R.B. and Dynan, W.S. (1988) Binding of cellular proteins to the regulatory region of BK virus DNA. Journal of Virology 62, 3388-3398.

Marshall, C.J. (1987) Oncogenes and growth control 1987. Cell 49, 723-725.

Marshall, T., Pater, A. and Pater, M.M. (1989) Transregulation and differential cell specificity of Human papillomavirus types 16, 18 and 11 cis-acting elements. Journal of Medical Virology 29, 115-126.

Martin, J.D., King, D.M., Slauch, J.M. and Frisque, R.J. (1985) Differences in regulatory sequences of naturally occuring JC virus variants. Journal of Virology 53, 306-311.

Mastui, T., Wakefield, L.M., Lechner, J.F., Lavick, M.A., Sporn, M.B. and Harris, C-C. (1986) Type beta transforming growth factor is the primary differentiation-inducing serum factor for normal human bronchial epithelial cells. Proceedings of the National Academy of Sciences USA 83, 2438-2442.

Masui, T., Segall, J., Weil, P.A. and Roeder, R.G. (1980) Multiple factors required for accurate initiation of

transcription by purified RNA polymerase II. The Journal of Biological Chemistry 255, 11992-11996.

Matthews, B.J., Levine, A.S. and Dixon, K. (1987) Deletion mutations in the small t-antigen gene alter the tissue specificity of tumors induced by simian virus 40. Journal of Virology 61, 1282-1285.

Matlashewski, G., Schneider, J., Banks, L., Jones, N., Murray, A. and Crawford, L. (1987) Human papillomavirus type 16 DNA cooperates with activated *ras* in transforming primary cells. The EMBO Journal 6, 1741-1746.

Matsukura, T., Kanda, T., Furuno, A., Yoshikawa, H., Kawana, T. and Yoshiike, K. (1986) Cloning of monomeric human papillomavirus type 16 DNA integrated within cell DNA from a cervical carcinoma. Journal of Virology 58, 979-982.

Maxam, A.M. and Gilbert, W. (1980) Sequencing end-labelled DNA with base-specific chemical cleavages. Methods in Enzymology. 65, 499-560.

Mayo K.E., Warren, R. and Palmiter, R.D. (1982) The mouse metallothionein-1 gene is transcriptionally regulated by cadmium following transfection into human and mouse cells. Cell 29, 99-108.

McBurney, M.W. and Rogers, B.J. (1982) Isolation of male embryonal carcinoma cells and their chromosome replication patterns. Developmental Biology 89, 503-503.

McCormick, F. (1989) ras GTPase activating protein: signal transmitter and signal terminator. Cell 56, 5-8.

McKnight, S. and Tjian, R. (1986) Transcriptional selectivity of viral genes in mammalian cells. Cell 46, 795-805.

McVey, D., Brizuela, L., Mohr, I., Marshak, D.R., Gluzman, Y. and Beach, D. (1989) Phosphorylation of large tumor antigen by cdc2 stimulates SV40 DNA replication. Nature 341, 503-507.

Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Research 12, 7035-7056.

Mermod, N., Williams, T.J. and Tjian, R. (1988) Enhancer binding factors AP-4 and AP-1 act in concert to activat.

SV40 late transcription *in vitro*. Nature (London) 557-561.

Mermod, N., O'Neill, E.A., Kelly, T.J. and Tjian, R. (1989) The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. Cell 58, 741-753.

Meyer, M.E., Gronemeyer, H., Turcotte, B., Bocquel, M-T, Tasset, D. and Chambon, P. (1989) Steroid hormone receptors compete for factors that mediate their enhancer function. Cell 57, 433-442.

Michalovitz, D., Fisher-Fantuzzi, L., Vesco, C., Pipas, J.M. and Oren, M. (1987) Activated Ha-ras can cooperate with defective simian virus 40 in the transformation of nonestablished rat embryo fibroblasts. Journal of Virology 61, 2648-2654.

Mitchell, P.J., Wang, C. and Tjian, R. (1987). Positive and negative regulation of transcription *in vitro*: Enhancer-binding protein AP-2 is inhibited by SV40 T-antigen. Cell 50, 847-861.

Mitchell P.J. and Tjian, R. (1989) Transcriptional regulation in mammalian cells by sequence- specific DNA binding proteins. Science 24, 371-378.

Montenarh, M., Vesca, C., Kemmerling, G., Muller, D. and Henning, R. (1986) Regions of SV40 large T-antigen necessary for oligomerization and complex formation with the cellular oncoprotein p53. FEBS Letters 204, 51-56.

Moran E. (1988) A region of SV 40 T-antigen can substitute for a transforming domain of the adenovirus Ela products. Nature (London) 334, 168-170.

