This is to certify that the

dissertation entitled

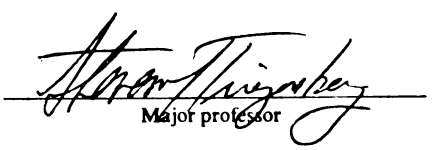
The effect of deletion mutations of the activation domain of VP16 upon herpes simplex virus type 1 lytic infection

presented by

Rath Pichyangkura

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Biochemistry



Major professor

Date 7/26/96

THE EFFECT OF DELETION MUTATIONS OF THE ACTIVATION DOMAIN OF VP16 UPON HERPES SIMPLEX VIRUS TYPE 1 LYTIC INFECTION

By

Rath Pichyangkura

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry

1996

ABSTRACT

THE EFFECT OF DELETION MUTATIONS OF THE ACTIVATION DOMAIN OF VP16 UPON HERPES SIMPLEX VIRUS TYPE 1 LYTIC INFECTION

By

Rath Pichyangkura

VP16 is an essential structural protein in the virion of herpes simplex virus type 1 which functions in the activation of the viral immediately early genes and in maturation of the virion particle. The transcriptional activation domain of the protein can be separated into two subdomains, the N-subdomain and the C-subdomain. Both of these subdomains are necessary for the full transcription activation function of the protein, but either one is able to activate transcription independently of the other. VP16 truncation mutations with deletion of either one or both subdomains were constructed in the context of the viral genome and tested for their effect upon viral lytic infection in cell culture.

Recombinant viruses with deletions in the activation domain of VP16 are viable in tissue culture and produce intact particles containing the truncated form of VP16 protein. However, these viruses demonstrated defects in their plaque characteristics, single step growth, and the number of particles per plaque

forming unit. A recombinant virus bearing a deletion of the C-subdomain gave a slightly small plaque with no other defective phenotype. Deletion of the N or both subdomains gave much smaller plaques and grew poorly in single step growth assays. These viruses also exhibit higher particle per plaque forming unit ratios, comparing to the wild type virus. The particle per plaque forming unit ratios of the viruses that bear deletions of the N subdomain or of both subdomains were approximately 30-fold and 100-fold greater than of the wild type virus. Despite their defects, all recombinant viruses with truncated VP16 genes were able to deliver their DNA to the nucleus of the host cell. Therefore, these defects involve the activation of immediate early gene expression.

Homologues of VP16 have been identified in herpes simplex virus type 2, bovine herpes virus type 1, equine herpes virus types 1 and 4, varicella zoster virus, and Marek's disease virus. Hydrophobic cluster analysis of deduced protein sequences was used to predict the location and critical amino acids of the activation domain of the varicella zoster virus VP16 homologue. This prediction was tested by mutational analysis which further emphasized the importance of large hydrophobic residues as an essential component of the activation domain of VP16 and its homologues. Hydrophobic cluster analysis also suggested the location and critical residues in other VP16 homologues.

ACKNOWLEDGMENTS

I wish to thank Dr. Steven Triezenberg for being an excellent graduate advisor and mentor. I also thank him for his patience, understanding and encouragements through the years. I also wish to thank the members of the Triezenberg laboratory, both past and present, namely Doug and Andrea Cress, Dave Pepperl, Jeff and Marty Regier, Lisa Ortquist, Fan Shen, Peter Horn, John Stebbins, Lee Alexander , Jaya Reddy, Susan Sullivan, Eugene Chung and Soren Jaglo-Ottosen. I thank them for their great friendship and creative discussions.

I acknowledge the contributions of the members of my guidance committee Drs. Laurie Kaguni, Lee Kroos, Lee Velicer and Pamela Green. I acknowledge the contributions of my collaborators Drs. Hiroyuki Moriuchi, Masako Moriyuchi and Jeff Cohen (NIAID), for the VZV ORF10 work, and Dr. David Koelle (University of Washington) for the $CD4^+$ T-cell work. I also wish to acknowledge the Anantamahidol foundation for their financial support.

I thank the members of the Thai community as well as the international community that kept my life interesting during my stay at MSU. I also thank my close friends Darunee Buripakdi and Takeshi Ueda for their encouragement and support during the final period of my dissertation. Finally, I thank my parents, brother and sister for their support, encouragement and patience over the years.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont'd)

TABLE OF CONTENTS (cont'd)

PAGE

vii

LIST OF TABLES

PAGE

CHAPTER II

CHAPTER III

APPENDIX B

LIST OF FIGURES

PAGE

CHAPTER I

CHAPTER II

LIST OF FIGURES (cont'd)

PAGE

CHAPTER III

LIST OF ABBREVIATIONS

BHV-1	bovine herpes virus type 1
CAT	chloramphenicol acetyltransferase
CD	circular dichroism
CPE	cytopathic effect
CTD	carboxy-terminal domain
E	early
EHV-1	equine herpes virus type 1
EHV-4	equine herpes virus type 4
FCS	fetal calf serum
GTFs	general transcription factors
HCA	hydrophobic cluster analysis
HMG	high mobility group proteins
HSV-1	herpes simplex virus type 1
HSV-2	herpes simplex virus type 2
hpi.	hours post infection
ICP	infected cell protein
IE	immediately early
L	late
MDV	Marek's disease virus
moi	multiplicity of infection
NMR	nuclear magnetic resonance

LIST OF ABBREVIATIONS (cont'd)

PHD	profile network system from Heidelberg
pfu	plaque forming unit
RNAPII	RNA polymerase II
SDS	sodium dodecyl sulfate
SDS PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
TAFs	TBP associated factors
TBP	TATA binding protein
TF	transcription factor
VP16	viral protein 16
VP16II	viral protein 16 from herpes simplex virus type 2
VZV	varicella zoster virus

Single letter abbreviation for the amino acids: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr

CHAPTER I

INTRODUCTION

Eukaryotic Transcription

The regulation of gene expression in eukaryotes is one of the most interesting and extensively studied aspects in biology. Gene expression is regulated at multiple steps from transcription initiation, elongation, and termination; mRNA processing, transport, and stability; translation, modification, and degradation of gene products. This multi-level control of regulation is essential for cell growth, development, maintenance, and the ability to respond to both internal and external stimuli.

Transcription of protein-encoding genes by RNA polymerase II (RNAPII) is the initial and major regulated step in gene expression. In contrast to prokaryotic RNA polymerase which is composed of 3 core peptides (α , β, β') and σ factor, that confer its specificity to the promoter, eukaryotic RNAPII is composed of 8-14 peptides, ranging from 10-220 kDa, and requires a variety of additional factors for selective promoter recognition and regulation of transcription initiation (Young, 1991). These factors are roughly divided into two major categories, general transcription factors (GTFs) and transcription regulatory factors.

The largest subunit of RNAPII has a unique heptapeptide repeat sequence, Tyr-Ser-Pro-Thr-Ser-Pro-Ser, at the carboxy-terminal domain (CTD) of the

protein (Cadena & Dahmus, 1987). In cells, the largest subunit of RNAPII is found in two forms, II_0 and II_a, in which the CTD is highly phosphorylated and unphosphorylated, respectively. The CTD is essential for cell viability, and appears to be involved in the response to transcriptional regulation, and pre-initiation complex formation of some promoters (Young, 1991). The highly phosphorylated CTD may also play a role in transcription initiation as a switch from pre-initiation complex to open complex formation and promoter clearance (Young, 1991; Zawel & Reinberg, 1993; Zawel & Reinberg, 1995).

The promoter structure of protein-encoding genes varies considerably among different gene families, nevertheless each of them share common core structures. RNAPII promoters have two possible core elements which can mediate the nucleation of the transcription initiation complex, either independently or synergistically. One of these elements, the TATA box, in mammals is located approximately 25-30 nucleotides upstream of the transcription start site. The other, the initiator (Inr) element is normally found encompassing the transcription start site (Smale & Baltimore, 1989; Weis & Reinberg, 1992; Zawel & Reinberg, 1995).

Most GTFs for RNAPII which are essential for a basal level of transcription have been purified, characterized, and cloned. The schematic order of assembly of these factors into a transcription preinitiation complex has been studied and laid out, however there are still some ambiguities concerning the precise order of assembly, as well as the basal factors needed for different promoters. Transcription initiation begins with the template commitment step in which transcription factor (TF) IID, a large multiple protein complex comprising the TATA binding protein (TBP), and the TBP associated factors (TAFs), binds to the TATA motif of the promoter. The formation of the TFIID·DNA complex can be further facilitated and stabilized by TFIIA. TFIIB then joins the complex

followed by the recruitment of RNAPII and TFIIF to the pre-initiation complex. TFIIE and TFIIH then join the complex. In the presence of ATP and dNTP's, extensive phosphorylation of the CTD by TFIIH occurs followed by open complex formation and the synthesis of the nascent mRNA chain (Zawel & Reinberg, 1995).

There is increasing evidence suggesting that these factors exist in a large multi-protein complex and the schematic order of assembly might not reflect the natural phenomena of complex assembly within a living cell. RNAPII was found to be associated with multiple transcription machinery SRB, SWI/SNF and GTFs in a RNAPII holoenzyme (Koleske & Young, 1994; Koleske & Young, 1995; Ossipow, et al., 1995; Chao, et al., 1996; Wilson, et al., 1996). TFIIF was initially found to be associated with RNAPII, and recent studies have established an interaction between TFIIB, TFIIF and RNAPII in solution (Fang & Burton, 1995; Ha, et al., 1993). It is plausible to assume that in the process of initiation complex assembly, RNAPII, TFIIB, and TFIIF could be recruited to the TFIID·TFIIA·DNA complex as a TFIIB·RNAPII·TFIIF, multi-protein complex instead of as individual factors. This could be the case for other assembly steps in the initiation complex as well.

On TATA-less promoters, initiation of complex assembly usually commences with the association of initiator binding factors such as YYI (Usheva & Shenk, 1994) or TFII-I (Roy, et al., 1993) to the Inr element nucleating preinitiation complex assembly, followed by the recruitment of GTFs, RNAPII and other components of the transcription machinery. Moreover some TAFs such as TAF150 have also been found to specifically recognize promoters containing an Inr element (Verrijzer, et al., 1994; Verrijzer, et al., 1995).

Regulated Transcription

Regulated transcription is accomplished by the binding of additional factors either upstream or downstream of the promoter element, which interact either directly or indirectly with the transcription initiation complex to regulate its activity. These factors could either stimulate or repress transcription from the promoter which they regulate. Transcriptional regulatory proteins consist of at least two functional domains. One of these domains confers specificity to the promoter, through specific DNA binding or protein-protein interactions. The other domain is the regulatory domain which functions to regulate the level of transcription (Johnson & McKnight, 1989; Triezenberg, 1995; Zawel & Reinberg, 1995). These regulatory factors are simply characterized by their functional role as transcription activators or repressors.

Despite the regulation by promoter specific factors, more general modes of gene regulation, such as those mediated by histone proteins, high mobility group proteins (HMG) and the methylation state of the DNA template could greatly contribute to the precise regulation of gene expression. It is well documented that packaging of a DNA template into chromatin by histone proteins renders the DNA template inert to transcription (Paranjape, et al., 1994). Both biochemical and genetic studies have demonstrated the significant role of histone proteins in transcriptional regulation (Paranjape, et al., 1994). Histone H1, in a complex with template DNA, has been shown to have a potent inhibitory effect on RNAPII (Croston, et al., 1991; Laybourn & Kadonaga, 1991). The N-terminal domains of histone H3 and H4 have been shown to play a role in the regulation of gene expression. The N-terminal domain of H4 is necessary for activation of the GAL1 promoter, as well as silencing of the yeast mating loci and other yeast promoters,

while H3 is required for the repression of most genes which require H4 for activation (Wolffe, 1994). Lysine residues in the N-terminal domain of histone H3 and H4 are also subjected to reversible acetylation, which has been found to have a positive effect upon transcription (Paranjape, et al., 1994).

HMG proteins are a group of non-histone proteins which were initially defined as proteins that could be extracted from nuclei with 0.35 M NaCl and were soluble in 2% trichloroacetic acid. Three groups of HMG proteins have been defined, HMG1/2, HMG14/17, and HMGI/Y (Paranjape, et al., 1994). HMG 1/2 was found to be able to stimulate RNAPII transcription and the binding of upstream stimulating factors *in vitro* (Watt & Molly, 1988). HMG17 can stimulate activation of transcription by VP16 on *in vitro* assembled chromatin (Paranjape, Krumm, & Kadonaga, 1995). HMGI/Y was also found to stimulate the binding of NF-κB to the PDRII promoter element (Thanos & Maniatis, 1992). In spite of these minor clues, the biological role of these proteins in transcription regulation remains to be clarified.

Methylation of the DNA, at the 5 position of cytosine residues in the sequence dCpdG, plays an important role in higher eukaryotic transcription regulation. There is a strong correlation between the presence of 5-methylcytosine (m^5C) in the DNA template and the repression of transcription (Cedar, 1988; Doerfler, 1983; Doerfler, et al., 1990; Dynan, 1989; Guentte, et al., 1992; Lewis & Bird, 1991; Peek, et al., 1991). Methylation of the DNA template can effect transcription presumably at two levels, by affecting the packaging of the DNA, and by interacting either directly or indirectly with the transcription machinery of the cell.

Methylated cytosine residues in chromatin are less accessible to nucleases such as micrococcal nuclease, suggesting that they occupy a portion of DNA that is tightly packed (Lewis & Bird, 1991). When chromatin is fractionated, the

nucleosome fraction which is tightly associated with histone H1 was found to contain the majority of the methylated cytosine (Lewis & Bird, 1991). There is also evidence indicating that histone H1 can associate more readily with methylated than unmethylated DNA (Johnson, Goddard, & Adams, 1995).

Methylation of specific sites in genes or flanking sequences, such as a promoter sequence (Guentte, et al., 1992; Peek, et al., 1991) or an intron (Graessmann, et al., 1994), can also repress transcription. It is believed that methylation inhibits activation by interfering with the binding of positive regulators to their recognition sites. There are also proteins that bind specifically methylated DNA sequences (Boyes & Bird, 1991) that might block transcription by RNAPII. Despite these inhibitory effects, methylation of some promoters or genes does not seem to effect transcription or the function of certain activators (Harrington, et al., 1988; Wölfl, et al., 1991).

Although these results do not imply that methylation is only involved in the inhibition of gene expression, it suggests an undeniable role of methylation on gene expression by RNAPII. The mechanism of how methylation effects transcription by RNAPII is very intriguing and still remains to be studied.

Transcription Activators

Transcription activators are characterized roughly according to the most prominent species of amino acid found in the primary structure of the activation domains. Four groups of activators have been initially defined as acidic-rich , proline-rich, glutamine-rich and serine and threonine-rich activators (Johnson, et al., 1993). Recent studies have demonstrated that the most apparent amino acids in these activation domains may not be the most critical residues that contribute to activation ability (Triezenberg, 1995). For example, in the case of the acidic-

rich activators, VP16 (Cress & Triezenberg, 1991b; Regier, Shen, & Triezenberg, 1993) and GCN4 (Drysdale, et al., 1995) or the glutamine-rich activators Sp1 (Gill, et al., 1994) and Oct-1 (Tanaka & Herr, 1994), the most critical residues for their function appears to be the bulky hydrophobic residues.

Transcriptional activators could function in multiple distinct fashions at multiple steps in the transcription process; however, the mechanism of how these proteins work is not well established. To understand the mechanism of how these activators work, attempts have been made to identify the "targets" of activation domains. Biochemical and biophysical studies have shown interactions between GTFs and transcription activators. The acidic activation domain of VP16 has been shown to interact with TBP (Ingles, et al., 1991; Stringer, Ingles, & Greenblatt, 1990), and TFIIB (Lin, et al., 1991). The 62 kDa subunit of TFIIH was also shown to interact with the acidic activators VP16 and p53 (Xiao, et al., 1994). Mutations in the activation domain that effect the transcription activation ability also effect the interaction between these factors suggesting a significant role for these interactions in transcription activation.

