TRANSCRIPTION FROM HPV-16 EARLY PROMOTER (P_{97}) IN AN HPV-16/HSV-1 RECOMBINANT VIRUS

by

Hashem Salloukh, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Sciences

McMaster University

November 1991

MASTER OF SCIENCE (1991) MCMASTER UNIVERSITY

(Biology) Hamilton, Ontario

TITLE: Transcription from HPV-16 Early Promoter (P_{97}) in An HPV-16/HSV-1 Recombinant Virus

AUTHOR: Hashem Salloukh, B.Sc. (American University of Beirut)

SUPERVISOR: Dr. James R. Smiley

NUMBER OF PAGES: x, 177

DEDICATION

TO THE MEMORY OF MY COUSINS .....

AMIN, NADA, and HUSSIEN.

MAY ALLAH'S ETERNAL PEACE BE BESTOWED UPON THEM.

ABSTRACT

Infection with Human papillomaviruses has been suggested to play an important role in the etiology of a number of human malignancies. The most investigated of HPVs is HPV-16. Infection with HPV-16, in association with other unknown factors, is implicated in the development of cervical cancer. Genetic analysis of HPVs has been hampered by the lack of an *in vitro* system in which the complete replicative cycle and the transcription of the various genes is possible. Replication and transcription of HPVs seem to be in part regulated by cellular factors expressed at different stages in the maturation of epithelial cells which constitute the normal hosts of those viruses. This study was primarily designed to address this difficulty. HPV-16 genome was introduced into a viral system, HSV-1, hoping that virally encoded factors can substitute for cellular factors required for the expression of HPV genes. This hypothesis was tested by analyzing the activity of the HPV-16 early promoter P_{97}, in the constructed recombinant, in cells which are unsusceptible to infection with HPV-16. In Vero cells infected with the recombinant, P_{97} was shown to be active. This suggests that HSV-1 encoded factors can influence transcription from endogenous papilloma promoters.

ACKNOWLEDGMENTS

I would like to thank my parents for teaching me principles and values not found in any textbook or heard in any lecture hall. I am indebted to Dr. W.E. Rawls for introducing me to research and to Dr. J.R. Smiley for coaching my journey in it. I am grateful to all of my co-workers in the Cancer Research Group, in particular Craig Smibert and others working in Dr. Smiley's lab, for their invaluble advice. I would like to thank Carol Lavery for her help and support especially after Dr. Rawls passed away. Dominique Barbeau was inspirational in explaining the intricacies of English grammer and finding the time to assist. Ramsi Haddad always surprised me in philosophical discussions and his expertise in computers. I would like to express my gratitude to Josie Maljar, Pat Hayward, and the rest of the administrative staff in Biology for their help. I am also grateful to the volunteer girl in the gift shop for allowing me to skim through magazines when things in the lab seemed to be going nowhere.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

1.1 Papillomavirus

1.1.1 Structural Classification

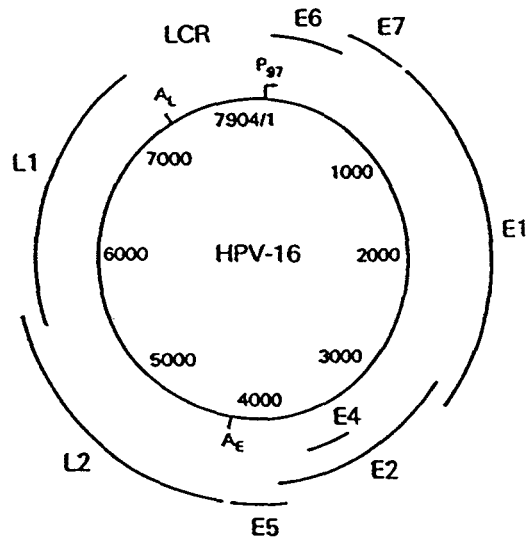
Papillomaviruses are small, nonenveloped, icosahedral DNA viruses that replicate in the nucleus of squamous epithelial cells. Historically, the papillomaviruses were members of the papova family along with the polyomaviruses. However, because of differences in the genomic organization and the biology of these viruses, each are considered independent subfamilies of the papova family (Howley, 1990). The virion particles consist of a single double-stranded circular DNA molecule, approximately 7900 base pairs in length, surrounded by a capsid consisting of 72 capsomers which form an icosohedral protein coat. Analysis of proteins in the virus particle reveals that the DNA of the viral genome is associated with cellular histones to form a chromatin like complex (Pfister et al., 1977). The papillomaviruses are widespread in

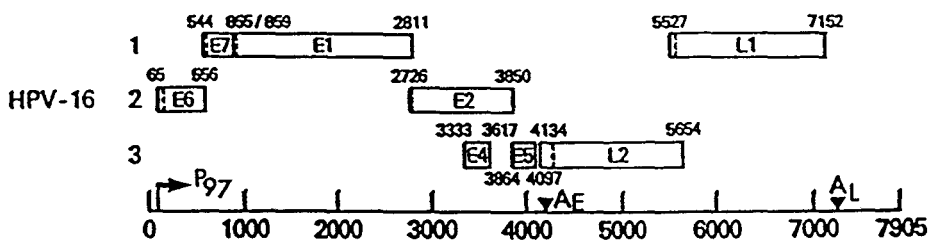
nature and have been detected primarily in higher vertebrates. Viruses have been isolated from humans, cattle, rabbits, horses, dogs, sheep, elk, deer and non-human primates. Papillomaviruses are highly species-specific in that papillomaviruses from one species do not cause a productive infection in another species (Howley, 1990).

mRNA transcription of the 9 to 10 genes which are thought to be encoded by the circular double-stranded DNA molecule, utilizes only one of the two strands. A typical papilloma genome (Fig.1) is organized into 7 or 8 early (E) open reading frames (ORFs), and 2 late (L) ORFs, based on their location in the genome and their temporal expression in productively and non-productively infected lesions (Lambert et al., 1988). Between the 5' end of the early region and the 3' end of the late region there is a 735 base pair DNA segment with no coding capability. This region is known as the Long Control Region (LCR), or the Upstream Regulatory Region (URR). Most of the known regulatory DNA sequences in the genome of papillomaviruses are located in the LCR region. This region contains the viral origin of replication and several sequences which interact with viral and cellular transacting factors (Lazo, 1988). The products of most ORFs, have been investigated

Figure 1 Human Papillomavirus Type 16 Genomic Organization

The structure of Human papillomavirus type 16 in its normal circular form (Howley, 1990) and a linearized form (Shah and Howley, 1990). Of special interest are the early and late polyadenylation signals and the early promoter P_{97}.


LCR
E6
E7
P97
AL
7904/1
7000
1000
L1
E1
6000
HPV-16
2000
5000
3000
4000
E4
AE
L2
E2
E5


HPV-16
1
2
3
544
865/859
2811
E7
E1
5527
7152
L1
65
656
2726
3850
E6
E2
3333
3617
4134
5654
E4
E5
L2
3864
4097
P97
AE
AL
0
1000
2000
3000
4000
5000
6000
7000
7905

and specific functions have been assigned to almost all of them. The early region contains the information necessary for viral replication and transformation functions. The late region (L) has two large ORFs, L1 and L2, coding for the structural polypeptides of the virions. L1 codes for the major capsid component and L2 codes for the minor capsid component (Ward et al., 1989). The E1 ORF plays a role in genome replication and in maintenance of the plasmid state (Lusky and Botchan, 1984). The E2 ORF codes for two trans-acting regulatory functions which influence E6 and E7 gene transcription (Lambert et al., 1987). A full length E2 gene product transactivates the E6 / E7 promoter while an N-terminally truncated form of the protein inhibits transcription. The E4 gene product plays a role in virus maturation and could be present in the virion. In the case of Bovine papillomavirus type one (BPV-1) E5, E6, and E7 ORFs are necessary for the transforming potential of this virus (Schiller et al., 1986).

In human papillomavirus type 16 (HPV-16), the most extensively researched among the papillomaviruses, E6 and E7 ORFs are the major transforming genes. The E6 and E7 viral proteins are expressed in cells transformed by HPV-16 DNA <u>in vitro</u> (Androphy et al., 1987a). Also, the E6 and E7

regions are transcribed and translated into the predicted viral proteins in cervical carcinoma cells (Boshart et al., 1984). The functions of E3 and E8 in BPV-1 are not yet known.

1.1.2 Pathogenesis of Infection

Biologically, papillomaviruses induce proliferative responses in epithelial cells which produce benign self-limiting lesions that regress after a period of time. These lesions consist of localized epithelial hyperplasia with a defined boundary and an intact basement membrane. All the layers of normal skin are represented in the lesion (Howley, 1990).

Infection probably occurs as a result of exposure of the basal cells to infectious virus particles after minor traumas to the epithelium (eg. during sexual intercourse or after minor damage to the skin). Papillomaviruses multiply in the nucleus of infected cells and utilize cellular transcriptional machinery to express their own genes (Shah and Howley, 1990). In these infected basal epithelial cells, expression of early messages only is observed. The viral genome usually exists in an extra-chromosomal stable plasmid found in multiple copies (Howley, 1990).

Replication is best studied in bovine papillomavirus type one (BPV-1) which was found to transform C127 mouse cells (Lambert et al., 1988). Plasmid-mode replication in BPV-1 depends on two viral sequences termed plasmid maintenance sequences (PMS1 and 2) and the E1 ORF. Mutations in the 3' end of this ORF leads to the integration of the viral genome into the host DNA (Lusky and Botchan, 1986). The E1 ORF is the most conserved early ORF in all papillomaviruses suggesting a common function. This is evident in cervical biopsies containing HPV-16 in an unintegrated form in which E1 mRNA was identified (Sousa et al., 1990).

Infection by papillomaviruses triggers the proliferation of the infected basal epithelial cells leading to the formation of a wart or a "papilloma". All the cells in a wart contain copies of the viral genome, but show different patterns of gene expression in response to the state of differentiation exhibited by the infected epithelial cells. Late viral gene expression, viral capsid antigen formation, and the assembly of viral particles take place only in differentiated and keratinized epithelial cells. Unlike the basal epithelial cells, cells permissive to late gene expression and viral production are terminally differentiated and can no longer multiply or divide (Howley, 1990). Gene expression in papillomaviruses is,

therefore, tightly linked to the state of differentiation of the infected cell. The switch from early to late gene expression is not fully understood in papillomaviruses. Analyzing papillomavirus gene expression is further complicated by the inability to propagate these viruses in culture.

BPV-1 was found to transform a mouse cell line (C127) and is, therefore, the most extensively researched among papillomaviruses. The switch from early to late gene expression in BPV-1, discussed later, seems to involve many genetic processes such as differential promoter utilization, transcription termination, differential selection of poly(A) signals, and RNA turn over (Lambert et al., 1988). Before going further into papillomavirus transcription, the pathological importance of papillomavirus infections and initiation of eukaryotic transcription will be discussed.

In rare cases lesions caused by papillomavirus infection develop into serious life threatening cancers. The progression of the benign lesions to carcinomas was first observed in rabbits (Krieder and Barlette, 1987). Out of six papillomaviruses that infect cattle, malignant transformation was associated with only one, namely bovine papillomavirus type 4 (BPV-4) (Jarrett et al., 1980). In

humans, very few of the 59 different virus types have oncogenic potential. The time interval between the initial infection and the development of invasive cancer may be decades. Progression to malignancy seems to require a number of cofactors, some of which are still unknown in many cancers involving infection with papillomavirus. The interaction between BPV-4 and an environmental carcinogen derived from braken fern, the normal diet of cattle in upland areas of Scotland, increases the number of papillomas in the alimentary tract which later progress to carcinomas (Jarrett et al., 1978).

In humans, many diseases which involve infection with papillomavirus may progress in their final stages towards malignancy. Epidermodysplasia verruciformus (EV) is a very rare skin disorder affecting young children where both HPV-5 and HPV-8 are detected in the lesions of such patients. Some EV lesions may progress to skin carcinomas after a long latency period. This is more frequently observed in lesions that have been exposed to sunlight (Lazo, 1988). Benign respiratory papillomas, which result from infection with HPV-6 or HPV-11, progress to a malignant state after being x-irradiated to reduce tumor size (Galloway et al., 1960).

1.1.3 HPV-16 and Cancer

In 1976 Harald zur Hausen reported that genital warts or condylomata acuminata, which were previously shown to contain papillomavirus (Oriel and Almeida, 1970), can become malignant in rare cases (zur Hausen, 1976). This finding enhanced the attempts of many laboratories to establish a correlation between the progression of a genital lesion to malignancy and infection with a specific human papillomavirus type.

DNA isolated from a biopsy sample of invasive cancer of the cervix hybridized with human papillomavirus type 11 DNA under non-stringent conditions only. Under stringent conditions, it hybridized weakly (less than 0.1%) with HPV types 10, 14 and 15. This new type of papillomavirus was thus designated HPV type 16 (Durst et al., 1983). In the same study, this new type of human papillomavirus was present in 60% of cervical carcinoma biopsies from German patients, whereas it was absent in benign papillomas arising in the same region.

The second most prevalent type of cancer afflicting women is cervical cancer. This has led to an increase in

the attempts to determine the possible role of papillomavirus infection in the etiology of this disease. Other than HPV-16, HPV-18, 31, 33, 35, 39, 42, 43, 44, 45, 51, 52, 56, 57, and 58 are also associated with cervical cancer or CIN (Cervical Intraepithelial Neoplasea) lesions which are widely believed to be precursors of cervical carcinomas (Howley, 1990). Epidemiological studies (Reeves et al., 1989), however, suggest that infection with any of the HPVs is not sufficient for the occurrence of cervical cancer. In fact, the progression of an HPV infection to cancer still remains rare in the pathology of such an infection (Shah and Howley, 1990).

The cervical cancer-derived cell line, CaSki, was found to contain several hundred copies of integrated HPV-16 DNA. The major viral RNA transcripts in these cells were from the E7 open reading frame (Smotkin and Wettstein, 1986).

Genomic DNA isolated from a human cervical cancer sample biopsy, which was found to contain HPV-16 sequences, was tested for transformation of the mouse fibroblast cell line NIH 3T3. The DNA sample was able to transform NIH 3T3 cells, and this activity was associated with HPV-16 DNA sequences and human specific Alu family sequences. Alu sequences are repetitive fragments which identify the DNA

as being of human origin and not mouse (Tsunokawa et al., 1986).

HPV-16 early genes driven by the simian virus 40 (SV40) promoter were tested for immortalization and transformation activity on primary rat cells (Kanda et al, 1988). E7 was shown to immortalize and transform primary rat cells. E6 was shown to be weakly immortalizing for the same cells. However, the product of E7 cannot transform primary human fibroblasts (Watanabe et al., 1989) or keratinocytes (Munger et al., 1989) without the presence of the E6 gene product.

The oncogenicity of the E7 gene product was further established when it was discovered that this product can form a stable protein-protein complex with the product of retinoblastoma gene RB1 (Dyson et al, 1989). The deletion or mutation of the retinoblastoma gene predisposes the carrier to a number of tumors which include retinoblastoma and osteosarcomas (Friend et al., 1986), small cell lung carcinomas (Harbour et al., 1988), breast tumors, and bladder carcinomas (Dyson et al., 1989). The RB1 gene product, p105-RB, plays a role in limiting the proliferative potential of certain cells. Thus, the binding of this anti-oncoprotein by oncoproteins such as E1A of adenovirus (Whyte et al., 1988) and the SV40 large T

antigen (DeCaprio et al., 1988) releases this block and results in cell proliferation.

E6 on the other hand, was shown to bind the product of p53 which, similar to Rb-1, is a cellular anti-oncogene (Werness et al., 1990). Wild type p53 is a cellular tumor suppressor gene whose loss or alteration can lead to the development of different human cancers (Nigro et al., 1989). Werness et al. suggested that binding of p53 by the product of E6 leads to a fast degradation of the p53 product contributing by that to the establishment of a transformed state (Werness et al., 1990).

An important aspect to consider in the progression of an infection with papillomavirus towards malignancy, is the state of the viral genome. The viral genome of BPV-4 present in the benign alimentary papillomas of cattle is lost as malignancy is reached (Campo et al., 1985). Viral genomes persist and remain extra-chromosomal in the case of skin carcinomas in EV patients (Howley, 1990). In pre-malignant lesions of the cervix, HPV-16 DNA is usually maintained as an extra-chromosomal plasmid (Crum et al., 1985). In cancers or cancer-derived cell lines, the HPV-16 sequences are found integrated (Schwartz et al., 1985). The pattern of HPV-16 integration is uniform; the upstream early region responsible for transformation is intact,

while the downstream region encoding early genes, namely E2, is disrupted (Baker et al., 1987). This disruption is thought to block the viral product responsible for the regulation of transcription of these transforming genes, thus leading to an irreversible path towards malignancy (Cripe et al., 1987).

1.2 Eukaryotic Transcription

Three different RNA polymerases are involved in eukaryotic transcription. RNA pol I transcribes the ribosomal RNA (rRNA) 45S precursor which gives rise to the 5.8S, 18S and 28S rRNA. RNA pol III is involved in the transcription of small RNAs such as transfer RNAs (tRNAs) and 5S rRNA. Most studies have focused on RNA pol II which is responsible for the transcription of all messenger RNAs (mRNAs), and therefore much more is known about this enzyme (Watson, 1987).

Transcription and translation are segregated in eukaryotic cells because of the presence of a nucleus. Transcription and processing of mRNA occurs within the nucleus. The primary transcript can be processed by a number of mechanisms before it becomes a mature mRNA. These mechanisms briefly include; 5' capping where a methylated

guanosine is added to the 5' terminus of mRNA, splicing where non-coding sequences (introns) are excised from the message leaving coding sequences (exons), and polyadenylation which consists of the addition of about 200 adenosine monophosphate (AMP) residues upon the recognition of the AAUAAA hexanucleotide. The mature mRNAs are then transported to the cytoplasm where they are translated by ribosomes into proteins (Watson, 1987). Of special importance is the process of transcription initiation which represents a high degree of complexity in terms of specific protein-DNA and protein-protein interactions.

1.2.1 Initiation of polII gene transcription

Initiation of transcription by RNA polymerase II (RNA pol II) is a complex process by which a DNA template is utilized by this enzyme to synthesize mRNA. This process occurs in a 5' to 3' direction and requires the four ribonucleoside triphosphates. Purified RNA polymerase II cannot recognize promoters *in vitro* (Wu, 1978). Development of cell-free transcription systems made it possible to show that RNA polymerase II accurately initiates transcription *in vitro* at the same sites as those used for *in vivo* transcription (Weil et al., 1979). *In*

vitro initiation by RNA polymerase II is made possible only by the addition of factors present in crude cell extracts (Davidson et al., 1983). The identification of such factors led to their classification into two basic groups.

The first is the general transcription factors which are necessary for the transcription of all class II genes, transcribed by RNA pol II. The second type of factors termed specific transcription factors influence the initiation of transcription of some genes only by acting on well defined regions close to the promoters of these genes. These factors are not required for "basal" levels of transcription, but instead influence transcription by enhancing the rate of initiation or the formation of the preinitiation complex which is made up of the RNA polymerase II and the general transcription factors. Initiation of transcription by RNA polymerase II remains largely controlled by short DNA sequences. These sequences seem to be the binding sites of nuclear proteins involved in the process of transcription. These cis-acting DNA sequences are commonly referred to as the promoter region or enhancer elements depending on their location and orientation with respect to the site of initiation of transcription (Mermelstein et al., 1989).

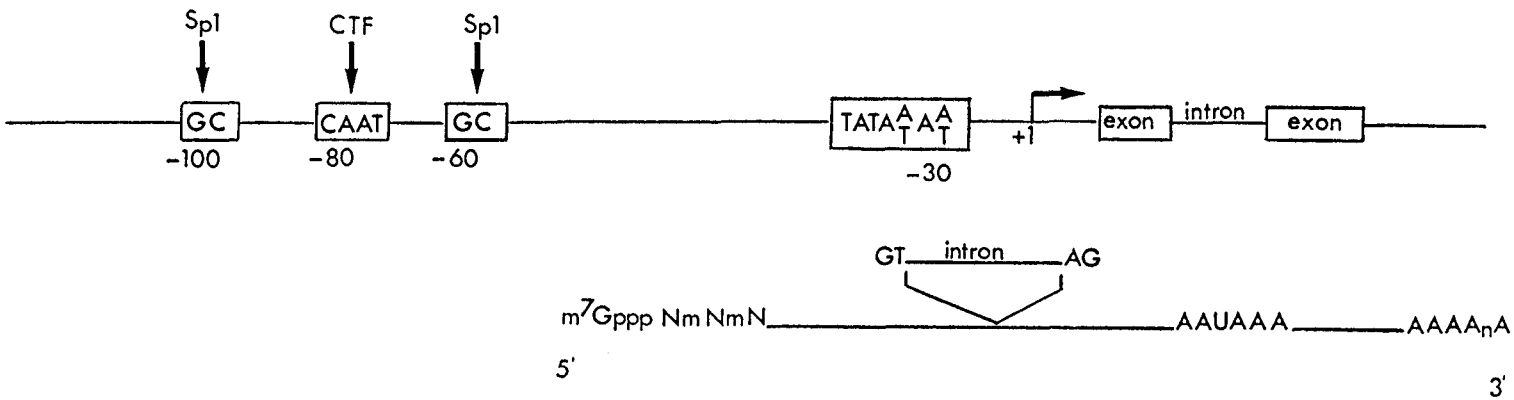
1.2.2 Cis-Acting Sequences in Eukaryotic Transcriptional Control Elements

Cis elements are the binding sites of RNA polymerase II and different transcription factors. They are scattered within several hundred base pairs from the initiation site, although some elements can exert their action from larger distances. The strategy employed to investigate cis-acting sequences required for transcription initiation in eukaryotic genes focused on searching for sequence similarities near the transcriptional start site in different genes, which were then mutated and reintroduced into cells to study their effect on transcription initiation (Dynan and Tjian, 1985).

Cis elements that control trancription initiation from a site close to the transcription start site (30-100 bps) (Fig.2) constitute the promoter region (Review by Mitchell and Tjian, 1989). The most conserved sequence found so far in polII promoters is the TATA box with a consensus sequence 5' TATAAA 3', located 25 to 30 base pairs upstream from the start site of transcription (Breathnach and Chambon, 1981). The function of this sequence appears to be related to the correct position at which transcription starts. Mutations in this region

Figure 2 A Eukaryotic Transcriptional Unit

A typical RNA polymerase II transcription unit is shown with the start site for transcription and the various cis-acting sequences constituting the promoter region. Also illustrated is a processed mRNA transcript with the 5' end capped, intervening sequences deleted, and poly A tail added. (Adapted from Watson, 1985)


Sp1
CTF
Sp1
GC
CAAT
GC
-100
-80
-60
TATA
-30
+1
exon
intron
exon
GT
intron
AG
m7Gppp Nm NmN
AAUAAA
AAAAnA
5'
3'

reduce the amount of transcripts produced . Transcripts produced in this case will have different 5' ends (Ghosh et al., 1981,)

Other common cis sequences that serve as recognition sites for DNA binding proteins are usually located upstream from the TATA box and appear to increase the efficiency of expression of a number of genes (Dynan and Tjian, 1985). Two such sequences are the CCAAT and the GGGCGG (Fig.2) homologies which, along with the TATA box , are found in the promoter of many RNA polII genes (Dierks et al., 1983).

The CCAAT box is usually located 70-90 nucleotides upstream from the initiation site of transcription (McKnight and Tjian, 1988). This sequence plays an important role in the transcription of many cellular and viral genes such as the β-globin gene and the herpes simplex thymidine kinase gene (Johnson and McKnight, 1989). A number of factors which recognize this sequence, termed CCAAT transcription factors (CTF), have been reported and will be discussed under "Specific Transcription Factors".

The GGGCGG sequence can occur in more than one copy 50-100 nucleotides upstream from the initiation site (McKnight, 1982). This sequence is essential for the binding of a transcriptional factor known as the promoter specific factor (Sp1), which was initially isolated from

the cervical carcinoma cell line, Hela, and was found to activate the SV40 early promoter (Dynan and Tjian, 1983).

The activity of many promoters can be affected by a separate regulatory element called an enhancer. While studying eukaryotic gene expression by injecting cloned sea-urchin histone genes into Xenopus oocytes, Grosschdl and Birstel came across sequences located more than 1000 bps upstream of the trancription start site of an histone H2A gene which could act in either orientation to restore efficient transcription (Grosschedl and Birnstiel, 1980). Enhancers must be on the same molecule of DNA as the promoter to exert their actions, but can be up to 1,000 bp or more downstream or upstream from a promoter. Some enhancers are only active in specific tissues, while others are responsible for transcription activation of certain genes by steroid hormones (Serfling et al., 1985).

1.2.3 General Transcription Factors

Five protein factors required for transcription have been identified using the simplest promoter units made up of a TATA box and CAP site. These factors are required for the transcription of all class II genes analyzed and are therefore called general transcription factors (Matsui et

al., 1980). Only one out of those five factors can bind to DNA, specificically to the TATA promoter element, and is called transcription factor IID (TFIID) (Davidson et al., 1983). TFIID binding to DNA is considered to be one of the early steps towards the formation of a DNA-protein complex called the rapid start complex, which is capable of initiating transcription (Fire et al., 1983). The binding of TFIID to the TATA sequence sets the stage for the entry of RNA polymerase II into the transcription cycle (Reinberg et al., 1987). A protein factor required for transcription and which acts by stabilizing the binding of TFIID to the DNA has been identified as TFIIA (Fire et al., 1983). Direct evidence of such a role for TFIIA came from analyzing the binding activity of yeast TFIID to the adenovirus major late promoter. It was demonstrated that the formation of a stable TFIID-promoter complex required TFIIA (Buratowski et al., 1989).

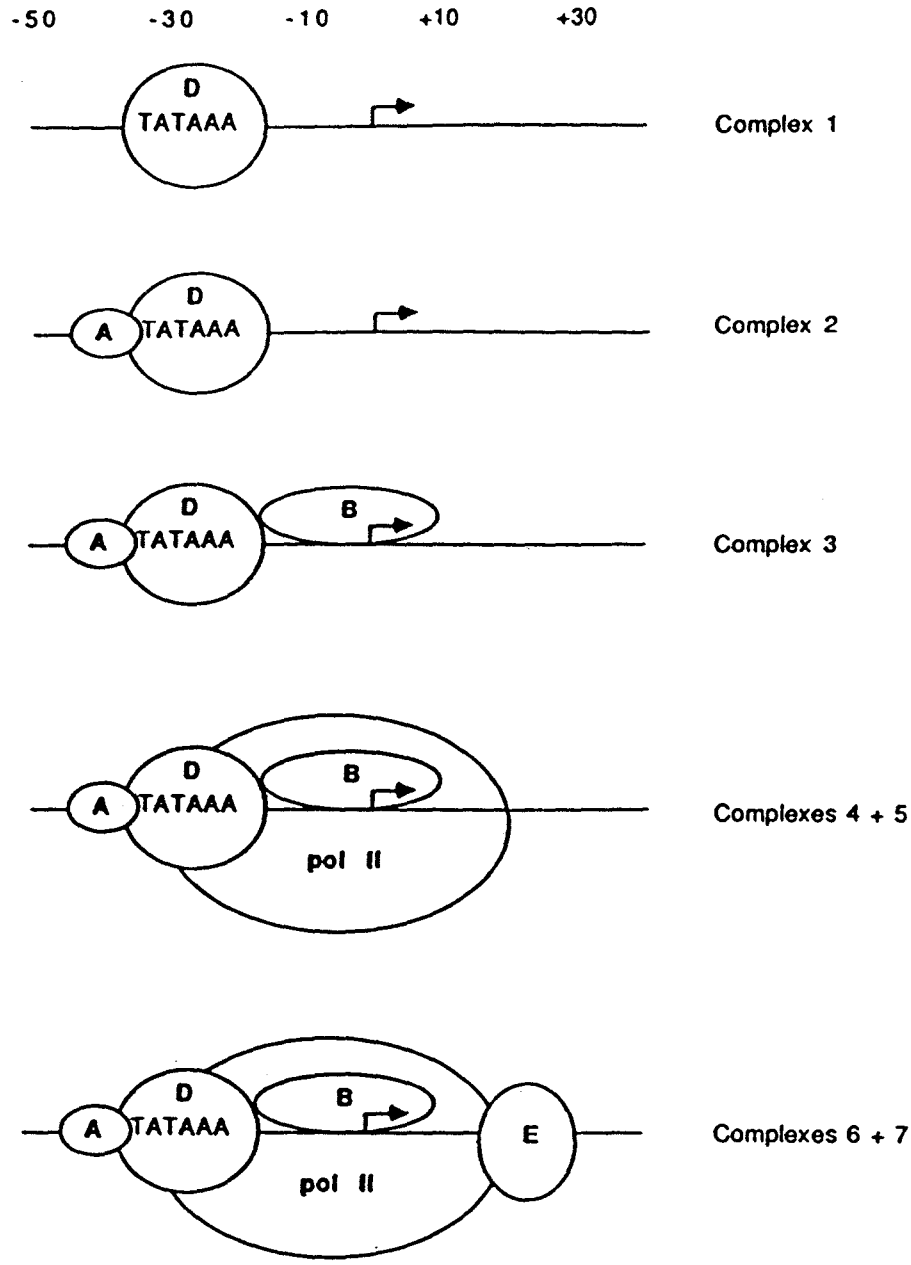
In the presence of a bound TFIID, RNA polymerase binds to the promoter and allows the entry of three other factors into the transcription cycle: TFIIB, TFIIE and TFIIF (Reinberg et al., 1987). Both TFIIE and TFIIF can independently interact with purified RNA polymerase II and the presence of both is essential for the formation of the rapid start complex (Flores et al., 1989). The isolation

of different DNA-protein complexes at different stages of transcription led to a model which explains the sequence by which general transcription factors enter or participate in the transcription cycle (Reinberg et al., 1987). In the first step, TFIIA is postulated to recognize the TATA element which leads to the formation of the preinitiation complex, although TFIIA has never been shown to contain any DNA binding ability. TFIID then recognizes the TATA box which results in the formation of the committed complex (TFIIA:DNA:TFIID). The committed complex is directly recognized by RNA polymerase II. TFIIB, TFIIE, and TFIIF enter the transcription cycle by associating with the committed complex to form the rapid start complex.