Moyle, M., Lee, J.S. Anderson, W.F. and Ingles, J.C. (1989) The c-terminal domain of the largest subunit of RNA polymerase II and transcription initiation. Molecular and Cellular Biology 9, 5750-5753.

Müller, M.M., Ruppert, S., Schaffner, W. and Matthias, P. (1988) A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. Nature (London) 336, 544-551.

Muller, P.R. and Wold, W. (1989) *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. Science 246, 780-786.

Mullér-strom, H-P, Sogo, J.M. and Schaffner, W. (1989) An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. Cell 58, 767-778.

Münger, K., Werness, B.A., Dyson, N., Phelps, W.E., Harlow, E. and Howley, P.M. (1989) Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. The EMBO Journal 8, 4099-4106.

Murre, C., McCaw, P.S. and Baltimore, D. (1989) A new DNA binding and dimerizaiton motif in Immunoglobulin enhancer binding, daughterless, MyoD and myc proteins. Cell 56, 777-783.

Nelbock, P., Dillon, P.J., Perkins, A. and Rosen, C.A. (1990) A cDNA for a protein that interacts with the human immunodeficiency virus TAT transactivator. Science 248, 1650-1653.

Nomiyama, H., Fromental, C., Xiao, J.H. and Chambon, P. (1987) Cell-specific activity of the constituent elements of the simian virus 40 enhancer. Proceedings of the National Academy of Sciences USA 84, 7881-7885.

Ohkuma, Y., Horikoshi, M., Roeder, R.G. and Desplan, C. (1990) Binding site-dependent direct activation and repression of *in vitro* transcription by drosophila homeodomain proteins. Cell 61, 475-484.

Ondek, B., Shepard, A. and Herr, W. (1987) Discrete elements within the SV40 enhancer region display different cell-specific enhancer activities. The EMBO Journal 6, 1017-1025.

Ondek, B., Gloss, L. and Herr, W. (1988) The SV40 enhancer contains two distinct levels of organization. Nature (London) 333, 40-45.

Oren, M., Maltzman, W. and Levine, A.J. (1981) Post-translational regulation of the 54 K cellular tumor antigen in normal and transformed cells. Molecular and Cellular Biology 1, 101-110.

Pabo, C.O. and Sauer, R.T. (1984) Protein-DNA recognition. Annual Review of Biochemistry 53, 293-321.

Padgett, B.L. and Walker, D.L. (1973) Prevalence of antibodies in human sera against JC virus, an isolate from a case of progressive multifocal leucoencephalopathy. Journal of Infectious Diseases 127, 667-670.

Padgett, B.L., Walker, D.L., ZuRhein, G.M., Hodach, A.E. and Chow, S.M. (1976) JC papovavirus in progressive multifocal leukoencephalopathy. Journal of Infectious Diseases 133, 686-690.

Padgett, B.L., Rogers, C.M. and Walker, D.L. (1977a) JC virus, a human polyomavirus associated with progressive multifocal leukoencephalopathy: additional biological characteristics and antigenic relationships. Infection and Immunity 15, 656-662.

Padgett, B.L., Walker, D.L., ZuRhein, G.M. and Varakis, J.M. (1977b) Differential neuro-oncogenecity of strain of JC virus, a polyoma virus, in newborn syrian hamster. Cancer Research 37, 718-720.

Pagnani, M., Negrini, M., Reschiglian, P., Corallini, A., Balboni, P.G., Schernck, S., Macino, R., Milanesi, G. and Barbanti-Brodano, G. (1986) Molecular and biological properties of BK virus IR, a BK virus variant isolated from a human tumor. Journal of Virology 59, 500-505.

Pallas, D.C., Shahrik, L.K., Martin, B.L., Jaspers, S., Miller, T.B., Brautigan, D.L. and Roberts, T.M. (1990) Polyoma small and middle T-antigens and SV40 small t antigen form stable complexes with protein phosphatase 2A. Cell 60, 167-176.

Paonessa, G., Gounari, F., Frank, R. and Cortese, R. (1988) Purification of a NF-1-like DNA-binding protein from rat liver and cloning of the corresponding cDNA. The EMBO Journal 7, 3115-3123.

Pater, A., Pater, M.M. and diMayorca, G. (1980) Arrangement of the genome of the human papovavirus, RF virus. Journal of Virology 36, 480-487.

Pater, M.M., Pater, A., diMayorca, G., Beth, E. and Giraldo, G. (1982) BK-virus transformed inbred hamster brain cells: Status of viral DNA in subclones. Molecular and Cellular Biology. 2, 837-844.

Pater, A., Pater, M.M., Chang, L-S, Slawin, K. and diMayorca, G.C. (1983) Multiple origins of the complementary defective genomes of RF and origin proximal sequences of GS, two human papovavirus isolates. Virology 131, 426-436.

Pater, M.M. and Pater, A. (1985) Human papillomavirus types 16 and 18 sequences in carcinoma cell lines of the cervix. Virology 145, 313-318.