In some cases the activation domains of activators do not directly interact with their targets, but rather through factor(s) which are termed adaptors, coactivators or mediators. A number of factors which serve as a bridge or link between activators and GTFs have been characterized. ADA2 in complex with ADA3 and GCN5 was found to be necessary for the acidic activators GAL4 and the chimeric GAL4-VP16 to work properly in yeast (Berger, et al., 1992; Horiuchi, Silverman, Marcus, & Guarente, 1995; Marcus, et al., 1994). Previously, several TAFs were also found to interact with transcription activators. Drosophila TAFII40, homologous to human TAFII32, was found to interact with the activation domains of VP16 (Goodrich, et al., 1993; Klemm, et al., 1995) and p53 (Thut, et al., 1995). Drosophila TAF110 interacts with the glutamine rich Sp1

(Hoey, et al., 1993), human TAF55 interacts with CTF (Zawel & Reinberg, 1995) and TAF150 has been shown to interact with NTF-1 (Chen, et al., 1994).

Furthermore, Hengartner *et al.* have demonstrated that the multi-protein RNAPII holoenzyme, comprising RNAPII, SRB, SWI/SNF and GTFs (Koleske & Young, 1994; Koleske & Young, 1995; Ossipow, et al., 1995; Chao, et al., 1996; Wilson, et al., 1996), was capable of responding to transcription activators. The acidic activation domain of VP16 was shown to specifically bind to the RNAPII holoenzyme, suggesting that transcription activators may function, in part, through direct interaction with the RNAPII holoenzyme (Hengartner, et al., 1995).

The mechanism by which these interactions could lead to the activation of transcription has been studied for some activators. Transcription activators were found to have the ability to disrupt or reconfigure chromatin to allow access of GTFs as well as other molecules of activators (Axelrod, Reagan, & Majors, 1993; Pazin, Kamakaka, & Kadonaga, 1995) to the promoter.

Recent studies by Wilson *et al.* demonstrated that SWI/SNF is a component of the RNAPII holoenzyme (Wilson, et al., 1996). This would provide the holoenzyme with the capability to disrupt nucleosomal DNA. Therefore, the recruitment of RNAPII holoenzyme by transcription activators, such as VP16, would facilitate the disruption of nucleosome and stable binding of the various transcription components of the initiation complex to the promoter.

Some activators are able to promote assembly of the preinitiation complex. For example the formation of the TFIID·TFIIB·TFIIA complex is enhanced by the activation domain of Zta (Lieberman, 1994; Lieberman & Berk, 1994; Pazin, et al., 1995). Moreover, in some cases activators can promote elongation of the mRNA chain (Yankulov, et al., 1994) or alleviate inhibitory effects of some repressors such as Dr1 or Dr2 (Kraus, et al., 1994; Merino, et al., 1993).

Previously it has been shown that the largest subunit of RNAPII has a highly conserved acidic domain that shares similarity with the acidic activation domain of VP16 (Xiao, et al., 1994). This domain could also act as a potent activator of transcription when fused to the GAL4 DNA binding domain. These findings together with others suggest an interesting model in which transcription activators could work at both steps, by recruiting the GTFs to the promoter and also displacing the RNAPII to promote transcription initiation and promoter clearance. This is a very plausible model, since the VP16 activation domain has been shown to work at steps following transcription initiation (Tiley, et al., 1992).

Another interesting aspect of how activators might work is by altering the conformation of a transcription factor. TFIIB undergoes a conformational shift upon assembly into the preinitiation complex (Roberts, et al., 1993). The activation domain of VP16 can induce that same conformational change in TFIIB (Roberts, 1994). Therefore, VP16 could promote preinitiation complex assembly by providing TFIIB with the proper conformation to enter the preinitiation complex.

Structure of activation domains

Presently there are no well defined molecular models for the structure of transcriptional activation domains, which prevents us from fully understanding the mechanism of how activators work. Biophysical data, using circular dichroism (CD) and nuclear magnetic resonance (NMR) , found that activation domains are mostly unstructured, though some activation domains can be induced to adopt an α-helical or β-sheet structure (Donaldson & Capone, 1992; Van Hoy, et al., 1993) by altering the pH of the solution. In other words, activation domains do not possess a defined conformation in solution at

physiological conditions. Nevertheless, emerging data from mutational analysis of various activators has provided some clues as to their structure (Triezenberg, 1995) and the function of some activators depends on specific residues (Cress & Triezenberg, 1991b), which indicates some degree of the structural requirement for function.

Recent biophysical studies on the activation domain of VP16 using fluorescence spectroscopy by Shen *et al.* was performed in the context of a GAL4 fusion protein. These studies demonstrated that the activation domain of VP16 is mobile and unstructured in solution (Shen, et al., 1996a). Shen *et al.* also showed that when the target protein, TBP or TFIIB, was added, the VP16 activation domain was restricted from segmental motion. The environment of the fluorescence probe used also changed to a more hydrophobic environment suggesting that there are interactions between the activation domain and its target (Shen, et al., 1996b). Thus transcription activators might form a distinct conformation once they interact with their targets. Co-crystallized complexes of activation domains with their targets could provide information concerning their structure.

A majority of transcription activators are found to comprise multiple domains or subdomains that can function as transcription activators (Triezenberg, 1995). These domains could work completely independently, dependently, or synergistically with each other. Different subdomains may also be functionally distinct from each other. In other words they are not always simply a duplication of one another. This is clearly shown in the case of VP16, which is composed of two distinct activation subdomains. These two subdomains can interact with different components of the transcription machinery (Goodrich, et al., 1993; Ingles, et al., 1991; Lin, et al., 1991; Shen, et al., 1996b) suggesting different mechanisms of activation. Therefore, a single

activator could function at multiple steps in the regulation of transcription.

VP16

Herpes simplex virus type 1 (HSV-1), is one of the most extensively studied viruses in the herpes family. The genome of HSV-1 is a large double-stranded DNA virus of approximately 150 kbp. The HSV-1 genome has been sequenced entirely (McGeoch, et al., 1986; McGeoch, Dalrymple, & Davison, 1988) and most of the genes and the origin of replication have been defined. The viral DNA is packed within an icosahedral capsid surrounded by an amorphous tegument and a spiked envelope derived from the host nuclear membrane (Figure 1). The virion contains at least 33 viral polypeptides which represent approximately one half of the 70 or so genes estimated to be encoded by the HSV-1 genome (Roizman & Sears, 1990).

HSV-1 has been used as a model system to study gene regulation in mammalian cells. HSV-1 gene expression is coordinately regulated in a sequentially ordered manner. Three major groups of genes, Immediate Early (IE) genes, Delayed Early (DE) genes and Late (L) genes, have been characterized based primarily on the order in which they are expressed. The IE genes are expressed early in infection, and reach a peak expression rate at 2-4 hours post infection. Their expression is independent of viral protein synthesis. Five IE genes have been identified in HSV-1; ICP0, ICP4, ICP22, ICP27 and ICP47. These gene products (except ICP47) were found to have regulatory functions influencing the expression of another group of genes. The DE genes are expressed later in infection. They are not expressed in the absence of IE gene expression but can be induced in the absence of viral DNA replication. The expression of DE genes peaks at about 5-7 hours post infection and they have

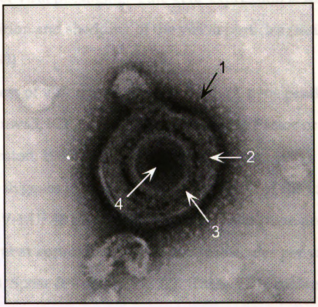


Figure 1. Transmission electron micrograph of HSV-1 (KOS) virion. Virions were purified from infected cells by ultracentrifugation. This sample on a 200 mesh, gilder grid, was stained with 1% ammonium molybdate and observed on a JEM-100CX II transmission electron microscope (JOEL). The arrows point to; (1) viral envelope, (2) viral tegument, (3) viral capsid and (4) viral core, containing the viral genome. Magnification: 58,000X

been found to encode primarily enzymes and proteins needed for viral replication. The last group of genes, L genes, are expressed last in the cascade of HSV-1 gene expression. The true L genes, such as the gC and L42 genes (Homa, et al., 1991), are completely dependent upon viral DNA replication, while expression of other viral L genes is reduced when viral DNA synthesis is inhibited. These genes mostly encode the viral structural proteins and essential proteins needed for viral packaging. Late gene expression is followed by the assembly of the virion and shedding of the virion particles (Roizman & Sears, 1990; Whitley, 1990).

In the process of virion assembly, one of the L gene products termed VP16 [also known as Vmw65, ICP25, and α-Trans Inducing Factor or α-TIF (Campbell, Palfreyman, & Preston, 1984; Post, Mackem, & Roizman, 1981)] that specifically activates the viral IE genes is packed into the tegument of the virion. VP16, the product of the HSV-1 UL48 gene functions as a transcription activator . VP16 proteins are present at approximately 500-1000 molecules per virion particle (Heine, et al., 1974; Spear & Roizman, 1972). These VP16 molecules will activate the viral IE genes and start the cascade of viral gene expression in an infection. VP16 was found to be essential for virion replication. Viral particles bearing genomes that lack the gene encoding VP16 can infect and replicate their DNA in the host cell, but fail to produce intact virions (Weinheimer, et al., 1992). Therefore, VP16 seems to serves two distinct functional roles, as a transcription activator for the viral IE genes and also as a component of the virion essential for particle assembly. The viral infection cycle as well as the order of viral gene expression are summarized in Figure 2.

Deletion mutagenesis has been used to roughly map VP16, a 490 amino acid residue protein, into two domains. The specificity domain which targets VP16 to the IE promoters maps in the N-terminal residues 1-410 (Ace, et al., 1988;

Figure 2. Schematic representation of the replication cycle of HSV-1 in permissive cells. (1): The virus initiates infection by attachment to the cell surface by the viral glycoproteins followed by the fusion of the envelope with the plasma membrane. The fusion of the viral envelope with the cell membrane releases at least two viral proteins into the cytoplasm, viral host shut off protein (vhs) which inhibits protein synthesis in the host cell by promoting the degradation of the polyribosomes, and VP16 which will migrate to the nucleus to activate the expression of HSV-1 IE genes. (2): The viral capsid containing the DNA is then transported to the nucleus, where the viral DNA is released into the nucleus and circularizes. (3): The transcription of the IE gene is induced by VP16. The IE transcripts are translated and the IE proteins are transported back into the nucleus. (4): The presence of the IE gene products induces transcription and translation of the DE genes. The viral DNA is then replicated. (5): L genes are then transcribed and translated. The L proteins consist primarily of structural proteins. (6): The capsid proteins form the viral capsids. The viral DNA is then packed into the empty capsid with other structural proteins. (7): Viral glycoproteins and tegument proteins accumulate and form patches in the nuclear membranes. The capsid containing the DNA and additional proteins attach to the underside of the membrane patch containing the viral protein, and are enveloped. (8): The envelope capsids accumulate in the endoplasmic reticulum and are transported into the extracellular space.

Polyribosomes
vhs
VP16
1
2
IE mRNA
3
IE Proteins
4
DE mRNA
DE Proteins
5
DE
L mRNA
L Proteins
6
7
L
8
IE

Figure 2

Friedman, Triezenberg, & McKnight, 1988; Triezenberg, Kingsbury, & McKnight, 1988a; Werstuck & Capone, 1989b). The activation domain which is both necessary and sufficient for VP16 to activate transcription maps at the C-terminal residues 410-490 (Sadowski, et al., 1988; Triezenberg, et al., 1988a; Werstuck & Capone, 1989b; Cousens, et al., 1989b). The activation domain of VP16 is one of the most potent activation domains tested (Carey, Leatherwood, & Ptashne, 1990; Cousens, et al., 1989; Sadowski, et al., 1988). VP16 does not directly bind DNA (Kristie & Sharp, 1990), but works through interactions with Oct-1 and at least one other host cell protein called HCF, CFF, C1 or VCAF-1 (Katan, et al., 1990; Kristie, LeBowitz, & Sharp, 1989; Xiao & Capone, 1990). VP16 specifically interacts with the POU domain of Oct-1 which binds to a recognition site, the TAATGARAT (where R is a purine) (Kristie & Roizman, 1984; Triezenberg, LaMarco, & McKnight, 1988b) element or the overlapping ATGCTAATGARAT in the HSV-1 IE gene promoters (Cleary & Herr, 1995; Preston, Frame, & Campbell, 1988; Stern & Herr, 1991).

VP16 was categorized as a member of the acidic activator group due to the prominence of acidic residues found in its activation domain. Its activation domain, the 80 C-terminal amino acids of VP16, have been divided into two subdomains, the N-subdomain (residues 410-456) and the C-subdomain (residues 456-490). These two subdomains have distinct primary structures, and evidence suggests that they can function independently and mechanistically differently from each other (Goodrich, et al., 1993; Ingles, et al., 1991; Lin, et al., 1991; Shen, et al., 1996b). Studies on the N-subdomain have shown that acidic residues are necessary, but that hydrophobic residues, especially a phenylalanine residue at position 442 (Phe 442), seem to be the key elements of this activation subdomain (Cress & Triezenberg, 1991b; Regier, et al., 1993). Despite the differences in the function and primary structure of the N and C subdomains,

random polymerase chain reaction and alanine scanning mutagenesis experiments of the C-subdomain, residues 455-487, suggested that bulky hydrophobic residues are also important for activity of the C-subdomain rather than the abundant acidic residues (P. Horn & S. Sullivan, unpublished results). As mentioned earlier, the mechanism of VP16 activation domain function is not totally understood. Nonetheless, data from numerous biochemical, genetic and biophysical studies on VP16 have provided important clues suggesting that the N-subdomain of VP16 activation domain is able to interact with TBP, TFIIB and TFIIH, and the C-subdomain with TAFII40 and TFIIB (Ingles, et al., 1991; Lin, et al., 1991; Goodrich, et al., 1993, Shen, et al., 1996b).

Some structural studies on the activation domain of VP16 have been done. Preliminary studies using ultraviolet spectroscopy, Fourier-transform infrared spectroscopy, and circular dichroism done by Doug Cress using the full length activation domain or N-subdomain of VP16, either in the form of a fusion protein with GAL4 residues 1-147, or the peptide fragment of the activation domain itself, suggest that the activation domain is likely to be a mixture of coil and α-helical structure. Initial hypotheses suggesting that these acidic activators might fold as an amphipathic α-helix or the negative noodle model have been refuted by mutational studies (Cress & Triezenberg, 1991b; Regier, et al., 1993).

Recent work using fluorescence spectroscopy techniques such as dynamic quenching, time-resolved fluorescence decay, and time-dependent anisotropy has been performed by Fan Shen. Chimeric proteins comprising the GAL4 DNA binding domain residues 1-147, fused to VP16 activation domain, in which Trp residues were substituted for Phe at either position 442 or 473 of VP16, providing a unique fluorescence probe, were used. The results suggest that both phenylalanine residues are well exposed and highly mobile in solution (Shen, et al., 1996a). A more ordered structure of the activation domain could be induced

upon addition of TBP and TFIIB (Shen, et al., 1996b).

Data from both Doug Cress' mutagenesis and Fan Shen's fluorescence spectroscopy studies suggest that the activation domain of VP16 is poorly structured in solution and that a more structured form of the activation domain is likely to be induced upon interaction with its target protein.

VP16 Homologues

HSV-1 is a member of the alphaherpesvirinae or the α herpes virus subfamily. The α herpes are classified on the basis of a variable host range, relatively short reproduction cycle, rapid spread in cell culture, efficient destruction of infected cells and finally their ability to establish latent infections primarily in the sensory ganglia (Roizman, 1990). Homologues of VP16 have been identified in other α herpes viruses such as varicella-zoster virus (VZV) (Davison, 1991; McKee, et al., 1990), bovine herpes virus type 1 (BHV-1) (Carpenter & Misra, 1992), equine herpes virus type 1 and 4 (EHV-1 and EHV-4) (Purewal, et al., 1992; Purewal, et al., 1994), and the Marek's disease virus (MDV) (Yanagida, et al., 1993; Koptidesova, et al., 1995; M. Boussaha, Ph.D. thesis, 1996). These VP16 homologue proteins were found to activate transcription of reporter genes linked to α gene promoters (Misra, 1994; Moriuchi, et al., 1993; Purewal, et al., 1994; M. Boussaha, Ph.D. thesis, 1996) similarly to HSV-1 VP16. This finding indicates that all these homologues of VP16 must contain domains responsible for targeting to the promoter and for activation. Significant sequence homology is shared among these VP16 homologues. Attempts were made to identify the activation domains of these proteins by comparing their primary structure with the well characterized HSV-1 VP16 protein. Unfortunately, the amino acid sequence of these proteins varies considerably at the C-terminus. Moreover, the

VZV gene 10 homologue of HSV-1 VP16 lacks the 80 amino acid C-terminus altogether. Deletion mutagenesis has been performed on the different VP16 homologues to map their activation domains. The activation domain of BHV-1 was mapped to the C-terminus of the protein, but the protein retains some transcription activation ability despite missing the C-terminal 87 amino acid residues (Misra, 1994). Activity of EHV-1 gene 12 seems to lie at the C-terminus as well, despite poor homology with HSV-1 VP16. Deletion of the last 7 amino acids abolishes its transcription activation ability and the methionine residue at position 476 is likely to be a critical element for its activation function (Elliott, 1994). These findings suggest that though these proteins share significant homology, their means of transcriptional activation might be different.