In another study Buratowski et al. suggested a series of complexes in the ordered assembly of different general transcription factors and RNA pol II on the major late promoter of adenovirus 2 (Buratowski et al., 1989). Each of these complexes represent an intermediate in the process of initiation (Fig. 3). Complex 1 is formed after the recognition of the TATA element by TFIID. This recognition might require the presence of TFIIA. Complex 2 is formed by the binding of TFIIA to complex 1. TFIIA seems to increase the binding activity of TFIID to the TATA box.

Figure 3 Complexes in the Initiation Process

The different complexes resulting from the loading of the general transcription factors and RNA pol II onto the late promoter region of Ad2. The site of transcription initiation is shown with an arrow, and the nucleotide positions relative to this site are represented by the numbers at the top. (From Buratowski et al., 1989)


-50
-30
-10
+10
+30
D
TATAAA
Complex 1
D
A
TATAAA
Complex 2
D
B
A
TATAAA
Complex 3
D
B
A
TATAAA
pol II
Complexes 4 + 5
D
B
A
TATAAA
E
pol II
Complexes 6 + 7

This has led some to suggest that TFIIA can act through nonspecific binding to the DNA template before TFIID is bound, though, more experiments are needed to verify this hypothesis. TFIIB binds to the initiation complex to form complex 3. TFIIB is essential for the binding of the TFIIA-TFIID-TATA complex to RNA pol II. Complexes 4 and 5 are formed after the binding of pol II to complex 3. Complexes 6 and 7 are generated by the addition of TFIIE to complexes 4 and 5. TFIIE can bind pol II in solution, and is suggested to be associated with pol II and TFIIB before binding the TFIIA-TFIID-TATA complex. TFIIE binds the DNA template downstream of pol II protecting an area of +20 to +30 from DNAase I, and has been shown to have a DNA-dependent ATPase activity. Hydrolysis of ATP or dATP by TFIIE will satisfy the energy requirement needed for transcription forming what is known as an "activated transcription complex". TFIIE can, therefore, act as a helicase unwinding the DNA template in front of the RNA polymerase. In the case of a linear DNA template, TFIIE will fall off forming complexes 8 and 9 which were found to contain accurately initiated transcripts in the presence of NTPs.

1.2.4 Specific Transcription Factors

Initiation of transcription from different class II genes is controlled by specific DNA binding proteins which recognize sequences upstream from the start site (Ptashne, 1986). The mechanism by which these proteins interact with the basic transcription machinery, RNA polymeraseII and the general transcription factors, is still ambiguous. It was demonstrated in adenovirus that the binding of upstream activating transcription factors (ATF) to the E4 promoter can affect TFIID binding to the TATA sequence (Horikoshi et al., 1988). This led to the belief that upstream binding proteins operate by modifying general transcription factors. Included in this category are proteins which bind to the two promoter elements, CCAAT and GGGCGG.

Nuclear factor I (NFI) which was originally defined with regards to its activity in the adenovirus in vitro replication system, was later shown to bind to CCAAT consensus sequence and is therefore, a CCAAT transcription factor (CTF) (Jones et al., 1987). Apart from NF1/CTF, a liver protein was shown to bind the CCAAT sequence of the HSV tk promoter and was subsequently called CCAAT-binding protein (CBP) (Graves et al., 1986). Other proteins distinct from NF1/CTF and which bind to the CCAAT consensus

of the major histocompatibility complex genes and the adenovirus E2 promoters have been identified (Dorn et al., 1987). Not all proteins that bind to CCAAT boxes are activation proteins. A protein in sea urchin embryos was found to bind a CCAAT sequence present in the promoter of the sperm histone H2b-1 gene. This protein, by doing so, prevents the binding of a positive-acting CCAAT-binding factor, thus inhibiting transcription (Barberis et al., 1987). The mechanism by which factors which bind the CCAAT sequence activate or inhibit transcription is not clear yet. It is possible that these factors accelerate a specific step during the initiation reaction which involves general transcription factors (McKnight and Tjian, 1988).

The GGGCGG sequence is essential for the binding of a transcription factor known as Sp1 (Promoter Specific factor 1) which as mentioned earlier, was found to activate the SV40 early promoter (Dynan and Tjian, 1983). Activation is made possible by the presence of a DNA sequence located 50-100 bps upstream from the start site which Sp1 can directly interact with. When the DNA sequence of the section of the promoter to which Sp1 binds was analyzed, it was found to contain three 21 bp repeats, each having two copies of the sequence 5' - GGGCGG - 3'. The deletion of the three 21 base pair repeats limits early transcription

in SV40 (Myers and Tjian, 1980). Out of six possible hexanucleotide sites, Sp1 interacts with only five of these because of steric hindrance (Gidoni et al, 1984).

Sp1 potential binding sites are also reported in promoters of cellular genes such as human adenosine deaminase (Valerio et al., 1985), mouse adenine phosphoribosyl transferase (Dush et al., 1985), as well as other promoters of viral genes such as retrovirus (Sanchez-Pescador et al., 1985). Sp1 binding sites are also present in promoters of many different genes which do not share a common function. It is possible that Sp1 binding sites provide a common transcriptional base to which regulatory proteins can bind and modulate transcription (Dynan and Tjian, 1985).

In addition to the CCAAT and GGGCGG sequences, specialized classes of upstream elements which mediate a response to a defined stimuli in some genes only, have been recognized (Mermelstein et al., 1989). One such class is the sequences to which the glucocorticoid receptor protein binds. (Payvar et al., 1983). The glucocorticoid receptor has the capacity to bind to specific sites located within or around genes which are transcriptionally responsive to the glucocorticoid hormone. Analysis of the sites to which the glucocorticoid receptor binds and acts as a positive

regulator of transcription in response to the glucocorticoid hormone, revealed a consensus sequence composed of AGA A/T CAG A/T and was named glucocorticoid response element (GRE) (Karin et al., 1984). The glucocorticoid receptor can also act as a hormone-dependent negative regulator. In this case the binding site is termed negative glucocorticoid response element (nGRE) (Miesfeld et al., 1987). nGRE and GRE consensus sequences are different, however, both share the same region of the receptor protein required for the binding of the two sequences. It is proposed that nGRE changes the shape of the receptor protein after it binds to DNA and by that causes it to act as a repressor of transcription in the presence of the glucocorticoid hormone (Meisfeld et al., 1987).

Another example of upstream factors which respond to a certain stimulus by recognizing specific upstream sequences was revealed in studies centered around a group of cellular transcription factors collectively termed activator protein 1 (AP-1). AP-1 was first isolated from HeLa cells and shown to activate transcription by binding specifically to the sequence 5' TGACTCA 3' in an enhancer of human metallothionein II_A gene ($hMTII_A$) and in the 72 bp enhancer region of SV40 (Lee et al., 1987a). Both SV40 and

$hMTII_A$ genes are induced by the tumour promoter phorbol ester 12-O-tetra decanoylphorbol-13 acetate (TPA) (Imbra and Karin, 1986). There appears to be a strong correlation, since those same genes which respond to TPA also contain this short AP-1 binding sequence, suggesting that TPA induction occurs through AP-1 activation (Fujita et al., 1986). Further evidence for this relationship has been demonstrated by Lee et al., where AP-1 binding sequences placed upstream and downstream from the chloramphenicol acetyl transferase (CAT) gene, were shown to act as TPA-inducible enhancer elements.

AP-1 recognition sites were recently shown to share a high degree of similarity with the binding sites of the purified AP-1 preparations which were found to consist of the cellular Jun protein, product of the proto-oncogene, as the major component. These preparations were also found to contain the product of the proto-oncogene c-fos and fos related antigens. The c-fos product is a phosphoprotein which binds DNA <u>in vitro</u>. The expression of c-fos is low in most cell types, nevertheless, this state is rapidly reversed under the effects of a whole range of extracellular stimuli. This has encouraged the concept of a coupling role by which fos responds to "short-term" signals as a result of stimuli to the cell surface with

"long-term" changes in the cell's phenotypes by the activation of specific genes in the cell (Franza et al., 1988). C-jun and c-fos can form a stable complex capable of binding to AP-1 consensus sequences which in turn can activate the transcription of reporter genes containing these AP-1 consensus sequences.

The current idea on how activation via AP-1 sites arises, involves external stimuli which are thought to lead to the activation and increase in synthesis of c-fos and c-jun products and other related proteins which then form a complex and bind to AP-1 sites either directly or via protein-protein interactions as in the case of fos. This binding to AP-1 consensus sequences activates genes containing such sequences leading to the "long-term" response to signals regulating the growth of cells (Rauscher et al., 1988). The mechanism by which activation of transcription occurs through proteins binding AP-1 sequences, is still ambiguous. In the case of SV40 and $hMTII_A$, both respond to Sp1 and to AP-1 binding proteins, which may cooperate together to form an initiation complex which is recognized by RNA polymerase II (Lee et al., 1987b).

1.3 Transcription of Papillomaviruses

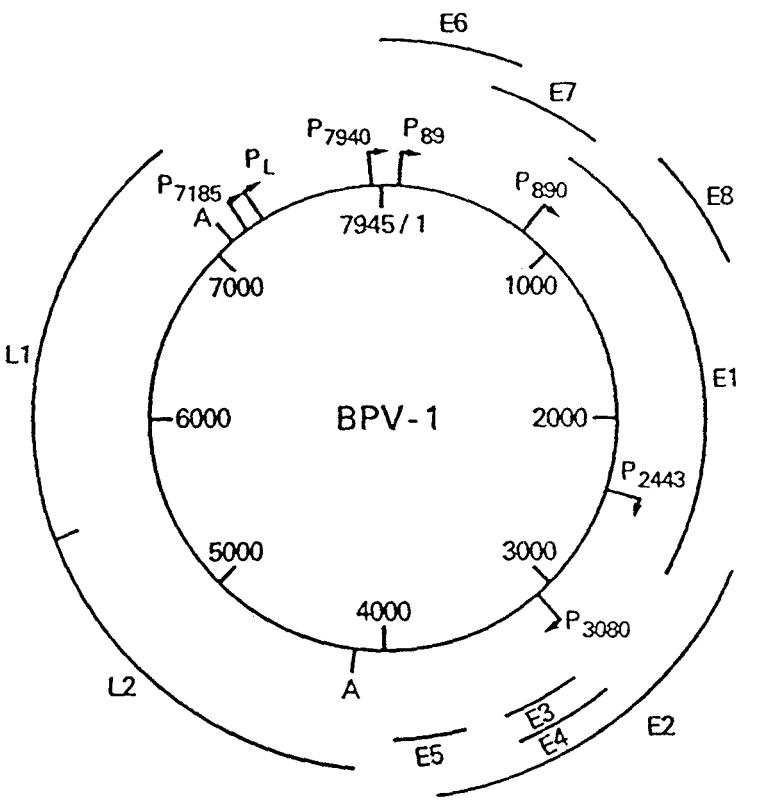
1.3.1 General Description and Difficulties

Studies on the transcription of papillomaviruses have been refractory to common genetic techniques due to the absence of an *in vitro* system in which they can be propagated (Lambert et al., 1988). Papillomavirus gene expression seems to be linked to the state of differentiation of the infected epithelial cells in that transcripts from the early region only are expressed in the primary infected epithelial cells and as this cell passes through different stages of development, more transcription is observed leading to the expression of late transcripts in the terminally differentiated keratinocytes (Stoler and Broker, 1986).

Studies on codylomas and lesions caused by papilloma infection revealed the over-production of early messages relative to the formation of virus capsid proteins or the assembly of mature virus particles. Viral activity such as DNA replication, production of viral capsid proteins, and the assembly of virions occur only in cells showing squamous differentiation located at the surface of lesions (Howley, 1983). The extent to which the early to late

Figure 4 The Genomic Organization of BPV-1

An illustration of the Bovine papillomavirus type one genome highlighting the different promoters and poly (A) signals. The most active early promoter is P_{89}, while P_L is the "wart specific" promoter from which transcription of late messages occurs. The poly (A) signal used by early messages is at 4203 while the late signal is at 7175. (Howley, 1990)


E6
E7
E8
E1
E2
E3
E4
E5
L1
L2
P_{7940}
P_{89}
P_L
P_{7185}
A
7945 / 1
P_{890}
7000
1000
6000
BPV-1
2000
P_{2443}
5000
3000
4000
P_{3080}
A

switch depends on differential promoter use, transcription termination, alterations in splicing, or regulated choice of poly(A) sites remains largely unknown.

A breakthrough in papillomavirus transcription occurred when bovine papillomavirus type 1 was shown to transform rodent cells, particularly mouse C127 cells (Dvoretzky et al., 1980). DNAs from different papillomaviruses share a high degree of homology and the genome is organized in a similar fashion, therefore, many of these viruses share common biological features.
Analysis of mRNA species in C127 rodent cells transformed by BPV-1 showed that most transcripts initiated from nucleotide 89. The promoter utilized to transcribe those mRNA species is designated as P_{89} (Fig.4). Other promoters are present in the early region but are far less active than P_{89} (Ahola et al., 1987). Multiple splice patterns result in different mRNAs all of which use a polyadenylation site at nucleotide 4203 (Fig.4). mRNA coding for late genes are not detected in transformed C127 cells (Amtmann and Saur, 1982). Transcription of BPV-1 late genes was observed using northern blot and cDNA analysis on mRNA isolated from lesions caused by this virus in cattle and called fibropapillomas (Baker and Howley, 1987). Both L1 and L2 mRNAs are transcribed from a "wart-

specific" promoter P_{7250}. Both transcripts use a polyadenylation site at nucleotide 7175 (P_L in Fig.4). Expression of papillomavirus late genes is yet to be observed in any tissue culture system . In BPV-1 infected lesions, P_{7250} seems to be only active in mature differentiating keratinocytes, and not in the primary infected basal epithelial cells, and seems to require cellular transcriptional factors found in these cells (Baker and Howley, 1987).

Differential promoter utilization does not seem to be the only factor involved in the regulation of late transcription in BPV-1. The late region lies downstream of the early genes in the same transcriptional direction. Multiple promoters located in the URR and early region and which are active in non-productively infected cells do not seem to drive transcription from the late region. This suggests that factors other than differential promoter utilization such as polyadenylation, mRNA turnover, and transcription termination are involved in late transcription (Heilman et al., 1982). Baker and Noe have shown that in C127 cells transformed with BPV-1 more than 90 % of polymerase transcriptional complexes transcribing into the late region terminate before reaching the late polyadenylation signal. Transcription termination between

late and early polyadenylation signals in transformed cells results in the preferential utilization of the early polyadenylation signal (Baker and Noe, 1989). Mutations in the polyadenylation signal utilized by the early genes designed to encourage the use of the late polyadenylation signal did not succeed in increasing late gene expression. The observed attenuation is not, therefore, due to the utilization of a strong polyadenylation signal, in this case the early signal, but rather suggests the existence of a complex regulatory mechanism leading to the block of late expression. In C127 transformed cells, messages coding for the late region cannot use the late polyadenylation signal and are, therefore, very unstable. A plausible mechanism addressing transcription termination in BPV-1 productively infected fibropapillomas where transcription from the late region is detected suggests that termination factors are filtered out during DNA replication allowing transcription from the late region (Lambert et al., 1988).

The untranslated region between L1 ORF and the late polyadenylation site in (BPV-1) contains cis-acting sequences which act as mRNA instability elements leading to the rapid degradation of L1 and L2 transcripts in BPV-1 transformed cells. Therefore, in non-terminally differentiated keratinocytes, transcription attenuation

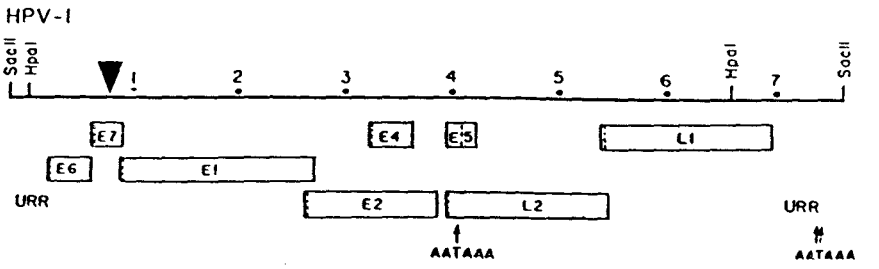
between the early and late polyadenylation signals significantly reduces the utilization of the late poladenylation site. Transcripts, however, which succeed in using this site are extremely unstable due to the cis-acting instability sequences and are rapidly degraded. Both phenomena suggest a role by a cellular factor in differentiated keratinocytes, where late gene expression is observed, to over-ride the two negative elements (Howley, 1990).

1.3.2 Transcription in Human Papillomaviruses

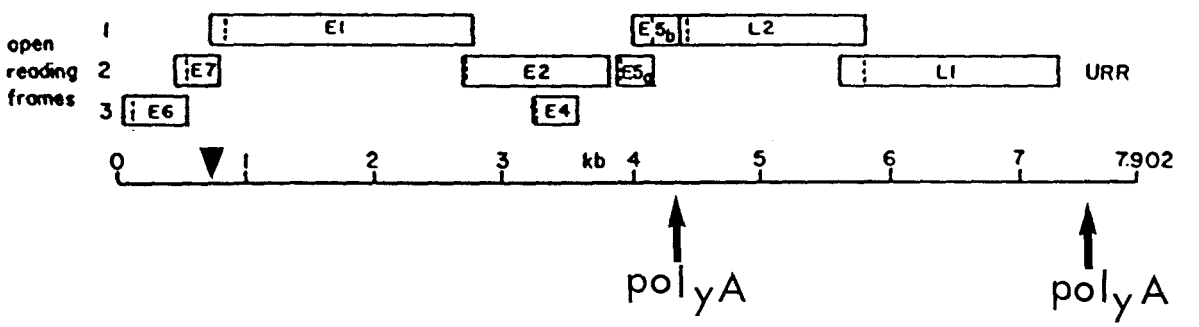
Analysis of mRNA species found in benign genital lesions, condylomata acuminata, caused by human papillomavirus types 6 and 11 revealed a major promoter in the early region corresponding to the E7 ORF located approximately at nucleotide 700 (Fig. 5). Other minor promoters are located in the early region just upstream from E6 and in E1. Two polyadenylation sites were utilized by the transcripts, one at the end of the early region and the other at the end of the late region. However, more than 99% of the mRNA transcripts utilized the "early" polyadenylation signal. Transcripts coding for structural proteins from the late region were very rare, which is in

Figure 5 Genomic organization of HPV-1 and HPV 6/11

The genomic organization of HPV-1 and HPV 6/11 showing the different ORFs, the major promoters, and the polyadenylation sites. (From Chow et al. 1987a,b).



HPV 6/11



▼ Major Promoter

agreement with the general observation that HPV infections produce few virus particles. A rare transcript coding for the L1 region was found to originate in the major early promoter at nucleotide 700.

Regulation of late gene expression in HPV 6/11 seems to depend on whether RNA polymerase II (or the transcriptional complex) utilizes the early or the late polyadenylation signal and transcribes beyond the early polyadenylation signal to the late one. Other control factors include mRNA processing or transport (Chow et al., 1987a).

HPV-1 is associated with the benign non-malignant plantar warts. Most transcripts from such warts initiate from a promoter in the E7 ORF at about nucleotide 700 (Fig. 5). Two other promoters utilized to a lesser extent are located just upstream of E6 and within the URR. The most adundant transcript has a 5' exon of 100 nucleotides representing the E7/E1 junction spliced to a 3' exon representing the region of the genome where the E2 and the E4 ORFs overlap. This transcript is polyadenylated at a poly(A) site located at the end of the early region (nt 3983), and codes for the E4 protein which is found in abundance in plantar warts. A transcript coding for L1 is very infrequent and initiates at the same position as the

major transcript but utilizes a polyadenylation signal at the end of the late region. Similar to the case of HPV 6/11, regulation of late gene expression seems to be dependent on the polyadenylation site used (Chow et al., 1987b).

In cancer cell lines containing HPV-16 such as the CaSki cells, most RNA transcripts code for the E6 and E7 transforming proteins (Androphy et al., 1987b). The 5' end of these transcripts corresponds to nucleotide 97, thus, the promoter responsible for their expression is called P_{97} (Baker et al., 1987). The non-coding URR of HPV-16 contains three late mRNA polyadenylation signals beginning at approximately positions 7260, 7320 and 7659. Experiments were conducted on HeLa cells to investigate whether late gene expression in HPV-16 is regulated by differential utilization of late polyadenylation sites. HeLa cells resemble undifferentiated keratinocytes in that they are non-permissive for late gene expression. <u>In vitro</u> and <u>in vivo</u> transfection experiments using these cells and the investigated sequences in HPV-16 confirmed that the second late polyadenylation signal (starts at nt. 7320) is functional. The absence of late gene expression in non-differentiated epithelial cells such as HeLa cells is, therefore, not solely due to regulation of late

polyadenylation signal utilization. In the same experiments, a negative regulatory element located immediately upstream of the late mRNA polyadenylation signal was shown to have no interference with 3' processing as indicated by the unaltered utilization of the second late polyadenylation signal (Kennedy et al., 1990).

Late gene expression is detected in differentiated keratinocytes. It is possible that in these cells 3' processing signals, not functional in other cell types, are utilized. The utilization of such "new" polyadenylation signals in differentiated keratinocytes can lead to splicing patterns different from those which are detected in undifferentiated epithelial cells (Kennedy et al., 1990). As in the case of BPV-1, these observations suggest a role by a cellular factor, present in differentiated keratinocytes, in the regulation of late gene expression.

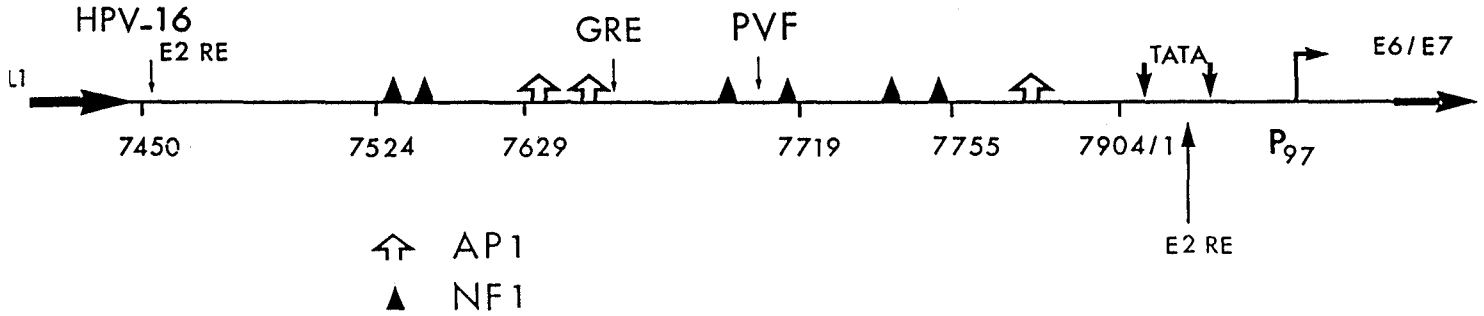
Due to their importance in the transformation process and the implications of a role by HPV-16 in the etiology of cervical cancer, E6 and E7 transcription in HPV-16 has been the focus of many studies.

1.3.3 Transcription from P_{97} promoter of HPV-16

P_{97} is located 3' of the upstream regulatory region

Figure 6 Keratinocyte Dependent Enhancer of HPV-16

Keratinocyte Dependent Enhancer (KDE) of HPV-16 is illustrated with the different DNA sequences which mediate the action of different factors on the early promoter P_{97}. These include the glucocorticoid responsive element, Nuclear Factor (NF-1) binding sequences, AP-1 binding sequences, sequences that respond to the the E2 viral regulatory protein (E2 RE), and the papilloma specific enhancer (PVF). (Taken from Chong et al., 1990 with modifications).


HPV-16
E2 RE
L1
GRE
PVF
TATA
E6/E7
7450
7524
7629
7719
7755
7904/1
P_{97}
E2 RE
AP1
NF1

(URR). A region approximately 400 nucleotides upstream from the P_{97} promoter plays an essential role in the transcription of the transforming genes. A glucocorticoid responsive element is located in the URR about 377 nucleotides 5' from P_{97} extending from nucleotides 7640 to 7655. This DNA sequence, 5'- TGTACATTGTGTCAT -3', has a 12/15 homology with the GRE of the human metallothionein II_A gene promoter (Gloss et al., 1987).

A 400 base pair fragment from HPV-16 free of TATA boxes or sequences responsive to virally encoded proteins, but containing the GRE was tested for stimulation of transcription after the hormone, glucocorticoid, was supplied (Gloss et al., 1987). The 400 base pair fragment was linked to the thymidine kinase (tk) promoter in a plasmid in which this promoter drives the transcription of the bacterial chloramphenicol acetyltransferase (CAT) gene. The resulting plasmid was transfected into Hela cells. When synthetic glucocorticoid was supplied, an increase in the expression of the CAT gene was reported. This phenomenon was independent of the orientation of the fragment carrying the GRE (Gloss et al., 1987).

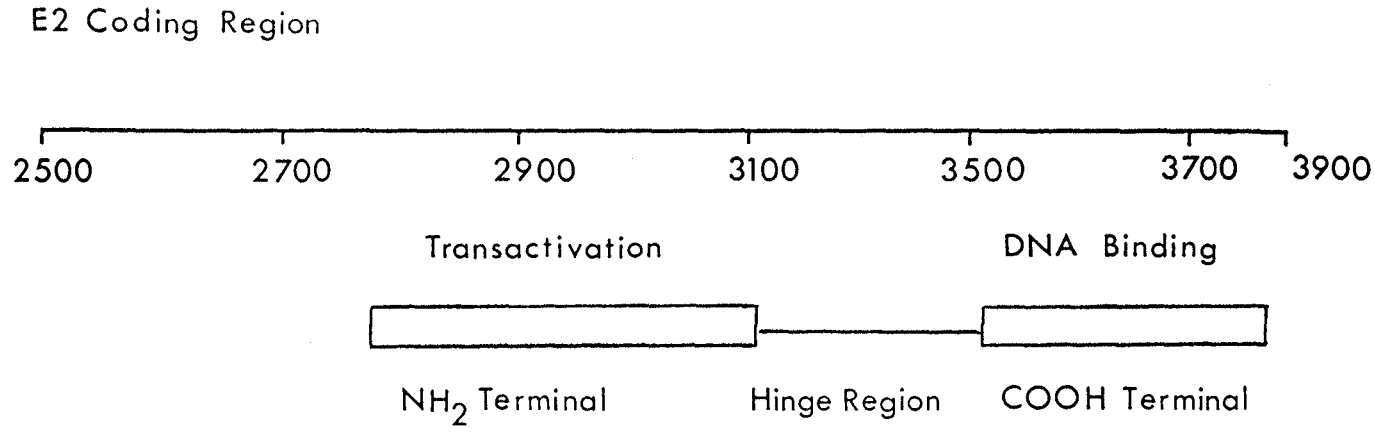
The GRE is part of an enhancer element extending from nucleotides 7524-7755 (Fig.5) (Cripe et al., 1987). The activity of this enhancer was tested in different cell

lines using the tk promoter and an indicator gene. The enhancer was active in human cancer cell lines (HeLa, SiHa, and CaSki) and uninfected foreskin keratinocytes, but was not in murine and monkey fibroblast lines, nor in human skin fibroblasts. The activity of the enhancer element was independent of its orientation. These results indicate that the enhancer element depends on factors present in human keratinocyte cells, and is therefore called the keratinocyte dependent enhancer (KDE) (Cripe et al., 1987).