Pater, A. and Pater, M.M. (1986) Transformation of primary human embryonic kidney cells to anchorage independence by a combination of BK virus DNA and the Harvey-ras oncogene. Journal of Virology 58, 680-683.

Pater, M.M., Hughes, G.A., Hyslop, D.E., Nakshatri, H. and Pater, A. (1988a) Glucocorticoid-dependent oncogenic transformation by type 16 but not type 11 human papillomavirus DNA. Nature (London) 335, 832-835.

Pater, A., Gardner, H., Respler, D.S., Jahn, A. and Pater, M.M. (1988b) Isolation and characterization of a varient of human papillomavirus type 11 from a nasal inverting (Schneiderian) papilloma. Journal of Medical Virology 25, 149-156.

Pater, A. and Pater, M.M (1988a) Expression of the control elements of BK and SV40 viruses in human cells exhibiting different transformed phenotypes. Virology 163, 625-628.

Pater, M.M. and Pater, A. (1988b) Expression of human papillomavirus type 16 and 18 DNA sequences in cervical carcinoma cell lines. Journal of Medical Virology 26, 185-195.

Pater, M.M., Marshall, T., Nakshatri, H. and Pater, A. (1989) Glucocorticoid-dependent transformation of primary cells by human papillomavirus type 16 DNA. Papillomaviruses (UCLA Symposia on Molecular and Cellular Biology) 124, 313-321.

Paucha, E., Kalderon, D., Harvey, R.W. and Smith, A.E. (1986) Simian virus origin DNA-binding domain on large T-antigen. Journal of Virology 57, 50-64.

Pecorano, G., Morgan, D. and Defendi, V. (1989) Differential effects of human papillomavirus type 6, 18 and 18 DNAs on immortalization and transformation of human cervical epithelial cells. Proceedings of the National Academy of Sciences USA 86, 563-567.

Pelham, H.R.B. (1982) A regulatory upstream promoter element in the Drosphila HSP70 heat-shock gene. Cell 517-528.

Peterson, M.G., Tanese, N., Pugh, B.F. and Tjian, R. (1990) Functional domains and upstream activation properties of cloned human TATA binding protein. Science 248, 1625-1630.

Phelps, W.C. and Howley, P.M. (1987) Transcriptional transactivation by the human papillomavirus type 16 E2 gene product. Journal of Virology 61, 1630-1638.

Phelps, W.C., Yee, C.L., Munger, K. and Howley, P.M. (1988) The human papillomavirus type 16 E7 gene encode transactivator and transformation functions to those of adenovirus E1a. Cell 53, 539-547.

Pietenpol, J.A., Holt, J.J., Stein, R.W. and Moses, H.L. (1990a) TGFβ1 suppression of c-myc gene transcription: role in inhibition of keratinocyte differentiation. Proceedings of the National Academy of Sciences USA 87, 3758-3762.

Pietenpol, J.A., Stein, R.W., Moran, E., Yaciuk, P., Schlegel, R., Lyons, R.M., Pittelkow, M.R., Munger, K., Howley, P.M. and Moses, H.L. (1990b) TGF-β1 inhibition of c-myc transcription and growth in keratinocytes is abrogated by viral transforming proteins with pRB binding domains. Cell 61, 777-785.

Pipas, J.M., Peden, K.W.C. and Nathans, D. (1983) Mutational analysis of simian virus 40 T-antigen: Isolation and characterization of mutants with deletions in the T-antigen gene. Molecular and Cellular Biology 3, 203-213.

Pirsi, L., Yasumoto, S., Feller, M., Doniger, J. and Dipaolo, J.A. (1987) Transformation of human fibroblasts and keratinocytes with human papillomavirus type 16 DNA. Journal of Virology 61, 1061-1066.

Prives, C., Covey, L., Scheller, A. and Gluzman, Y. (1983) DNA-binding properties of simian virus 40 T-antigen mutants defective in viral DNA replication. Molecular and Cellular Biology 3, 1958-1966.

Prives, C. (1990) The replication functions of SV40 T-antigen are regulated by phosphorylation. Cell 61, 735-738.

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

Ptashne, M. (1988) How eukaryotic transcriptional activators work. Nature (London) 335, 683-689.

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

Purchio, A.F. and Fareed, G.C. (1979) Transformation of human embryonic kidney cells by human papovavirus BK. Journal of Virology 29, 763-769.

Ransone, L.J., Visvadr, J., Wamsley, P. and Verma, I.M. (1990). Trans-dominant negative mutants of Fos and Jun. Proceedings of the National Academy of Sciences USA 87, 3806-3810.

Raymondjean, M., Cereghini, S. and Yaniv, M. (1988) Several distinct "CCAAT" box binding proteins coexit in eukaryotic cells. Proceedings of the National Academy of Sciences USA 85, 757-761.