Overview

Despite considerable information emerging from mutational, biochemical and biophysical studies on VP16 and its activation domain, very little is known regarding the function of VP16 in viral lytic infection. VP16 plays two major roles in viral infection, as a transcription activator for the viral IE genes and a structural protein essential for assembly and maturation of the viral particle (Ace, et al., 1988; Ace, McKee, Ryan, Cameron, & Preston, 1989; Campbell, et al., 1984; Weinheimer, et al., 1992). VP16 has recently been found to interact with other components in the tegument, the viral host shut off protein (Simibert, et al., 1994) and VP22 (Elliott, et al., 1995), which might be important in the maturation or disassembly of the particle during infection. Therefore it is becoming increasingly interesting to know how different mutations in VP16 would effect HSV-1 replication within the context of the viral genome.

The significance of the transcriptional activation function of VP16 in lytic infection was assessed by constructing recombinant viruses bearing mutant VP16 genes, truncated in their activation domains. Viruses bearing either of the two subdomains (N or C subdomains) are weakened in lytic growth. Viruses lacking the activation domain altogether are viable in culture, but significantly debilitated. These results are described in chapter II.

Despite the available amino acid sequences of VP16 homologues from other α herpes viruses, the activation domains of other VP16 homologues remain poorly defined. Linear sequence comparison does not show any obvious homology between the activation domain of VP16 (C-terminal 80 amino acids) and the C-terminal portion of the other VP16 homologues. However, these homologues were found to be able to activate transcription from their α promoters, which indicates that they should possess a transcription activation domain(s).

In chapter III, I have examined these sequences using a technique termed Hydrophobic Cluster Analysis (HCA) (Lemesle-Varloot, et al., 1990; Woodcock, Mornon, & Henrissat, 1992). This analysis presents the amino acid sequence of proteins in a two dimensional α helical representation. The similarity is based on comparing conserved clusters of hydrophobic residues presumed to constitute the hydrophobic core of globular proteins rather than the similarities in the amino acid sequences (Lemesle-Varloot, et al., 1990; Woodcock, et al., 1992). I propose that these herpes viruses do share similar transcription activation domains. Interestingly, with the exception of herpes simplex virus type 2 (HSV-2), the position of the N-subdomain in the homologue protein from BHV-1, VZV, EHV-1, EHV-4 and MDV all reside in the N-terminal portion of the protein. The C-subdomain, however, still residues at the C-terminus, with the exception of VZV gene 10 and MDV UL48 which lack the C-subdomain altogether. This

hypothesis is supported by introducing single amino acid substitutions in the proposed activation domain of VZV gene 10, replacing the critical phenylalanine residue at position 28 (Phe28) and other hydrophobic residues in the vicinity thereof. This result has been published in the Proceedings of the National Academy of Sciences USA (1995, Vol. 92, pp. 9333-9337). At the end of chapter III, I have used several other comparison and prediction methods recently become available to predict the secondary structure of the N-subdomain of VP16. The profile network system from Heidelberg program (PHD) (Rost & Sander, 1994) and Sequery (Collawn, et al., 1990; Collawn, et al., 1991), were used in this study.

CHAPTER II

EFFECTS OF C-TERMINAL DELETION MUTATIONS IN VP16 ON THE LYTIC INFECTION OF HSV-1 KOS

Introduction

To understand the structural and functional role of VP16 in the lytic infection of HSV-1, the gene encoding VP16 has been identified and cloned from the viral genome (Campbell, et al., 1984). VP16 binds to its responsive element through interactions with at least two host cell factors including Oct-1 and HCF (Katan, et al., 1990; Kristie, et al., 1989; Xiao & Capone, 1990). The DNA element necessary for binding of VP16 and host cell factors of VP16 has been defined (Bzik & Preston, 1986; Cordingley, Campbell, & Preston, 1983; Kristie & Roizman, 1984; Mackem & Roizman, 1982a; Mackem & Roizman, 1982b). The functional dissection of VP16 has been performed using short-term transfection assays. In these assays plasmids encoding the native or mutated derivative of VP16 were co-transfected into host cells together with a reporter plasmid containing the HSV-1 α promoter linked to a CAT or thimidine kinase gene (Ace, et al., 1988; Ace, et al., 1989; Campbell, et al., 1984; Triezenberg, et al., 1988a).

Deletion mutations, linker insertions, and point mutations of VP16 have been generated and studied (Ace, et al., 1988; Cress & Triezenberg, 1991b; Regier, et al., 1993; Triezenberg, et al., 1988a; Peter Horn and Susan Sullivan unpublished

results). Based on these data, the activation domain of VP16, residues 419-490, is divided into two subdomains: the N-subdomain, residues 410 -456; and the C-subdomain, residues 456-490. Either of these domains is sufficient to activate transcription in the context of a GAL4-VP16 fusion protein, but both domains are necessary for full activation activity (Cress & Triezenberg, 1991b; Regier, et al., 1993; Triezenberg, et al., 1988a; P. Horn and S. Sullivan unpublished results). Walker et al. demonstrated that the C-subdomain of the VP16 activation domain in the context of a GAL4-VP16 fusion protein is critical for its activity on a promoter containing a single GAL4 binding site. Deletion of the C-subdomain or point mutations of bulky hydrophobic residues in the C-subdomain abolish activity of the GAL4-VP16 fusion protein. The reduction of the activity could be complemented by using a promoter with multiple GAL4 binding sites and duplication of the N-subdomain, but not the C-subdomain (Walker, et al., 1993). This indicates that the N and C subdomains are not simply redundant subdomains. The N and C subdomains seem to have distinct primary structures and biochemical properties and may serve different roles in transcription activation (Goodrich, et al., 1993; Ingles, et al., 1991; Shen, et al., 1996b; Xiao, et al., 1994; P. Horn and S. Berger unpublished results).

Very little is known about how mutations in VP16 would affect its function as a component of the virion and the overall lytic infection. Earlier work demonstrated that VP16 has at least two functional roles in the virus life cycle, activation of IE genes and a structural role as a component of the virion particles (Ace, et al., 1988; Ace, et al., 1989; Campbell, et al., 1984; Spear & Roizman, 1972). To study the effect of VP16 mutations on HSV-1 infection, recombinant HSV-1 bearing mutations of the VP16 gene were generated.

Ace *et al.* introduced a twelve nucleotide insertion in the VP16 gene, producing a HSV-1 bearing VP16 with a four amino acid insertion at residue 379.

This HSV-1 strain, *in1814*, is defective in infection and expression of IE genes (Ace, et al., 1989). In the Ace *et al.* study, *in1814* contains the entire VP16 coding sequence, but bears an insertion which renders the VP16 protein defective in complex formation on IE promoters. This was illustrated by a gel mobility shift using a IE promoter responsive DNA sequence. The *in*1814 virus was defective in infection and produced viruses with a high particle per pfu ratio. However, at high multiplicity of infection (100-1000 pfu/cell) the expression of HSV-1 genes was not affected. This was determined by measuring the level of thymidine kinase gene expression.

Care should be taken in interpretation of these data. Though the *in*1814 mutant VP16 was inactive in the transient expression assay at low moi, stimulation by wild type VP16 was only 5-10 fold in these experiments . Residual activity of the mutant VP16 could easily be below the detectable levels under such circumstances. Detection of mutant VP16 participating in complex formation on IE promoters is also limited by the sensitivity of the assay. At a high multiplicity of infection, approximately 10^5-10^6 molecules of VP16 are being introduced into cells. In this case the cumulative effect of residual activity might be sufficient to allow normal levels of gene expression in the *in*1814 infection.

Poon *et al.* constructed point mutations substituting various cysteine residues with glycine residues at different positions in VP16, generating temperature sensitive mutations of VP16 as well as recombinant viruses bearing those mutations (Poon & Roizman, 1995). The temperature sensitive viruses bearing the mutant VP16 generated by Poon *et al.* yielded very low titers at non-permissive temperatures. This effect could either be due to the inability of mutated VP16 to initiate infection by activating IE genes or a defect in the maturation of the virus. The particle per pfu ratio of these viruses was not determined under non-permissive conditions, but it is likely that the overall

production of viral particles would dramatically decrease. At non-permissive temperatures the mutated VP16 proteins are likely to be improperly folded, resulting in proteins that could not function in the maturation process. Results from Poon *et al.* indicated that these viruses are affected late in infection implying that these viruses are defective in maturation of the particles. The effect of these VP16 mutations on IE gene expression was only tested on the ICP4 promoter. A slight reduction of transcription was observed for the virus bearing mutant VP16 containing substitutions at Cys120 and Cys176 (Poon & Roizman, 1995).

An experiment where the gene encoding VP16 was deleted altogether was also performed (Weinheimer, et al., 1992). The 8MA virus was incapable of replicating in normal cell lines. Instead, a complementing cell line constitutively producing VP16, Vero16-8, was used to replicate this virus. This demonstrates that although these viruses carry the VP16 (produced from the Vero16-8 cell they were replicating in) they cannot complete their lytic infection cycle since they lack the capability to produce new VP16 protein during the late stages of infection. This result reinforced the essential role of VP16 in maturation of the viral particle and as a structural component of the virion.

Changes made in VP16 could affect either the maturation or IE activation of the virus. To fully understand the role of VP16 in gene expression within the context of the viral genome, we must be able to separate the role of VP16 in transactivation of the viral IE genes from its role in maturation as a structural component of the virus. Though Ace *et al.* attempted to use insertion mutagenesis to prevent VP16 from activating IE genes, there are still concerns about the residual activity of mutant VP16. Work by Poon *et al.* did not provide any new insights concerning the role of VP16 in activation of the IE genes or maturation of the virus, though, the temperature sensitive viruses may prove to be very useful in future experiments.

Earlier work in our laboratory showed that the last 80 amino acids of VP16 are necessary and sufficient for VP16 activation of IE genes. The truncated VP16 protein, 1-410, though unable to activate transcription, can form a complex with host proteins on the IE promoters (Triezenberg et al., 1988a; Werstuck & Capone, 1989a; Greaves & O'Hare, 1989). If the truncated VP16 could also support the structural role of VP16 in the virion, it would be possible to produce mutant HSV-1 viruses bearing VP16 proteins that are only defective in their activation function. These viruses could then be used to study the specific effect of IE gene expression on replication of the virus. Moreover, the function of the different subdomains of the VP16 activation domain could be determined by using different VP16 deletions.

In this chapter of my thesis I have shown that recombinant viruses carrying deletion mutations in the activation domain of VP16 could be generated, and demonstrated that these viruses are viable in cell culture. I have used the viral strain 8MA, in which the gene encoding VP16 has been replaced by the gene encoding for β-galactosidase (Weinheimer, et al., 1992), as a parental strain to generate recombinant viruses via homologous recombination (Figure 2). The 8MA strain can only replicate in a complementing cell line (Vero16-8) that constitutively expresses VP16, Vero16-8 cells (Weinheimer, et al., 1992). The 8MA system provides a powerful screen and selection for the recombinant viruses.

I have studied the effects of a set of deletion mutations of the VP16 activation domain which lack either the N or C subdomain, or the whole activation domain, upon HSV-1 lytic infection. The ability to form plaques, plaque characteristics, single step replication, and some aspects of the effect of mutant VP16 on structure and IE gene activation were studied.

Methods

Cell line and viral strains

Vero cells were grown in Dulbecco's modified Eagle's medium (GIBCO BRL) (DMEM) supplemented with 10% Fetal calf serum unless indicated otherwise. Vero16-8 cells (Weinheimer, et al., 1992) were grown in DMEM supplemented with 10% Fetal calf serum and 450 µg/ml G418.

8MA virus, kindly provided by Steve Weinheimer, was propagated on Vero16-8 cells. HSV-1 KOS and recombinant viruses were propagated on Vero cells unless indicated otherwise. Viral stocks were prepared by infecting host cells at 0.01 pfu/cell and harvesting cells and media after an extensive cytopathic effect was apparent. The cells were sonicated in a TC75 tissue culture flask (Corning) to disrupt the cell membrane and release the virions. The cell debris were removed by centrifugation and the media containing virions were aliquoted and frozen at -80° C.

Plasmids

Plasmid pKOS-VP16 (Weinheimer, et al., 1992) was digested with restriction endonuclease *Eco*RI and *Xba*I, then re-ligated, thus removing *Sac*I, *Kpn*I, *Sma*I, and *Bam*HI restriction endonuclease sites in the multiple cloning region of the plasmid. A *Bam*HI site was created 19 base pair downstream from the VP16 TAG stop codon, using site directed mutagenesis with the oligonucleotide 5'-CCGGACCCGGATCCCCCGT-3' (ST35). Double-stranded oligonucleotide encoding stop codons in all three reading frame, generated by annealing the oligonucleotide 5'-GATCCTAGTTAATTGA-3' and 5'-

GATCTCAATTAACTAG-3', was inserted at the *Bam*HI site, regenerating the *Bam*HI site at the open reading frame proximal end of the oligonucleotide, resulting in pKOS-VP16.2 (Figure 1).

A series of pKOS-VP16.2 derivatives were constructed as follows. pKOS-VP16.2 was digested with *Eco*RV and *Bam*HI to remove the wild type VP16 gene. Truncated VP16 gene fragments were isolated from VP16 expression plasmids, pMSVP16Δ456 (VP16 with deletion of codons 456-490), pMSVP16ΔSmaI (VP16 with deletion of codons 413-452), and pMSVP163δ50N (VP16 with deletion of codons 413-490) with *Eco*RV-*Bam*HI digestion and subcloned into pKOS-VP16.2. Fragments were subcloned into the prepared vector resulting in pKOS-VP16Δ456, pKOS-VP16ΔSmaI, and pKOS-VP163δ50N.

A derivative of pKOS-VP16.2 containing the VP16 gene from HSV-2 (VP16II) (Cress & Triezenberg, 1991a) was constructed. pMSVP16-2(ΔE11), an expression plasmid containing the VP16II gene (A. Cress unpublished results), was digested with *Xho*I, treated with T4 DNA polymerase to create blunt ends, and digested with *Bam*HI. The *Xho*I-*Bam*HI fragment from pMSVP16-2(ΔE11), containing the VP16II gene fragment with 33 nucleotides upstream of the start codon and 45 nucleotides from the stop codon, was then subcloned into pKOSVP16.2 replacing the *Eco*RV-*Bam*HI VP16 gene.