The KDE has 5 NFI sites (Gloss et al., 1989a) and 2 AP-1 binding sites (Gloss et al., 1989b) (Fig.6). Two AP-1 binding sites, (7631-7637) and (7810-7816), when linked to a CAT expression vector were able to induce CAT expression in the presence of phorbol esters (Chong et al., 1990). In the same study a third site (7648-7754) for AP-1 binding overlapping with the GRE did not show any response to phorbol esters. Deletion and mutation studies indicate that those sites collaborate in their contribution to the activity of the KDE (Chong et al., 1990). DNase I protection experiments in the URR region revealed a sequence of 10 base pairs (AAGGACATAC) repeated five times (Gloss et al., 1989b). A 10 base pairs related sequence (AGGCACATAT) in the KDE is also protected by nuclear extracts (Gloss et al., 1989b). When a mutation was

Figure 7 E2 Transactivator

A schematic representation of the E2 viral transactivator. In its complete form, E2 transactivates transcription from the P_{97} promoter (Baker and Howley, 1987).


E2 Coding Region
2500
2700
2900
3100
3500
3700
3900
Transactivation
DNA Binding
NH2 Terminal
Hinge Region
COOH Terminal

introduced in this protein binding sequence, the KDE lost 45% of its activity (Chong et al., 1990). The factor binding to these sequences contributes to the enhancer's activity. The sequences bound by this factor do not resemble any known enhancer element, therefore this factor is termed papillomavirus enhancer associated factor (PVF) (Chong et al., 1990) (Fig. 6). Factors binding to the KDE are believed to largely influence the expression of the E6/E7 genes from P_{97}.

AP-1 sites are of special interest considering the tumor causing nature of genital papillomaviruses. These sites can bring the activation of the viral tumor genes (E6/E7) under the influence of proto-oncogenes such as the genes of the jun and fos families (Chan et al., 1990). AP-1 is thought to couple the activity of promoters to extra extracellular signals (Sharma et al., 1989). Papillomaviruses are non-lytic viruses, and hence require a stable genomic number and transcriptional activity. AP-1 sites can, in this case, regulate gene expression in response to extracellular signals such as epidermal growth factor or other unknown factors (Squinto et al., 1989).

Two TATA boxes are located immediately 5' to the E6 open reading frame at positions 17 and 65. Between those two boxes and further upstream at an EcoRI site (at

nucleotide 7454) are DNA sequences which bind the E2 viral product (Hirochica et al., 1987) (Fig. 6).

The E2 ORF codes for both a positive and a negative regulatory factor. DNA sequences to which E2 proteins can bind and mediate their action are called E2-responsive elements. Each element contains one or two consensus cores made up of ACC $(N)_6$ GGT where N is any nucleotide (Haugen et al., 1987). The E2 protein is made up of three domains two of which are biologically active (Fig. 7). The carboxyl terminal domain is required for DNA binding, while the amino terminal domain is nessary for transactivational properties (Lambert et al., 1987).

The complete form of the E2 protein acts as a transactivator of transcription from the P_{97} promoter. Another form of E2 protein can occur resulting from translational initiation from an internal ATG codon within the E2 ORF to give an N-terminally truncated protein. This protein can bind to the same responsive element by virtue of its carboxyl terminus, but acts as a repressor of transcription (Lambert et al., 1987). In the latter case, the responsive element acts as a silencer of KDE's response to cellular factors (Cripe et al., 1987). The binding of the repressor form could inhibit sliding of cellular factors from their entry sites to the promoter. This model

is not favoured since the KD enhancer element can function from large distances by other mechanisms than sliding (Lambert et al., 1987). It is possible that the repressor form of E2 shares unknown functional domains with the E2 transactivator. The bound repressor in this case may interact with cellular transcription factors at the promoter, but lack the ability to form an active transcription initiation complex (Lambert et al., 1987).

It is observed that in invasive carcinomas, HPV-16 is found integrated in the cellular genome (Baker et al., 1987). The upstream early region (E6/E7) is intact while the downstream early region (E2, E4, E5) is disrupted (Schneider-Gadicke and Schwarz, 1986). The disruption of E2 will lead to the uncontrolled E6/E7 transcription under the control of cellular factors acting through the KDE resulting in a change towards a transformed phenotype in infected cells.

1.3.4 Structure of the P_{97} Transcripts

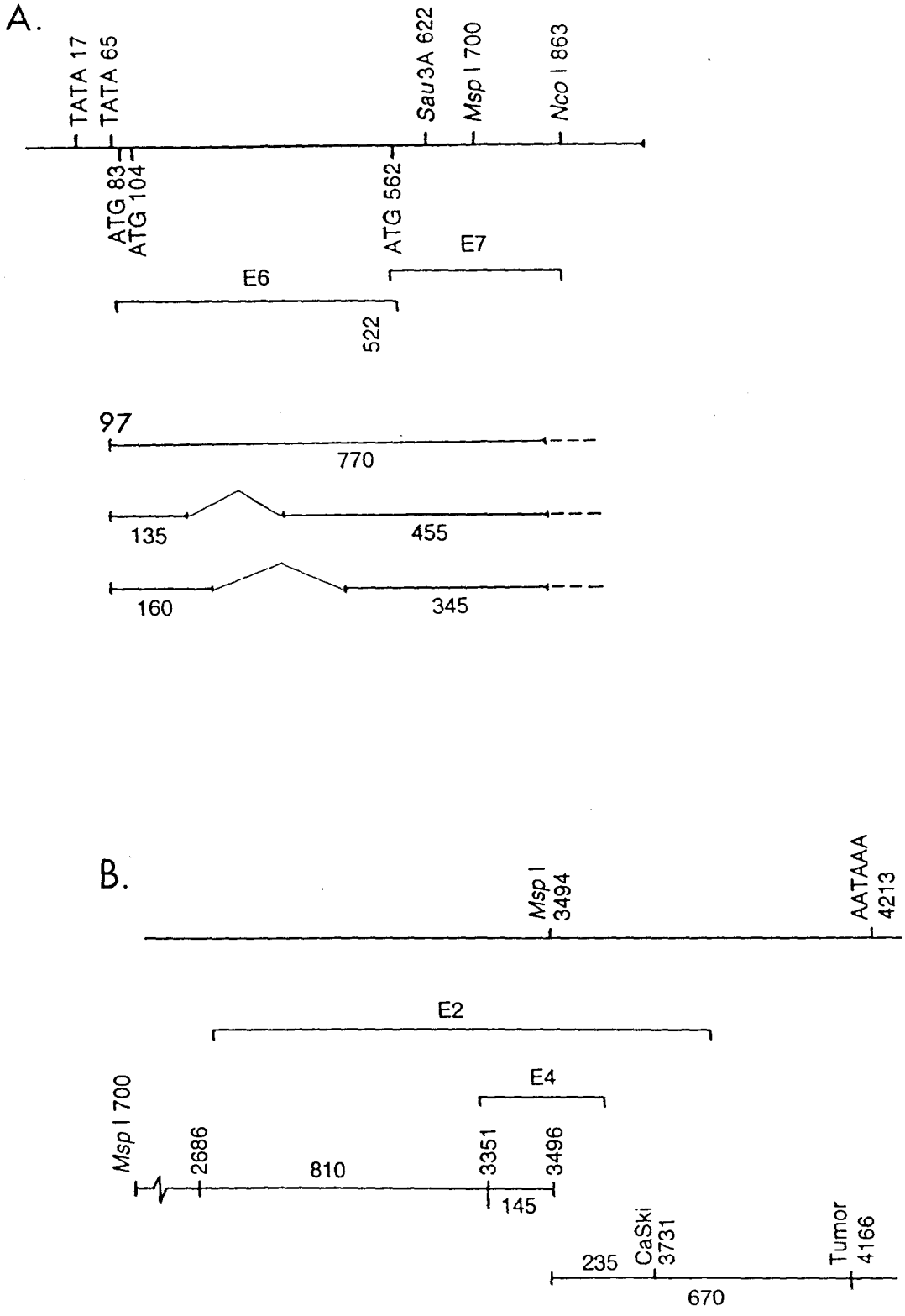
RNA isolated from CaSki cells and hybridized with HPV-16 specific probes corresponding to the early region yielded a fixed pattern consisting of three bands 1.5, 2.3, and 4.5 Kbs in size. All three transcripts are transcribed

Figure 8A Structure of Early HPV-16 Messages in CaSki Cells

The structures of RNA transcripts in CaSki cells are shown. Transcription initiation in all three is from nucleotide 97. The first transcript is considered minor and codes for the E6 ORF. The second (major) and third (minor) code for the E7 protein utilizing different splicing signals to splice out sequences coding for E6 ORF (Smotkin and Wettstein, 1986).

Figure 8B The 3' End of HPV-16 Transcripts in CaSki Cells

Analysis of the 5' end of transcripts coding for the E2/E4 region in CaSki and cervical tumor cells revealed a major (3351) and a minor (2686) 5' ends. In CaSki cells the 3' end of those transcripts was found at position 3731 and in tumor cells at position 4166 both well short of the poly(A) signal at 4213 (Smotkin and Wettstein, 1986).


A.
TATA 17
TATA 65
Sau3A 622
Msp I 700
Nco I 863
ATG 83
ATG 104
ATG 562
E7
E6
522
97
770
135
455
160
345
B.
Msp I
3494
AATAAA
4213
E2
E4
Msp I 700
2686
810
3351
3496
145
235
CaSki
3731
Tumor
4166
670

from the P_{97} promoter but have different coding capacities although two are similar by having sequences in E6 spliced out at different splice signals (Smotkin and Wettstein, 1986). The major transcript in CaSki cells (1.5 Kb) has sequences coding for E6 spliced out. This transcript utilizes a splice donor signal at position 232, and a splice acceptor signal at position 412 (Fig. 8a). One of the other two minor transcripts (2.3 Kb) also has sequences in the E6 region spliced out utilizing a splice donor signal at 257, and a splice acceptor signal located at position 518. Both splice acceptor signals are located upstream of the translation initiation site, encoded by the sequence ATG (position 562), of the E7 open reading frame (Fig.8a). These two transcripts are designated as the mRNA for the E7 protein. A third transcript (4.5 Kb) which is also considered as minor in terms of its abundance, does not have any sequence from the E6 region spliced out. This transcript consists of an exon extending from nucleotide 97 to the 3' end of the E6/E7. Mapping the 3' end of those transcripts revealed two positions at which the transcripts end and those are nucleotides 884 and 940 (Smotkin and Wettstein, 1986).

In the same study, mRNA from a cervical tumor positive for HPV-16 was analyzed. Only the major

transcript was detected. In another study, (Smotkin et al., 1989) in which a similar tumor was analyzed, both the major transcript and the minor transcript containing a spliced out sequences in E6 were detected. The minor mRNA transcript lacking any spliced out sequences in E6, and coding for an intact E6 protein but not for an E7 protein, was detected (Smotkin et al., 1989). Oncogenic HPVs such as HPV-18, 31, and 33 have splicing signals similar to those found in HPV-16 for the E6 intron. These signals are absent in non-oncogenic HPVs such as HPV6b and HPV-11 which are usually the causative agents of benign genital condylomata commonly known as genital warts (Chow et al., 1987a).

When CaSki cells and a cervical tumor were analyzed in parallel, transcripts originating from the E6/E7 region were found to splice into exons coding for the E2/E4 region (Fig. 8b). These E2/E4 exons have two 5' ends: a major and a minor. The major 5' end is situated within E4 sequences at position 3351. The minor 5' end is located further upstream 5' to E2 at position 2686. When the 3' ends were analyzed in both CaSki and tumor RNA, they were found to be different. In CaSki, the 3' ends of the exons corresponded to sequences within E2 but beyond E4 at position 3731. The 3' ends of the same exons in RNA extracted from the tumor

were located downstream of E2 at position 4166. The transcripts in the tumor, unlike those in CaSki cells, have the capacity to code for an intact E2 protein. HPV genomes in cervical cancer cell lines have been found to be integrated into cellular sequences. HPV-18 integrates in the HeLa cell line just 5' to c-myc proto-oncogene. HPV-16 derived from two cervical carcinomas was found to integrate proximal to c-src-1 and c-raf-1 proto-oncogenes. It is suggested that papilloma sequences can act as enhancers and/or promoters which over-express nearby proto-oncogenes facilitating by that the emergence of a malignant state (Durst et al., 1987).

1.4 Project Rationale

As discussed previously, the mechanism of the switch from early to late gene expression in papillomaviruses is still largely ambiguous. In differentiated keratinocytes infected with HPVs where late genes are expressed, studies have failed to locate the late promoter used. The switch to late gene expression in HPVs seems to be contributed by a variety of mechanisms such as differential promoter utilization, change in the selection of splice signals and 3' polyadenylation sites, as well as transcription

termination and mRNA turnover. Cellular factors in differentiated epithelial cells are largely believed to orchestrate the switch from early to late gene expression. The normal differentiation process of epithelial cells is yet to be mimicked in vitro. The tight link between early to late gene expression switch and the state of differentiation of infected cells makes it crucial to delineate those factors involved in the differentiation process from those which influence transcription of HPVs.

One way to facilitate the study of HPV gene regulation is by introducing HPV into the nuclei of different cells by the utilization of a viral vector. The focus of this thesis is to test the feasibility of this idea. Herpes simplex virus (HSV-1) was the vector of choice for several reasons. First, the 152 Kb genome of HSV was demonstrated to accept DNA fragments as large as the rabbit beta-globin gene to generate functional mutants (Smiley et al., 1987). Second, efficient methods have been developed to generate viral recombinants without interfering with the lytic cycle of HSV (Ligas and Johnson, 1988; Smiley, 1980) Third, herpes simplex virus seems to target the same epithelial cells which HPV infects and has a broad host range (Roizman and Sears, 1990). Fourth, HSV DNA is delivered to the nucleus of infected cells where it

is transcribed by the host's RNA polymerase. The short term goal of this project was to investigate whether a viable HSV recombinant carrying the entire HPV-16 genome could be constructed. If a viable recombinant was constructed, then later studies would try to learn whether HPV-16 promoters were active during lytic infection with the virus. The long term goals, not the subject of this thesis, were to determine whether papilloma early and late gene expression were subject to the same cell specificity as in natural HPV infection and whether such a regulation is influenced by any HSV encoded proteins.

Infection by HSV-1 is lytic in nature yielding a large quantity of progeny virus (Costanzo et al., 1977). Genes encoded by the viral genome belong to three different groups which are activated sequentially to high levels by viral proteins (Honess and Roizman, 1974). It was widely accepted that during viral infection transcription of cellular genes is completely shut-off (Fenwick and Roizman, 1977). Lately it has become apparent that some cellular genes are also expressed after infection with herpes particularly when these genes are newly introduced into cells by transfection or as part of the HSV genome. The promoter of such a limited group of cellular genes may contain sequences which can be recognized by viral proteins

resulting in the expression of such genes (Smiley et al., 1990). The first group of viral genes to be transcribed, the immediate-early (IE) genes, are five in number and their transcription is independent of new viral protein synthesis (Mackem and Roizman, 1980). Transcription of immediate-early genes is controlled by upstream sequences which are recognized by a structural component from the virion termed Vmw65 (Batterson and Roizman, 1983). The immediate-early gene products play an important role in activating the transcription of the other two classes of genes, the early and the late. More specifically, the product of the immediate-early gene, Vmw175, seems to be required for efficient early and late gene expression.

Dr. J.R. Smiley and co-workers constructed a non-defective virus mutant which contains the entire rabbit beta-globin gene inserted into HSV-1 thymidine kinase gene (Smiley et al., 1987). During productive infection in Vero cells, the inserted globin gene produced messages which were normally spliced and processed indicating that the globin promoter was being used as opposed to neighboring viral promoters. The globin gene was expressed in a fashion similar to the early TK gene of HSV since the level of globin RNA transcripts did not drop when DNA synthesis was blocked. When cycloheximide was used to block viral

immediate-early protein synthesis, the expression of globin RNA was greatly reduced.

These results showed that a cellular promoter could be recognized as an HSV early promoter and transcription could, therefore, be activated from such promoters during productive infection alongside other early HSV genes. This further strengthened the view that activation of transcription by HSV immediate-early proteins does not require the recognition of virus-specific promoter elements. The fact that the products of two immediate-early genes, Vmw175 and Vmw110, interact with a wide spectrum of viral and non-viral DNA sequences makes it difficult to identify a specific mechanism by which immediate-early proteins activate transcription of the globin gene. Possible mechanisms include attracting cellular transcriptional factors or changing the RNA polymerase II so that the globin gene is recognized as a herpes virus early gene (Smiley et al., 1987).

In another study, McCusker and Bacchetti showed that Vmw175 enhances transcription of the upstream regulatory region (URR) of both HPV-1 and 16, but not the URR of HPV-11. Other experiments conducted in the same study showed that Vmw110 is capable of activating transcription of the URR of HPV-16. In this experiment the promoter used

to investigate the reactivity of URR to Vmw175 and Vmw110 was not a papilloma promoter, nevertheless, this result is encouraging to further investigate HSV/HPV interactions in a system containing the entire genome of HPV-16 inserted into HSV-1 (Bacchetti and McCusker, 1988).

In this thesis I describe the construction of an HSV-1 recombinant with the entire HPV-16 genome inserted in the sequences coding for the dispensable glycoprotein I. I demonstrated that the early HPV-16 promoter (P_{97}) was active during lytic infection of Vero cells with the recombinant. The constructed recombinant may be useful in future studies aimed at delineating HPV-16 late gene expression.

II- Materials and Methods

2.1 Bacterial Culture

Plasmids were maintained in E. coli strain HB101. These cells were grown in Luria Broth (LB), or on Luria agar plates. Suspension cultures in LB were grown overnight with shaking at 37°C. LB was made up of 1% (w/v) yeast extract (Difco), 125 mM NaCl, 10 mM Tris HCl pH 7.6, and 0.4% (w/v) glucose. Ampicillin or kanamycin (Sigma) was added to a final concentration of 35 μg/ml.

Luria agar plates were prepared from LB to which 1.5% bacto-agar (Difco) was added. Cells were grown at 37°C in a bacterial cell incubator in inverted petri dishes (Fisher) containing agar-based medium. Ampicillin or kanamycin (Sigma) was added to a final concentration of 35 μg/ml after the autoclaved medium was allowed to cool down to 55°C, and just before pouring into petri dishes.

2.2 Transformation of E. coli with plasmid DNA

E. coli HB101 cells were made competent by the following procedure: viable bacterial cells were retrieved

by scraping the frozen surface (-70°C) of liquid cultures previously diluted 1:1 with 40% glycerol (v/v) with a sterile inoculation loop. The cells were grown overnight at 37°C with constant shaking. 0.5 ml of the overnight culture was added to 100 mls of Luria broth and grown at 37°C to an OD_{550} between 0.35-0.6. The cells were chilled on ice, and subsequently centrifuged at 13000 rpm for 15 minutes. The cell pellet was then resuspended in 2 ml 0.1 M $CaCl_2$, and left for 20 minutes on ice followed by centrifugation as above. The cell pellet was again resuspended in 2 ml 0.1 M $CaCl_2$. The cells were then competent for transformation.

To 200 μl of freshly competent cells, 15 μl of plasmid DNA and 100 μl of (100 mM NaCl, 33.3 mM $CaCl_2$) solution were added. The mixture was placed on ice for 20 minutes and then heat shocked at 42°C for 2 minutes. The mixture was placed on ice for one hour, shaking gently every 20 minutes. Each mixture was transfered to a plastic tube (Falcon 2059) containing 4 ml LB. Each sample was incubated at 37°C for 45 minutes with shaking.

20, 80, and 200 μl aliquots of the transformed cells were applied on solid LB plates containing either ampicillin or kanamycin. After the liquid was absorbed, the plates were inverted and incubated at 37°C for 16-20 hours.

2.3 Isolation of Plasmid DNA

2.3.1 Large Scale Isolation of Plasmid DNA

500 ml of LB containing ampicillin or kanamycin was inoculated using a saturated overnight culture of HB101 bacteria. The culture was allowed to grow to OD_{600} 0.7-0.8, after which chloramphenicol (Sigma) at a final concentration of 200 mg/ml was added. Amplification of plasmid copy number was allowed overnight at 37°C.

Overnight cultures were centrifuged at 5000 rpm for 15 minutes at 4°C. The cell pellet was resuspended in 12 ml of freshly made mixture containing: 25 mM Tris-Cl pH 7.5, 10 mM EDTA, 15% sucrose, and 2 mg/ml lysozyme (Boehringer Mannheim). The resuspended cell pellets were left at room temperature for 5 minutes with mixing by carefully inverting the mixture. 24 ml of freshly mixed alkaline sodium dodecyl sulfate (SDS) consisting of 0.2 M NaOH and 1% SDS were added to the cell mixture which was then incubated for 10 minutes. 15 ml of 5 M KAc, pH 4.8, was then added with further incubation for 20 minutes on ice.

The resulting mixture was centrifuged at 20000 rpm for 20 minutes at 4°C using SW 28 Beckman rotor. The

supernatant containing DNA was retrieved and 2.5 volumes of isopropanol were added to it. The precipitated DNA was pelleted by centrifugation at 12000 rpm for 15 minutes at 4°C. DNA was then resuspended in 6 ml of TE (10 mM Tris pH 7.5, 1 mM EDTA) buffer. An equal volume of 5 M LiCl was added to precipitate the RNA. The mixture was left on ice for 10 minutes, and then centrifuged at 12000 rpm for 15 minutes at 4°C. 2.5 volumes of ice-cold 95% ethanol was added to the supernatant. The precipitated DNA was pelleted by centrifugation at 12000 rpm for 15 minutes at 4°C. The pelleted DNA was resuspended in 0.5 ml TE buffer and treated with 50 μl of 1 mg/ml RNase A for 15 minutes at 37°C. An equal volume of polyethylene glycol (PEG) solution (13% PEG, 1.6 M NaCl) was added, and the mixture was incubated for 30 minutes on ice. The DNA precipitate was pelleted by centrifugation at 13000 rpm for 5 minutes in an Eppendorf centrifuge. The DNA was later extracted with an equal volume of a phenol-chloroform (1:1, v/v) mixture.

The phenol was prepared as described in Maniatis et al., 1982. Chloroform was prepared by mixing chloroform and isoamyl alcohol (24:1, v/v). Phenol-chloroform was added to the DNA solution, the mixture was vortexed, and the aqueous layer was re-extracted as above until no precipitate was observed at the interphase. 1 ml of 95%

ethanol and a final concentration of 0.15 M NaCl was added to the aqueous phase to precipitate the DNA. The DNA precipitate was centrifuged in an Eppendorf centrifuge at 13000 rpm for 15 minutes, and the pellet was resuspended in 0.5 ml of 0.3 M Sodium acetate (NaAc). The DNA was then precipitated by adding 2.5 volumes of ethanol and then centrifuged at 13000 rpm for 15 minutes. The DNA pellet was then washed by adding one ml of 70% ethanol, vortexing, and then centrifuging in an Eppendorf centrifuge. Subsequently, the pellet was washed in the same way but with 95% ethanol. The DNA pellet was then dried in a speedVac and redissolved in 0.6 ml of autoclaved double-distilled water.

2.3.2 Small Scale Isolation of Plasmid DNA

The procedure used is a modified version of the one reported by Holmes and Quigley, 1981. 4 ml of antibiotic containing medium was inoculated with a single bacterial colony. This was subsequently incubated at 37^{o}C overnight with agitation. 3 ml of the overnight culture was divided into two Eppendorf tubes, and the remainder of the culture was stored at 4^{o}C. The two tubes were centrifuged for 2 minutes at 13000 rpm in an Eppendorf centrifuge. The medium was removed by aspiration, and the bacterial pellets

were both resuspended together in 0.30 ml sucrose solution made up of 20% sucrose (w/v) and 50 mM Tris-HCl, pH 8.0. 30 μl of freshly prepared 10 mg/ml lysozyme (Boehringer Mannheim) were added, and the mixture was left on ice for 15 minutes. 150 ml of a Triton lysis buffer (0.3% Triton X-100, 188 mM EDTA, 150 mM Tris-HCl pH 8.0) was added to the mixture. The sample was then incubated in a boiling water bath for 2 minutes. Subsequently, the sample was pelleted by centrifugation at 13000 rpm in an Eppendorf centrifuge for 10 minutes. The supernatant was removed into a new tube to which an equal volume of ice cold isopropanol was added followed by a 30 minutes incubation at -20^{o}C. After 30 minutes, the sample was centrifuged at 4^{o}C for 15 minutes, and the pellet dried in a SpeedVac lyophilizer. The dried pellet was then resuspended in 50 μl of autoclaved double-distilled water.

2.4 Gel Electrophoresis

2.4.1 Analytical Purpose

DNA fragments were separated by electrophoresis through agarose (Sigma) or polyacrylamide gels (BioRad). Agarose gels were prepared by melting 1% agarose (w/v) in TAE buffer (0.04 M Tris-acetate 0.002 M EDTA, pH 8.0). The

gel was then poured on a removable glass plate. The plate was installed on the platform of a horizontal gel tank such that the gel was submerged just beneath the surface of the electrophoresis buffer (TAE). Vertical polyacrylamide gels were prepared in TBE buffer (0.089 M Tris, 0.089 M Boric acid, 0.002 M EDTA, pH 8.3), and the final concentration of the polyacrylamide solutions ranged from 5% to 10%. Polyacrylamide gels were made by mixing 40% (w/v) acrylamide, 1.37% (w/v) N_1N_1-methylene-bis-acrylamide (BIS) (Bio-Rad) solutions with electrophoresis buffer, polymerization was achieved by adding 0.1 ml TEMED (N, N, N', N'-tetramethylethylene diamine) (Bio-Rad). Polyacrylamide gels were prepared using 19 cm and 32 cm long glass plates. Polyacrylamide gels were run in TBE buffer. DNA Samples were loaded after being mixed with a dye made up of: 5% (v/v) glycerol, 0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanol and 10 mM EDTA.

Electrophoresis was carried out at constant voltage. The running buffer was circulated constantly throughout electrophoresis by means of a peristaltic pump. After the separation of bands, DNA fragments were stained by immersing the gels in electrophoresis buffer containing 0.5 μg/ml ethidium bromide. The bands were visualized by illuminating the gels with ultra-violet light. In the case of labelling with a radioactive isotope (Southern blots,

primer extension or S1 nuclease mapping), the gels were exposed to Kodak XAR-5 fast x-ray films overnight at -70°C.

2.4.2 Purification of DNA fragments

Purification of DNA fragments was performed as described by Maniatis et al., 1982. The DNA band of interest was excised from the polyacrylamide gel, placed in an Eppendorf centrifuge tube, and crushed using a silicon rod. 700 µl of elution buffer (500 mM Ammonium acetate, 10 mM Magnesium acetate, 1 mM EDTA, and 0.1% (w/v) SDS), were added and the mixture was incubated at 37°C overnight. The next day residual polyacrylamide was pelleted by centrifugation in an Eppendorf centrifuge for 10 minutes. The supernatant was removed and the residual polyacrylamide washed with 200 µl of the elution buffer and re-centrifuged. The two supernatants were pooled together and DNA was precipitated by adding 800 µl of 100% ethanol. The precipitated DNA was pelleted as described before and reprecipitated out of 0.3 M NaAc and then washed with ethanol and dried.

2.5 Enzymes

2.5.1 Restriction Analysis and Recombinant Plasmids

Restriction endonucleases used to digest DNA fragments or plasmids were obtained from Bethesda Research Laboratories (BRL) and Pharmacia. Digestions were carried out in small volumes (10-200 μl) as described in Maniatis et al., 1982. Generation of recombinant plasmids, and reactions using T_4 DNA ligase (BRL) and alkaline phosphatase were carried out as described in Maniatis et al., 1982.

2.5.2 Nick Translation

A nick translation kit along with recommended procedures for labeling DNA by nick translation was purchased from BRL. To an Eppendorf tube, 5 μl containing 0.2 mM of dNTP (dATP, dGTP, and dTTP), 500 mM Tris-HCL (pH 7.8), 50 mM $MgCl_2$, 100 mM 2-mercaptoethanol, and 100 μg/ml nuclease-free bovine serum albumin (BSA) was added. 1 μg of plasmid DNA, 100 μCi alpha ^{32}P labelled dCTP (specific activity 3000 Ci/mmol, New England Nuclear) was then added to the same tube in which the volume was brought up to 45

μl by the addition of distilled H_2O. 5 μl of DNA polymerase I/DNase I (100 units) (0.4 U/μl E. coli DNA polymerase I, 40 pg/μl DNase I, 50mM Tris-HCl (pH 7.5), 5mM $MgAc_2$, 1 mM 2-mercaptoethanol, 0.1 mM phenyl-methyl-sulfonyl-fluoride (PMSF), 50% (v/v) glycerol, and 100 μg/ml nuclease-free BSA) was then added to the mixture which was incubated at 15°C for 60 minutes. The reaction was then stopped by adding 5 μl stop buffer (300 mM Na_2EDTA pH 8). Labelled DNA was separated from unincorporated $dNTP_s$ by centrifugation through a small column of Sephadex G-50 (Pharmacia) as described in Maniatis et al., 1982.