Respler, D.S., Jahn, A., Pater, A. and Pater. M.M. (1987) Isolation and characterization of papillomavirus DNA from nasal inverting (Schneiderian) papillomas. Annals of Otology, Rhinology & Laryngology 96, 170-173.

Robertson, M. (1988) Homeoboxes, POU proteins and the limits of promiscuity. Nature (London) 336, 522-524.

Rockford, R., Campbell, B.A. and Villarrial, L.P. (1987) A pancreas specificity results from the combination of polyomavirus and Moloney murine leukemia virus enhancer. Proceedings of the National Academy of Sciences USA 84, 449-453.

Ronde, A.D., Mannens, M., Slater, R.M., Hoovers, J., Heyting, C., Bluker-wagemakers, E.M., Leschot, N.J., Stein, A.V., Sol, C.J.A., Schegget, J.T. and van der Noordaa, J. (1988) Morphological transformation by early region human polyomavirus BK DNA of human fibroblasts with deletions in the short arm of one chromosome 11. Journal of General Virology 69, 467-471.

Rosales, R., Vigneron, M., Macchi., M., Davidson, I., Xiao, J.H. and Chambon, P. (1987) *In vitro* binding of cell-specific and ubiquitous nuclear proteins to the octamer motif of the SV40 enhancer and related motifs present in other promoters and enhancers. The EMBO Journal 6, 3015-3025.

Rosenthal, N., Kress, M., Gruss, P. and Khoury, G. (1983) BK viral enhancer element and a human cellular homology. Science 222, 749-755.

Rossi, P., Karsenty, G., Roberts, A.B., Roche, N.S., Sporn, M.B. and Crombrugghe, B. (1988) A nuclear factor 1 binding site mediates the transcriptional activation of a type 1 collagen promoter by transforming growth factor-β. Cell 52, 405-414.

Rotenberg, M.O., Chow, L.T. and Broker, T.R. (1989) Characterization of rare human papillomavirus type 11 mRNAs coding for regulatory and structural proteins, using the polymerase chain reaction. Virology 172, 489-497.

Rudnicki, M.A. and McBurney, M.W. (1987) Cell culture methods and induction of differentiation of embryonal carcinoma cell lines. In: E-J. Robertson (ed.), Teratocarcinoma and embryonic stem cells: a practical approach. IRL press, Oxford, 19-49.

Rundell, K., Major, E.O. and Lampert, M. (1981) Association of cellular 56,000 and 32,000 molecular weight proteins with BK virus and polyoma virus t-antigen. Journal of Virology 37, 1090-1093.

Saffer, J.D., Jackson, S.P. and Thurston, S.J. (1990) SV40 stimulates expression of the trans-acting Sp1 at the mRNA level. Genes & Development 4, 659-666.

Sager, R. (1989) Tumour suppressor genes: the puzzle and the promise. Science 246, 1406-1412.

Samuels, M., Fire, A. and Sharp, P.A. (1982) Separation and characterization of factors mediating accurate transcription by RNA polymerase II. The Journal of Biological Chemistry 257, 5321-5326.

Santoro, C., Mermod, N., Andrews, P.C. and Tjian, R. (1988) A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature (London) 334, 218-224.

Sarkar, G. and Sommer, S.S. (1989) Access to a messenger RNA sequence or its protein product is not limited by tissue or species specificity. Science 244, 331-334.

Sarnow, P. Ho. Y.S., Williams, J. and Levine, A.J. (1982) Adenovirus E1b-58 Kd tumor antigen and SV40 large tumor antigen are physically associated with the same 54 Kd cellular protein in transformed cells. Cell 28, 387-39[illegible].

Sassone-Corsi, P., Wildeman, A. and Chambon, P. (1985) A trans-acting factor is responsible for the simian virus 40 enhancer activity *in vitro*. Nature (London) 313, 458-463.

Sassone-Corsi, P., Lamph, W.W., Kamps, M. and Verma, I.M. (1988) fos-associated cellular p39 is related to nuclear transcription factor AP1. Cell 54, 553-560.

Sawadgo, M. and Roeder, R.G. (1985) Interaction of a gene specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell 43, 165-175.

Schaffner, W. (1989) How do different transcription factors binding the same DNA sequence sort out their jobs? Trends in Genetics 5, 37-39.

Schartz, C. and Chatton, B. (1990) Stereoalignment requirements for activation of transcription by the simian 40 enhancer. Nucleic Acid Research 18, 421-427.

Scheidereit, C., Heguy, A. and Roeder, R.G. (1987) Identification and purification of a human lymphoid-specific octamer binding protein (OTF-2) that activates transcription of an immunoglobulin promoter *in vitro*. Cell 51, 783-793.