Viral DNA extraction and purification

Vero cells were grown in TC120 tissue culture flasks (Corning) to 80% confluency and inoculated with virions at 0.01 multiplicity of infection (moi). Cells were incubated at 37° C until 90% of the cells showed cytopathic effects (CPE). The cells were scraped, transfered to a TC75 tissue culture flask and sonicated, with 30 second bursts three times using a cup horn, to release the

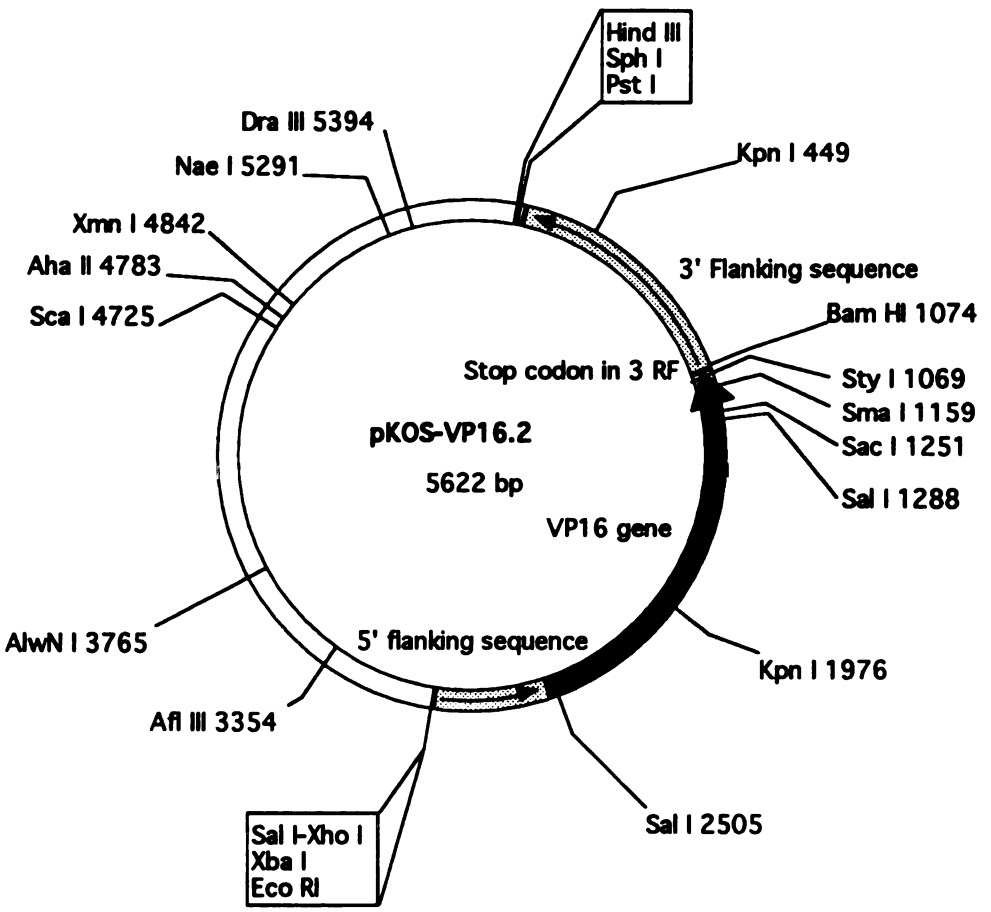


Figure 1. Plasmid map of pKOSVP16.2. The pKOSVP16 (Weinheimer, et al., 1992) was modified by removing restriction endonuclease sites between the *Eco*RI and *Xba*I site. Mutation of a single nucleotide in the 3' untranslated region creates a unique *Bam*HI site 19 base pairs from the stop codon of the VP16 gene. An oligonucleotide containing stop codons in all three reading frames was then inserted at the *Bam*HI site providing a stop codon for the carboxy-terminal deletion derivatives of VP16, which were subcloned into pKOSVP16.2 replacing the wild type VP16 gene (described in detail in the methods section).

virion particles. The sample was separated by centrifugation at 12K rpm in SS34 rotor, at 4° C for 30 minutes. The supernatant which contains the virion particles was removed, overlayed on 10 ml of 20% glycerol in PBSa (137 mM NaCl, 10 mM Na_2HPO_4, 1.5 mM KH_2PO_4, and 2.7 mM KCl) cushion and centrifuged at 25K rpm in SW28 rotor, at 4° C for 90 min. The supernatant was discarded and the virion particle pellet was resuspended in 2 ml of virion lysis buffer (100 mM NaCl, 10 mM EDTA, 50 mM Tris-HCl pH 8.0, and 0.5% lauryl sarcosinate). The sample was treated with 250 µl proteinase K (2 mg/ml) and incubated at 37° C for 1 hour twice. The DNA was then isolated by ultracentrifugation. Equal volumes of sample and cesium trifluoroacetate (2 g/ml, CsTFA™, Pharmacia LKB) were combined and mixed well, and 25 µl of ethidium bromide (10 mg/ml) was added. The sample was transferred into 5 ml ultracentrifuge tube, sealed and centrifuged at 45K rpm in VTi 65.2 rotor, at 15° C overnight. The purified DNA sample was removed, extracted with isoamyl alcohol to remove ethidium bromide, and dialyzed in TE (10 mM Tris-HCl, 1 mM EDTA) overnight. The DNA concentration was determined by absorbance at 260 nm.

Generating Recombinant viruses

pKOS-VP16.2, pKOS-VP16II, pKOS-VP16Δ456, pKOS-VP16ΔSmaI, and pKOS-VP163δ50N were digested with *Eco*RI and *Pst*I to liberate the VP16 gene fragment (Figure 2). Vero cells were seeded in P60 tissue culture plates at $5.0x10^5$ cells per plate. Each plate was transfected with a mixture of 5 µg of 8MA viral genomic DNA and 5 µg of digested pKOS-VP16.2 or one of its derivatives. The DNA mixture was transfected via $CaPO_4$ DNA co-precipitation (Graham & Eb., 1973). The cells were then overlaid with DMEM (GIBCO, BRL) supplemented with 5% fetal calf serum (FCS, Hyclone Laboratories) and 1% low melting-point

Figure 2. Fragments containing the VP16 gene and derivatives used to generate recombinant viruses.. The viral strain and the fragments used for the construction of the recombinant viruses are shown. The HSV-1 genome and the *Bam*HI fragment containing the VP16 gene and the approximate location of the VP16 gene in the HSV-1 genome are shown at the top of the panel. The fragments used to generate the recombinant viral strains RP1, RP2, RP3, RP4 and RP5 from pKOSVP16.2 (wild type VP16), pKOSVP16Δ456 (VP16Δ456-490), pKOSVP16ΔSmaI (VP16Δ413-452), pKOSVP163δ50N (VP16Δ413-490) and pKOSVP16II (VP16 gene from HSV-2) are shown in the lower half of the panel. The viral strains produced are listed at the right hand side of each DNA fragment.

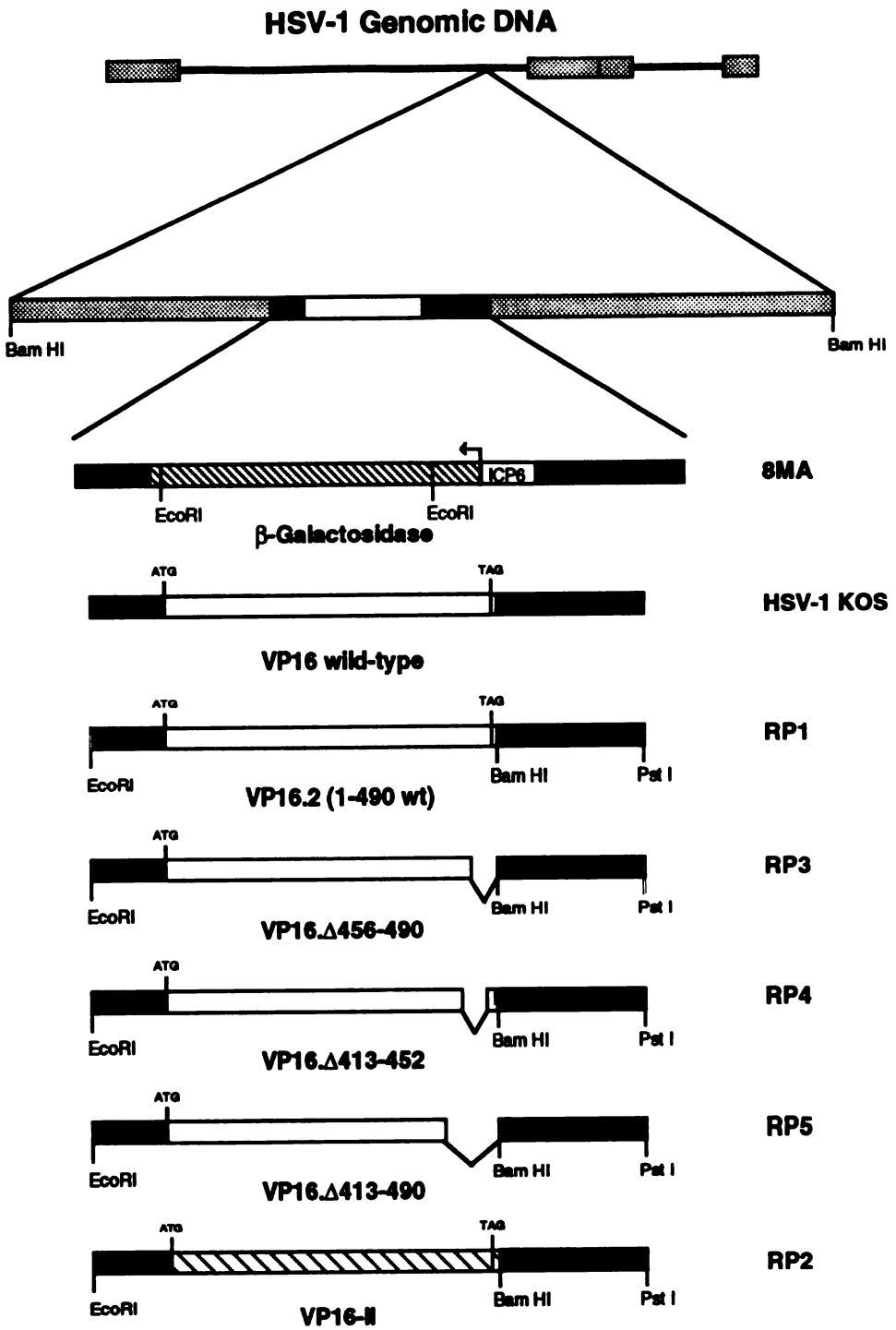

HSV-1 Genomic DNA
Bam HI
Bam HI
ICP6
EcoRI
EcoRI
β-Galactosidase
8MA
ATG
TAG
VP16 wild-type
HSV-1 KOS
ATG
TAG
EcoRI
Bam HI
Pst I
VP16.2 (1-490 wt)
RP1
ATG
EcoRI
Bam HI
Pst I
VP16.Δ456-490
RP3
ATG
EcoRI
Bam HI
Pst I
VP16.Δ413-452
RP4
ATG
EcoRI
Bam HI
Pst I
VP16.Δ413-490
RP5
ATG
TAG
EcoRI
Bam HI
Pst I
VP16-II
RP2

Figure 2

agarose (Sea Plaque, FMC), and incubated at 37° C for 3-4 days or until plaques formed. At least three plaques from each plate were picked and subsequently plaque purified twice. DNA preparations of each were analyzed by Southern blot to verify that the isolated viruses contain the proper form of the VP16 gene. Primary stocks were prepared and titered to determine the plaque forming unit (pfu).

Revertant virus strains were generated by co-transfecting Vero cells with DNA from each mutant virus and a wild type VP16 DNA fragment from pKOS-VP16.2. Three plates were transfected for each DNA mixture. Plates were overlaid with DMEM supplemented with 5% FCS without agarose. After plaques formed (3-4 days) 500 µl of media was removed from each plate, diluted 100 fold, and inoculated on Vero cells at 80-90% confluency. These plates were overlaid with DMEM supplemented with 5% FCS and 1% agarose and incubated at 37° C. Plaques which showed a revertant phenotype (plaque formation within 2 days and large plaque size) were picked and plaque purified twice. The presence of the wild type VP16 gene in the revertants was verified by Southern bolt analysis. Viral stocks were generated and titered.

Southern blot

Recombinant viral DNA (approximately 1 µg) was digested to completion with *Bam*HI and separated by gel electrophoresis in 1% agarose using TAE buffer. The DNA was stained with ethidium bromide, denatured, and transfered to nitrocellulose membrane using standard methods. The *Eco*RI-*Pst*I VP16 gene fragment from pKOS-VP16.2 was used as a probe. Random primed DNA labeling kit (Boehringer Mannheim Biochemica) and α-[^{32}P] dCTP (Amersham Life Science, 3000Ci/mmol) was used. The nitrocellulose blot was pre-

hybridized for two hours in 5X Denhardt's solution (Sambrook, et al., 1989), 6X SSC (Sambrook, et al., 1989), 0.5% SDS, and 100 μg/ml salmon sperm DNA at 65° C. The synthesized probe was added and incubated overnight at 65° C. The blot was subsequently washed twice in 2X SSC, 0.5% SDS at 67° C, twice in 0.5X SSC, 0.5% SDS at 67° C, and twice in 0.2X SSC, 0.5% SDS at 67° C or until the background was eliminated. After the blot was exposed to X-ray (X-OMAT, AR, Kodak) film for 3-4 hours, the film was developed using standard methods.

Western blot

Virions were inoculated at 0.1 moi on Vero or Vero16-8 cells grown in TC120 tissue culture flasks to 80-90% confluency. The infected cells were incubated until 90% of the cells showed CPE. The virions were harvested by sonication and centrifugation as described above. Intact virions were purified by centrifuging the particles through a 20% glycerol PBSa cushion at 25K rpm in a SW28 rotor, at 4° C for 90 minutes. Pelleted virions were resuspended in 100 μl PBSa. 10 μl of each purified virion sample was solubilized in SDS sample loading buffer by heating the samples to 100° C for 10 minutes, and separated by SDS PAGE in a 12% gel. The proteins were transferred to a nitrocellulose membrane using standard protocols. Western blot analysis was performed using the biotin/avidin system, Vectastain® ABC kit, Vector laboratories, and the anti-VP16 polyclonal antibody, C-8.

Single step replication assay

Vero cells were seeded into a six-well tissue culture plate at $3.5x10^5$ cells per well and incubated overnight at 37° C. Each plate was washed with DMEM

without fetal calf serum. The number of cells per plate was determined, and the cells were subsequently infected with 5-10 pfu/cell of RP1, RP3 or RP4 and overlayed with 3 mls of DMEM supplemented with 5% FCS. Infected cells from each well were scraped and collected from each plate. At 4, 8, 12, 18, 24, and 36 hours post infection (hpi.) the infected cells were scraped and transfered into 15 ml tubes, and frozen at -70° C for later titering. The infected cells were thawed and sonicated with three 30 second bursts to release the virion particles. The titer of each sample was then determined on Vero cells. The titer of the virus from each strain was plotted against time, hpi.. The single step replication curve of each strain was determined at least three times.

Penetration curve

Each strain of recombinant virus, RP1, RP3, RP4 and RP5, was inoculated, 100 pfu per plate, on fifteen P60 plates of Vero cells at approximately 90% confluency. Three plates from each strain were removed at 15 minutes intervals for a period of 1 hour and washed once with 2 mls of DMEM without fetal calf serum followed by 2 mls of Glycine saline buffer (8 g of NaCl, 0.38 g of KCl, 0.10 g of $MgCl_2 \cdot 6H_2O$, 0.10 g of $CaCl_2 \cdot 2H_2O$ and 7.5 g of glycine per liter, pH adjusted to 3.0), followed by two subsequent washes with DMEM without FCS. The plates were overlayed with DMEM supplemented with 5% FCS and 1% low melting point agarose, and incubated until plaques were formed. The agarose was then removed from each plate and the cells were stained with 1% methylene blue in 70% isopropanol. The number of plaques in each plate was counted, recorded and plotted against the time allowed for absorption.

Determination of particle per pfu/ml ratio

Virus stocks were generated by inoculating Vero cells, approximately 80% confluent, in TC120 tissue culture flasks at 0.01 pfu/cell. The inoculum was removed after 1 hour. Cells were washed with DMEM without FCS then overlayed with DMEM supplemented with 2% FCS. The cells were incubated at 37° C until 90% of the cells show CPE. The cells were scraped, spun down, and resuspended in 10 mls of DMEM. The cell suspension was sonicated to release the virion particles. Cell debris were removed by centrifugation at 4,000 rpm in a table top centrifuge. The supernatant was frozen in 1 ml aliquots and used for titering and particle count. The number of particles per ml was determined by transmission electron microscopy, using standard latex bead (Latex spheres of known concentration, 0.176 μm, Ernest F. Fullam, Inc.), and negative staining with 1% ammonium molybdate. The number of latex beads and virions from 20 viewing screen were recorded from each quarter of the grid (total of 80 samplings per stock). The numbers of latex beads and virions were used to calculate the number of virion particles in each viral stock. The titer and number of particles in each stock was then used to calculate the particle per plaque forming unit ratio.