2.6 Southern Blots

The procedure followed is a modification of the protocol developed by Southern, 1975. After electrophoresis, the agarose gel was soaked in the following solutions: 0.25 M HCl for 30 minutes, 0.5 M NaOH for 60 minutes, 1.5 M NaCl / 1 M Tris-HCl pH 7.5 for 30 minutes, and 20 x SSC for 30 minutes. The gel was then placed on top of 4 pieces of Whatman 3MM paper presoaked in 20 x SSC and resting on a clean glass plate. A nitrocellulose sheet, presoaked in distilled water and then in 20 x SSC, was placed over the gel. 2 pieces of 3 MM paper presoaked in 20 x SSC and 2 dry 3 MM pieces were

placed over the nitrocellulose sheet. The edges of the set were covered with parafilm around the gel, and air bubbles were removed from the gel surface. A stack of paper towels was placed over the dry 3MM paper, and a weight applied on top of it. The DNA fragments were allowed to transfer onto the nitrocellulose overnight. The nitrocellulose was then baked at 80°C for 2 hours.

The baked nitrocellulose was placed in a plastic bag and 20 ml of prehybridization solution (2 ml 20 x SSC, 2 ml Denhardt's, 500 µl Salmon Sperm DNA, and 15.5 ml double-distilled water) was added. The blot was prehybridized for 5 hours at 60°C. The prehybridization solution was then discarded, and replaced with 20 ml of hybridization solution (2 ml 20 x SSC, 4 ml 10% Dextran Sulphate, 1 ml Denhardt's, 200 µl 1 M Tris pH 7.6, 250 µl Salmon Sperm DNA, 11.5 ml dd H_2O, and 1 x 10^6 cpm/ml of denatured nick translated probe in 1 ml). Hybridization was allowed to proceed overnight in a 60°C water bath.

The next day, the nitrocellulose was washed three times with 2 x SSC and 0.1% SDS at 67°C, each wash lasting 15 minutes. This was followed by three similar washes with 0.1 x SSC and 0.1% SDS at 54°C. The blot was then air-dried and exposed to Kodak XAR-5 x-ray film.

2.7 Tissue Culture

African Green Monkey Kidney (Vero) cells were grown in monolayers in plastic tissue culture flasks and plates. Cells were kept at 37°C in a humidified atmosphere containing 5% CO_2. Vero cells were grown in modified alpha-minimal essential medium (α-MEM, GIBCO) containing 5% (v/v) fetal bovine serum (FBS,GIBCO), 1% (v/v) L-glutamine (GIBCO), 1% (v/v) penicillin-streptomycin (GIBCO), and 2% $NaHCO_3$. Cells were sub-cultured as needed by washing in phosphate buffered saline (PBS) (137 mM NaCl, 3 mM KCl, 1 mM KH_2PO_4, and 20 mM Na_2HPO_4, pH 7.3) followed by trypsinization.

CaSki cells were obtained from American Type Culture Collection. Those cells are derived from a human cervical epidermoid carcinoma. CaSkis are positive for HPV-16, and are believed to contain more than 600 gene copies of integrated viral DNA. RNA species coding for the early region of HPV-16, and the E7 protein have been identified in this cell line (Smotkin and Wettstein, 1986).

VD60 cells were maintained in Eagle MEM (GIBCO) lacking histidine (MEM-his) supplemented with 0.3 to 1.2 mM histidinol (Sigma). VD60 cells were passed at least once

in α-MEM containing 7% FCS before infection. VD60 is a cell line capable of expressing high levels of gD glycoprotein constitutively, and therefore has the ability of complementing the replication of gD^- mutant herpes viruses (Ligas and Johnson, 1988).

All viruses except F-gDβ were propagated and titered on Vero cells. Flat 150 cm^2 flasks, with a monolayer of about 2 x 10^7 cells, were infected with 1 x 10^6 plaque forming units (pfu) of virus in a volume of 10 ml of serum free α-MEM containing 1% l-glutamine, 1% penicillin-streptomycin, and 2% $NaHCO_3$. Adsorption was allowed for two hours, after which 20 ml of pre-warmed α-MEM with 5% (v/v) FBS was added to the infected cells. After 2-3 days when cytopathic effects (cpe) were observed, the old medium was discarded and cells were harvested by scraping or tapping flask walls. Cells were then resuspended in PBS, and recovered by centrifugation at 2700 rpm at 4°C. Cells were then resuspended in α-MEM containing 5% FBS and sonicated. Cellular debris was removed by centrifugation, and the supernatant containing the released virus was aliquoted and stored at -70°C.

Viruses were titered by infecting confluent monolayers of Vero cells growing in 6 well (Corning) dishes with serial dilutions (ranging from 10^{-2} to 10^{-7}) of the viral stock in serum free α-MEM. Two hours post-infection,

the wells were overlaid with pre-warmed (37°C) α-MEM containing 5% FBS and 0.05% human immune serum (Connaught Laboratories) promoting by that the growth of independent viral plaques. Infected cells were incubated at 37°C for 3 days after which plaques were counted under a light microscope.

2.8 Preparation of Infectious Viral DNA

All infections were performed at an moi (multiplicity of infection) of 10 where ten plates or flasks of confluent Vero cells or VD60 cells were used for infection with virus. Cells were harvested at 4^+ CPE. Infections and harvests were carried out as described before. Cell pellets were resuspended in 10 ml 0.2 M EDTA pH 8.0. SDS was then added to a final concentration of 0.5%. The mixture was gently mixed by inverting. Proteinase K (Boehringer Mannheim) was then added to a final concentration of 100 μg/ml, and the mixture was extracted with phenol 5 times, and placed in a 15.9 mm diameter standard cellulose dialysis tubing. Dialysis was performed in a 2 liter plastic bucket containing 0.1 x SSC in the cold room with a magnetic stirrer. The solution was changed many times over a period of 3 days. Infectious DNA was then collected in a 50 ml Corning tube, and kept at

4°C.

Transfections took place on Vero cells in all cases with different volumes of infectious DNA being used to determine the best titer yield.

2.9 Construction of the Recombinant Virus RI

Infectious viral DNA and plasmid DNA were co-transfected into Vero cells as described by Smiley (1980) to generate recombinant viruses. More specifically, F-gDβ infectious DNA was co-transfected into Vero cells with a plasmid carrying a BamHI fragment of HSV-1 containing the essential glycoprotein D (gD) and the non-essential glycoprotein I (gI) genes. The gI gene in this plasmid had the HPV-16 genome inserted at the Bgl II site. F-gDβ contains the β-galactosidase sequences of E. coli replacing gD and part of gI yielding a mutant virus which cannot infect gD non-complementing cells such as Veros. A double recombination event between F-gDβ infectious DNA and the plasmid DNA will result in a virus which has an intact gD gene and the HPV-16 genome in gI. Such a virus can infect Vero cells, and DNA from resulting plaques can be analyzed for papilloma sequences.

2.9.1 Transfection of Vero Cells

The cell transfection method of Graham and VanderEb (1973) was followed with some modifications. The transfection reaction consists of 0.250 ml of HEPES buffered Saline (42 mM HEPES, 27 mM NaCl, 10 mM KCl, 1.4 mM Na_2HPO_4, 11 mM dextrose, pH 7.1), 20-60 μl infectious viral DNA, 500 ng of plasmid DNA, 10 μg denatured Salmon testes DNA, where the final volume is brought up to 0.475 ml with H_2O. 2.5 μl of 2.5 M $CaCl_2$ was subsequently added and the mixture was briefly shaken and left for 20 minutes at room temperature. The mixture was then added to 50% confluent Vero monolayers in 6 well dishes (Corning) containing 4 ml α-MEM enriched with 5% FBS. The dishes were then incubated at 37^{o}C for 5 hours. The medium on the transfected cells was aspirated and the cells were glycerol-shocked by adding 1 ml of serum-free α-MEM containing 15% glycerol. This medium was left on for 1 minute after which 4 ml of serum-free α-MEM medium was added and the whole mixture aspirated. The glycerol-shocked cells were washed 5 times by adding 4 ml of serum-free α-MEM and aspirating. Finally, 5 ml of α-MEM containing 5% FBS was added to the cells which were incubated at 37^{o}C. Confluent cpe was observed one week after transfections. Viruses picked from

transfection reactions were propagated and titered as previously described.

2.9.2 Isolation of gD^{+} recombinants

Viruses growing on Vero cells as a result of transfection with FgD-β and plasmid DNA must have acquired the sequences coding for gD on the plasmid through a recombination event since gD is required for growth. A double recombination event resulting in the acquisition of further upstream sequences including HPV-16 was very probable and required confirmation by Southern analysis. Plaques were picked and used to infect Nunc 24 well dishes in which monolayers of Vero cells were growing. Following complete infection, viral DNA was extracted and analyzed.

2.10 Viral DNA Extraction

When confluent cpe was observed in Nunc 24 well dishes, the cells were scraped, and resuspended in the old medium. Half of the volume was stored frozen at -70^{o}C. The other half was collected in an Eppendorf tube. The cells were pelleted by centrifugation in an Eppendorf centrifuge. The cell pellet was resuspended in 0.5 ml pronase buffer (10 mM Tris-HCl pH 7.8, 10 mM EDTA, and 0.5%

SDS). Pronase was then added to a final concentration of 1 mg/ml, and incubated for 3 hours in a 37°C water bath. The mixture was then extracted with phenol 4-5 times, and the DNA precipitated and washed as described before.

2.11 Extraction of Total Cellular RNA

RNA extraction was performed according to the RNAzol™ method described by Chomczynski (Chomczynski and Sacchi, 1987). RNase contamination was reduced by treating solutions with 0.1% diethyl pyrocarbonate (DEPC) and autoclaving.

Infected cells in 75cm^2 flasks (Corning) were washed with PBS, and then scraped into the PBS solution. The cells were then centrifuged at 2700 rpm at 4°C, and the cell pellet resuspended in 1ml RNAzol in an Eppendorf tube. 0.1 ml of chloroform was added, and the mixture quickly vortexed. The solution was then incubated on ice for 15 min.. The mixture was then centrifuged at 12,000 rpm in an Eppendorf centrifuge at 4°C for 15 min. The aqueous phase was then removed to a new centrifuge tube, and an equal volume of isopropanol was added. The RNA was left to precipitate O/N at -20°C. The next day RNA was pelleted by centrifugation at 12,000 rpm for 15 min. at 4°C. The isopropanol was carefully removed, and the RNA pellet

redissolved in 0.4 ml DEPC treated water. 40 μl of DEPC treated 3 M NaAc (pH 5.2) were then added to the dissolved RNA. The RNA was precipitated by the addition of 0.8 ml of 100% ethanol followed by incubation at -70°C for 20 min. The RNA was then pelleted by centrifugation at 12,000 rpm for 15 min. at 4°C, and washed twice with 75% and 95% ethanol. The RNA was then dried using a SpeedVac lyophilizer without heating. The dry RNA pellet was resuspended in 0.4 ml DEPC treated water and stored at -70°C. RNA concentration was measured by UV absorbance spectroscopy at 260 nm.

2.12 5' Labelling of Synthetic Oligonucleotides

5' end labelling of synthetic oligonucleotides was carried out according to Maxam and Gilbert, 1978. Oligos in a total volume of 5 μl were mixed with 35μl of Kinase I buffer (0.1 mM EDTA, 20 mM Tris pH 9.5, 1 mM Spermidine). 5 μl Kinase 2 buffer (500 mM Tris pH 9.5, 100 mM $MgCl_2$, 50 mM DTT, 50% glycerol), 100 μCi gamma 32P labelled dATP, and 2 μl (10 units) T4 polynucleotide kinase were added. The reactants were thoroughly mixed, and then incubated in a 37°C water bath for one hour. 200 μl of 2.5 M Ammonium Acetate, 1 μl of 3 mg/ml tRNA, and 750 μl 95% ethanol were added. The mixture was then placed on dry ice for 15

minutes, after which the labelled oligos were pelleted by centrifugation in an Eppendorf centrifuge at 4°C for 15 minutes. The supernatant was removed, and the pellet was re-dissolved in 250 μl 0.3 M Sodium Acetate. The labelled oligos were then reprecipitated and pelleted as above. The oligo was then dried in a speedVac for 15-20 minutes, and resuspended in 50 μl H_2O.

2.13 Gel Purification of Oligonucleotides

Labelled oligonucleotides were run on an 8% sequencing gel containing 7 M urea. The gel was then exposed to Kodak-5 X-ray film. Gel fragments containing labelled oligonucleotides detectable on the film were excised from the gel. The excised bands were incubated in 0.5 ml elution buffer (0.5 M NaAc, 10 mM $MgCl_2$, 1mM EDTA, and 0.1% SDS) containing 15 μg *E. coli* tRNA. After an overnight incubation at 37°C, the polyacrylamide was pelleted by centrifugation in an Eppendorf centrifuge. The supernatant was collected, and the oligonucleotides precipitated with 3 volumes of ice cold 95% ethanol. The oligonucleotides were pelleted by centrifugation, and resuspended in 300 μl of 0.3 M NaAc, pH 5.6. The oligonucleotides were reprecipitated using 3 volumes of ice cold 95% ethanol, and then pelleted by centrifugation. The

pellet was then washed with 70% and 95% ethanol, dried, and resuspended in TE (10 mM Tris-HCl pH 7.9, 1mM EDTA).

2.14 Primer extension

The method was reported by Smiley et al., 1987. 10-30 μg of cytoplasmic RNA was precipitated in 0.3 M NaAc pH 5.2 with 2.5 volumes ethanol, and dried in a SpeedVac for one hour without applying heat. Each pellet was resuspended in 8 μl of TE (10 mM Tris-HCl pH 7.9, 1 mM EDTA) containing 30000-50000 cpm of 5' labelled oligonucleotide. 2 ml of TKE (10 mM Tris-HCl PH 7.9, 1 mM EDTA, 1.25 M KCl) were added to the mixture. Promotion of hybridization between RNA and the oligonucleotide primer was done by incubating the mixture at 60°C for 1 hour. The mixture was allowed to cool down to room temperature. 25 ml of reverse transcriptase buffer (20 mM Tris-HCl pH 8.7, 10 mM $MgCl_2$, 5 mM DTT, 0.33 mM of dATP, dTTP, dGTP, and dCTP, 10 μg/ml actinomycin D) containing 10 units of avian myeloblastosis virus reverse transcriptase (Life Sciences) was added to the mixture. After incubation at 37°C for one hour, the extension products were precipitated by adding 0.3 ml cold 95% ethanol. The products were pelleted by centrifugation, washed with 70% and 95% ethanol, dried, and resuspended in 4 μl H_2O. To the extension product, 8 μl of sequencing dye

(90 % formamide v/v, 0.02 % w/v bromophenol blue, 0.02 % xylene cyanol, 0.089 M Tris, 0.089 Boric acid, 0.002 M EDTA) was added and the products were run on an 8% polyacrylamide sequencing gel containing 7 M urea. The gel was then exposed to Kodak XAR-5 X-ray film.

2.14.1 E97 oligo

This synthetic oligo was used to investigate transcription initiation from the early promoter P_{97}. The oligo is 25 base pairs long extending from nucleotides 152 to 177, and reads in the 5' to 3' direction as such: 5'-TCATGTATAGTTGTTTGCAGCTCTG - 3'. If the P_{97} promoter is utilized, then the product of a primer extention reaction using this primer should be 80 base pairs long.

2.14.2 AB166 oligo

This synthetic primer, 5'-ACGTGAGATATACGCCGGGCGGGTA-3', was obtained from the Central Facility of the Institute of Molecular Biology and Biotechnology in McMaster University. The 5' end labelled primer was used in primer extension experiments to monitor the expression of glycoprotein B (gB) of HSV-1. The expected extension product using this primer is 120 nucleotides long.

2.15 S1 nuclease mapping

The procedure for S1 mapping was reported by Smiley et al. (1987). 15 μg of RNA and 10^4 Cerenkov cpm of labelled probe were ethanol precipitated out of 0.3 M sodium acetate (pH 5.6). The pellets were washed with 70% and 95% ethanol after which they were dried in a speedVac for one hour without applying heat. The pellets were resuspended in 30 μl of hybridization buffer consisting of 0.4 M NaCl, 40 mM PIPES [piperazine N,N'-bis(2-ethanesulfonic acid)] (pH 6.4), and 1 mM EDTA and containing 80% (v/v) recrystallized, deionized formamide. The mixtures were heated at 80°C for ten minutes to denature the DNA probe. The mixtures were immediately incubated at 45°C for 16 hours. Each hybridization mixture was digested by adding 300 μl of S1 buffer made up of 150 mM NaCl, 50 mM NaAc (pH 4.6), and 5 mM $ZnSO_4$ containing 66.7 new units of S1 nuclease (Boehringer Mannheim). The digests were transferred, without delay, to a water bath at 37°C where they were left for one hour. The digests were cooled by incubating them in an ice bath for ten minutes. The reactions were then stopped by adding 50 μl of 4 M Ammonium Acetate and 30 μl of 0.2 M EDTA. Followig extraction with phenol/chloroform, 10 μg carrier tRNA was

added to each mixture. The products of S1 nuclease digestion were then precipitated by adding 750 μl of 95% ethanol. Pellets were washed with 70% and 95% ethanol and then dried. Each pellet was resuspended in 12 μl of sequencing dye, and run on an 8% polyacrylamide sequencing gel containing 7 M urea. The gel was then exposed to Kodak XAR-5 X-ray film.

2.16 Preparation of the S1 Probe

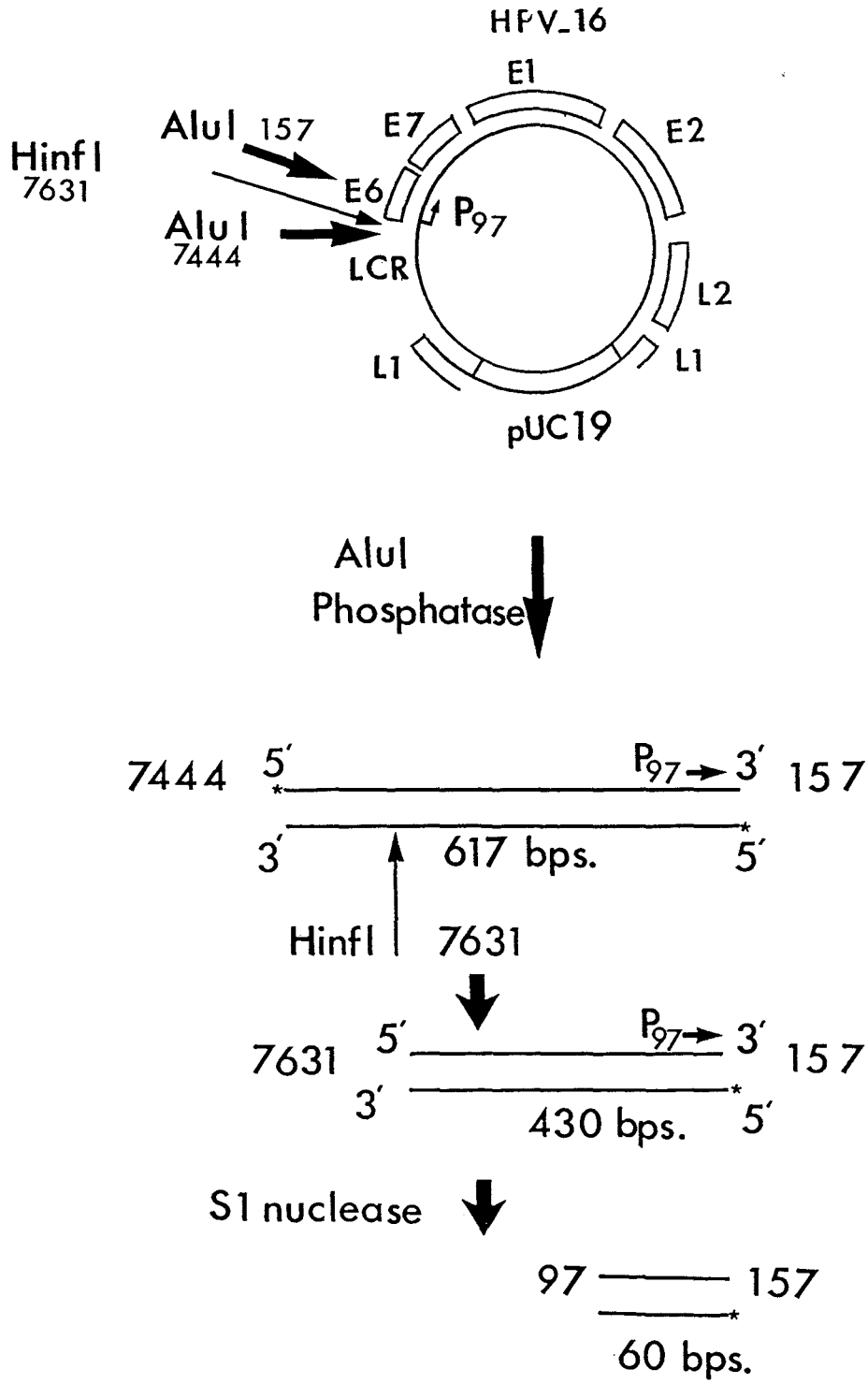
100 μg of HPV-16/pUC19 DNA was digested with AluI. After digesting for five hours, Tris (pH 9) was added to a final concentration of 50 mM. Plasmid DNA was treated with phosphatase in order to remove the 5' phosphate groups from both ends of the restricted fragments in preparation for 5' end labelling with T4 polynucleotide kinase. This was achieved by adding ten units of calf intestinal alkaline phosphatase to the mixture which was then incubated at 37°C for one hour. The 32 fragments resulting from this digest were resolved on a 5% medium sized acrylamide gel. A fragment 617 base pairs long was isolated and eluted as described before. This fragment of interest is the result of cleavage within the HPV sequences at positions 7444 and 157 (Fig. 9) (HPV-16 is 7904 base pairs long). This fragment is therefore ideal to investigate the activity of

the papilloma early promoter denoted P_{97}.

1 μg of the purified fragment was 5' end labelled using gamma ^{32}P labelled ATP and T4 polynucleotide kinase in three separate reactions as described in section 2.12 of this chapter. The three kinase reactions were then pooled together and digested with HinfI which cuts once in the 617 base pair fragment at position 7631 of the papilloma genome. The smaller of the two recovered fragments is 187 base pairs long and is located upstream of the P_{97} promoter (Fig. 9). The other fragment is 430 base pairs long and overlaps the promoter region 60 base pairs in the downstream 3' direction and 370 base pairs in the upstream 5' direction. The HinfI digest of the kinased fragment was run on a 5% medium sized acrylamide gel. The gel was then exposed to a Kodak XAR X-ray film for 20 minutes. The autoradiogram revealed the positions of the two fragments, of which the larger 430 base pair fragment was excised and purified. Transcription initiating from the P97 promoter should therefore result in a 60 base pair protected fragment in an S1 nuclease assay using the 430 base pair fragment as a DNA probe.

Figure 9 Preparation of S1 Nuclease probe

Preparation of the probe fragment used in S1 nuclease mapping was achieved by starting with an AluI digest of 100 μg of an HPV-16/pUC 19 plasmid the digest was then phosphatased and the fragment of interest (7444-157) was eluted out of a 5% acrylamide gel. 1 μg of this fragment was 5' end labelled and then cut with HinfI. A fragment 430 bps. large, overlapping the P_{97} promoter, and labelled at one end was utilized in S1 mapping. In the case of transcription from P_{97} this probe will protect a fragment which is 60 bps. large.


HPV_16
E1
E7
E2
AluI 157
HinfI
7631
E6
AluI
7444
P97
LCR
L2
L1
L1
pUC19
AluI
Phosphatase
7444 5′
P97→ 3′ 157
3′
617 bps.
5′
HinfI 7631
7631 5′
P97→ 3′ 157
3′
430 bps.
5′
S1 nuclease
97
157
60 bps.

III- Results

3.1 Construction of Recombinant Plasmids:

3.1.1 pFHS-1 and pFHS-2: HPV-16 in gI of HSV-1

In order to facilitate the study of HPV-16 gene expression and look more closely at the different factors involved in the switch from early to late gene expression, it was decided to insert the HPV-16 genome into a non-defective HSV-1 recombinant. The targeted gene in HSV-1 will be the one coding for a non-essential glycoprotein (gI) and will, therefore, not alter the infectivity and the host range of HSV-1. A system aimed at rescuing foreign DNA into HSV was developed in David Johnson's laboratory and utilizes a replication defective HSV-1 mutant (F-gDβ) which has the gD gene and part of the gI gene replaced by E. coli β-galactosidase gene. Such a defect can be rescued by transfecting F-gDβ into cells with plasmids containing intact gD sequences. At the same time, sequences such as

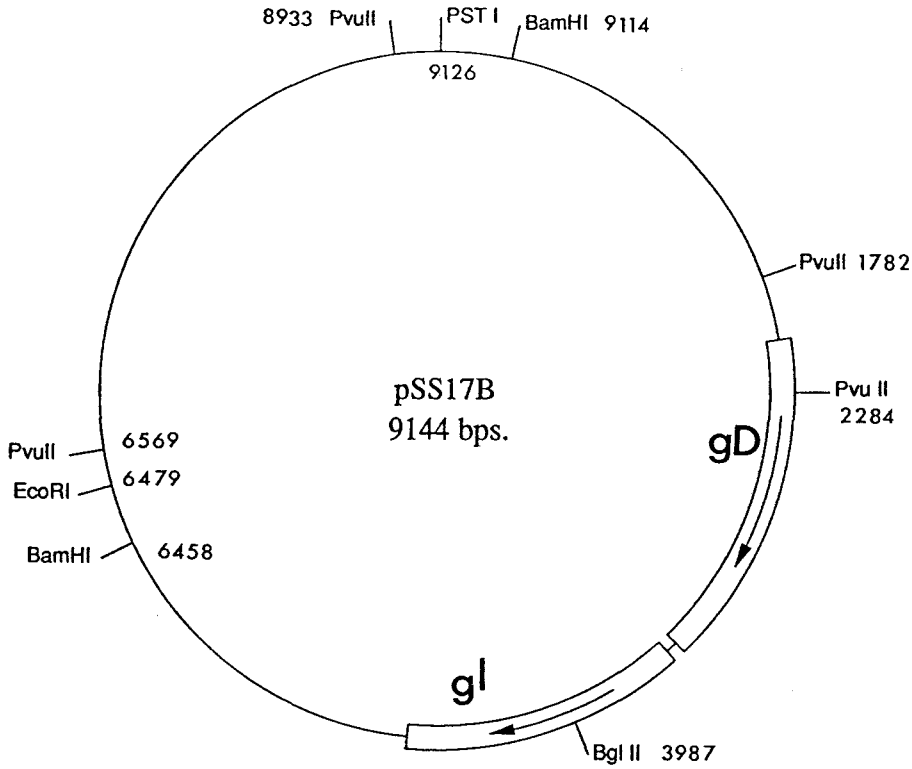
HPV-16 can be rescued into HSV by including these sequences in a non-essential herpes gene such as gI in the same plasmid carrying intact gD. The resultant virus, hence, will have intact gD, gI gene disrupted by HPV-16, and replicates normally on gD non-complementing cells such as Vero cells as detailed in section 3.2 (Ligas and Johnson, 1988). It was hoped that such a recombinant virus would allow the introduction of HPV-16 to a variety of cells permitting by that the evaluation of the effect of different cellular factors on the transcription of HPV-16 genes.