Schirm, S., Jiricny, J. and Schaffner, W. (1987) The SV40 enhancer can be dissected into multiple segments, each with a different cell type specificity. Genes and Development 1, 65-74.

Schlegel, R., Phelps, W.C., Zhang, Y-L and Barbosa, M. (1988) Quantitative keratinocyte assay detects two biological activities in human papillomavirus DNA and identifies viral types associated with cervical carcinoma. The EMBO Journal 7, 3181-3187.

Schmidt, M.C., Zhou, Q. and Berk, A.J. (1989) Sp1 activates transcription without enhancing DNA-binding activity of the TATA box factor. Molecular and Cellular Biology 9, 3299-3307.

Schneider-Gadicke, A. and Schwarz, E. (1986) Different human cervical carcinoma cell lines show similar transcription patterns of human papillomavirus type 18 early genes. The EMBO Journal 5, 2285-2292.

Schüle, R., Umesono, K., Mangelsdorf, D.J., Bolado, J., Pike, J.W. and Evans, R.M. (1990) Jun-Fos and receptors for vitamin A and D recognize a common response element in the human osteocalcin gene. Cell 61, 497-504.

Schwartz, L.B. and Roeder, R.G. (1975) Purification and subunit structure of deoxyribonucleic acid-dependent ribonucleic acid polymerase II from the mouse plasmacytoma, MOPC 315. The Journal of Biological Chemistry 250, 3221-3228.

Schwarz, E., Freese, U.K., Gissmann, L., Mayer, W., Roggenbuck, B., Strimlace, A. and Zur Hausen, H. (1985) Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. Nature (London) 314, 111-114.

Sealey, L. and Chalkley, R. (1987) At least two nuclear proteins bind specifically to the rous sarcoma virus long terminal repeat enhancer. Molecular and Cellular Biology. 7, 787-798.

Seedorf, K., Krammer, G., Durst, M., Suhai, S. and Rowkamp, W.G. (1985) Human papillomavirus type 16 DNA sequence. Virology 145, 181-185.

Segawa, K. and Yamaguchi, N. (1987) Induction of c-Ha-ras transcription in rat cells by simian virus 40 Large T-antigen. Molecular and Cellular Biology 7, 556-559.

Seif, I., Khoury, G. and Dhar, R. (1979) The genome of human papovavirus BKV. Cell 18, 963-977.

Shah, K.V. and Howley, P.M. (1990) Papillomaviruses. In: Virology (Second edition), edited by Fields, B.N. Raven Press, New York. 1651-1676.

Shaw, P.E., Schroter, H. and Nordheim, A. (1989) The ability of a ternary complex to form over the serum responsive element correlates with Serum inducibility of the human c-fos promoter. Cell 56, 563-572.

Shinohara, T., Mastuda, M., Yasui, K. and Yoskiike, K. (1989) Host range bias of the JC virus mutant enhancer with DNA rearrangement. Virology 170, 261-263.

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

Silberstein, G.B. and Daniel, C.W. (1987) Reversible inhibition of mammary gland growth by transforming growth factor-beta. Science 237, 291-293.

Simmons, D.T., Takemoto, K.K. and Martin, M.A. (1977) Relationship between the methionine tryptic peptides of simian virus 40 and BK virus tumor antigens. Journal of Virology 24, 319-325.

Simon, M.C., Fisch, T.M., Benecke, B.J., Nevins, J.R. and Heintz, N. (1988) Definition of multiple, functionally distinct TATA elements, one of which is a target in the hsp70 promoter for E1A regulation. Cell 52, 723-729.

Singh, H., LeBowitz, J.G., Baldwin, A.S., Jr. and Sharp, P.A. (1988) Molecular cloning of an enhancer binding protein: isolation by screening of an expression library with a recognition site DNA. Cell 52, 415-423.

Sleigh, M.J. (1987) Differential regulation of viral and cellular genes in F9 mouse embryonal carcinoma cells. Nucleic Acids Research 15, 9379-9393.

Sleigh, M.J., Lockett, T.J., Kelly, J. and Lewy, D. (1987) Competition Studies with repressors and activators of viral enhancer function in F9 mouse embryonal carcinoma cells. Nucleic Acids Research 15, 4307-4324.

Smale, J.A., Khoury, G., Jay, G., Howley, P.M. and Scangos, G.A. (1986) Early regions of JC virus and BK virus induce distinct and tissue-specific tumors in transgenic mice. Proceedings of the National Academy of Sciences USA 83, 8288-8292.

Smale, S.T. and Baltimore, D. (1989) The "initiator" as a transcription control element. Cell 57, 103-113.

Smotkin, D. and Wettstein, F.O. (1986) Transcription of human papillomaviruses type 16 early genes in a cervical and a cancer-derived cell line and identification of the E7 protein. Proceedings of the National Academy of Sciences USA 83, 4680-4684.