Southern blot of infected cell nuclei

Vero cells in P100 plates were infected, in triplicate, with the recombinant viruses. Cells were infected with equivalent amounts of particle, equal to the amount of particles present in 10 moi of wild type RP1 strain. At 3 hours post infection, nuclei of the infected cells were harvested. The cells were washed once with 5 mls PBSa, removed from the plates by adding 5 ml of ice cold RBS

solution (10 mM NaCl, 10 mM Tris-HCl pH 7.4, and 3 mM $MgCl_2$) supplemented with 0.1% (v/v) Nonidet P-40 and harvested by scraping the cells with a rubber policeman. The cell suspension was collected and homogenized, 3 strokes, in a ice-chilled glass homogenizer. The samples were transferred into 10 ml tubes and chilled on ice. Nuclei were collected by centrifugation in a table top centrifuge at 1,000 rpm for 10 minutes. The integrity of the nuclei was checked using a phase-contrast microscope.

The nuclei were then resuspended in 500 µl of a buffer containing 100 mM Tris-HCl, 10 mM EDTA, and 0.1% SDS, and were treated with proteinase K (100 µg per ml, pH 7.5) at 37° C for 2 hours. The sample was treated with 50 µl of 3 M sodium acetate pH 5.3 and extracted twice with phenol:chloroform (1:1) saturated with TE. The aqueous phase was transferred to a new tube and 2.5 volumes of ethanol were added. The DNA was then precipitated by centrifugation and resuspended in 100 µl of TE. Approximately 1 µg of DNA (25 µl) of each sample was cut with 5 units of *Bam*HI and separated on a 1% agarose gel and subjected to Southern blot analysis, probing with the *Eco*RI-*Pst*I fragment of pKOSVP16.2 containing the VP16 gene.

Results

Isolation and identification of recombinant HSV-1

Homologous recombination using 8MA virus in Vero cells was performed to generate recombinant HSV-1. Plaques were isolated and analyzed by Southern blot analysis to verify that the recombinant viruses had acquired the appropriate wild type or mutated VP16 gene.

The wild type HSV-1 KOS parental strain cut with *Bam*HI and probed with the VP16 gene fragment produces a single 8.5 kbp band. In contrast the recombinant viral DNA cut with *Bam*HI and probed in the same manner would yield two distinct bands, since all the recombinant viruses contain an additional *Bam*HI site within the fragment containing the VP16 gene.

Two distinct bands were observed when the *Bam*HI digested recombinant viral DNA was probed with the *Eco*RI-*Pst*I VP16 gene fragment from pKOSVP16.2 (one of approximately 4.5 kbp and another of 4.0-3.7 kbp band reflecting the different deletion mutations in the VP16 gene) replacing the single 11.8 kbp fragment of 8MA parental strain, as shown in Figure 3. The slightly larger size of the *Bam*HI fragment from 8MA is due to the larger size of the lacZ gene compared to the VP16 gene in HSV-1 KOS.

Because the difference in size of the fragments from strains RP1 and RP4 was not clearly resolved, a second Southern blot was performed using *Sac*I digestion (data not shown). This method clearly showed that RP4 contains the appropriate VP16 mutation.

Five recombinant viral strains were isolated. RP1 and RP2 express wild type VP16 from HSV-1 and HSV-2, respectively. Viral strains RP3, RP4 and RP5 express truncated mutant VP16 with deletions of residues 456-490, residues 413-452, and residues 413-490, respectively.

To confirm that these recombinant viruses express and contain the mutant VP16 molecules in their particle, purified virions were subjected to Western blot analysis. Polyclonal antibody against VP16 was used to probe for VP16. An approximately 65 kD polypeptide corresponding to full length VP16 protein was detected for the control HSV-1 KOS as well as RP1 and RP2 (Figure 4). RP3, RP4 and RP5 particles contained a slightly faster migrating VP16 protein as expected from the truncated genes.

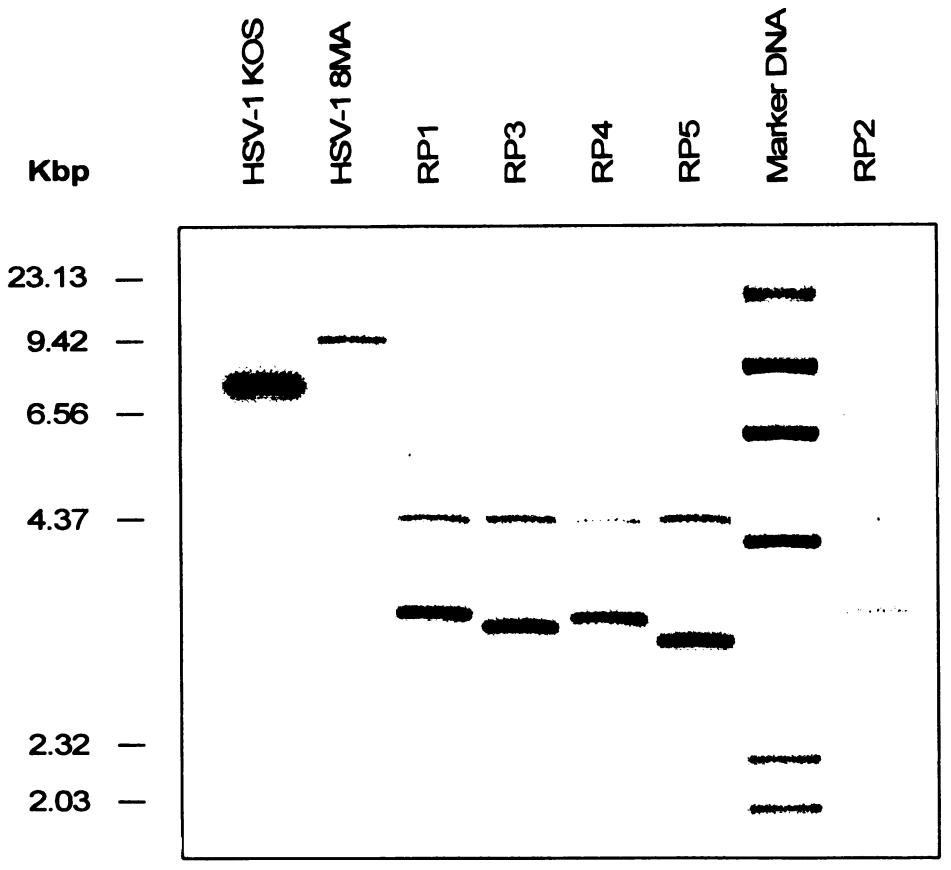


Figure 3. Southern blot analysis of recombinant viral DNA. The recombinant viruses were isolated, purified, their DNA extracted, and analyzed by Southern blot analysis to confirm that they have acquired the proper VP16 gene construct. The viral DNA was cut with *Bam*HI to completion and separated on a 1% agarose gel. The DNA was then transferred to a nitrocellulose membrane and probed for the VP16 gene as described in detail in the methods section.

Southern and Western blot analyses clearly confirm that each of the recombinant viruses contain and express the appropriate mutant VP16 gene and protein in their particles. These results indicate that virus strains carrying a deletion mutation in the activation domain of VP16 could be generated. The complementing cell line was not required for generation or isolation of the viruses, suggesting that the truncated VP16 was sufficient to support maturation and the structural role of VP16 in the virion particle.

Another interesting question is whether or not the virus has a preference for packaging of wild type over truncated versions of VP16 into the particles. If the truncated VP16 was unable to fully function in particle assembly or maturation, we might see some preference for packaging the wild type protein over the truncated form. I have chosen RP5 for this experiment since it lacks the entire 80 amino acid portion of the activation domain. If there is any defect in the packaging of truncated VP16 into the particle at all, RP5 would be the most likely candidate to reveal the defect.

A Western blot of RP5 virions produced in Vero 16-8 cells was performed. Two bands of VP16 were observed, one migrating similarly to wild type VP16 protein, and the other at a lower molecular weight, similar to the truncated VP16 protein carried by RP5 particles (Figure 4). This result demonstrates that when RP5 was grown in Vero16-8 cells, the virions contain comparable amounts of both the wild type and the truncated versions of VP16.

To be confident that the signals from VP16 protein seen on the Western blot reflects the VP16 proteins that were packaged in the virion, not contamination from host cells, RP5 particles produced from Vero16-8 cells were also purified from the media and subjected to Western blot analysis. This would eliminate any contamination of the wild type VP16 protein from the host cells. When the signals from the Western blot were compared between both virion

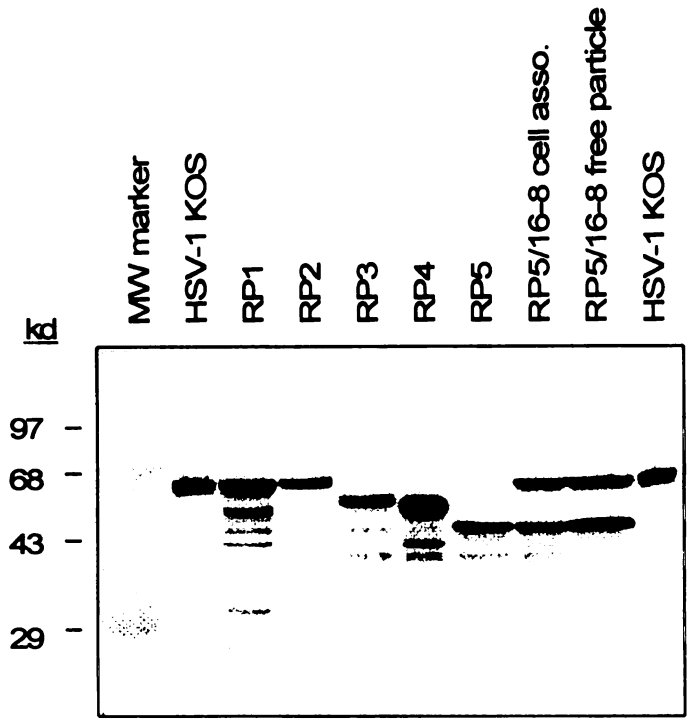


Figure 4. Western blot analysis of recombinant viral particles. Viral particles were purified from recombinant viral strains RP1, RP2, RP3, RP4 and RP5. RP5 was grown both in Vero and Vero16-8 cells to study the packaging preference of full length versus truncated VP16 protein into the viral particle, as described in the methods section. The viral particles were subjected to SDS PAGE and Western blot analysis. The results demonstrate that the recombinant viral particles contain the proper form of VP16.

preparations, similar ratios of the wild type and truncated form of VP16 were found (Figure 4).

To further confirm that the ratio of the signals observed from the Western blot reflects the accurate ratio of the two forms of VP16 protein, lower amounts (10 5, and 2.5 µl) of purified virions were loaded on the SDS PAGE, 10, 5 and 2.5 µl, for Western blot analysis, as shown in Figure 5. The results demonstrated that the signal we observed on the previous gel was in linear range, reflecting the accurate ratio of the two proteins in the particles.

These results demonstrated that there is no preference for packaging of the wild type VP16 protein over the truncated form in the process of virion assembly. They also suggest that the truncated form of VP16 is sufficient for viral maturation and structural requirement. Therefore any effects of these mutations on viral lytic infection should reflect defects in early events of infection.

The VP16 homologue from HSV-2 was previously characterized in our laboratory. Since VP16 from HSV-1 and HSV-2 share very high homology, it is interesting whether or not VP16 from HSV-2 can fully complement HSV-1 growth. To answer this question, I have recombined the gene encoding VP16 from HSV-2 (VP16II) into 8MA, resulting in viral strain RP2. RP2 forms plaques within two days of inoculation and gives rise to similar size plaques as RP1. Slightly different plaque characteristics can be observed under the phase-contrast microscope in cells infected with RP2. The cells in RP2 plaques do not show as extensive CPE as in RP1 plaques. These data demonstrated that VP16II can complement HSV-1 for growth on cell culture. This indicates that VP16II shares similar functional properties to VP16 of HSV-1 and that these functions are conserved between HSV-1 and HSV-2.

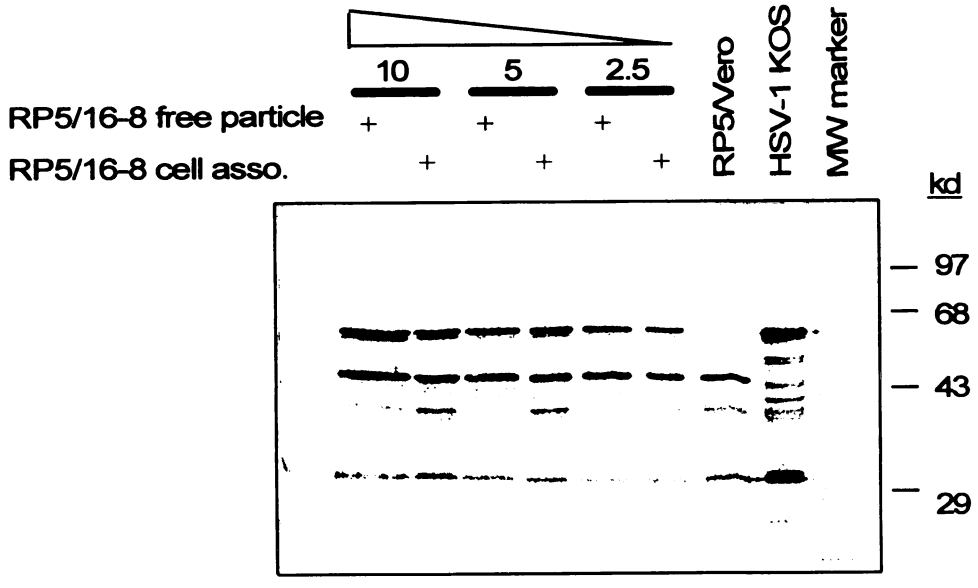


Figure 5. Western blot analysis of RP5 particles generated in Vero16-8 cells. RP5 particles were purified either from the media (free particle) or cell associated viruses (cell asso.) as described in the methods section. The purified particles were then subjected to SDS PAGE followed by Western blot analysis. The number above the bar indicates the volume (in microliters) of sample loaded in the lanes below. The "+" indicates the sample loaded, "cell asso." or "cell free".

Plaque and growth characteristics of recombinant HSV-1

We expected that recombinant viruses carrying the different truncated derivatives of the VP16 activation domain would demonstrate different growth characteristics, since activation of the viral IE genes is one of the most important steps in the replication of HSV-1. RP1 which bears wild type VP16 should replicate similarly to the wild type virus, HSV-1 KOS. RP3 and RP4 should replicate slightly slower since they contain only one of the activation subdomains. I expected RP3 to grow slightly better than RP4 since it contains the N-subdomain which was shown to be a stronger activator than the C-subdomain, in mammalian cells (Walker, Greaves, & O'Hare, 1993). I expected RP5 to grow very slowly since RP5 carries the VP16 gene that lacks the entire activation domain, a mutant which should abolish activation of IE genes entirely.

These recombinant viral strains produced different plaque characteristics on Vero cells. RP1, RP3 and RP4 formed visible plaques within 2-3 days post infection, varying in size from the larger RP1 plaques to smaller plaques for RP3 and RP4. RP5 did not form visible plaques until 5-7 post infection and the plaques were still very small compared to the other recombinant viruses, as shown in Figure 6.

The titer of the virus is dramatically affected by the VP16 mutations. For RP4 the titer was reduced approximately 10 fold and RP5 approximately 100 fold, as shown in Table 1. To determine whether the lower titers of RP4 and RP5 are due to inability to produce intact particles, the particles per pfu ratio of each strain was determined. The results in Table 1 suggest that RP4 and RP5 were able to produce as many particles as RP1 and RP3. However the particle per pfu ratio of RP4 and RP5 were approximately 36 fold and 120 fold higher, respectively, than that of RP1 and RP3. This suggested that the low titers

Figure 6. Plaque characteristics of RP1, RP3, RP4, and RP5 on Vero cells. Viruses were inoculated on Vero cells to perform a standard plaque assay, as described in the methods section. The plates were incubated for 3 days at 37° C then the agarose was removed and the plates stained with 1% methylene blue. A set of RP5-infected cell plates was also allowed to incubate further to 5 days. The plaque characteristics of RP1, RP3, RP4 at 3 days post infection and RP5 at 3 days and 5 days post infection are shown.