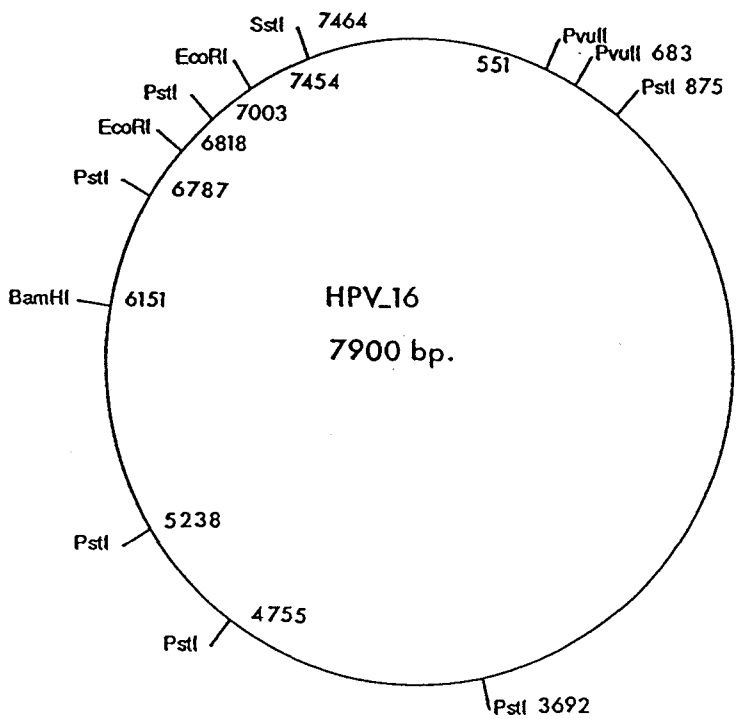
Preliminary experiments aimed at constructing a plasmid consisting of HPV-16 flanked by HSV-1 sequences which will be later rescued into HSV-1. A plasmid (pSS17B) consisting of an HSV-1 BamHI J fragment (136290-142747) and pUC19 containing, in addition to other HSV sequences, the essential gD gene and the non-essential gI gene was utilized for this purpose. The following is a description of the insertion of HPV-16 into the gI sequences of pSS17B.

The starting plasmid consisted of HPV-16 inserted into the BamHI site on pUC9 (Fig.9). This plasmid was cut with BamHI, and run on a 4% polyacrylamide gel. Subsequently, the band corresponding to HPV-16 was excised

Figure 10 Concise Restriction Maps of pSS17B and HPV-16.

A BamHI site (6150) in HPV-16 and a BglII site (3987) in pSS17B were utilized to construct the two recombinants pFHS-1 and pFHS-2.


8933 PvuII
PST I
BamHI 9114
9126
PvuII 1782
Pvu II
2284
pSS17B
9144 bps.
gD
gI
PvuII
6569
EcoRI
6479
BamHI
6458
BglII 3987

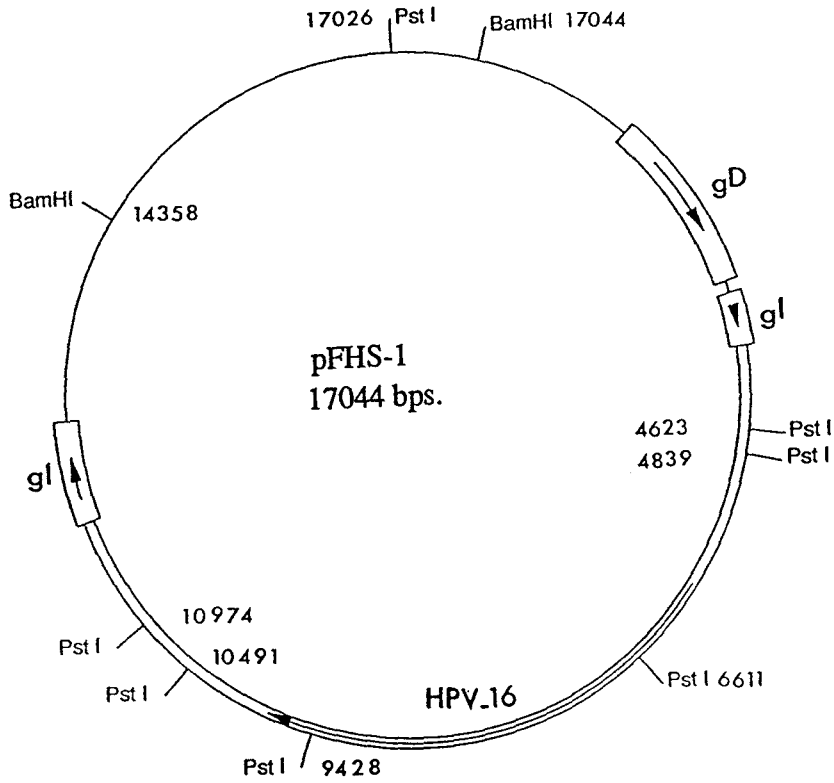

SstI
7464
EcoRI
7454
PstI
7003
EcoRI
6818
PstI
6787
BamHI
6151
PstI
5238
PstI
4755
PstI 3692
PvuII
551
PvuII 683
PstI 875
HPV_16
7900 bp.

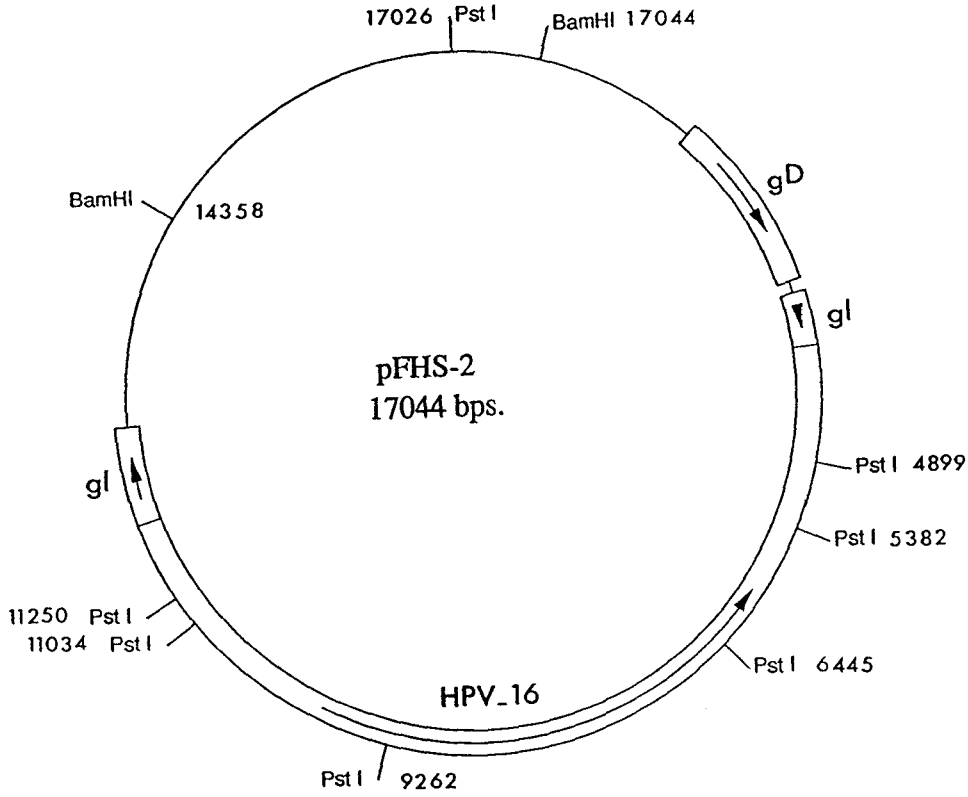
and purified. BamHI cuts HPV-16 at residue 6150, therefore, the HPV-16 fragment used contains an intact URR and early region but disrupted in the late region.

pSS17B (constructed by Mike Ligas) is a derivative of pSS17 which is a plasmid containing the BamHI J fragment of the Paa^r5 strain of HSV-1 inserted into pUC19. The BamHI J fragment contains, in addition to other HSV sequences, sequences coding for the gD and gI herpes glycoproteins. pSS17B (Fig.9) was constructed by inserting a BglII linker into the gI sequences of pSS17 at the BalII site. This BglII site at position 3987 of pSS17B was utilized to insert the purified BamHI HPV-16 fragment. BamHI and BglII sites are compatible and their ligation generates a site which cannot be digested by either enzyme. The fragment generated by BglII is 5' A GATCT 3', and that generated by BamHI is 5' G GATCC 3'. The resulting hybrid sequence is 5' AGATCC 3'. The ligation reaction resulting from the insertion of the purified BamHI HPV-16 fragment into the BglII site of pSS17B, was cut with BglII to reduce the background created by pSS17B self-ligation. Since the recombinant plasmids of interest have this site destroyed due to ligation, they are not affected by the digestion reaction. The digestion mixture was used to transform E. coli HB101 cells.

Figure 11 The Two Recombinant Plasmids pFHS-1 and pFHS-2.

In pFHS-1, HPV-16 sequences are in the same orientation with respect to the remaining pSS17B sequences. pFHS-2 has HPV-16 in the opposite direction with respect to pSS17B.


17026
Pst I
BamHI 17044
gD
gI
BamHI
14358
pFHS-1
17044 bps.
4623
4839
Pst I
Pst I
gI
10974
Pst I
10491
Pst I
HPV_16
Pst I 6611
Pst I
9428

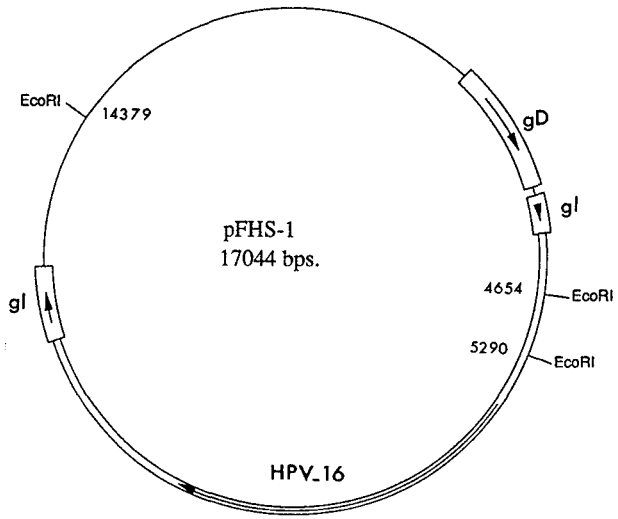

17026
Pst I
BamHI 17044
gD
BamHI
14358
gI
pFHS-2
17044 bps.
gI
Pst I 4899
Pst I 5382
11250
Pst I
11034
Pst I
Pst I 6445
HPV_16
Pst I
9262

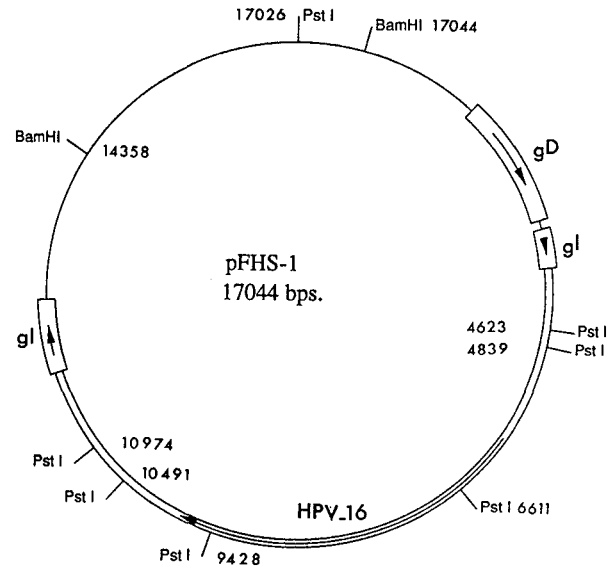
Transformed cells were grown on agar plates containing ampicillin. 20, 80, and 100 μl of the transformed culture were streaked on these plates which were incubated at 37^{o}C for 16 hours. The 20 μl plate was the only plate to give a countable number of colonies. Out of 108 colonies 12 were randomly selected. DNA from small plasmid preparations of these 12 colonies was digested with PstI. DNA from two colonies was positive and gave expected banding patterns for both forward and reverse insert orientations. This difference is due to the two possible orientations by which HPV-16 can be inserted into pSS17B. In one plasmid the papilloma sequences were inserted in the same orientation as the HSV gD and gI genes in pSS17B (forward orientation). The other plasmid, in contrast, consisted of the papilloma genome inserted in an anti-sense (reverse) orientation with respect to the herpes sequences (Fig.10). The first plasmid was named pFHS-1, and the latter pFHS-2.

A large scale DNA preparation of both plasmids was made. Subsequently, both were digested with a set of restriction enzymes to confirm their identity. The four enzymes used were: EcoR1, XbaI, PvuII, and PstI . HPV-16 DNA (linearized at BamHI) and pSS17B were also cut with the same enzymes as controls. In all digestion reactions 1 μg of DNA was digested with one of the four restriction

Figure 12 Restriction Maps of pFHS-1.

The positions of EcoRI, PstI, and PvuII restriction sites are shown. These sites were used to verify the identity of pFHS-1 and pFHS-2.


EcoRI
14379
gD
gI
pFHS-1
17044 bps.
gI
4654
EcoRI
5290
EcoRI
HPV_16


17026
Pst I
BamHI 17044
BamHI
14358
gD
gI
pFHS-1
17044 bps.
gI
4623
4839
Pst I
Pst I
10974
Pst I
10491
Pst I
HPV_16
Pst I 6611
Pst I
9428

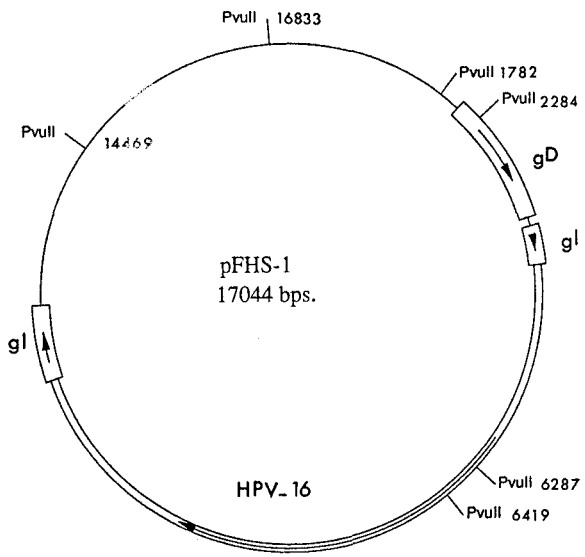
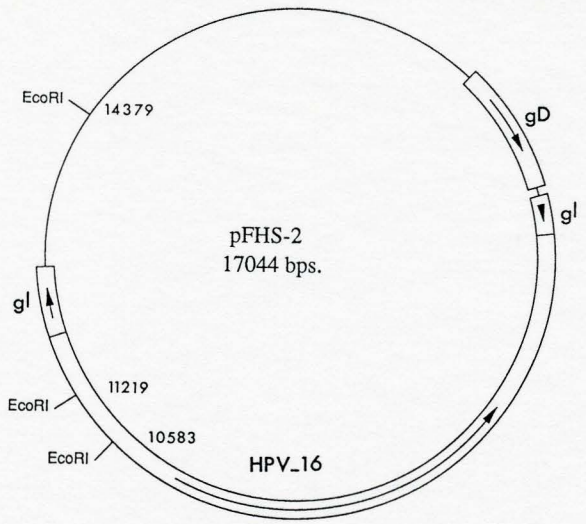
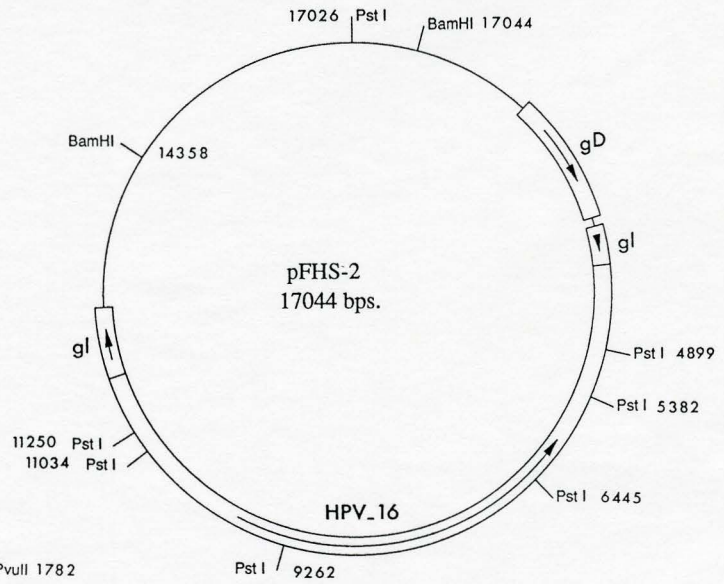

PvuII
16833
PvuII 1782
PvuII 2284
PvuII
14469
gD
gI
pFHS-1
17044 bps.
gI
HPV_16
PvuII 6287
PvuII 6419

Figure 13 Restriction Maps of pFHS-2.

The positions of EcoRI, PstI, and PvuII restriction sites are shown. These sites were used to verify the identity of pFHS-1 and pFHS-2.


EcoRI
14379
gD
gI
pFHS-2
17044 bps.
gI
11219
EcoRI
10583
EcoRI
HPV_16


17026
Pst I
BamHI 17044
BamHI
14358
gD
gI
pFHS-2
17044 bps.
gI
Pst I 4899
Pst I 5382
11250 Pst I
11034 Pst I
Pst I 6445
HPV_16
Pst I
9262

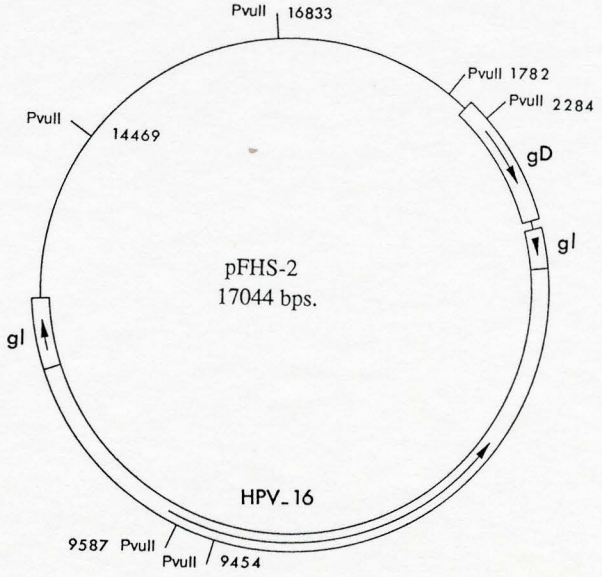

PvuII
16833
PvuII 1782
PvuII 2284
PvuII
14469
gD
gI
pFHS-2
17044 bps.
gI
HPV_16
9587 PvuII
PvuII
9454

Figure 14 Restriction Analysis of pFHS-1 and pFHS-2.

pFHS-1, pFHS-2, HPV-16, and pSS17B were analyzed using the four indicated enzymes. In each reaction 1 μg of DNA was cut with 10 units of enzyme. The digested DNA were run on a 1% agarose mini-gel for two hours at 90 volts. The gel was then stained in Ethidium Bromide for 20 minutes, after which, DNA fragments were visualized and photographed using an ultra-violet light source. The marker DNA used is lambda phage DNA digested with HindIII giving the indicated fragment sizes.

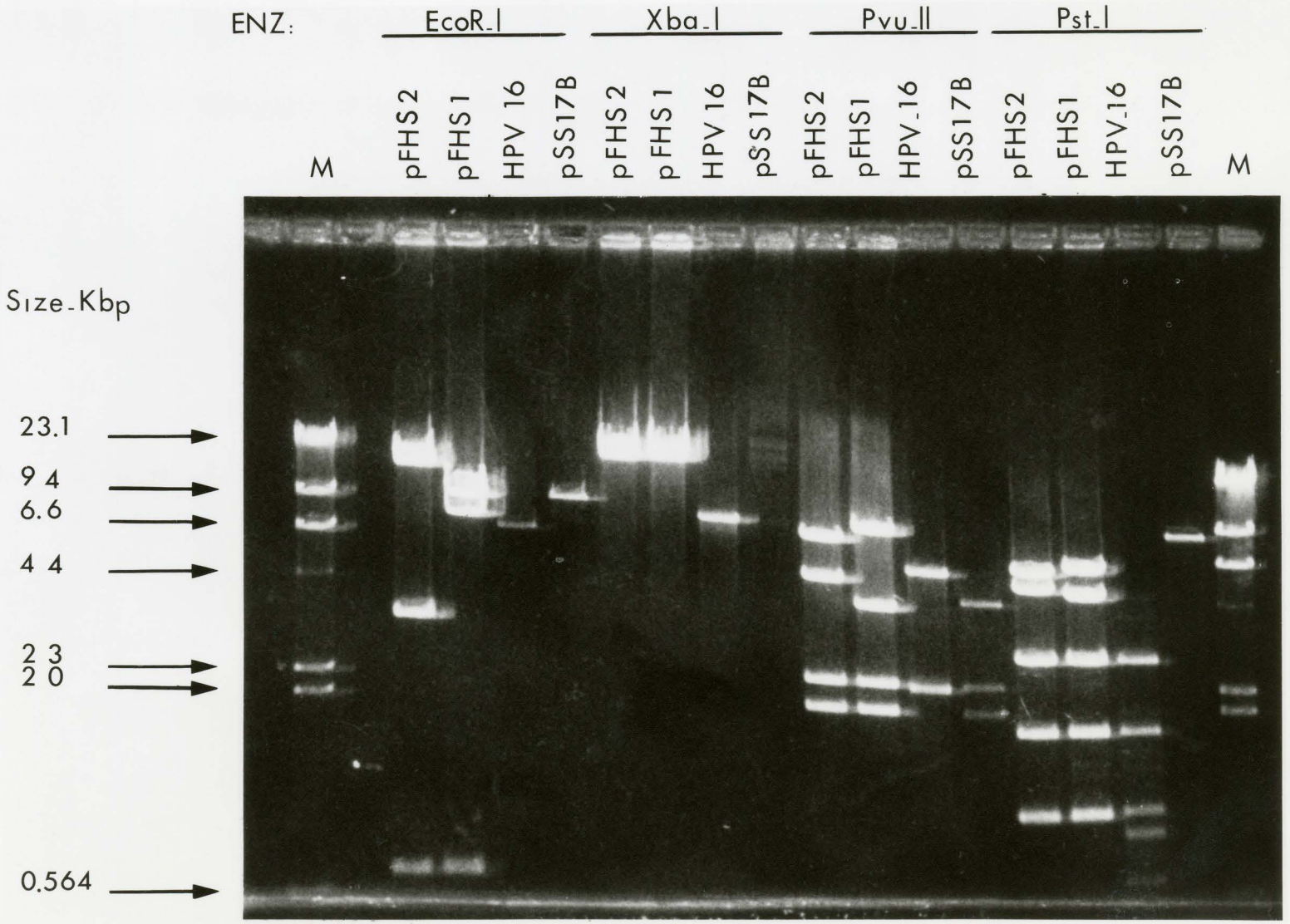
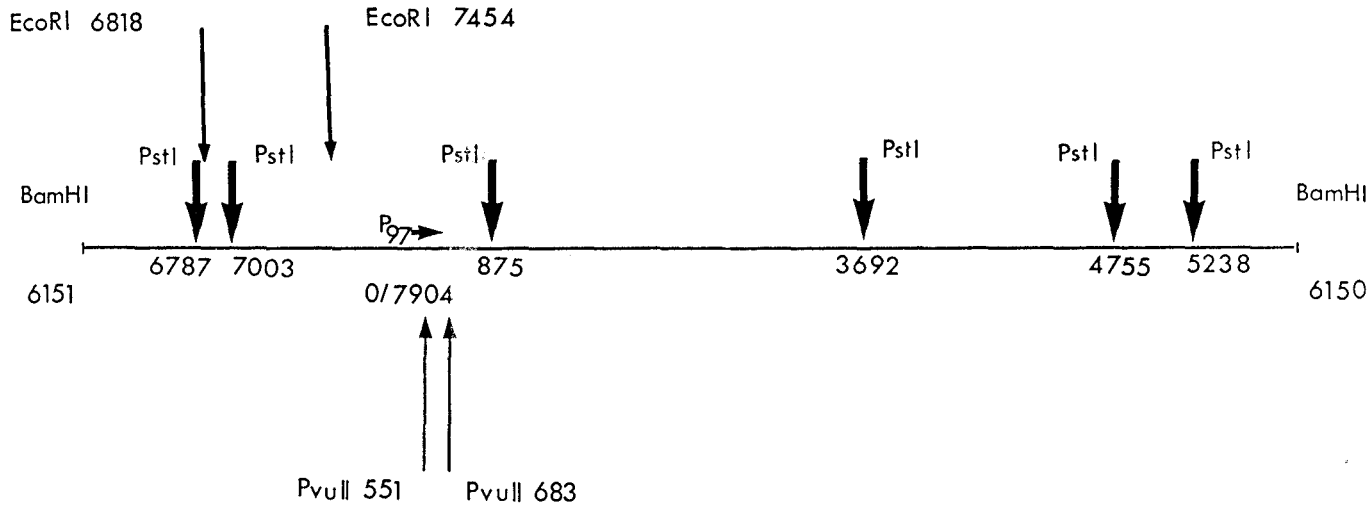

ENZ:
EcoR_I
Xba_I
Pvu_II
Pst_I
M
pFHS2
pFHS1
HPV 16
pSS17B
pFHS2
pFHS1
HPV 16
pSS17B
pFHS2
pFHS1
HPV_16
pSS17B
pFHS2
pFHS1
HPV_16
pSS17B
M
Size_Kbp
23.1
9 4
6.6
4 4
2 3
2 0
0.564

Figure 15 Restriction Map of Linearized HPV-16

HPV-16 linearized at the BamHI site (6150) showing the positions of EcoRI, PstI, and PvuII restriction sites.


EcoRI 6818
EcoRI 7454
PstI
PstI
PstI
PstI
PstI
PstI
BamHI
BamHI
P97
6787
7003
875
3692
4755
5238
6151
0/7904
6150
PvuII 551
PvuII 683

HPV_16

enzymes, after which the digested DNA was run on a 1% agarose wide mini-gel (Fig.14).

The three plasmids and HPV-16 DNA do not have any XbaI sites and therefore the bands observed corresponded to their actual (undigested) sizes. pFHS-1 and 2 are 17044 base pairs long, pSS17B is 9144 bps., and HPV-16 is 7904 bps.

Linearized HPV-16 has six PstI sites (Fig.15) which give a unique pattern of bands consisting of the following (in bps): 216, 483, 636, 912, 1063, 1776, and 2817; all of which were observed except for the smallest two bands which probably ran off (Fig.14).

A PstI site exists in pSS17B at position 9126 (Fig.10). After the insertion of HPV-16, this PstI site shifts to position 17026 in both pFHS-1 and 2 (Figs.12 and 13). Cutting pSS17B with PstI yields one band corresponding to the size of the plasmid (Fig.14).