Smotkin, D., Prokoph, H. and Wettstein, F.O. (1989) Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms. Journal of Virology 63, 1441-1447.

Sompayrac, L. and Danna, K.J. (1984) Less than 40% of the simian virus 40 large T-antigen-coding sequences is required for transformation. Molecular and Cellular Biology 4, 1661-1663.

Soprano, K.J., Galanti, N., Jonak, G.J., McKercher, S., Pipas, J.M., Peden, K.W.C. and Baserga, R. (1983) Mutational analysis of simian 40 T-antigen: Stimulation of cellular DNA synthesis and activation of rRNA genes by mutants with deletions in the T-antigen gene. Molecular and Cellular Biology. 3, 214-219.

Sopta, M., Zachary, F.B. and Greenblatt, J. (1989) Structure and associated DNA-helicase activity of a general transcription initiation factor that bind to RNA polymerase II. Nature (London) 341, 410-414.

Southern, E.M. and Berg, P. (1982) Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. Journal of Molecular and Applied Genetics 1, 327-341.

Spandidos, D.A. and Wilkie, N.M. (1984) Malignant transformation of early passage rodent cells by a single mutated human oncogene. Nature (London) 310, 469-475.

Speck, N.A. and Baltimore, D. (1987) Six distinct nuclear factors interact with the 75-base-pair repeat of the Moloney murine leukemia virus enhancer. Molecular and Cellular Biology. 7, 1101-1110.

Stehelin, D., Varmus, H.E., Bishop, J.M. and Vogt, P.K. (1976) DNA related to the transforming gene(s) of avian sarcoma viruses are present in normal avian DNA. Nature (London) 260, 170-173.

Steinberg, B.M., Auborn, K.J., Brandsma, J.L. and Taiehman, L.B. (1989) Tissue site-specific enhancer function of the upstream regulatory region of human papillomavirus type 11 in cultured kirotinocytes. Journal of Virology 63, 957-960.

Stern, S., Tanaka, M. and Herr, W. (1989) The Oct-1 homeodomain directs formation of a multiprotein-DNA complex with the HSV transactivator VP16. Nature (London) 341, 624-630.

Stoler, M.H., Wolinsky, S.M., Whitbeck, A., Broker, T.R. and Chow, L.T. (1989) Differentiation-linked human papillomavirus types 6 and 11 transcription in genital condylomata revealed by *in situ* hybridization with messenger-specific RNA probes. Virology 172, 331-340.

Storey, A., Pim, D., Murray, A., Osborn, K., Banks, L. and Crawford, L. (1988) Comparison of the *in vitro* transforming activities of human papilloma types. The EMBO Journal 7, 1815-1820.

Stringer, J.R. (1982) Mutant of simian virus 40 large T-antigen that is defective for viral DNA synthesis, but competent for transformation of cultured rat cells. Journal of Virology 42, 854-864.

Struhl, K. (1986) Constitutive and inducible saccharoymces cervisiae promoters: evidence for two distinct molecular mechanisms. Molecular and Cellular Biology 6, 3847-3853.

Struhl, K., Brandl, C.J., Chen, W., Harbury, P.A.B., Hope, I.A. and Mahadevan, S. (1988) Transcriptional activation by yeast GCN4, a functional homolog to the jun oncoprotein. Cold Spring Harbor Symposium on Quantitative Biology LIII, 701-710.

Struhl, K. (1989) Helix-turn-Helix, zinc finger and leucine zipper motifs for eukaryotic transcriptional regulatory proteins. Trends in Biochemical Sciences 14, 137-140.

Sturn, R.A. and Herr, W. (1988) The POU domain is a bipartite DNA-binding structure. Nature (London) 336, 601-604.

Sugimoto, G., Hara, K., Taguchi, F. and Yogo, Y. (1989) Growth efficiency of naturally occuring BK virus variants *in vivo* and *in vitro*. Journal of Virology 63, 3195-3199.

Sukhatme, V.P., Cao, X., Chang, L.C., Tsai-Morris, C.H., Stamenkoyich, D., Ferreira, P.C.P., Cohen, D.R., Edwards, S.A., Shows, T.B., Curran, T., LeBeau, M.M. and Adamson, E.D. (1988) A zinc finger-encoding gene co-regulated with c-fos during growth and differentiation, and after cellular depolarization. Cell 53, 37-43.

Swift, F.V., Bhat, K., Younghusband, B. and Hamada, H. (1987) Characterization of a cell-specific enhancer found in the human papillomavirus type 18 genome. The EMBO Journal 6, 1339-1344.

Tabin, C.J., Bradley, S.M., Bargmann, C.I., Weinberg, R.A., Papageorge, A.G., Scolnick, E.M., Dhar, R., Lowy, D.M. and Chang, E.H,. (1982) Mechanism of activation of a human oncogene. Nature (London) 300, 143-149.