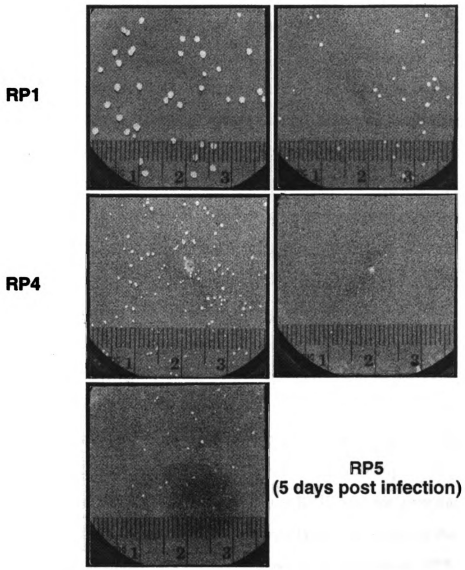


Figure 6

observed for RP4 and RP5 were a result of a failure of the particles to infect cells productively.

The phenotypes observed for RP3, RP4 and RP5 could result from other mutations inadvertently introduced into the HSV-1 genome during the recombination event or subsequent growth of these viruses. Thus, the phenotypes observed could result from a cumulative effect of more than one mutation. To verify that the phenotype observed was only caused by the truncation in the VP16 gene, I generated revertants by introducing the wild type VP16 gene back into RP3, RP4 and RP5. If the defective phenotype was only due to the truncation in VP16, these viruses should exhibit the wild type phenotype. On the other hand, if the virus has picked up other mutations, they should continue to show the defective phenotype.

Wild type VP16 gene was recombined back into RP3, RP4 and RP5 to create revertant strains, RP3R, RP4R and RP5R. Southern blot analysis was performed to verify that the revertants acquired the wild type VP16 gene (Figure 7). Each revertant gave a single band when probed with the VP16 gene fragment, demonstrating that all the revertants have acquired the wild type VP16 gene. Titer, particle per pfu ratio and plaque characteristics (data not shown) of RP4R and RP5R were determined (Table 1). RP3 had a similar particle per pfu ratio as wild type, therefore, the particle per pfu of RP3R was not determined. All revertants exhibited the phenotype of wild type viruses which demonstrates that the defective phenotype of the parental recombinant viruses, RP3, RP4 and RP5, were due to the truncations in the VP16 activation domain.

To further characterize these viruses, single step replication kinetics of RP1, RP3 and RP4 were determined. The single step replication curve of RP5 could not be determined because the titer of the virus was too low for this experiment. Figure 8 shows that within the first 4 hours of infection the infected

Table 1

Particle count and particle per pfu ratio

Recombinant HSV-1 Strain	Titer of stock (pfu/ml)	Number of particles per ml	particle/pfu Ratio
RP1	3.9×10^8	3.4×10^9	8.6
RP3	5.8×10^8	5.3×10^9	9.1
RP4	3.3×10^7	1.2×10^{10}	360
RP5	3.1×10^6	3.6×10^9	1200
RP4R	1.1×10^9	6.1×10^9	5.7
RP5R	1.1×10^9	6.3×10^9	5.5

Table 1. Particle count and particle per pfu ratio. Recombinant virus stocks were generated, titered and intact particles in each stock were counted as described in the methods section. The particle per pfu ratio of each stock was calculated. These results demonstrated that all the recombinant viruses were able to produce comparative amount of intact particle despite their low ability to initiate viral infection.

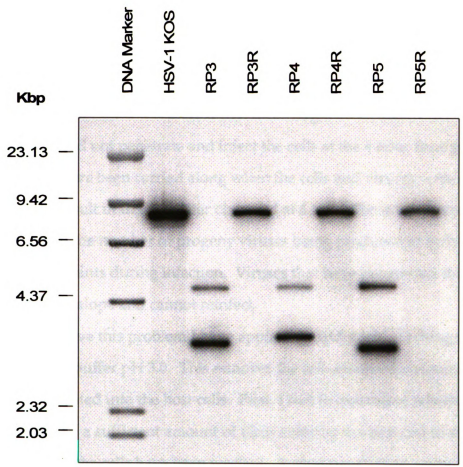


Figure 7. Southern blot analysis of revertant viral strains. The purified viral DNA from RP3R, RP4R and RP5R was subjected to Southern blot analysis. The purified DNA was digested with *Bam*HI and probed with the *Pst*I-*Eco*RI VP16 gene fragment from pKOSVP16.2, described previously. Viral DNA from RP3, RP4 and RP5 was used as a control in this experiment.

cells yielded 10^5-10^6 pfu/ml for all three strains tested. RP1 and RP3 titers rose to approximately 10^7 pfu/ml at 16 hpi.. The RP4 titer was slightly lower than that of RP1 and RP3. The titer of RP4 was about $5x10^6$ pfu/ml at 16 hpi.. These results suggested that there might be differences in the growth curves, but there was a very high initial titer of the recombinant viruses at 4 hpi.. It was surprising to see such high titers at 4 hpi., since it was unlikely that any new particles were produced by that time. I reasoned that there might be viruses which adsorbed to the cells but did not penetrate and infect the cells at the 4 hour time point and would then have been carried along when the cells and viruses were collected. This would result in the high titer observed at 4 hpi.. These viruses would mask differences in the number of progeny viruses being produced at early as well as at later time points during infection. Viruses that have penetrated the host cell loose their envelope and cannot reinfect.

To resolve this problem I have applied an additional washing step with glycine-saline buffer pH 3.0. This removes the cell-adsorbed viruses which have not yet penetrated into the host cells. First, I had to determine whether or not there would be a sufficient amount of virus entering the host cell to start the infection once the cells have been washed. A viral penetration assay was performed where cells were inoculated with 100 pfu per plate of recombinant or wild type virus, and subsequently washed with the glycine-saline buffer pH 3.0 at different time intervals. The number of plaques on each set of plates at different times was determined and compared with the unwashed plates, with the results shown in Figure 9. At the end of the first hour approximately 50 percent of the RP1, RP3 and RP4 particles have absorbed and were able to start infection. Therefore in the single step replication experiments where viruses were inoculated at 5-10 pfu per cell, there should be at least 2-5 pfu per cell after the wash.

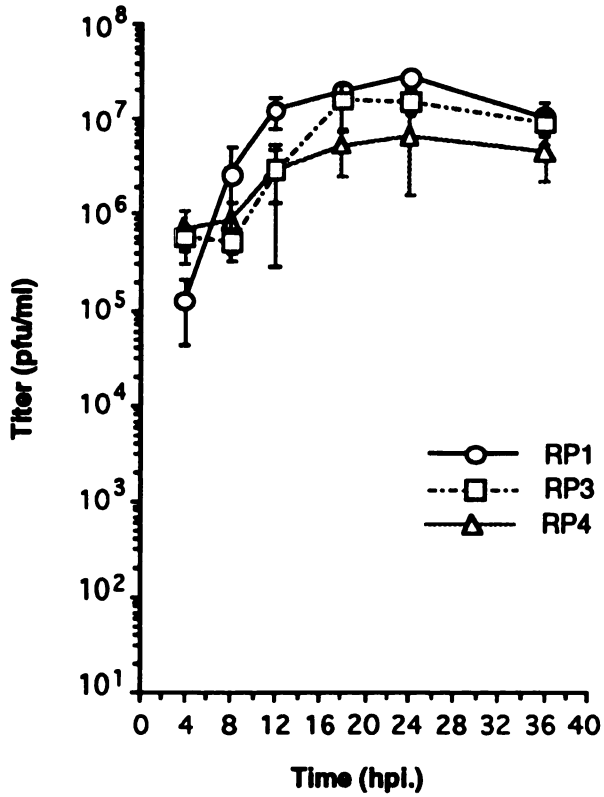


Figure 8. Single step growth curve of recombinant virus strains RP1, RP3, and RP4 (Standard condition). Vero cells were inoculated with recombinant virus RP1, RP3 or RP4 at 5-10 moi. At 4, 8,12, 16,24 and 36 hpi, the cells were harvested and the titer at each time point was determined and plotted against time, hpi., as described in the methods section.

Single step replication experiments with the additional wash with glycine-saline pH 3.0 after absorption were then performed, as shown in Figure 10. The titer at 4 hpi. dropped dramatically and a difference between RP1 and RP4 could be seen clearly. Despite the difference in the plaque forming characteristic there seems to be no difference in the single step replication curve of RP1 and RP3. In contrast, RP4 replicated poorly and did not produce an outburst of the particles at 8-12 hpi., as seen for RP1 and RP3, but rather replicated slowly throughout the course of infection.

Recombinant HSV-1 are defective in infection

Together, the plaque characteristics, particle per pfu ratio and single step replication experiments show that RP3, RP4 and RP5 are defective in the overall viral infection. Although there were no significant differences in the single step replication experiment or the particle per pfu of RP3, the smaller plaque size of RP3 is clearly distinguishable from RP1.

To roughly determine the steps during infection that might be defective in these viruses, I used Vero16-8 cells which constitutively express wild type VP16 protein. If the viruses are defective in the activation of their IE genes, Vero16-8 cells should rescue their defective phenotype by providing the wild type protein which could function in IE activation.

The results in Figure 11 indicate that when RP3, RP4 and RP5 were titered on Vero16-8 cells, VP16 present in the Vero16-8 cells did not complement the infection of RP4 and RP5. A significantly higher number of plaques was not found on Vero16-8 compared to Vero cells (panel A). On the other hand, RP5 produced in Vero16-8 cells gave a titer as high as the RP1 strain (panel B). Further more, the plaque characteristics of all the recombinant viruses strains are

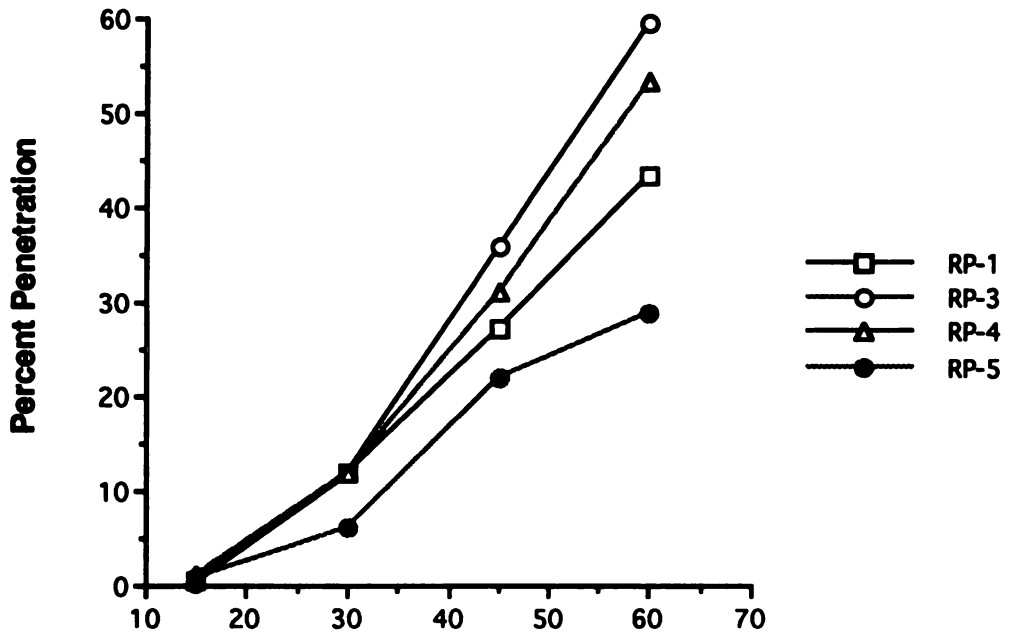


Figure 9. Penetration curve. Vero cells in p60 tissue culture plates were inoculated at 100 pfu per plate. At 15, 30, 45, and 60 minutes after inoculation the plates were washed with glycine-saline, pH 3.0, solution to remove the viruses that have adsorbed but not penetrated into the host cells. The plates were overlaid with media containing agarose and incubated until plaques formed. The number of plaques formed on the plates (expressed as a percentage compared to unwashed plates) was plotted against the time allowed for penetration, as described in the methods section.

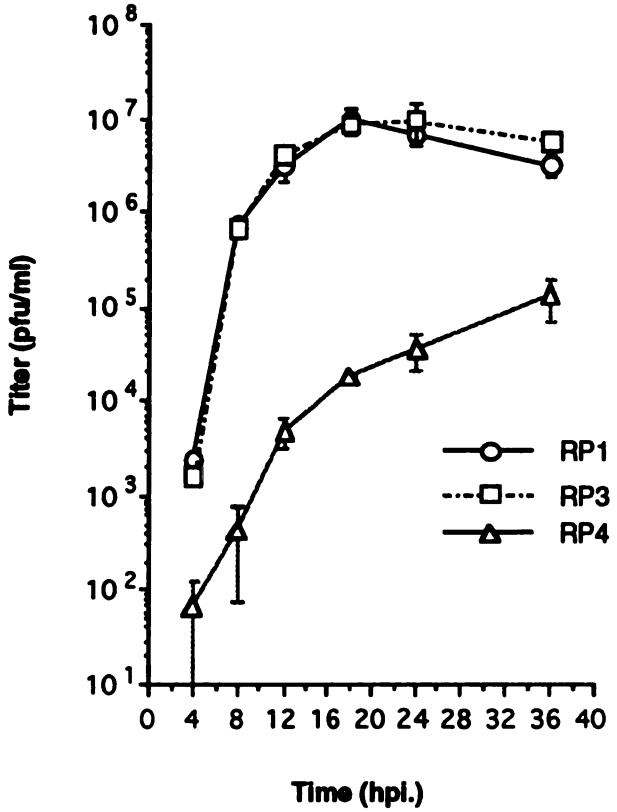


Figure 10. Single step growth curve of recombinant virus strains RP1, RP3 and RP4 (Acid wash condition). Recombinant viruses, RP1, RP3, and RP4 were inoculated on Vero cells in a 6-well plate at 5-10 moi. After the viruses were allowed to adsorb and penetrate for 1 hr. the cells were washed with glycine-saline buffer pH 3.0 to remove the excess viral particles. The infection was allowed to proceed and the cells were collected at different time points and assayed for the presence of infectious virus as described in Figure 7.

similar on Vero16-8 cells. The similar plaque characteristics of all the viral strains on Vero16-8 cells suggests that the progeny viruses produced were able to spread and infect with similar kinetics. These results indicate that wild type VP16 can rescue the defective strains only when the VP16 protein is present in the particle, not when VP16 is supplied by the infected host cell.

The absence of the whole activation domain or one of the subdomains might also affect the structure of the virion in such a way that the uncoating or migration of the virus is blocked. Therefore the growth defect we observed could result from problems in these steps, not from a defect in transcription activation. To test this hypothesis I examined how well viral DNA reaches the cell nucleus. Equal amounts of particles of the recombinant viruses, equivalent to 10 pfu/cell of RP1, were used to infect Vero cells. At 3 hpi. the nuclei were harvested, DNA was extracted and subjected to Southern blot analysis.

Figure 12 shows the Southern blot of DNA extract from nuclei of cells infected with the recombinant viruses. The results clearly indicate that RP3, RP4, and RP5 can deliver their DNA into the host cell nucleus within the first 3 hours of infection. This demonstrated that RP3, RP4 and RP5 do not have any detectable defect in their ability to deliver their DNA to the host nucleus. These results imply that the steps prior to DNA delivery, such as adsorption, penetration, uncoating and migration of the particle, should also be normal in these recombinant viruses.