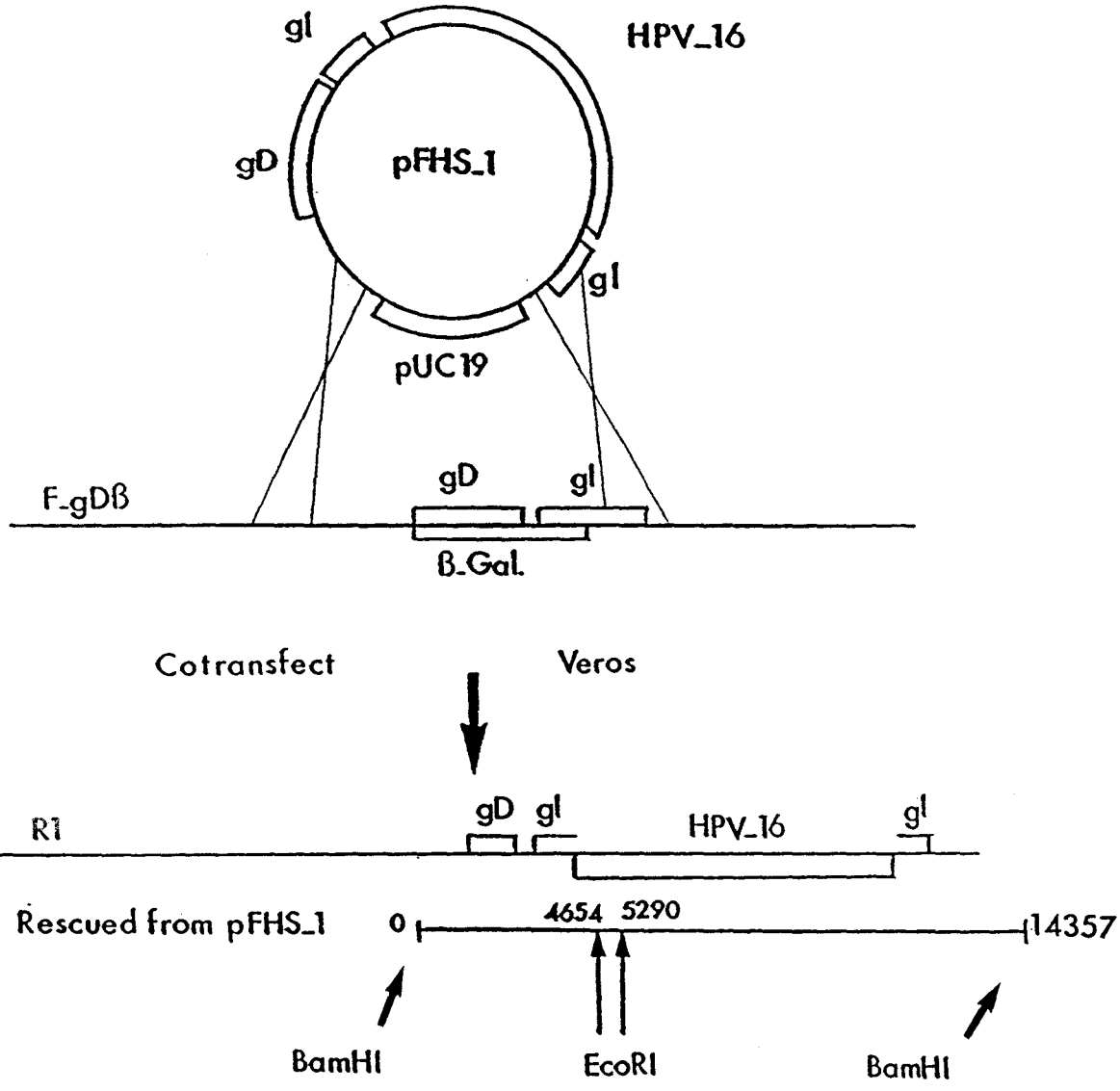
Digesting pFHS-1 with PstI gives the following bands (in base pairs): 216, 483, 1063, 1772, 2817, resulting from cuts within HPV sequences, 4641, and 6052 which represent the junction fragments between papilloma sequences and pSS17B. Only the smallest two bands did not appear on the gel (Fig.14). Digesting pFHS-2 with the same enzyme yields the following bands (Fig.14): 216, 483, 1063, 1772, 2817, resulting from cuts within papilloma sequences, 4872, and

5770 represent junction fragments. Only the first two bands were not visualized on the gel (Fig.14). As expected the bands resulting from cuts within the papilloma genome were unaltered by the difference in orientation in which HPV-16 was inserted with respect to the herpes sequences in pSS17B. Restriction sites situated at the two ends of the papilloma fragment yielded different junction bands in the two plasmids due to the insert orientation.

pSS17B has four PvuII sites (Fig.10) which give the following bands (in bps.): 502, 1993, 4285, and 2364. All the bands except the smallest were accounted for in a pSS17B digest with PvuII (Fig. 14). Two out of the four PvuII sites in pSS17B are located upstream of the BglII site (Fig. 10) and were therefore shifted to new positions which are 14469 and 16833 in both pFHS-1 and pFHS-2 (Figs. 12 and 13), whereas the other two sites remained unaltered. When digested with PvuII, pFHS-1 yields the following bands (in bps): 132 from cuts within papilloma sequences, 502, 1993, 2364 resulting from cuts in pSS17B sequences only, 4003, and 8050 which are junction fragments. The expected pattern from a PvuII digestion of pFHS-2 consisted of the following bands (in bps): 132 representing papilloma sequences, 502, 1993, 2364 resulting from cuts in pSS17B

Figure 16 Co-transfection of pFHS (1 or 2) and F-gDβ into Vero Cells.

Plasmid DNA and infectious F-gDβ DNA were co-transfected into Vero cells. Plasmid DNA contains HSV-1 sequences including those coding for the gI and gD glycoproteins. When plasmid DNA is co-transfected with F-gDβ, a mutant herpesvirus, defective in gD production, into Vero cells a number of single or double cross-over events can take place. A double recombination event between plasmid and viral herpes sequences occuring upstream of gD sequences and downstream of the inserted HPV-16 in the plasmid, produces a recombinant virus which has an intact gD gene and an interupted gI gene due to HPV-16 insertion. DNA from the recombinant virus RI, which should carry the HPV-16 insert, was analyzed with a BamHI/EcoRI double-digest which cuts out a BamHI fragment of herpes which in turn was initially used to construct pSS17B. EcoRI cuts twice internally in the papilloma sequences.


gI
HPV_16
gD
pFHS_1
gI
pUC19
F-gDß
gD
gI
ß-Gal.
Cotransfect
Veros
R1
gD
gI
HPV_16
gI
Rescued from pFHS_1
0
4654
5290
14357
BamHI
EcoRI
BamHI

sequences, 4883, and 7170 which are junction fragments. In both digests the first two bands were not detected.

Linearized HPV-16 has two PvuII sites (Fig.15) giving three bands (in bps.): 132, 2304, and 5467. In lanes containing linearized HPV-16 digested with PvuII, the largest two bands could be detected while the smaller had probably run off.The last restriction analysis was done using the enzyme EcoRI which cuts at three sites in both pFHS-1 and pFHS-2 (Figs. 12 and 13). pFHS-1 produced all expected bands which are: 636 from cuts within papilloma sequences, and the junction fragments which are 7316, and 9089 base pairs in size (Fig. 14) pFHS-2 also produced the expected bands which consist of in (bps.): 636 representing papilloma sequences 3160, and 13,248 which are junction fragments (Fig. 13).

pSS17B has one EcoRI site (Fig. 10) and, therefore, yields one band equivalent to its size (Fig. 14). Linearized HPV-16 has two EcoRI sites, which when restricted produce three bands (in bps.): 636, 667, and 6600. Only the largest band was observed (Fig. 14).

The results of the different restriction analyses carried out confirm the construction of two recombinant plasmids. pFHS-1 carries HPV-16 sequences in the same orientation as the herpes sequences, while pFHS-2 carries HPV-16 in an opposite orientation to those same

sequences. In both plasmids HPV-16 was inserted in the BglII site in the sequences coding for the gI gene.

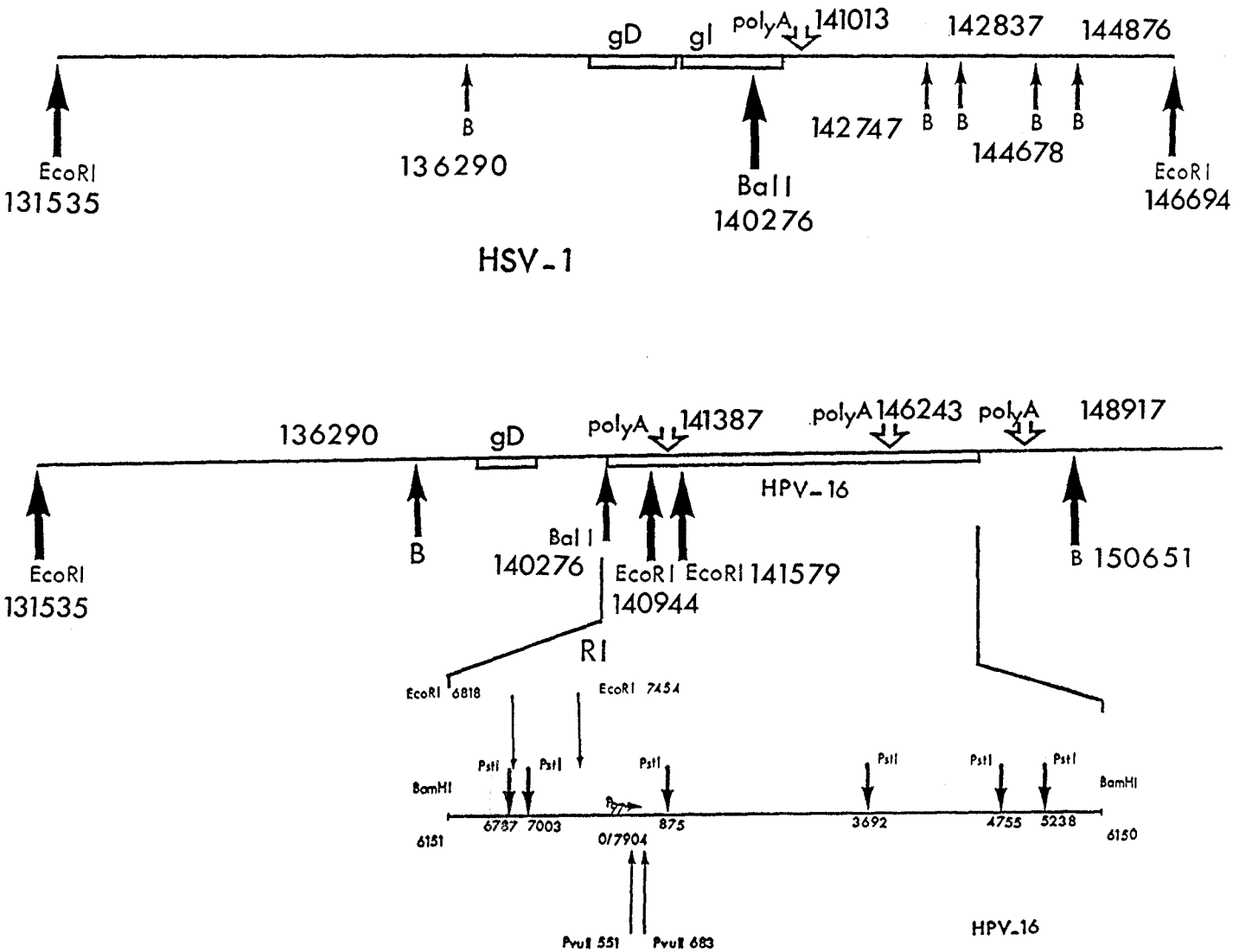
3.2 Construction of Recombinant virus RI

The immediate objective was to insert HPV-16 into non-defective HSV-1. A mutant herpesvirus, F-gDβ, was utilized for this purpose. In F-gDβ the gD gene and part of the gI gene of herpes are replaced by the E. coli β-galactosidase gene. gI is a non-essential gene of herpes while gD is indispensible for viral replication. F-gDβ cannot complete a full round of infection on Vero cells but is able to replicate normally on VD60 cells which express HSV-1 gD (Ligas and Johnson, 1988). A plasmid carrying intact gD but disrupted gI sequences can restore normal replication of F-gDβ when co-transfected into non-complementing cells such as Vero cells. pFHS-1 and 2 have an intact gD gene while gI is disrupted by the HPV-16 insert.

Infectious DNA from the herpesvirus mutant F-gDβ, which is defective in the production of gD glycoproteins, was co-transfected with either pFHS-1 or pFHS-2 into Vero cells (Fig. 16). A double recombination event between the

Figure 17 Diagram of the Region in HSV-1 into which HPV-16 was inserted.

HPV-16 was inserted into the BalI site of HSV-1 forming RI, as a result, all the restriction sites in HSV downstream from BalI increased in number by 7904 bps. The numbering of restriction sites in the papilloma sequences of RI (papilloma sequences are represented as a rectangular annex in RI) was established by calculating the number of sequences between a given site on HPV-16 and the BalI site of HSV-1 and adding this number to 140276 which represents the position of BalI in HSV-1. Also illustrated is the inserted HPV-16 genome with restriction sites as they appear on HPV-16 (from 0 to 7904) and the unique BamHI site (6150) which was utilized to insert the whole genome into HSV-1. The positions of the poly A signals are given in both HSV-1 and RI. BamHI, in this figure, is represented by B.


gD
gI
polyA
141013
142837
144876
B
136290
142747
B
B
B
B
144678
EcoRI
131535
Ball
140276
EcoRI
146694
HSV-1
136290
gD
polyA
141387
polyA
146243
polyA
148917
HPV-16
EcoRI
131535
B
Ball
140276
EcoRI
140944
EcoRI
141579
B
150651
RI
EcoRI 6818
EcoRI 7454
PstI
PstI
PstI
PstI
PstI
PstI
BamHI
BamHI
6787
7003
875
3692
4755
5238
6151
6150
0/7904
PvuII 551
PvuII 683
HPV-16

herpes sequences in the plasmid and the mutant virus, occuring upstream of the gD sequences and downstream of the HPV-16 insert in the sequences coding for gI, will generate a recombinant virus with an intact gD gene capable of growing on Vero cells and carrying the HPV-16 insert.

Cytopathic effects were first observed on Vero monolayers three days post-transfection with pFHS-1 co-transfections only and not pFHS-2. Individual plaques were picked on the sixth day and used to infect Vero monolayers prepared in 24 well dishes. After complete infection, cells were harvested and DNA was extracted.

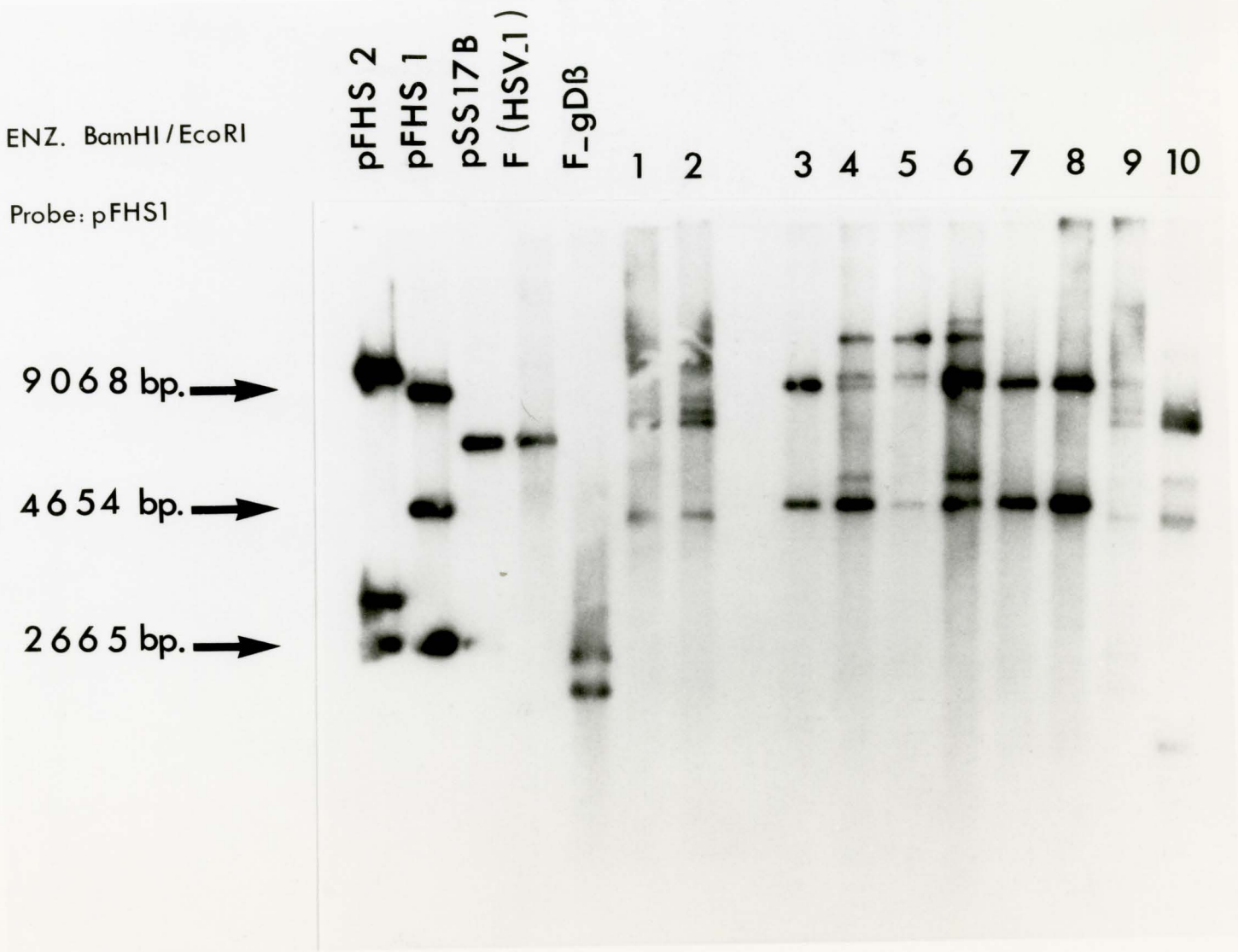
The DNA fragment which was rescued into virus is the BamHI J fragment of the Paa^r5 strain of HSV-1 containing the HPV-16 genome inserted at the unique BglII site which in turn lies within the gI sequences of this fragment. The part of the HSV-1 genome into which the HPV-16 sequences were rescued is depicted in Figure 17. The BamHI J fragment constituting the herpes sequences of pSS17B extends from positions 136290-142747 in HSV-1 (Fig. 17). HPV-16 was inserted into a BglII site of a polylinker located at the BalI site (140276) of HSV-1. A successful double cross-over, as described previously, will result in HPV-16 being inserted at the BalI site of HSV-1 (Fig. 17). In RI, herpes sequences upstream from the BalI site remain

unaltered, whereas herpes sequences and restriction sites downstream from this site shift by 7904 bps. HPV-16 was inserted as a linearized BamHI fragment at the BglII site of pSS17B which, as mentioned before, leads to the destruction of both sites. In RI, therefore, papilloma sequences start at position 6150 (BamHI site in HPV-16) of HPV-16 (140276 of HSV-1 and RI) increase to 7904/0 and terminate at 6151 of HPV-16 which becomes 148180 in RI. An EcoRI site at position 6818 of HPV-16, for example, becomes 6818 - 6150 = 668 + 140276 = 140944 in RI (Fig. 17). An EcoRI/BamHI double-digest of RI viral DNA which was then probed with radiolabelled pFHS-1, a derivative of pSS17B containing HPV-16 sequences, should produce the following detectable bands (bps.): 4654 (136290-140944), 636 (140944-141579), 9072 (141579-150651).

A southern blot was prepared from EcoRI/BamHI digested viral DNA in attempts to reveal the presence of the HPV-16 genome insert (Fig. 18). pFHS-1 digested with BamHI and EcoRI was used as a marker in the southern analysis (Fig. 18). Such a digest yields four bands of sizes (in bps.): 636, 2665, 4654, and 9068 (Fig. 12). A fragment produced by a BamHI and an EcoRI site located 20 bps apart does not interfere with the observed pattern. pFHS-2 cut with BamHI/EcoRI was also used as a marker

Figure 18 Southern Blot Analysis of Preliminary Recombinant Isolates.

Viral DNA isolated from plaques resulting from co-transfecting pFHS-1 and F-gDβ into Vero cells, were digested with BamHI and EcoRI. The digested DNA was electrophoresed on a 1% agarose gel along with pFHS-1, pFHS-2, pSS17B, F strain of HSV-1, and F-gDβ which were similarily digested. The separated DNA were transferred onto nitrocellulose filters and probed with ^{32}P dCTP nick translated pFHS-1. Positive recombinant viruses were identified by the presence of two bands co-migrating with the 4654 bp. and 9068 bp. bands in the pFHS-1 lane.


ENZ. BamHI / EcoRI
Probe: pFHS1
pFHS 2
pFHS 1
pSS17B
F (HSV-1)
F-gDβ
1
2
3
4
5
6
7
8
9
10
9068 bp.
4654 bp.
2665 bp.

(Fig.18). This digest gives rise to four bands (Fig. 13) (bps.): 10583, 3139, 2665, and 636. A third marker used in this analysis, consisted of a BamHI/EcoRI digest of pSS17B (Fig. 18) which gives the following bands (Fig. 10) (in bps.): 6458, 2665, and 20. F-gDβ and its parental strain (F-strain of HSV-1) were both cut with BamHI and EcoRI and included in this analysis as controls for recombinant viruses which revert back to one of the two original viruses. Radiolabelled pFHS-1 was used as a probe. pFHS-1 has the same sequences as pFHS-2 and is a derivative of pSS17B, thus, all BamHI/EcoRI bands generated by these three marker plasmids will hybridize to the probe and can be detected. The 2665 band corresponding to pUC19 sequences was observed in all three plasmids but in the case of a double cross-over, as described previously, these sequences will not be incorporated in RI and such a band will, therefore, remain undetected in RI viral DNA double-digested with BamHI and EcoRI. Of the ten analyzed viruses, three (Clones 1, 2, and 10) appear to have undergone rearrangements leading to a different pattern of bands than that expected for RI. Clone 9 seems to have undergone rearrangements and also appears to be a mixture of the expected pattern and F-gDβ (Fig. 18). In clones 3-8, two of the three expected bands (9072 and 4654) are

detected, where the smallest band (636) has run off the gel and therefore remains unobservable. In clones 4 and 6, an additional band larger than 4645 also appears. In clones 4,5, and 6 an additional band runs close to the 9072 band giving it the appearance of a doublet. In the same clones, a band larger than the 9072 band is also detected which in the case of clone 6 is preceded by a larger additional band (Fig. 18). As described below, many of the additional bands observed in isolates 4, 5, and 6 can be accounted for assuming that the EcoRI digestion did not proceed to completion with these samples. The positive viruses (3-8) was passed once on Vero cells as described before.

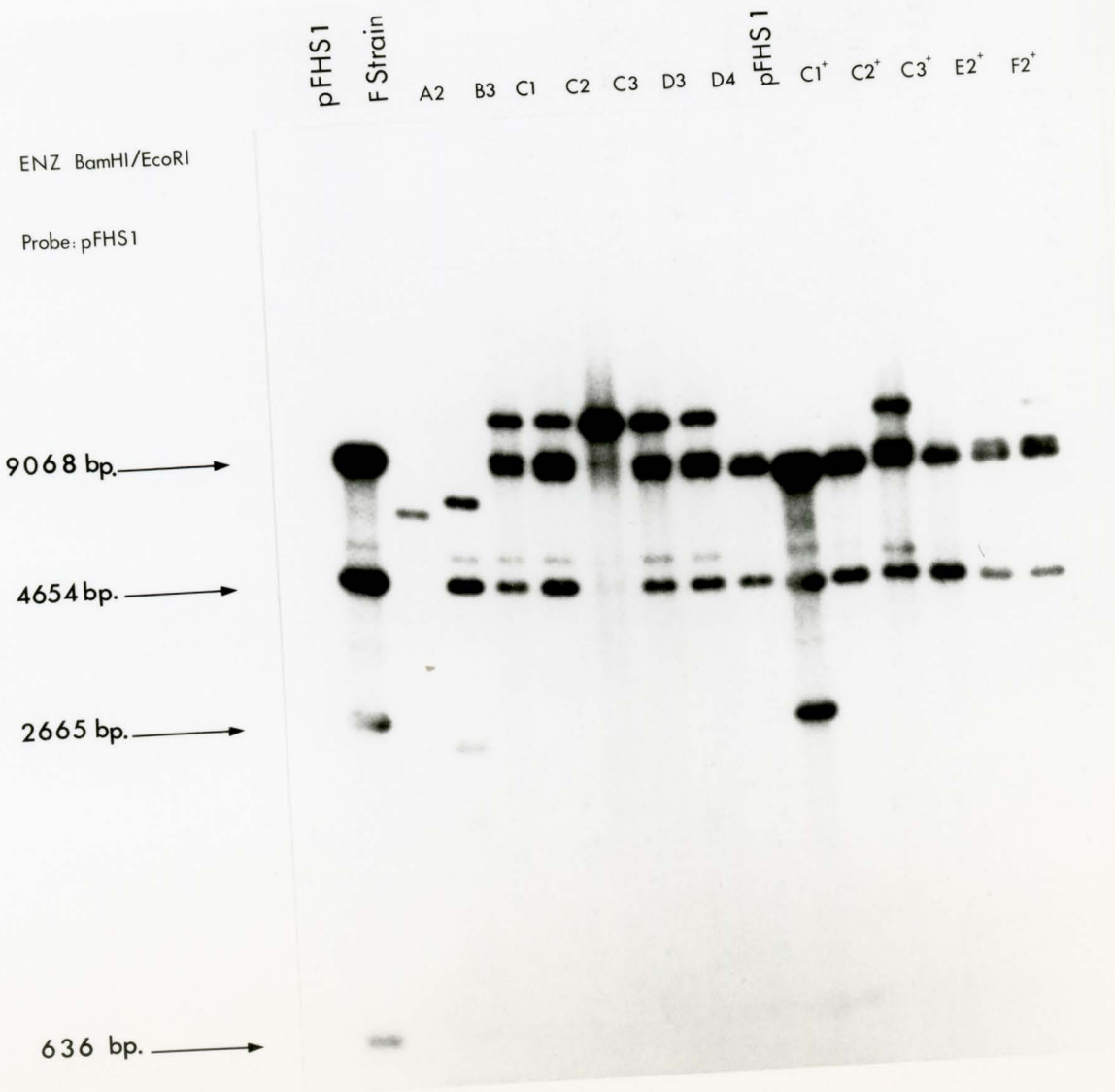
To ensure the stability and to purify the recombinant viruses, six individual clones designated A, B, C, D, E, and F were selected from one of the passed viruses (clone 3), and were simultaneously passed four times. Viral DNA extracted from these clones at different passages were double-digested with EcoRI and BamHI. A southern blot analysis of the extracted DNA was again prepared using a BamHI/EcoRI double-digest of pFHS-1 as a marker, F strain of HSV-1 DNA cut with EcoRI and BamHI as a control, and radiolabelled nick translated pFHS-1 as a probe (Fig.19). The numerical value used in the naming of the different clones represents the number of passages. On the same

southern blot, some of the viral DNA was digested twice with the second reaction containing double the amount of enzymes used in the first to test the possibility of incomplete digestion.

In the southern blot analysis of the passed viruses (Fig. 19), A2 seems to have undergone a rearrangement which affects the position of the BamHI site at position 33493 leading to the loss of the 9072 band and the appearance of two smaller bands. In all of the remaining viruses analyzed, the two largest bands expected were detected, with the 9072 band appearing as a doublet. In addition, a band larger than 4654 bps. and another larger than 9072 appear in the southern analysis. These additional bands are most likely partial digestion products. Examining the restriction map of the region in RI where the HPV-16 genome was inserted (Fig.17), and assuming an EcoRI partial digest occurs at different positions when cutting with BamHI/EcoRI, reveals a number of patterns that can arise when pFHS-1 is used as a probe. If EcoRI does not cut at position 131535 the expected pattern is not affected due to a BamHI cut just upstream at 136290. If EcoRI does not cut at 140944, two bands 5289 (bps.) and 9072 (bps.) would be detected. If EcoRI does not cut at 141579, two bands 4654 (bps.) and 9707 (bps.) would be detected. If EcoRI does

Figure 19 Southern Blot Analysis of Positive Recombinant Clone 3 Isolated at Different Passages.

Clone 3 was passed several times, and DNA from different passages was extracted from infected cells. Viral DNA was analyzed in a similar fashion as in figure 17. pFHS-1 and F strain DNA digested with EcoRI and BamHI were used as a marker and control respectively. The following clones: $C1^+$, $C2^+$, $C3^+$, $E2^+$, and $F2^+$ were digested with double the amount of enzymes compared to the other lanes.


pFHS1
F Strain
A2
B3
C1
C2
C3
D3
D4
pFHS 1
C1+
C2+
C3+
E2+
F2+
ENZ BamHI/EcoRI
Probe: pFHS1
9068 bp.
4654 bp.
2665 bp.
636 bp.

not cut at 140944 and 141579, one band 14361 (bps.) would be detected. The pattern obtained in clones 4, 5, and 6 (Fig.18) and all clones in Figure 19 except A2 can therefore be explained by incomplete EcoRI digestion where the bands expected for complete digestion (9072 and 4654) are present, as well as partial digestion products due to incomplete cleavage at the following EcoRI sites: 140944 and/or 141579. The occurrence of partial digestion at position 140944 gives rise to the 5289 (bps.) band just above the expected 4654 (bps.) band, whereas when it occurs at position 141579, a 9707 (bps.) band appears which forms a doublet with the expected 9072 (bps.), however, failure of cleavage at both these sites (140944 and 141579) results in the appearance of a 14361 (bps.) band above the doublet, as observed in both figures. Assuming the same situation occurs with BamHI sites in the case of a BamHI/EcoRI digest, does not provide an explanation for the two extra bands observed above the 4654 (bps.) and 9072 (bps.) bands. DNA from C1, C2, and C3 clones was digested with twice the amount of enzymes and analyzed on the same southern, where these are represented by the corresponding $C1^+$, $C2^+$ and $C3^+$ lanes. The expected pattern for complete digestion predominates in these latter lanes, whereas the extra bands observed earlier seem to disappear in most cases or at

least decrease in intensity. Thus, the data suggest that the submolar bands observed are the results of partial digestion with EcoRI. Presumably the same argument also applies to isolates 4-6 in Fig. 18. However, isolate 6 (Fig. 18) displays another band which cannot be accounted for in this way. In the case of clone 6 (Fig. 18), the largest band can be accounted for if EcoRI does not cut at all, and BamHI cuts at 136290 (Fig. 17), cleaves partially at position 150651 and cleaves at either 152582 or 152779 to give a 16292 bps. or 16489 bps. both of which can be detected using pFHS-1 as a probe. The viral clone denoted as D4 in Fig. 19 was used to grow viral stocks which were used in subsequent experiments under the common name RI.