Tack, L.C., Wright, J.H. and Gurney, E.G. (1988) Characterization of simian virus 40 large T-antigen by using different monoclonal antibodies: T-p53 complexes are preferentially ATPase active and adenylated. Journal of Virology 62, 1028-1037.

Tada, H., Lashgari, M., Rappaport, J. and Khalili, K. (1989) Cell type-specific expression of JC virus early promoter is determined by positive and negative regulation. Journal of Virology 63, 463-466.

Takemoto, K.K., Howley, P.M. and Miyamura, T. (1979a) JC human papovavirus replication in human amnion cells. Journal of Virology 30, 384-389.

Takemoto, K.K., Linke, H., Miyamura, T. and Fareed, G.C. (1979b) Persistent BK papovavirus infection of transformed human fetal brain cells. Journal of Virology 29, 1177-1185.

Tamura, T., Miura, M., Ikenaka, K. and Mikoshiha, K. (1988a) Analysis of transcriptional control elements of the mouse myelin basic protein gene in HeLa cell extracts: demonstration of a strong NF-1 binding motif in the upstream region. Nucleic Acid Research 16, 11441-11459.

Tamura, T., Inoue, T., Nagata, K. and Mikoshiba, K. (1988b) Enhancer of human polyoma JC virus contains nuclear factor 1 binding sequences; Analysis using mouse brain nuclear extracts. Biochemical and Biophysical Research Communications 157, 419-425.

Taparowsky, K., Shimizu, K., Goldfarb, M. and Wigler, M. (1983) Structure and activation of the human N-ras gene. Cell 34, 581-586.

Thierry, F. and Yaniv, M. (1987) The BPV1-E2 trans-acting protein can be either an activator or a repressor of the HPV 18 regulatory region. The EMBO Journal 6, 3391-3397.

Thierry, F., Heard, J.M., Dartmann, K. and Yaniv, M. (1987) Characterization of a transcriptional promoter of human papillomavirus 18 and modulation of its expression by simian virus 40 and adenovirus early antigens. Journal of Virology 61, 134-142.

Timblin, C., Battey, J. and Kuehl, W.M. (1990) Application for PCR technology to substractive cDNA cloning: identification of genes expressed specifically in murine plasmacytoma cells. Nucleic Acids Research 18, 1587-1593.

Tooze, J. (ed.) (1981) DNA tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Tornow, J., Polvino-Bodnar, M., Santangelo, G. and Cole, C.N. (1985) Two separate functional domains of simian virus 40 large T-antigen: Carboxyl-terminal region of simian virus 40 large T-antigen is required for efficient capsid protein synthesis. Journal of Virology 53, 415-424.

Trahey and McCormick, F. (1987) A cytoplasmic protein stimulates normal N-ras p21 GTPase but does not affect oncogenic mutants. Science 238, 542-545.

Triesman, J., Gonczy, P., Vashishtha, M., Harris, E. and Desplan, C. (1989) A single amino acid can determine the DNA binding specificity of homeodomain proteins. Cell 59, 553-562.

Trimble, W.S., Johnson, P.W., Hozumi, N. and Roder, J.C. (1986) Inducible cellular transformation by a metallothionin-ras hybrid oncogene leads to natural killer cell susceptibility. Nature (London) 321, 782-784.

Turner, R. and Tjian, R. (1989) Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-Jun heterodimers. Science 243, 1689-1694.

Uchida, S., Watanabe, S., Aizawa, T., Furuno, A. and Muto, T. (1979) Polyoncogenicity and insulinoma-inducing ability of BK virus, a human papovavirus, in Syrian golden hamsters. Journal of National Cancer Institute 63, 119-126.

van der Noordaa, J., Stein, A.V. and Sol, C.J.A. (1986) Persistence of BK virus in human foetal pancreas cells. Journal of General Virology 67, 1480-1490.

Van der Hoorn, F.A. (1987) C-mos upstream sequence exhibits species-specific enhancer activity and binds murine-specific nuclear proteins. Journal of Molecular Biology 193, 255-266.

Varmus, H.E. (1985) Viruses, Genes and Cancer. Cancer 59, 2324-2328.

Vass-Marengo, J., Ratiarson, A., Asselin, C. and Bastin, M. (1986) Ability of a T-antigen transport-defective mutant of simian virus 40 to immortalize primary cells and to complement polyomavirus middle T in tumorigenesis. Journal of Virology 59, 655-659.

Vinson, C.R., Sigler, P.B. and McKnight, S.L. (1989) Scissors-grip model for DNA recognition by a family of leucine zipper proteins. Science 246, 911-916.

Virshup, D.M., Kauffman, M.G. and Kelly, T.J. (1989) Activation of SV40 DNA replication *in vitro* by cellular protein phosphatase 2A. The EMBO Journal 8, 3891-3898.