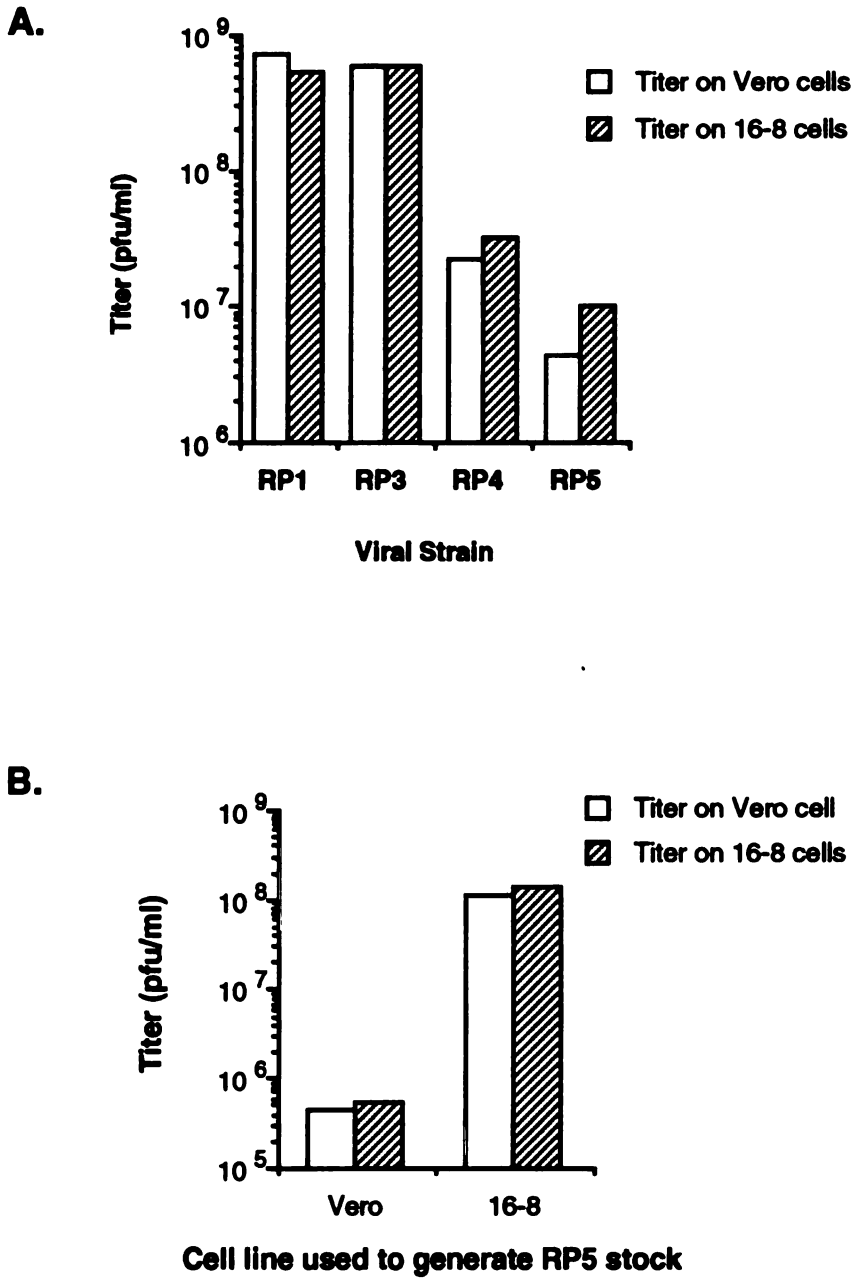


Figure 11. Titer of recombinant viral stocks in Vero and Vero 16-8. Recombinant viral stocks RP1, RP3, RP4 were generated in Vero cells, and RP5 was generated in Vero and Vero 16-8. (A) The titer of viral stocks titered on Vero or Vero 16-8 are compared. (B) The titer of RP5 stocks generated in Vero or Vero 16-8 cells was determined on Vero and Vero 16-8 cells.

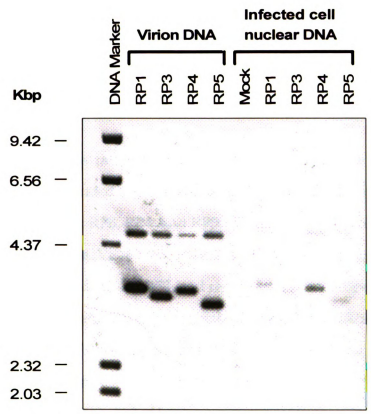


Figure 12. Southern blot analysis of nuclear localized recombinant HSV-1 DNA. Vero cells were infected with an equal number of particles of RP1, RP3, RP4, and RP5. At 3 hpi. the nuclei from the infected cells were harvested and DNA extracted. The DNA was digested with *Bam*HI to completion and Southern blot analysis was performed, as described in the methods section. The purified DNA from virions was used as a positive control in this experiment.

Discussion

Previous work suggested that the VP16 gene is essential for viral replication (Weinheimer, et al., 1992). Mutation of the VP16 gene within the context of the viral genome results in defective particles (Ace, et al., 1989; Poon & Roizman, 1995). Attempts have been made without success to generate deletion mutations of the VP16 gene within the context of the viral genome (Ace, et al., 1989).

I have demonstrated that viruses with deletions in the VP16 activation domain can be generated and are viable in cell culture. Five viral strains were generated: RP1, RP2, RP3, RP4, and RP5. These derivatives of HSV-1 contain the full length VP16 gene, VP16 from HSV-2, and VP16 with a deletion of amino acid residue 456-490, residues 413-452, and residues 413-490, respectively. Similar frequencies of generation of RP1 (recombining wild type VP16) and RP2, RP3, RP4 and RP5 were observed (data not shown), indicating that introduction of HSV-2 VP16 or the truncated VP16 genes is not a rare recombination event. A complementing cell line, Vero16-8, was not necessary for either the construction or isolation of these recombinant viruses. Moreover, Western blot analysis demonstrates that the appropriate derivative of VP16 is produced and assembled into the virion in each of the recombinant viral strains. These results suggested that truncated VP16 was able to support viral maturation and produce infectious virions.

Recent studies have shown that VP16 interacts with other viral proteins present in the tegument such as the viral host shutoff protein (Smibert, et al., 1994) and VP22 (Elliott, et al., 1995). VP16 interacts with VP22 via its activation domain, and deletions and point mutations in the VP16 activation domain greatly affect the interaction between VP22 and VP16. Mutations in the

activation domain resulted in alteration of the localization of the two proteins in the host cell during viral maturation (Elliott, et al., 1995). Our results suggest that this interaction must not be essential for viral maturation or particle assembly, since recombinant viruses bearing deletions of the activation domain were able to grow and produce intact particles.

Though viable, these viruses exhibit some defects in replication. RP1 (carrying the full length VP16 gene) has an identical phenotype to wild type HSV-1 KOS. RP3 produced significantly smaller plaques when compared to RP1, though single step replication and particle per pfu ratio were similar to that of RP1. The smaller plaques of RP3 could reflect the cumulative effect of slightly slower replication and cell to cell spreading of the virus observable only upon multiple rounds of infection. RP4 and RP5 are greatly defective in their ability to form plaques. Both gave higher particle per pfu ratio and slower replication as exhibited by their plaque characteristics. Single step replication kinetics studies also demonstrated that RP4 replicates much slower than RP1 and RP3 (not determined for RP5).

Particle count by electron microscopy demonstrated that these viruses are normal in their ability to produce intact particles. The next reasonable step to take was to test whether or not their phenotype was caused by the loss of transcription activation activity of truncated VP16. Therefore, the ability of Vero16-8 cells to rescue the phenotype of these recombinant viruses was tested. Vero16-8 cells are expected to rescue the defects in IE gene expression, by providing wild type VP16 in trans for the incoming virus. We would expect RP4 and RP5 to give a comparable titer with RP1 when titered on Vero16-8 cells. Results from Friedman *et al.* demonstrated that a cell line constitutively expressing a truncated form of VP16 which lacks the C-terminal 78 amino acid was shown to be resistant to HSV-1 infection (Friedman, et al., 1988). This result

clearly demonstrates that a VP16 derivative provided in trans by the host cell can interact with incoming virus during early infection.

However, our results show no significant differences in the titer of RP4 and RP5 on Vero16-8 compared to Vero cells (Figure 11). There were no differences in the plaque characteristics of RP1, RP3, RP4 and RP5 on Vero16-8 cells (data not shown), indicating that Vero16-8 cells were able to complement the growth defect of the progeny viruses. This was also observed when RP5 was grown in Vero16-8 cells. We observed almost a 1000 fold increase in the titer of the RP5 stock produced in Vero16-8 versus Vero cells, shown in Figure 11B. The presence of wild type VP16 protein in the RP5 particle was able to complement the low titer phenotype.

The failure of Vero16-8 cells to complement these viruses in plaque assays could result from an inadequate level of VP16 in these cells early in infection. Western blot analysis using a whole cell extract from Vero16-8 cells did not detect VP16 protein, suggesting that there is not a high level of VP16 protein expressed in this cell line (data not shown). It could also be the result of an aberrant localization of the VP16 in Vero16-8 cells or different post-translational modification of VP16 in Vero16-8 cells. The localization of VP16 could be verified by indirect immuno fluorescence, and the mobility of VP16 protein from the virion and Vero16-8 cells could be compared by isoelectric focusing to test for differences in phosphorylation states.

The above considerations did not rule out the possibility that Vero16-8 can complement for the activation of IE transcription, but that the recombinant viruses failed to penetrate, uncoat and deliver the viral DNA properly into the host cell nucleus.

I have shown by Southern blot analysis of the nuclear localized viral DNA that RP3, RP4 and RP5 are not defective in their ability to deliver DNA into the

host nucleus. This implies that the absorption, penetration and uncoating of these viruses is normal or not greatly effected by the mutations in VP16. Since the viral DNA was delivered into the host nucleus, the failure of RP4 and RP5 to initiate infection properly is most likely a result of failure of truncated VP16 to activate transcription of IE genes.

Preliminary experiments in our laboratory done recently by Eugene Chung, repeating the titering experiment of RP4 and RP5 using a BHK3378#2 cell line constitutively expressing VP16 protein (kindly provided by Dr. Paul Hippenmeyer, Monsanto) demonstrated that RP4 and RP5 can be complemented by this new cell line. These results confirm that the defective phenotype observed in RP3, RP4 and RP5 is most likely to be a result of aberrant IE gene expression.

The reason why Vero16-8 cannot complement RP4 and RP5 while BHK3378#2 can is still obscure. The expression level of VP16 protein in both cell lines is low and cannot be detected by Western blot analysis (data not shown). The differences between these two cell lines could be a result of different localization or modification of the VP16 protein.

Eugene has also shown by transient expression assay that RP4 and RP5 were unable to activate transcription from a reporter plasmid containing the HSV-1 ICP4 promoter. Curiously, RP3 was able to activate transcription from the reporter plasmid similarly or even better than RP1, which carries the wild type VP16. Preliminary quantitative reverse transcriptase polymerase chain reaction (RTPCR) experiments by Eugene, on the IE genes, ICP0, ICP4, ICP22 and ICP27 also demonstrated that RP4 and RP5 are defective in their ability to activate these genes. These results support the hypothesis that RP4 and RP5 are defective in the activation of IE genes which could greatly contribute to the growth defect.

We have clearly demonstrated that the activation domain of VP16 is necessary for normal viral growth. The recombinant viruses were able to produce intact particles, but the particles are defective in infection as shown by the high particle per pfu ratio. The migration of the viral DNA into the nucleus of the host cell is normal, suggesting that adsorption, penetration, uncoating and migration of the viral particle are normal for RP3, RP4 and RP5. A transient expression assay using the ICP4 promoter as well as RTPCR results demonstrated that these viruses are defective in IE gene expression. Therefore, the growth defects observed in these recombinant viruses are most likely to be a result of the defect in IE gene activation and overall gene expression. Our results also suggest that the interaction between VP22 and VP16 during maturation is not essential for viral replication.

Future studies

Since we have recombinant viruses bearing the different deletions in the VP16 protein we should be able to measure the activity of these mutant VP16 proteins in the context of the viral genome. IE gene mRNA products could be measured and the activity of the derivative VP16 proteins could be determined. Transient expression assays performed by transfecting reporter plasmids, with a reporter gene linked to IE promoters or their derivatives, followed by infection with these recombinant viruses would also be interesting. It could give us information on activation by virion delivered VP16 and derivatives and their ability to activate IE promoters. Mutations in the promoters of the IE genes could also be construct and tested in this fashion to investigate the effects of different deletion mutations on the different promoters.

Other mutations in VP16 that have been tested in transient expression assays could also be transferred in to the virus to study the effect upon viral replication. Especially single point mutations that greatly decreased VP16 activity, such as the Phe 422 to Ala change in the N-subdomain and substitutions of other important residues in the C-subdomain, should be tested in the viral context, both to determine the effects on the overall lytic infection of the virus and the effect on viral gene expression.

CHAPTER III

IDENTIFYING THE ACTIVATION DOMAINS IN VP16 HOMOLOGUES

Introduction

VP16 of HSV-1 is known for its ability to activate IE transcription and plays an essential role in HSV-1 maturation and particle assembly. Homologues of VP16, HSV-2 VP16 (Cress & Triezenberg, 1991a), BHV-1 UL48 (Carpenter & Misra, 1991), VZV ORF10 (Davison, 1991; McKee, et al., 1990), MDV UL48 (Koptidesova, et al., 1995; Yanagida, et al., 1993), EHV-1 Gene12 and EHV-4 GeneB5 (Purewal, et al., 1994; Purewal, et al., 1992) have been identified and the amino acid sequences have been deduced.

Despite the availability of the deduced protein sequences, the activation domains of these homologues of VP16 have not yet been identified (with the exception of VP16 from HSV-2). Linear sequence comparisons of the homologous proteins do not reveal any obvious similarity between the VP16 activation domain and the sequence at the C-terminus of other homologues. In some VP16 homologues such as VZV ORF10 and MDV UL48 the entire 80 amino acid C-terminal portion of the protein is missing.

VP16 was characterized as an acidic activator because the activation domain of the protein is rich in acidic residues. However, Cress *et al.* 1991 and Regier *et al.* 1993 have shown that hydrophobic residues, especially Phe residue

442 (Phe442) and others in its vicinity, are critical for its function. Considering the essential role of hydrophobic residues for activity in mind, we have employed a sequence analysis program termed hydrophobic cluster analysis (HCA) to analyze VP16 and its homologues, and search for their activation domains. HCA presents the protein sequence in a two dimensional α helical representation. A contour is then drawn grouping the adjacent hydrophobic residues, which presumably form the hydrophobic core of globular proteins, producing clusters of hydrophobic residues. These clusters were then visually compared to the patterns obtained for putative homologues (Lemesle-Varloot, et al., 1990; Woodcock, et al., 1992).

This analysis has revealed activation domains in VP16 homologues that were not obvious when the linear amino acid sequences were compared. In this study I and my collaborators have demonstrated, by site directed oligonucleotide mutagenesis substitutions of predicted critical residues, that the activation domain in VZV ORF10 is indeed homologous to the N-subdomain of the VP16 activation domain. These results were published in the Proceedings of the National Academy of Sciences USA (1995, Vol. 92, pp. 9333-9337).

We have identified the activation domain of at least three homologues of VP16 from HSV-1 (Triezenberg, et al., 1988a), HSV-2 (A. Cress unpublished results) and VZV (Moriuchi, et al., 1995). The activation domains of BHV-1, MDV, and EHV-1 & 4 are also being mapped and identified in other laboratories, with results supporting my predictions (Elliott, 1994; Misra, 1994; Mekki Boussaha, manuscript in preparation). These results demonstrated that we have successfully predicted the location and critical residues for the activity of the VP16 homologous.

To study structural aspects of the VP16 activation domain, we have used several different algorithms to predict the secondary structure of the N-

subdomain of the activation domain of VP16 and its homologues. We expected that the secondary structural analysis of the N-subdomain from the different homologues would demonstrate a common preferred secondary structure and would thus provide insights on the structure of the activation N-subdomain of VP16. The Chou-Fasman (Chou & Fasman, 1978), Garnier-Osguthorpe-Robson (Garnier & Osguthorpe, & Robson, 1978) and two new methods, Sequery (Collawn, et al., 1991; Collawn, et al., 1990) and profile network system from Heidelberg (PHD) (Rost & Sander, 1994) were applied in these studies. The predictions by Chou-Fasman, Garnier-Osguthorpe-Robinson and PHD suggested that the region around the predicted critical Phe residue does not have a strong preference for any secondary structure. However, results from predictions by Sequery suggested that the secondary structure of the segment is likely to be loop or extended structure.

Methods

DNA and Plasmids

Plasmid GAL4-ORF10(5-79), expressing the chimeric protein GAL4-ORF10, which contains the DNA binding domain GAL4 residue 1-147 fused to the N-terminal amino acid residues 5-79 of VZV open reading frame 10 (VZV-ORF10) was provided by J. Cohen.

pGAL4-ORF10(5-79) was digested to completion with *Eco*RI. The *Eco*RI fragment containing the sequence encoding amino acid residues 5-79 was then subcloned into M13mp18. The fragment orientations of several of the subclones were identified by standard complementation test and complementary M13

clones were sequenced to obtain the phage with the correct orientation, resulting in M13mp18-ORF10.