3.3 Transcription from P_{97} in Vero Cells

After verifying the identity of the recombinant virus, RI, initial experiments were designed to investigate the possibility of transcription from papilloma promoters in this recombinant. Vero cells were infected with RI and transcription from the defined early promoter P_{97} of HPV-16 was tested. It is conceivable that transcription from papilloma promoters in RI may require a cellular factor not

present in Vero cells. Another possibility is a read-through from upstream herpes promoters in RI into papilloma sequences.

As mentioned before, the initial studies focused on the transcription of the two transforming genes, E6 and E7, of HPV-16. The main question asked was whether the P_{97} promoter was being utilized in Vero cells infected with the recombinant virus RI. Primer extension experiments and S1 nuclease mapping were performed on viral RNA extracted from Vero cells following infection with RI to address this question.

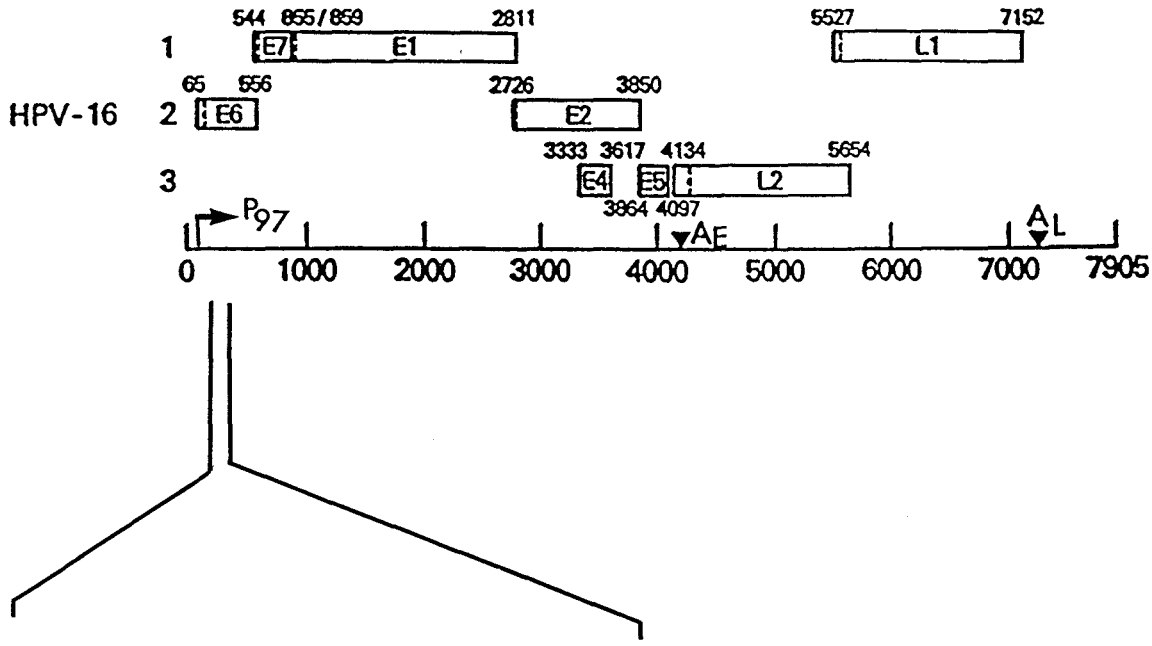
3.6.1 Primer Extension Analysis

Confluent Vero cells were infected with either RI or the F strain of HSV-1 at a multiplicity of infection (moi) of 10. Infected Vero cells were harvested 6 and 12 hours post-infection. Total cellular RNA was extracted, and primer extension experiments were performed using two primers E97 (Fig. 20) and AB166.

AB166 is a primer specific for HSV glycoprotein B (gB) transcripts and yields a primer extension product 112 bps long. This primer was used to monitor viral RNA levels and normal infectivity by the HSV vector. The presence of

Figure 20 Primer E97

The location of the synthetic oligonucleotide primer used to monitor transcription from P_{97} is shown. The primer is 25 nucleotides long extending from position 177 to 152.

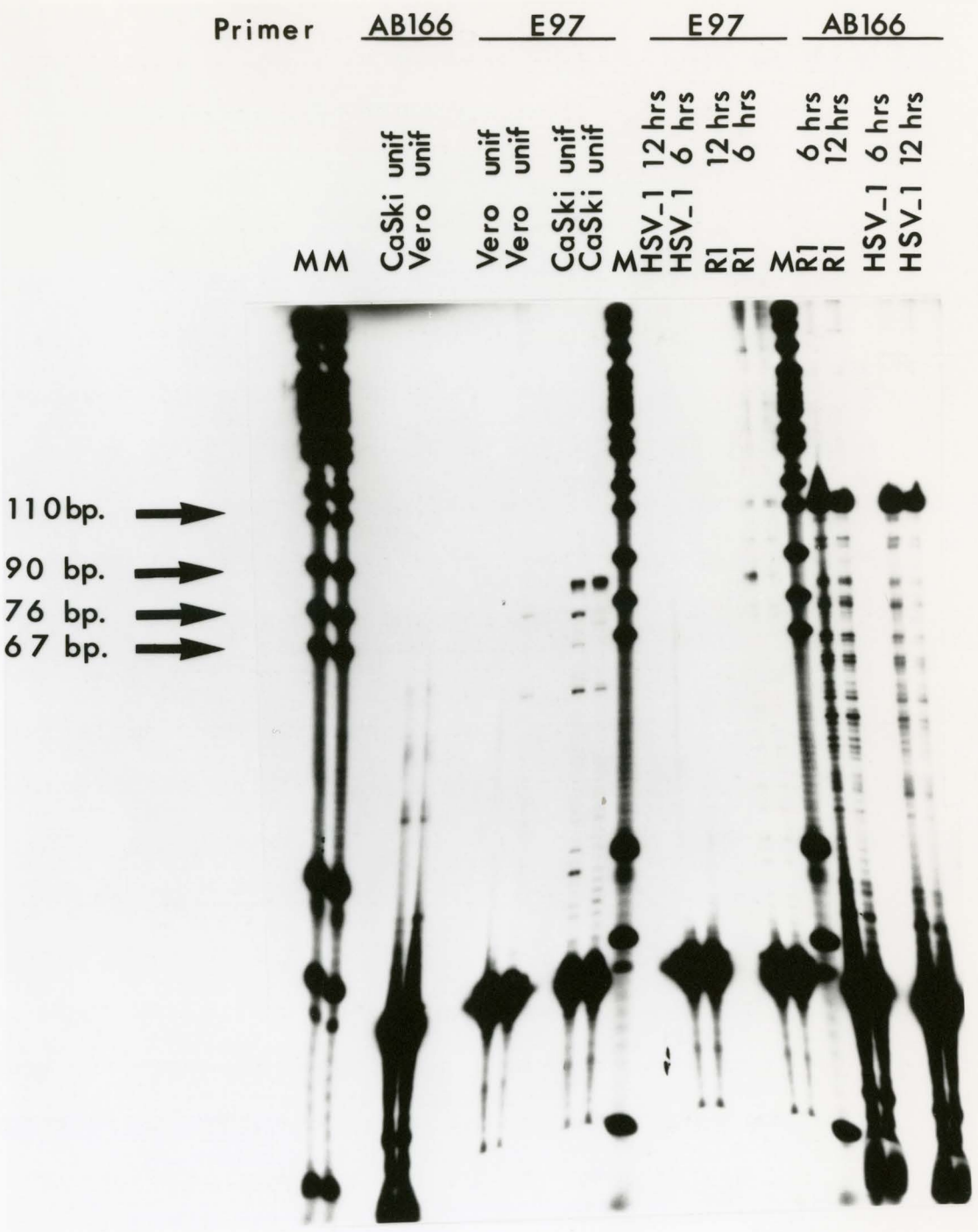


5′ TCATGTATAGTTGTTTGCAGCTCTG 3′

177 152

Figure 21 Investigation of Transcription Initiation from the P_{97} Promoter of HPV-16 using Primer Extension.

Total cellular RNA (15 μg) was isolated from Vero cells infected (at moi of 10) with RI (6 and 12 hours post-infection), HSV-1 (6 and 12 hours post-infection), uninfected Vero cells, and uninfected CaSki cells. Two different primers were used in this analysis. AB166 is a primer specific to the herpes glycoprotein B (gB) message and gives rise to a primer extension product of 112 bps. In the case of transcription initiation from P_{97}, the E97 primer will yield an 80 bps. product in Vero cells infected with RI. In uninfected CaSki cells where the P_{97} promoter was shown to be active (Smotkin and Wettstein, 1986) this product is also expected. 30000 counts per minutes (Cerenkov) of each 5' end labelled primer were hybridized to RNA extracted from infected or uninfected cells. Reverse transcriptase was added to the mixture, and the extension products were subsequently precipitated, denatured and run on an 8% polyacrylamide sequencing gel containing 7 M urea. The marker (M) used in this experiment is a 3' end labelled HpaI digest of pBR322.

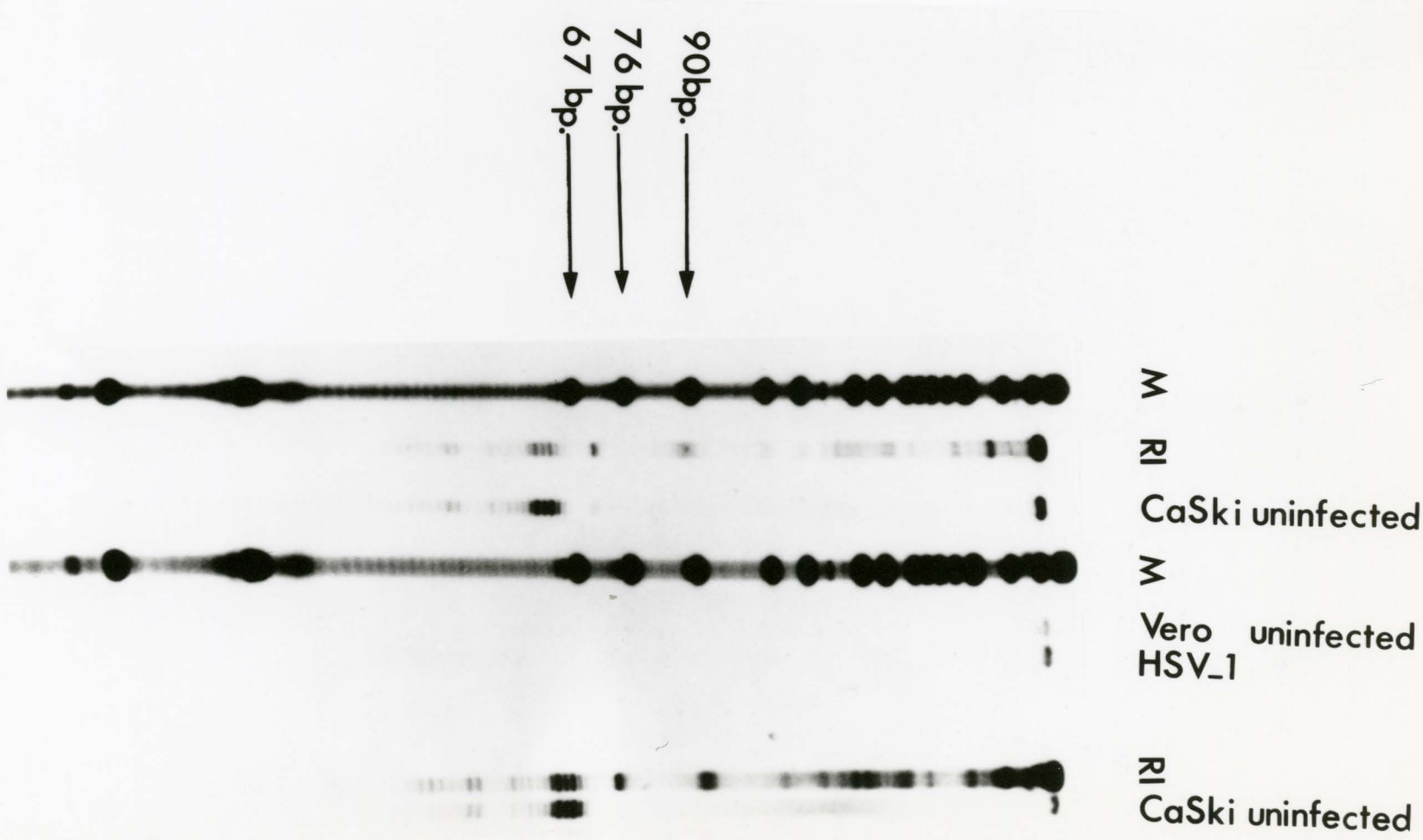

Primer
AB166
E97
E97
AB166
MM
CaSki unif
Vero unif
Vero unif
Vero unif
CaSki unif
CaSki unif
M
HSV_1 12 hrs
HSV_1 6 hrs
R1 12 hrs
R1 6 hrs
M
R1 6 hrs
R1 12 hrs
HSV_1 6 hrs
HSV_1 12 hrs
110bp.
90 bp.
76 bp.
67 bp.

a 112 bps. primer extension product using AB166 with RNA from RI infections confirms the normal expression of HSV-1 genes despite the papilloma insert. As expected, primer extension experiments using AB166 with RNA isolated from Vero cells 6 and 12 hours post-infection with both RI and the F strain of HSV-1 resulted in the 112 bp primer extension product (Fig. 21). The intensity of this band was equally strong in cells infected with RI as those infected with wild type HSV-1. This band was not obtained with RNA from uninfected Vero and CaSki cells, where both these cell lines serve as negative controls for the presence of any herpes sequences.

In the case of transcription from P_{97}, the E97 primer should give rise to an 80 bp primer extension product as observed with CaSki cells in which transcription from P_{97} results in transcripts coding for E6 and E7 ORFs (Smotkin and Wettstein, 1986). The same primer extension product was observed in infections with RI at 12 hours and to a lesser extent at 6 hours post-infection (Fig.21). The band corresponding to the 80 bps. primer extension product in RI at 6 hours is barely visible while at 12 hours the intensity of this band is not as strong as the positive control which is represented by uninfected CaSki cells. A negative result was observed with Vero cells infected with

Figure 22 S1 Nuclease Mapping of Trancripts Initiating at P_{97} in RI Infected Vero cells.

15 μg of RNA isolated from RI infected Vero cells, HSV-1 infected Vero cells, uninfected CaSki cells, and uninfected Vero cells were hybridized to a denatured 5' end labelled probe (10^4 cpms) spanning the P_{97} promoter region from nucleotides 7444 to 157 (Fig. 9 page 90)). The hybridization mixture was then subjected to S1 nuclease digestion, and the protected fragments were resolved on an 8% acrylamide gel containing 7 M urea. The expected length of the protected fragment resulting from transcription initiation at the P_{97} promoter in RI infected cells and uninfected CaSki cells is 60 bps. long.


67 bp.
76 bp.
90bp.
M
RI
CaSki uninfected
M
Vero uninfected
HSV_1
RI
CaSki uninfected

the F strain of HSV-1 and uninfected Vero cells, since these do not contain or code for any papilloma sequences. Other primer extension products were observed below the two expected products (80 and 112 bps.). This can be the result of the instability of the reverse transcriptase enzyme which pauses or falls off the RNA template while copying it, resulting in a ladder-like phenomenon. In Vero cells infected with RI, bands which are larger than the expected product are also observed. This can either be a result of RNA messages initiating at herpes promoters upstream of P_{97} reading through papilloma sequences or due to initiation from papilloma sequences upstream from P_{97}.

3.3.2 S1 Nuclease Mapping

To confirm the result obtained in the primer extension experiment, S1 nuclease mapping of RNA extracted from Vero cells infected with RI was undertaken. The probe used in this experiment should, as described before, protect a 60 bp. fragment (Fig. 9) in the event of transcription initiating from the P_{97} promoter. In the same experiment, RNA extracted from Vero cells infected with the F strain of HSV-1, uninfected Vero cells, and uninfected CaSki cells was analyzed for the presence of the protected fragment using the same S1 probe. RNA obtained

from CaSki cells served as a positive control, and a band corresponding to the protected fragment was detected. The same pattern of protection was observed with RNA from Vero cells infected with RI. RNA from Vero cells infected with the F strain of HSV-1, and that of uninfected Vero cells did not produce any protected fragments, as expected for negative controls (Fig.22).

The pattern of protection actually consisted of a cluster of bands different in size by a few bps. This pattern can be explained by slightly different start positions from which transcription can initiate leading to different sizes of protected fragments. Another possibility is that the S1 nuclease may not digest unprotected single stranded DNA fragments all the way to the site of initiation of transcription leading to different sizes of the protected fragment.

In RI infected Vero cells, S1 nuclease mapping of RNA revealed the existence of bands larger than the expected protected fragment. The 5' end labelled DNA probe used in this experiment extends 270 bps. upstream from the P_{97} promoter (Fig. 9 page 89) and has the potential to bind and protect RNA transcripts initiating upsteam from P_{97} resulting from either papilloma or herpes sequences (Fig.17) in RI infected cells.

The sizes of the additional bands observed in both

primer extension and S1 nuclease analyses were determined by referring to a standard graph where the logarithm of the sizes (in bps.) of the different fragments of the marker (HpaI digest of pBR322) was plotted against the distance (in cm) travelled by each of these fragments starting from a common reference point at the top of the gel. The sizes of the additional bands were then determined by measuring the distance each band travelled from the reference point on the gel and converting this number, by using the plotted graph, to the corresponding size (in bps.).

In the case of the primer extension experiment, 80 bps which represent the expected distance extending from the position where the primer binds up to P_{97}, were subtracted from each calculated fragment to find how far upstream from P_{97} each fragment originated (Fig. 17). The expected 80 bp extension product was calculated according to the graph to be 84 bp. The following are the sizes of the additional bands (in bps.) and the calculated position of initiation on the HPV-16 genome, respectively: (90,87), (93,84), (98,79), (102,75), (106,71), (113,64), (115,62), (130,47), (140,37), 310,7771). (Table 1)

In addition to detecting transcription from P_{97}, the E97 primer has the potential to detect those transcripts which initiate from upstream herpes promoters reading through papilloma sequences. The observed additional bands

could, therefore, also be due to initiation of transcription from gI and gD, where the start sites for both these transcripts are located at position 139694 and 138333, respectively, in both herpes and RI (Fig. 17). Acknowledging the fact that reverse transcriptase has the tendancy to pause or fall-off during the primer extension process can explain the pattern obtained with RI infections as being caused by transcription from gD or gI promoters, although it cannot be confirmed.

In the case of SI nuclease mapping, the fragment used as a protective probe (Fig. 9) was 430 bps. long, but only gave rise to a 60 bp protected fragment due to transcription from P_{97}. The existence of larger bands can be caused by initiation upstream from P_{97}. The distance upstream from P_{97} from which initiation might occur can, therefore, be calculated by subtracting 60 bps. from each of the extra bands observed. The calculated sizes and the position of initiation on the papilloma genome of the cluster of bands representing the expected protected fragment size (60 bps.) in the case of transcription from P_{97} are: (59,98), (60,97), (62,95), (63,94), (64,93), (66,92). The sizes of the larger additional bands and the position from where they initiate on the papilloma genome are: (72,85), (92,65), (119,38), (133,24), (141,16), (145,12), (150,7), (168,7893), (182,7882), (215,7846),

(320,7834). (Table 2)

It was not possible to calculate the size of the largest additional band by using the standard graph since the relationship becomes non-linear beyond a certain point and, therefore, it was estimated using the closest co-migrating bands of the marker and found to have a size between 404 and 527 bps.

If the hypothesis pertaining to initiation from upstream herpes promoters is valid then S1 nuclease mapping will give, in addition to the expected band (60 bps.), another band corresponding to the total size of the DNA probe used (430 bps.), protecting the proposed transcripts initiating from gD and gI promoters. The largest band in the lane corresponding to RI in the S1 mapping experiment (Fig.22) fits this size, yet this experiment cannot confirm that this band is due to transcription initiation from upstream herpes promoters rather than the result of the self rehybridization of the double stranded DNA probe used. In addition, other smaller bands in the same lane are observed suggesting that initiation from herpes promoters alone can not explain the presence of the additional bands in the case of infection with RI. The same rationale would apply to an argument suggesting that a late promoter similar to P_{7250} in BPV-1 exists in HPV-16 and is responsible for some of the additional bands. Such a

hypothetical promoter, if it exists, will probably be located approximately at the same position as P_{7250} similar to the situation found in BPV-1 with P_{89} and its counterpart P_{97} in HPV-16.

Three of the additional bands seem to initiate from the same positions in HPV-16 according to both primer extension analysis and S1 mapping, and these are: positions 84-85, 64-65, and 37-38. Because of the potential existence of other initiation sites, as indicated at least by the S1 mapping experiment, the authenticity of these three positions as initiation sites is questionable.

Table 1

Summary of Expected Bands and Additional Bands Observed in the Primer Extension Analysis Using E97 on RNA from Vero Cells Infected with RI. (Sizes of bands are in bps.)

Size of Expected band(s)	Size of Additional bands	Position of transcription start site on HPV-16
80 (observed = 84)	-	P97
	90	87
	93	84
	98	79
	102	75
	106	71
	113	64
	115	62
	130	47
	140	37
	310	7771

Table 2

Summary of Expected Bands and Additional Bands Observed in the S1 Nuclease Analysis on RNA from Vero Cells Infected with RI. (Sizes of bands are in bps.)

Expected bands	Additional bands	Position of transcription start site on HPV-16
60		97
Observed:		
59	-	98
60	-	97
62	-	95
63	-	94
64	-	93
66	-	92
	72	85
	92	65
	119	38
	133	24
	141	16
	145	12
	150	7
	168	7893
	182	7882
	215	7846
	320	7834

VI-Discussion

The two immediate conclusions that can be drawn from this work are first, a viable HSV-1 mutant virus carrying the entire HPV-16 genome (RI) was successfully constructed and second, transcription from the early papilloma promoter P_{97} of RI occurred when Vero cells, which are not the natural hosts of papilloma infection, were infected with RI.

As predicted, the 152 Kb HSV genome was able to accomodate the 8 Kb large HPV-16 genome to produce a recombinant virus which can be stably propagated in culture. Southern blot analyses confirmed the structure of the recombinant by revealing a unique pattern resulting from the insertion of the HPV-16 genome into gI sequences in the sense orientation. Additional bands appearing on these analyses resulted from partial digestion with one of the restriction enzymes used (EcoRI). In the constructed recombinant, the replicative and the infectious characteristics of HSV were not compromised as a result of the insertion of HPV-16 in the sequences coding for the non-essential glycoprotein I (gI). In the case of HPV-16,

insertion into HSV-1 forced the linearization of the circular papilloma genome at the late region thereby limiting the characterization of possible HPV-16 late messages. However, this system is permissive to investigations targeted at analyzing transcription from the early region of HPV-16, as well as delineating different viral and cellular factors which influence the function of the URR.

In primer extension experiments (Fig. 21) and S1 nuclease analysis (Fig. 22), transcription from P_{97} in RI was compared to the activity of the same promoter in uninfected CaSki cells which contain integrated copies of HPV-16 and in which P_{97} is known to be active (Androphy et al., 1987). Quantitatively, transcription from P_{97} in unifected CaSki cells was greater than in RI infected Vero cells. Also, primer extension experiments using the herpes specific promoter (AB 166) revealed the abundance of herpes transcripts in RI infected Vero cells as compared to the product of transcription from P_{97}. Transcription from P_{97} in RI infected cells was observed to a higher degree at late times (12 hrs.) compared to early times (6hrs.). Previous work has shown that temporal regulation of genes inserted into the HSV genome depend on the composition of its promoter and not on the region of the viral genome targeted by the inserted gene (Smiley et al., 1990).

Confirmation of the temporal regulation of P_{97} can be addressed in an experiment in which viral DNA replication, which is required for late gene expression in HSV, is chemically inhibited. If P_{97} was recognized in a fashion similar to a late herpes promoter, activation would not occur in the presence of such inhibitors. The additional bands observed in the primer extension experiment and the S1 nuclease analysis were different in size than the expected bands resulting from initiation at P_{97}. Additional experiments are required to determine the significance of those bands and whether they are the result of initiation at upstream herpes or papilloma sequences.

P_{97} is the only promoter that has been shown to be active in HPV-16. Transcription from P_{97} is primarily affected by elements in the URR 5' to this promoter. As detailed in the Introduction, this region contains many DNA sequences which can be recognized and bound by cellular and viral regulatory proteins suggesting a complex mode of transcription activation (Chan et al., 1990). The keratinocyte dependent enhancer located in the URR, for example, activates transcription from P_{97} in epithelial cells only. The exact mechanism by which different cellular proteins which can potentially bind cis-acting sequences in the enhancer region contribute to transcription activation is still largely unknown (Chong et

al., 1990). In addition to this cellular regulation, a viral regulation mediated by DNA sequences (E2 responsive elements) to which the papilloma transactivator E2 can bind also exists. Such sequences are located 5' and 3' from the keratinocyte dependent enhancer (Fig. 6). It is well established that the complete form of E2 protein activates transcription from E2 responsive elements, and the N-terminally truncated form represses transcription from the same elements (Lambert et al. 1988, Cripe et al., 1987).

In this work the activity of P_{97} was monitored in Vero cells which are not epithelial in nature and, therefore, do not produce the complete set of cellular factor(s) responsible for the activation of the keratinocyte dependent enhancer. Also, the role played by the E2 transactivator in transcription from P_{97} in RI infected Vero cells is not clear. For E2 to activate transcription from P_{97} it should be present prior to infection. In BPV-1 transformed cells, E2 is transcribed from an early promoter (P_{2443}) (Lambert et al., 1988), however such a promoter has not yet been characterized in HPV-16 infected cells or cancer cell lines containing copies of HPV-16. An RNA transcript originating from P_{97} analyzed from a cervical tumor positive to HPV-16 has the capacity to code for an intact E2 (Durst et al., 1987) (Fig. 8b). If E2 transcription initiates from P_{97}, as

suggested from available data, this would imply that E2 is not involved, at least, in the initial stages of transcription activation from P_{97} unless it is packaged in the RI virus particle. Evidence of a role by E2 in the packaging of virion particles, or being packaged itself as part of the virion does not exist. The constructed recombinant (RI) presents an opportunity to dissect the role of E2 in transactivation/repression mechanisms. Experiments directed towards analyzing different factors involved in the activation of transcription from P_{97} were mostly undertaken in cancer cell lines in which HPV-16 is in the integrated form and the E2 ORF is disrupted (Gloss et al., 1987) or in transient transfection experiments in which transcription from the E2 gene was driven by heterologous promoters (Cripe et al., 1987). If transcripts coding for E2 were detected in RI infected epithelial cells, where the keratinocyte dependent enhancer is active, then this would permit an interesting study on the interaction between viral and cellular regulatory proteins in the regulation of transcription from P_{97}.

Numerous CTF/NFI sites are present in the URR of HPV-16. The biochemical significance of such sites are not fully delineated. However, deletion experiments show that these sites contribute to the activity of the keratinocyte dependent enhancer (Chong et al., 1990). CTF/NFI proteins,

similarly to Sp1 are thought to harbour biologically active domains other than the one responsible for specific binding to DNA sequences. One simple rationale for this assumption is the large size of these proteins suggesting functions other than DNA binding. CTF/NFI is considered an activating transcription factor whose exact function is still ambiguous but several hypotheses have been outlined. One of these hypotheses suggests that activators can bind nuclear proteins which are not capable of binding the promoter in a sequence-specific manner, as in the transcription of the human ribosomal RNA promoter where a DNA binding protein confers selectivity to the promoter by binding and recruiting other proteins (McKnight and Tjian, 1986). Other experiments suggest a direct interaction between mammalian activators and TFIID. The conformation of TFIID bound to the TATA element in the promoter was altered by activators-bound DNA. This change in conformation extended the area of DNA covered by TFIID which might facilitate the establishment of a pre-initiation complex (Ptashne, 1988).

A second HSV-1/HPV-16 recombinant (R2), constructed in Dr. Rawl's laboratory, carried a disrupted URR region at the unique SphI site (nt 7464 of HPV-16). Transcription from P_{97} in this recombinant was tested in a primer extension experiment using the E97 primer. While CaSki

uninfected cells showed the expected extension product characteristic of transcripts initiating from P_{97}, RNA from R2 infected Vero cells failed to show any signal (unpublished results). This suggests a role by the URR in transcription from P_{97}, possibly through sequence specific transcriptional protein activators which can interact with one or more general transcription factors such as TFIID or with RNA polymerase II, or even with HSV-1 encoded regulatory proteins as discussed later. One hypothesis is that transcriptional activators can interact with the carboxyl terminal of the largest subunit of RNA polII which is rich in hydroxyl groups and shows no interactive specificity. Hydrogen bonds will permit the engagement between the activation domain of the DNA bound activators and the polymerase molecule. Subsequent phosphorylation of the hydroxyl groups, probably catalyzed by the activator protein, results in the disengagement of the polymerase and initiation of transcription (Mitchell and Tjian, 1989).