Ward, P., Coleman, D.V. and Malcolm, A.D.B. (1989) Regulatory mechanisms of the papillomaviruses. Trends in Genetics 5, 97-99.

Wasylyk, B., Wasylyk, C., Augereau, P. and Chambon, P. (1983) The SV40 72 bp repeat preferentially potentiates transcription starting from proximal natural or substitute promoter elements. Cell 32, 503-514.

Watanabe, S. and Yoshiike, K. (1982) Change of DNA near the origin of replication enhances the transforming capacity of human papovavirus BK. Journal of Virology 42, 978-985.

Watanabe, S., Soeda, E., Uchida, S. and Yoshiike, K. (1984) DNA rearrangements affecting expression of the BK virus transforming gene. Journal of Virology 51, 1-6.

Watanabe, S. and Yoskiike, K. (1989) Inefficient expression in rat cells of transforming gene of BK virus with 68 bp tandem repeats in the viral promoter enhancer. Virology 169, 204-212.

Wefold, F.C., Devlin, B.H. and Williams, S.R. (1990) Functional heterogenuity of mammalian TATA-box sequences revealed by interaction with a cell-specific enhancer. Nature (London) 334, 260-262.

Weil, P.A., Luse, D.S., Segall, J. and Roeder, R.G. (1979) Selectible and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified DNA polymerase II and DNA. Cell 18, 469-484.

Weis, R., Teich, N., Varmus, H. and Coffin J. (ed.) (1985) RNA tumour viruses. Cold Spring Harbor Laboratory, New York.

Werness, B.A., Levine, A.J. and Howley, P.M. (1990) Association of human papillomavirus types 16 and 18 E6 proteins with p53. Science 248, 76-79.

Whal, G.M., Stern, M. and Stark, G.R. (1979) Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization using dextran sulfate. Proceedings of the National Academy of Sciences USA 76, 3683-3687.

Whyte, P., Buchkovich, K.J., Horowitz, J.M., Friend, S.H., Raybuck, M., Weinberg, R.A. and Harlow, E. (1988) Association between an oncogene and an antioncogene: the

adenovirus Ela proteins bind to the retinoblastoma gene product. Nature (London) 334, 124-129.

Wieland, I., Bolger, G., Asouline, G. and Wigler, M. (1990) A method for difference cloning: Gene amplification following substractive hybridization. Proceedings of the National Academy of Sciences USA 87, 2720-2726.

Williams, G.T. and Morimoto. (1990) Maximal stress-induced transcription from the human HSP 70 promoter requires interaction with the basal promoter elements independent of rotational alignment. Molecular and Cellular Biology 10, 3125-3136.

Wingender, E. (1988) Compilation of transcription regulating proteins. Nucleic Acid Research 16, 1879-1902.

Wirth, T., Staudt, L. and Baltimore, D. (1987) An octomer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. Nature (London) 329, 174-178.

Xiao, J.H., Davidson,I., Ferrandon,D., Rosales, R., Vigneron, M., Macchi, M., Ruffenach, F. and Chambon, P. (1987) One cell-specific and three ubiquitous nuclear proteins bind *in vitro* to overlapping motifs in the domain B1 of the SV40 enhancer. The EMBO Journal 6, 3005-3013.

Yamamoto, K.K., Gonzalez, G.A., Biggs III, W.H. and Montiminy, M.G. (1988) Phosphorylation-induced binding and transcriptional efficacy of nuclear factor CREB. Nature 334, 494-498.

Yang, Y-C, Hearing, P. and Rundell, K. (1979) Cellular proteins associated with simian virus 40 early gene products in newly infected cells. Journal of Virology 32, 147-154.

Yang, R.C.A. and Wu, R. (1979) BK virus DNA: complete nucleotide sequence of a human tumor virus. Science 206, 456-462.

Zehring, W.A., Lee, J.M., Weeks, J.R., Jokerst, R.S. and Greenleaf, A.L. (1988) The c-terminal repeat domain of RNA polymerase II largest subunit is essential *in vivo* but is not required for accurate transcription initiation *in vitro*. Proceedings of the National Academy of Sciences USA 85, 3698-3702.

Zenke, M., Grundstrom, T., Mathes, H., Wintzerith, M., Schatz, E., Wildeman, A. and Chambon, P. (1986) Multiple

Sequence motifs are involved in SV40 enhancer function. The EMBO Journal 5, 387-397.

zur Hausen, H. and Schneider A. (1987) The role of papillomaviruses in human anogenital cancer. Salzman, N.P., Howley, P.M. eds. In: The papillomaviridae, Vol. 2: The papillomaviruses. New York Plenum Press; 245-263.

zur Hausen, H. (1986) Papillomaviruses in human cancer. Cancer 59, 1692-1696.