HSV immediate-early proteins such as ICP0 and ICP4 have been suggested to activate the HPV-16 URR (McCusker and Bacchetti, 1987). The exact mechanism of this activation is yet to be established. The concept of P_{97} being recognized as a herpes promoter regulated by herpes immediate-early proteins relegates the establishment of a mechanism explaining transcription from P_{97} to regulation

by herpes immediate proteins. Both proteins, ICP0 and ICP4, can bind DNA in a non-specific manner (Everett, 1988), suggesting that these proteins activate transcription in an indirect fashion by binding a certain cellular factor involved in transcription. Along this direction, studies on the TK promoter of herpes under ICP4-induced conditions have shown a more pronounced role by the TATA region of the promoter compared to the two other promoter elements, the CCAAT homology and the GC hexanucleotide region. This emphasizes that ICP4 can act through the TATA element on cellular transcription factors such as RNA polII accessory factors, mainly TFIID, which bind the TATA box (McKnight and Tjian, 1986). Furthermore, Johnson and Everett found that the TATA box/cap site is sufficient for activation by immediate-early proteins of a late gene during herpes infection (Johnson and Everett, 1986). In the same context, the Chinese hamster aprt promoter lacking a TATA box was not activated by immediate-early products when inserted into HSV-1 during infection with this recombinant (Dush et al., 1985). Immediate-early proteins, therefore, can activate transcription by either recognizing and binding the TATA sequence or interact with transcriptional factors which are known to specifically bind the TATA homology (Everett, 1988).

ICP4 was also shown to bind DNA in a sequence

specific manner binding generally to a core sequence made up of ATCGTC although binding is also possible at sites where this sequence is not present (Michael et al., 1988). The core sequence occurs more than seventy times on the HSV genome most of which are not proximal to a cis control region (Everett, 1988). This observation is in agreement with a hypothesis pertaining to activation by ICP4 which suggests that promoters introduced into the HSV genome come under the effect of ICP4 from a distance after the binding of ICP4 to viral sequences surrounding the inserted genes. This hypothesis assumes that viral sequences in HSV contain many ICP4 binding sites (Smiley et al., 1990). These can either be the ATCGTC or a related version which are abundantly present on the HSV genome (Everett, 1988). In short, ICP4 and maybe other herpes immediate-early proteins may activate transcription either by binding to specific DNA sequences and acting from a distance, or by being introduced indirectly to the promoter region by a cellular factor (Smiley et al., 1990).

With respect to activation of transcription from P_{97}, two TATA boxes are located at positions 17 and 65. The structure of the TATA box is important for activation by herpes immediate-early proteins which can interact with transcriptional factors such as TFIID whose binding to the TATA consensus sequence starts the formation of the

transcription pre-initiation complex. The level of transcription from P_{97} in RI infected cells is quantitatively less than the level of transcription in CaSki cells or the level of transcription from the herpes promoter of the gB gene (Fig. 20 & 21). A plausible explanation for this phenomenon may lie in the work conducted by some investigators where they have found that activation of transcription by immediate-early proteins in both transfection and infection experiments is dependent on the cell type (Everett, 1988). Experiments to clarify this point can, therefore, be carried out in keratinocytes where the KDE of HPV-16 is known to be active. It is hard to predict the relationship between cellular transcriptional activating factors which can potentially bind in the URR upstream from P_{97} and the herpes transactivating immediate-early proteins. Such proteins (activators) have been shown to interact with transcriptional factors such as TFIID. The presence of the URR is essential for activation as shown in the R2 construct. It is possible that the herpes immediate-early proteins such as ICP4 mediate their activation through cellular factors binding at this region. At this stage it is not possible to determine which immediate-early protein is involved in the activation of P_{97}. Rescuing HPV-16 into herpes mutants with deletions in different genes coding for herpes immediate-early proteins

and monitoring the transcriptional activity of P_{97} can be informative as to which immediate-early protein is involved in the mechanism of activation.

The long term goal of this work was to study late transcription in HPV-16, and to delineate the different viral and cellular factors involved in the switch from early to late gene expression. In RI, the disruption of the late region of HPV-16 greatly affects the evaluation of the early to late switch in papilloma transcription. Experiments can be designed to find out whether the early poly(A) signal is utilized by messages starting at P_{97}. Due to the disruption of the late region in RI, it is beyond experimental feasibility to determine whether transcriptional complexes initiating in the early region of HPV-16 utilize the late poly(A) signals. Yet, it is possible to examine transcription termination downstream of the early sequences of papilloma in RI and compare it to the utilization of the early poly(A) to formulate a hypothesis pertaining to the function of these two regulatory mechanisms in the complex early to late transcription switch. Inhibition of late transcription in RI infected Vero cells, if this is the case, might be due to the utilization of the early poly(A) signal rather than the late one, or under the negative regulation of a transcription termination signal downstream of the early

poly(A) signal. It is conceivable that transcriptional complexes initiating upstream from HSV promoters, such as the gD promoter, can recognize the late poly(A) signals implying the involvement of other mechanisms such as differential promoter utilization in the switch to late transcription. However, it is more informative to examine the function of the late poly(A) signals in the control of late transcription in the context of an intact late region. R2, in which the early and late regions are intact, was designed to explore such a situation.

Preliminary experiments on R2 have raised considerable doubts on transcription initiating from P_{97}. More experiments are needed to determine the precise role of a disrupted URR in the initiation of transcription from P_{97}. The structure of R2 presents an opportunity to examine the regulatory processes orchestrating late transcription. In addition to differential utilization of the late and early poly(A) sites and transcription termination, late transcription might be dependent on the activation of a late promoter which is only active in epithelial cells. Such cellular regulation of late transcription can be investigated by infecting epithelial cells at different maturation stages with R2 to identify and delineate the possible role of such cellular factor(s).

In summary, transcription was detected from P_{97} in

an HSV-1/HPV-16 recombinant virus. As concluded in other studies in which heterologous genes have been expressed as part of the herpes genome during infection, immediate-early genes such as ICP4 are important in activating such promoters. The URR of the HPV-16, in this case, seems to play an important role in the activation of P_{97}. The possible interaction between herpes immediate-early proteins and cellular transcription proteins binding at the URR is as complex and open to speculation as the mechanism by which immediate-early proteins activate early, late, and heterologous promoters.

REFERENCES

Ahola,H., Stenlund, A., Moreno-Lopez, J., and Pettersson, U., 1987, Promoters and processing sites within the transforming region of bovine papillomavirus type 1. J. Virol., 61: 2240-2244.

Amtmann, E., and Saur, G., 1982, Bovine papillomavirus transcription: polyadenylated RNA species and assessement of the direction of transcription. J. Virol., 43: 59-66.

Androphy, E.J., Lowy, D.R., and Schiller, J.T., 1987a, Bovine papillomavirus E2 trans-activating gene product binds to specific sites in papillomavirus DNA. Nature, 325: 70-73.

Androphy, E.J., Hubbert, N.L., Schiller, J.T., and Lowy, D.R., 1987b, Identification of the HPV-16 E6 protein from tranformed mouse cells and human cervical carcinoma cell lines. EMBO J., 6: 989-992.

Baker, C.C., and Howley, P.M., 1987, Differential promoter utilization by the bovine papillomavirus in transformed cells and productively infected wart tissue. EMBO J., 6: 1027-1035.

Baker, C.C., Phelps, W.C., Lindgren, V., Braun, M.J., Gonada, M.A., and Howley, P.M., 1987, Structural and functional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. J. Virol., 61: 962-971.

Baker, C.C., Phelps, W.C., Lindgren, V., Braun, M.J., Gonada, M.A., and Howley, P.M., 1987, Structural and functional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. J. Virol., 61: 962-971.

Baker, C.C., Noe, J.S., 1989, Transcription termination between bovine papillomavirus type 1 (BPV-1) early and late polyadenylation sites blocks late transcription in BPV-1 transformed cells. J. Virol., 63: 3529-3534.

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

Batterson, W., and Roizman, B., 1983, Charecterization of the herpes simplex virion-associated factor responsible for the induction of α genes. J. Virol., 46: 371-377.

Boshart, M., Gissmann, L., Ikenberg, H., Kleinheinz, A.,

Scheurlen, W., and zurHausen, H., 1984, A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer. EMBO J., 3: 1151-1157.

Breathnach, R., and Chambon, P., 1981, Organization and expression of eukaryotic split genes coding for proteins. Ann. Rev. Bioch., 50: 349-383.

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.

Campo, M.S., Moar, M.H., Sartirana, M.L., Kennedy, I.M., Jurrett, W.F.H., 1985, The presence of bovine papillomavirus type 4 DNA is not required for the progression to, or the maintenance of the malignant state in cancers of the alimentary canal in cattle. EMBO J., 4: 1819-1825.

Chan, W., Chong. T., Bernard, H., Klock, G., 1990, Transcription of the transforming genes of the oncogenic human papillomavirus-16 is stimulated by tumor promoters through AP1 binding sites. Nucleic Acids Res., 18: 763-769.

Chomczynski, P., and Sacchi, N., 1987, Single step method of

RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Bioch., 162: 156-159.

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

Chow, L.T., Nasseri, M., Wolinsky, S.M., and Broker, T.R., 1987a, Human papillomavirus types 6 and 11 mRNAs from genital condylomatas acumanata. J. Virol., 61: 2581-2588.

Chow, L.T., Reilly, S.S., Broker, T.R., and Taichman, L.B., 1987b, Identification and mapping of human papillomavirus type 1 RNA transcripts recovered from plantar warts and infected epithelial cell cultures. J. Virol., 61: 1913-1918.

Corey, L., Spear, P.G., 1986, Infections with herpes simplex virus. New Eng. J. Med., 314: 686-691.

Costanzo, F., Campadelli-Fiume, G., Foa-Tomas, L., Cassai, E., 1977, Evidence that herpes simplex virus DNA is transcribed by cellular RNA polymerase II. J. Virol., 21: 996-1001.

Cripe, T.P., Haugen,T., Jeffrey, P.T., Tabtabai, 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 viral E2 trans-activator and repressor gene products: implications for cervical carcinogenesis. EMBO J., 6: 3745-3753.

Crum, C.P., Mitao, M., Levine, R.V., and Silverstein, S., 1985, Cervical papillomaviruses segregate within morphologically distinct precancerous lesions. J. Virol., 54: 675-681.

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

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 tumor antigen forms a specific complex with the product of the retinoblastoma susceptibility gene. Cell, 54: 275-283.

Dierks, P., Ooyen, A.V., Cochran, M.D., Dobkin, C., Reiser, J., and Weissmann, C., 1983, Three regions upstream from the cap site are required for efficient and accurate transcription of

the rabbit ß-globin gene in mouse 3T6 cells. Cell, 32: 695-706.

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

Durst, M., Groce, C.M., Gissmann, L., Schwarz, E., and Huebner,K., 1987, Papillomavirus sequences integrate near cellular oncogenes in some cervical carcinomas. Proc. Natl. Acad. Sci. USA, 84: 1070-1074.

Durst, M., Gissmann, L., Ikenberg, H., and zur Hausen, H., 1983, A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions, Proc. Natl. Acad. Sci. USA, 80: 3812-3815.

Dush, M.K., Sikela, J.M., Khan, S.A., Tischfield, J.A., Stambrook, P.J., 1985, Nucleotide sequence and organization of the mouse adenine phosphoribosyltransferase gene: presence of a coding region common to animal and bacterial phosphoribosyl transferase that has variable intron/exon arrangement. Proc. Natl. Acad. Sci. USA, 82: 2731-2735.

Dvoretzky, I., Shober, R., Chattopadhyay, S.K., and Lowy, D.R., 1980, A quantitative in vitro focus assay for bovine papilloma

virus. Virology, 103: 369-375.

Dynan, W.S., and Tjian, R., 1983, Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. Cell, 32: 669-680.

Dynan, W.S., and Tjian, R., 1985, Control of eukaryotic messenger RNA synthesis by sequence-specific DNA binding proteins. Nature, 316: 774-777.

Dyson, N., Howley, P.M., Munger, K., Harlow, E.D., 1989, The human papillomavirus 16 E7 oncoprotein is able to bind to the retinoblastoma gene product. Science, 243: 934-936.

Everett, R.D., 1988, Promoter sequence and cell type can dramatically affect the efficiency of transcriptional activation induced by herpes simplex virus type 1 and its immediate-early products Vmw175 and Vmw110. J. Mol. Biol., 203: 739.

Fenwick, M., Roizman, B., 1977, Regulation of herpesvirus macromolecular synthesis IV. Synthesis and modification of viral polypeptides in enucleated cells. J. Virol., 22: 720-725.

Fire, A., Samuels, M., and Sharp, P., 1983, Interaction between RNA polymerase II, factors, and template leading to accurate

transcription. J. Biol. Chem., 259: 2509-2516.

Flores, O., Maldonado, E., Reinberg, D., 1989, Factors involved in specific transcription by mammalian RNA polymerase II. J. Biol. Chem., 264: 8913-8921.

Franza, R.B., Rauscher, F.J., Josephs, S.F., Curran, T., 1988, The fos complex and fos-related antigens recognize sequence elements that contain AP-1 binding sites. Science, 239: 1150-1153.

Friend, S.H., Bernards, R., Rogelj, S., et al., 1986, The human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature, 323: 643-646.

Fujita, T., Shibuya, H., Ohashi, T., Yamanishi, K., and Taniguchi, T., 1986, Regulation of human interleukin-2 gene: functional DNA sequences in the 5' flanking region for the gene expression in activated T lymphocytes. Cell, 46: 401-407.

Galloway, T.C., Soper, G.R., Elsen G., 1960, Carcinoma of the larynx after irradiation for papilloma. Arch. Otolaryngol. 72: 289-294.

Ghosh, P.K., Lebowitz, P., Frisque, R.J., and Gluzman, Y., 1981,

Identification of a promoter component involved in positioning the 5' termini of simian virus 40 early mRNAs. Proct. Natl. Acad. Sci. USA, 78: 100-104.

Gidoni, D., Dynan, W.S., Tjian, R., 1984, Multiple specific contacts between a mammalian transcription factor and its cognate promoter. Nature, 312: 409-413.

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

Gloss, B., Chong, T., and Bernard, H.U., 1989b, Numerous Nuclear proteins bind the long control region of human papillomavirus type 16: a subset of 6 of 23 DNase I protected segments coincides with the location of the cell-type-specific enhancer. J. Virol., 63: 1142-1152.

Gloss, B., Yeo-Gloss, M., Meisterenst, M., Rogge, L., Winnacker, E., Bernard, H., 1989a, Clusters of nuclear factor I binding sites identify enhancers of several papillomavirus but alone are not sufficient for enhancer function. Nucleic Acids Res., 17: 3519-3533.

Graham, F.L., and Van der Eb, A.J., 1973, A new technique for the assay of infectivity of adenovirus 5 DNA. Virol, 52: 456.

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

Grosschedl, R., and Birnstiel, M.L., 1980, Identification of regulatory sequences in the prelude sequences of an H2A histone gene by the study of specific deletion mutants in vivo. Proc. Natl. Acad. Sci. USA, 77: 1432-1436.

Harbour, J.W., Lai, S., Whang-Peng, J., Gazdar, A.F., Minna, J.D., Kaye, F.J., 1988, Abnormalities in structure and expression of the human retinoblastoma gene in SCLC. Science, 241: 353-357.

Haugen, T.H., Cripe, T.P., Ginder, G.D., Karin, M., and Turek,P., 1987, Trans-activation of an upstream early gene promoter of bovine papillomavirus-1 by a product of the viral E2 gene. EMBO J., 6: 145-152.

Heilman, C.A., Engel, L., Lowy, D.R., Howley, P.M., 1982, Virus-specific transcription in bovine papillomavirus type I. Virol 119: 22-34.

Holmes, D.S., and Quigley, M., 1981, A rapid boiling method for the preparation of bacterial plasmids. Anal. Chem., 114: 193

Hirochica, H., Brocker, T.R., and Chow, L.T., 1987, Enhancers and trans-acting E2 transcriptional factors of papillomavirus. J. Virol., 61: 2599-2606.

Honess, R.W., and Roizman, B., 1974, Regulation of herpesvirus macromolecular synthesis I. Cascade regulation of the synthesis of three groups of viral proteins. J. Virol., 14: 8-19.

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.

Howley, P.M., 1983, The molecular biology of papillomavirus transformation. Am. J. Path., 113: 413-421.

Howley, P.M., 1990, Papillomavirinae and their replication. In: Virology, (Fields, B.N., Knipe, D.M., et al., eds.), pp. 16250-1650, Raven Press, Ltd., N.Y.

Imbra, R.J., and Karin, M., 1986, Phorbol ester induces the transcriptional stimulatory activity of the SV40 enhancer.

Nature, 323: 555-558.

Jarret, W., McNeil, P., Grimshaw, W., Selman, I.E., and McIntyre W.I.M., 1978, High incidence area of cattle cancer with a possible interaction between an enviromental carcinogen and papillomavirus. Nature, 274: 215-217.

Jarrett, W., McNeil, P., Laird, H.,et al., 1980, Virus induced papillomas of the alimentary tract of cattle. Inter. J. Cancer, 22: 323-328.

Johnson, P.A., Everett, R.D., 1986, The control of herpes simplex virus type-1 late gene transcription: a "TATA-box"/cap site region is sufficient for fully efficient regulated activity. Nucl. Acids Res., 14: 8247-8264.

Johnson, P.F., and McKnight, S.L., 1989, Eukaryotic transcriptional regulatory proteins. Ann. Rev. Bioch., 58: 799-839.

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

Kanada, T., Watanabe, S., and Yushike, K., 1988, Immortalization

of primary rats cells by human papillomavirus type 16 subgenomic DNA fragments controlled by the SV40 promoter. Virology, 165: 321-325.

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

Kennedy, I.M., Haddow, J.K., and Clements, J.B., 1990, Analysis of human papillomavirus type 16 late mRNA 3' processing signals in vitro and in vivo. J. Virol., 64: 1825-1829.

Kreider, J.W., and Barlett, G.L., 1987, Shop papilloma-carcinoma complex of rabbits. Adv. Cancer Res. 35: 81-110.

Lambert, P.F., Baker, C.C., and Howley, P.M., 1988, The genetics of bovine papillomavirus type 1. Ann. Rev. Genet. 22: 235-258.

Lambert, P.F., Spalholz, B.A., and Howley, P.M., 1987, A transcriptional repressor encoded by BPV-1 shares a common carboxyl-terminal domain with the E2 transactivator. Cell, 50: 69-78.

Lazo, P.A., 1988, Human papillomavirus in oncogenesis. Bio. Essays, 9: 158-162.

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

Lee, W., Mitchell, P., and Tjian, R., 1987b, Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell, 49: 741-752.

Ligas, M.W., and Johnson, D.C., 1988, A herpes simplex virus mutant in which glycoprotein D sequences are replaced by beta-galactosidase sequences binds to but is unable to penetrate into cells. J. Virol., 62: 1486-1494.

Lusky, M., Botchan, M.R., 1984, Characterization of the bovine papilloma virus plasmid maintenance sequences. Cell, 36: 391-401.

Lusky, M., Botchan, M.R., 1986, Transient replication of bovine papillomavirus type 1 plasmids: cis and trans requirements. Proc. Natl. Acad. Sci. USA, 83: 3609-3613.

Mackem, S., and Roizman, B., 1980, Regulation of herpesvirus macromolecular synthesis: transcription initiation sites and domains of α genes. Proc. Natl. Acad. Sci. USA, 77: 7122-7126.

Maniatis, T., Fritsch, E., Sambrook, J., 1982, Molecular Cloning, a Laboratory Manual. Cold Spring Harbour Laboratory, Cold Spring Harbour.

Matsai, T., Segall, J., Weil, A., Roeder, R., 1980, Multiple factors required for acurate initiation of transcription by purified RNA polymerase II. J. Biol. Chem., 255: 11992-11996.

Maxam, A.A., and Gilbert, W., 1977, A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA, 74: 560-564.

McCusker, C.T., and Bacchetti, S., 1988, The responsiveness of human papillomavirus upstream regulatory regions to herpes simplex virus immediate early proteins. Virus Res., 11: 199-207.

McKnight, S.L., 1982, Functional relationship between transcriptional cotrol signals of the thymidine kinase gene of herpes simplex virus. Cell, 31: 355-365.

McKnight, S., and Tjian, R., 1986, Transcriptional selectivity of viral genes in mammalian cells. Cell, 46: 795-805.

McKnight, S., and Tjian, R., 1988, Transcriptional selectivity of viral genes in mammalian cells. Cell, 46: 795-805.

Mermelstein, F.H., Flores, O., Reinberg, D., 1989, Initiation of transcription by RNA polymerase II. Biochimica et Biophysica, 1009: 1-10.

Michael, N., Spector, D., Mavromara-Nazos, P., Kristie, T., Roizman, B., 1988, The DNA-binding properties of the major regulatory protein alpha 4 of herpes simplex virus. Science, 239: 1531-1534.

Meisfeld, R., Godowski, P.J., Maler, B.A., Yamamoto, K.R., 1987, Glucocorticoid receptor mutants that define a small region sufficient for enhancer activation. Science, 263: 423-427.

Mitchell, P.G., Tjian, R., 1989, Transcription regulation in mammalian cells by sequence-specific DNA binding proteins. Science, 245: 371-378.

Munger, K., Phelps, W.C., Bubb, V., Howley, P.M., 1989, The E6 and E7 genes of the human papillomavirus together are necessary

and sufficient for transformation of primary human keratinocytes. J. Virol., 63: 4417-4421.

Myers, R.M., Tjian, R., 1980, Construction and analysis of simian virus 40 origins defective in tumor antigen binding and DNA replication. Proc. Natl. Acad. Sci. USA, 77: 6491-6495.

Nigro, J.M., Baker, S,J., Preising, A.C., Jessup, J.M., Hostetter, R., Clearly, K., Binger,S.H., Davidson, N., Baylin, S., Devilee, P., Glover, T., Collins, F.S., Weston, A., Modali, R., Harris, C.C., and Vogelstein, B., 1989, Mutations in p53 gene occur in diverse human tumor types. Nature 342: 705-708.

Oriel, J.D., and Almeida, J.D., 1970, Demonstration of virus particles in human genital warts. Brit. J. Venereal Diseases, 46: 37-42.

Payvar, F., DeFranco, D., Firestone, G.L., Edgar, B., Wrange, O., Okret, S., Gustafsson, J.A., and Yamamoto, K.R., 1983, Sequence-specific binding of glucocorticoid receptor to MMTV DNA at sites within and upstream of the transcribed region. Cell, 35: 381-392.

Pfister, H., 1987, Papillomaviruses: general description, taxonomy, and classification. In: The Papovaviridae vol 2

(Salzman NP, Howley PM, eds.), pp. 1-38. Plenum Press, New York.

Ptashne, M., 1986, Gene regulation by proteins acting nearby and at a distance. Nature, 322: 697-701.

Ptashne, M., 1988, How eukaryotic transcriptional activators work. Nature, 335: 683-689.

Rauscher, F.J. III, Cohen, D.R., Curran, T., Bos, T.J., Vogt, P.K., Bohmann, D., Tjian, R., Franza, R. Jr., 1988, Fos-associated protein P39 is the product of the jun proto-oncogene. Science, 240: 1010-1016.

Reeves, W.C., Rawls, W.E., Brinton, L.A., 1989, Epidemiology of genital papillomaviruses and cervical cancer. Rev. Infect. Dis., 11: 426-439.

Reinderg, D., Horikoshi, M., and Roeder, R., 1987, Factors involved in specific transcription in mammalian RNA polymerase II. J. Biol. Chem., 252: 1375-1380.

Roizman, B., and Sears, A.E., 1990, Herpes simplex viruses and their replication. In: Virology (Fields B.N., Knipe D.M., et al. eds.), pp. 1795-1841. Raven Press Ltd., New York.

Sanchez-Pescador, R., Power, M.D., Barr, P.J., Steimer, K.S., Stempien, M.M., Brown-Shimer S.C., Gee, W.W., Renard, A., Randolph, A., Levy, J.A., 1985, Nucleotide sequence and expression of an AIDs-associated retrovirus (ARV-2). Science, 227: 484-492.

Schiller, J.T., Vass, W.C., Vousdan, K.H., Lowy, D.R., 1986, The E5 open reading frame of bovine papillomavirus type 1 encodes a transforming gene. J. Virol. 57: 1-6.

Schneider-Gadicke, A., and Schwarz, E., 1986, Different human cervical carcinoma cell lines show similar transcription patterns of human papillomavirus type 18 early genes. EMBO J., 5: 2285-2292.

Schwarz, E., Freese, U.K., Gissmann, L., Mayer, W., Roggenbuck, B., Stremlan, A., and zur Hausen, H., 1985, Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. Nature, 314: 111-114.

Serfling, E., Jasin, M., and Schaffner, W., 1985, Enhancers and eukaryotic gene transcription. Trends in Genet., 1: 224-230.

Shah, K.V., and Howley, P.M., 1990, Papillomaviruses. In: Virology (Fields B.N., Knipe D.M., et al. eds.), pp. 1651-1676.

Raven Press Ltd., New York.

Sharma, A., Bos, T., Pekkala-Flagan, A., Vogt, P.K., and Lee, A., 1989, Interaction of cellular factors related to the jun oncoprotein with the promoter of a replication-dependent hamster histone H3.2 gene. Proc. Natl. Acad. Sci. USA, 86: 491-495.

Smiley, J.R., 1980, Construction in vitro and rescue of a thymidine kinase-deficient deletion mutation of herpes simplex virus. Nature, 285: 333-335.

Smiley, J.R., Smibert, C.A., and Everett, R.D., 1987, Expression of cellular gene cloned in herpes simplex virus: rabbit beta-globin is regulated as an early viral gene in infected fibroblasts. J. Virol., 61: 2368-2377.

Smiley, J.R., Panning, B., Smibert, C.A., 1990, Regulation of cellular genes by HSV products. In: Herpesvirus transcription and its regulation (Wagner, E.K., ed.), pp. 151-179. CRC press, Boca Raton, Florida.

Smotkin, D., Prokoph, H., and Wettstein, F.O., 1989, Oncogenic and non-oncogenic human genital papillomaviruses generate the E7 mRNA by different mechanism. J. Virol., 63: 1441-1447.

Smotkin, D., and Wettstein, F.O., 1986, Transcription of human papillomavirus type 16 early genes in a cervical cancer and cancer-derived cell lines and identification of E7 protein. Proc. Natl. Acad. Sci. USA, 83: 4680-4684.

Sousa, R., Dostatni, N., and Yaniv, M., 1990, Control of papillomavirus gene expression. Biochimica et Biophysica Acta., 1032: 19-37.

Stoler, M.H., and Broker, T.R., 1986, In situ hybridization detection of human papillomavirus DNAs and messenger RNAs in genital condylomas and a cervical carcinoma. Hum. Path., 47: 1250-1258.

Squinto, S.P., Block, A.L., Doucet, J.P., 1989, Epidermal growth factor induction of cellular proliferation and proto oncogene expression in growth-arrested rat H4 IIE hepatoma cells: role of cyclic adenosine monophosphate. Mol. Endocrin., 3: 433-446.

Tsunokawa, Y., Takebe, N., Kasamatsu, T., Terado, M., and Sugimura, T., 1986, Transforming activity of human papillomavirus type 16 DNA sequences in a cervical cancer. Proc. Natl. Acad. Sci. USA, 83: 2200-2203.

Valerio, D., Duyvesteyn, G.C., Dekker, B.M.M., Weeda, G.,

Berkvens, T.M., vander Voorn, L., van Ormondt, H., and vander Eb, A.J., 1985, Adenosine deaminase: characterization and expression of a gene with remarkable promoter. EMBO J., 4: 437-443.

Ward, P., Coleman, D.V., and Malcolm, A.D., 1989, Regulatory mechanism of the papillomavirus. Trends in Genet., 5: 97-99.

Watanabe, S., Kanada, T., Yoshiike, K., 1989, Human papillomavirus type 16 transformation of primary human emberyonic fibroblasts requires expression of open reading frames E6 and E7. J. Virol., 63: 965-969.

Watson, J.D., Hopkins, N.H., Roberts, J.W., Steitz, J.A., Weiner, A.M., Molecular Biology of the Gene, ch. 3 pp. 65-94. The Benjamin/Cummings Publishing Company Inc., California.

Weil, P.A., Luse, D.S., Segall, J., Roeder, R.G., 1979, Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. Cell, 18: 469-484.

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.

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 anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. Nature, 334: 124-129.

Wu, G.J., 1978, Adenovirus DNA-directed transcription of 5.5s RNA in vitro. Proc. Natl. Acad. Sci. USA, 75: 2175-2179.

zur Hausen, H., 1976, Condylomata acuminata and human genital cancer. Cancer Res., 36: 794.