The reporter plasmid, p5GAL4-E1b-CAT, containing five GAL4 binding sites linked to the CAT gene was described previously (Lillie & Green, 1989).

Site directed oligonucleotide mutagenesis

Oligonucleotides used for site-directed mutagenesis are shown in Table 1. The oligonucleotides were provided by J. Cohen's laboratory. M13mp18-ORF10 was mutated using standard protocols (Zoller and Smith 1982; Kunkel, 1985; Bio-Rad catalog #170-3571). M13mp18-ORF10 was grown in *E. coli* host CJ 236 [(dut1, ung1, thi1, relA1; pCJ105(Cm^r)] which incorporates uridine nucleotides at approximately 1% of the normal thymidine residues. Uridine-containing phage DNA purified by precipitation with PEG8000, extraction with phenol and chloroform, and then ethanol precipitation. The mutagenic oligonucleotide primers were annealed to template DNA in 20 mM Tris-HCl (pH 7.4), 2 mM $MgCl_2$, 50 mM NaCl by heating to 65 °C and cooling, 1 °C/minute to 4 °C. A standard primer/template ratio of 20:1 was used (Bio-Rad catalog #170-3571). Synthesis of the complementary strand DNA was performed in 23 mM Tris-HCl (pH 7.4), 5 mM $MgCl_2$, 35 mM NaCl, 1.5 mM dithiothreitol, 0.4 mM deoxynucleotide triphosphates, 0.75 mM ATP, 0.2 unit/µl T4 DNA ligase, and 0.1 unit/µl T4 DNA polymerase. The synthesis was carried out for 5 minutes at 4 °C, 5 minutes at 25 °C and 60 minutes at 37 °C. The reaction was terminated by heating at 65 °C for 10 minutes. The synthesis reactions were then used to transform competent dut^+ ung^+ MV1193 cells {(Δlac-pro AB), thi, supE, Δ(sr1-recA)306::Tn10(tet^r) [F':traD36,proAB lac $I^qZ\Delta M15$]}. Transformation of the synthesized duplex DNA into MV1193 cells results in the selection against the

Table 1

Mutagenic Oligonucleotides

F28S	5'-GTT GTG GAC GCA T**C**T GAT GAA TCG TTG-3'
F28Y	5'-GTT GTG GAC GCA T**A**T GAT GAA TCG TTG-3'
F28L	5'-G GTT GTG GAC GCA **C**TT GAT GAA TCG TT-3'
F28I	5'-G GTT GTG GAC GCA **A**TT GAT GAA TCG TT-3'
F28V	5'-G GTT GTG GAC GCA **G**TT GAT GAA TCG TT-3'
F28A	5'-G GTT GTG GAC GCA **GC**T GAT GAA TCG TT-3'
F28P	5'-G GTT GTG GAC GCA **CC**T GAT GAA TCG TT-3'
F28T	5'-G GTT GTG GAC GCA **AC**T GAT GAA TCG TT-3'
V24A	5'-ACG GAA CAA GCG G**C**T GTG GAC GCA TTT-3'
V25A	5'-GAA CAA GCG GTT G**C**G GAC GCA TTT GAT-3'
L32A	5'-A TTT GAT GAA TCG **GC**G TTT GGT GAT GT-3'
F33S	5'-GAT GAA TCG TTG T**C**T GGT GAT GTA GCA-3'
F33Y	5'-GAT GAA TCG TTG T**A**T GGT GAT GTA GCA-3'

Table 1. Mutagenesis oligonucleotides. These oligonucleotides were synthesized in J. Cohen's laboratory. They were used in the mutagenesis of Phe28 and the flanking hydrophobic residues, Val24, Val25, Leu32, and Phe33, in the activation domain of VZV ORF10 in the context of M13mp18-ORF10(5-79). The mutated fragment was then subcloned into pGAL4-ORF10(5-79). The bold-underlined letters indicates the nucleotides that have been substituted to create the amino acid changes.

uridine-containing wild type DNA strand. Transformant MV1193 plagues were screened for mutations by dideoxy sequencing (Sanger, Nicklen, & Coulson, 1977). Replicative form DNAs of the mutant phage were prepared using standard protocols. The double stranded phage DNA was cut with *Eco*RI releasing the fragment encoding the VZV ORF10 residues 5-79. The *Eco*RI fragment was then subcloned back into pGAL4-ORF10(5-79) and the correct orientation of the fragment was confirmed by *Sty*I digestion. The mutated derivatives of pGAL4-ORF10(5-79) were amplified in DH5α cells. The plasmid DNA was extracted and purified twice by centrifugation in CsCl gradient and used in transient expression assays to determine the activity of the mutant relative to wild type GAL4-ORF10(5-79).

Transient expression assays were performed by Hiroyuki & Masako Moriuchi in the laboratory of J. Cohen. Plasmids encoding mutant derivatives of GAL4-ORF10(5-79) were co-transfected with a reporter plasmid p5GAL4-E1b-CAT (Lillie & Green, 1989) into Vero cells. CAT activity was determined for wild type GAL4-ORF10(5-79) and each of the derivatives. The relative CAT activity of the mutant relative to the wild type protein was determined.

HCA

The peptide sequences of VP16 homologues were acquired from the GenBank Genetic Sequence Data Bank (Access numbers: HSV-1 KOS VP16, S. J. Triezenberg's unpublished results; HSV-2 VP16, M60050; BHV-1 UL48, Z11610; VZV ORF10, X04370; EHV-1 Gene12, M86664; EHV-4 GeneB5, L16590; MDV UL48, X73370). The primary sequences were then processed through the hydrophobic cluster analysis program, HCA-plot-V2, kindly provided by Dr. B. Henrissat (Centre National de la Recherche Scientifique, Université Joseph

Fourier, Grenoble France). The HCA patterns of each of the homologues were converted to postscript format and printed. The hydrophobic cluster patterns of all the homologous were then aligned and compared by eye. The linear sequence alignment of the homologues was also compared using Pileup, a multiple sequence alignment program from the Genetic Computer Group, Inc. (GCG), 1994, and used in conjunction with the data from HCA to align and compare the primary sequences. The algorithm used for Pileup was according to the work of Feng & Doolittle, 1987.

Sequery

The computer program termed Sequery searches a database of proteins with known structures (the Brookhaven protein data bank) for a specific amino acid sequence (the query sequence) from a protein of interest (Collawn, et al., 1991; Collawn, et al., 1990). Sequery searches the data base for similarity in the sequence independent of the protein origin or function. The program output is a list of the proteins in the database which contain a sequence matching the query sequence. The amino acid portion of the matching amino acids is also listed. The protein structures of the matching proteins were then retrieved and displayed. The structure of the protein from the database containing the query sequence represents a predicted protein structure for the query sequence.

The sequence from residues 439-447 of HSV-1 VP16 was used as a query sequence using a sliding window of four residues. The amino acids designed as matching at each position were the original residues presented in HSV-1 VP16 and other conserved residues found at that position from the linear amino acid alignment of the N-subdomain (Figure 4B). The amino acid residues allowed at each position are shown in Figure 1. Some residues were not allowed and some

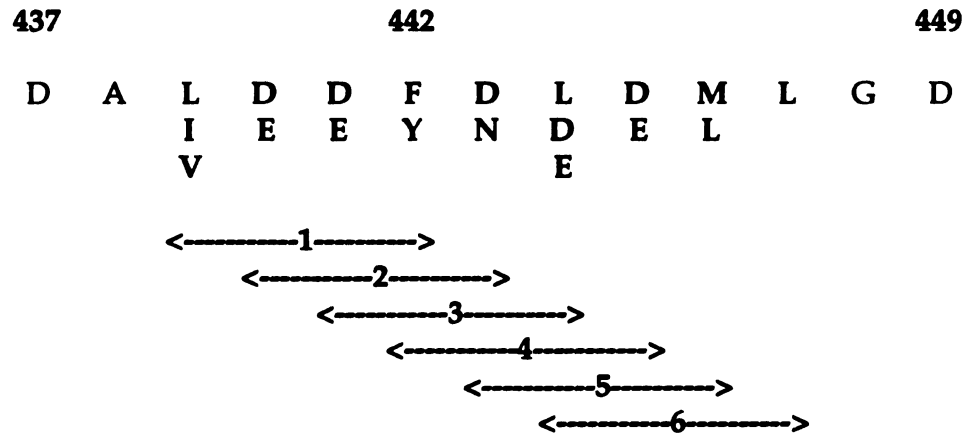


Figure 1. Residues allowed at each position for the Sequery program search and the six windows used in the search. A sliding window of four amino acid residues, starting from amino acid 439 to 447, was used to search for matching sequences in the Brookhaven protein data bank. The residues allowed at each position were selected based on the conserved residues found in other VP16 homologues. Some residues were avoided and some were added to minimize picking up unrelated sequence and yield a workable amount of structures from each search. The position of each window is indicated by "<-->".

residues (known to retain the function of VP16 activation ability) were added to avoiding picking up irrelevant sequences and to yield a workable sample size.

PHD

The computer program termed PHD (Rost & Sander, 1994), predicts the secondary structure of proteins by using evolutionary information of protein structures in combination with computer neural networks. The program takes into account not only the local information present in the primary sequence of the protein, but also picks up similar sequences from the data base and uses the information from other sequences to assist in the prediction. This method has improved the accuracy of the protein structural prediction to above 70% for the three-state prediction, α helix, β sheet and loop, of a given protein sequence.

The sequences of the VP16 homologues from the linear alignment of the N-subdomain in Figure 3 were analyzed by the PHD program. Sequences were sent for analysis via world wide web (http:/www.embl-heidelberg.de/predictprotein /predictprotein.html) The PHD results were returned via electronic mail. The output shows the prediction of the three-state structure as well as the similar sequences picked out from the data base.

Chou-Fasman and Garnier-Osguthorpe-Robson

Secondary structure predictions by the Chou-Fasman and Garnier-Osguthorpe-Robson algorithms use the information from the statistical survey of a selected set of proteins from the protein data base previously described (Chou & Fasman, 1978; Garnier, Osguthorpe, & Robson, 1978). The sequence of VP16 homologues were subjected to secondary structure prediction by the Peptidestructure program in the GCG, which employs both the Chou-Fasman

and Garnier-Osguthorpe-Robson methods. The surface probability was calculated according to the formula of Emini et al., 1985.

Results

HCA

The activation domain of VP16 was mapped to the C-terminal 80 amino acid residues of the protein (Triezenberg, et al., 1988a). Homologues of the VP16 protein in other α herpes viruses have been identified and their amino acid sequences have been deduced from cloned gene sequences. Sequence comparison of the deduced amino acid sequences of VP16 homologues has demonstrated that the amino acid sequences are mostly conserved in the middle portion of the protein which corresponds to residues 50-410 of HSV-1 VP16. The N and C terminus of the protein, however, were divergent in the amino acid sequence as well as length, except for HSV-1 and HSV-2 VP16 which have a high degree of similarity throughout the molecule (Figure 2). The sequence of VZV ORF10 and MDV UL48 are shorter than the rest of the homologues and lack the residues corresponding to the activation domain of VP16, residues beyond 410 of HSV-1 VP16. EHV-1 Gene12, EHV-4 GeneB5, and BHV-1 UL48 contain residues corresponding to residues beyond residues 410 in HSV-1 VP16, but these residues were not conserved in sequence when compared by conventional linear alignment methods (Figure 2). Therefore, comparison of these proteins using conventional linear sequence alignment to identify activation domains in these VP16 homologues was unfruitful.

Though the activation domain of VP16 was characterized as an acidic activation domain, Cress *et al.* 1991 and Regier *et al.* 1993 have shown that in

```
                                                    .   .       .
  HSV-1VP16  .......... .......... ......MDLL VDELFADMDA DGASPPPPRP
  HSV-2VP16  .......... .......... ......MDLL VDDLFADR.. DGVSPPPPRP
EHV-1gene12  MCLLHISLPY LSCALLPGWY FDARPAASIV MFAAAEENDD PYPGKSGYND
EHV-1gene5B  .......... .......... ....MAANIA MFADIEDYDD TRSCEYGY.G
  BHV-1UL48  .......... .......... .......... MSGRIKTAGR ALASQCG.GA
   VZVORF10  .......... .......... .......... ...MECNLGT EHPSTDTWNR
     MDVUL48 .......... ....MEANMS FENDYYSPIQ LFAEIEAYAN TM.......D

                 ..      .  .. ....             ...      . ...... ...
  HSV-1VP16  AGGPKNTPAA PPLYATGRLS .......... .......QAQ LMPSPPMPVP
  HSV-2VP16  AGGPKNTPAA PPLYATGRLS .......... .......QAQ LMPSPPMPVP
EHV-1gene12  TCELMDMDGA VASFDEGMLS ...A....IE SVYSIPTKKR LALPPPKAAS
EHV-1gene5B  TCELMDVDGV VASFDEGMLS ...A....SE SIYSSPAQKR LALPPPKATS
  BHV-1UL48  AAATMDPYDA IEAFDDSLLG SPLA....AG PLYDGPSPAR FALPPPRPAP
   VZVORF10  S...KTEQAV VDAFDESLFG DVASDIGFET SLYSHAVKT. .APSPPWVAS
     MDVUL48 KSPDLDILRT IEEFDETLLS EIEVRTQSIP SPLVAPSVTK MSLPSPSPAP

             . .:....   .:.: .:... .......:.: .:: .: ... .: . ..:::.
  HSV-1VP16  PAALFNRLLD DLGFSAGPAL CTMLDTWNED LFSALPTNAD LYRECKFLST
  HSV-2VP16  PAALFNRLLD DLGFSAGPAL CTMLDTWNED LFSGFPTNAD MYRECKFLST
EHV-1gene12  PGALYQRLQG ELGFPEGQTL LSAMEKWNED MFSALPGHVD LYTEIALLST
EHV-1gene5B  PTALYQRLQA ELGFPEGQAM LFAMEKWNED MFSAIPVHVD LYTEIALLST
  BHV-1UL48  LAALLERMQA ELGFPDGPAL LRAMERWNED LFSCLPTNAD LYADAALLSA
   VZVORF10  PKILYQQLIR DLDFSEGPRL LSCLETWNED LFSCFPINED LYSDMMVLSP
     MDVUL48 PNSLYTRLLH ELDFVEGPSI LARLEKINVD LFSCFPHNKH LYEHAKILSV

              . . . .   ..          ........   .: .:.   ..: :. .
  HSV-1VP16  LPSDVVEWGD AYVPERA... .QIDIRAHGD VAFPTLPATR DGLGLYYEAL
  HSV-2VP16  LPSDVIDWGD AHVPERS... .PIDIRAHGD VAFPTLPATR DELPSYYEAM
EHV-1gene12  SVDEVVRAGL DSLPTPSHYS PEVDLNAHGD EPFPEVPALE DDLEIYVISA
EHV-1gene5B  SVNEVVKAGL DSLPIPTNYI PEVDLNAHGS EPFPEVPALE DELETYVISA
  BHV-1UL48  DADAVVGAMY LAVPGDAE.. .RLDLNAHAN QPLPAPPASE EGLPEYVAGV
   VZVORF10  DPDDVI.... .STVSTKDHV EMFNLTTRGS VRLPSPPKQP TGLPAYVQEV
     MDVUL48 SPSEV....L EELSKNTWTY TALNLNEHGE MALPMPPTTK ADLPSYVDDI

             . ... ::.. ::. :  ..  ..: ..... . ....   .  .. . .  .
  HSV-1VP16  SRFFHAELRA REESYRTVLA NFCSALYRYL RASVRQLHRQ AHMRGRDRDL
  HSV-2VP16  AQFFRGELRA REESYRTVLA NFCSALYRYL RASVRQLHRQ AHMRGRNRDL
EHV-1gene12  QRFYLSELRT REEHYARLLR GYCVALLHYL YGSAKR...Q LRGSGSDASL
EHV-1gene5B  QRFYLSELRA REEHYSRLLR GYCVALLHYL YGSAKR...Q LRGAGSDSAL
  BHV-1UL48  QAHFLAELRA REERYAGLFL GYCRALLQHL RATAAR...G .RGAAGAGAQ
   VZVORF10  QDSFTVELRA REEAYTKLLV TYCKSIIRYL QGTAKRTTIG LNIQNPDQKA
     MDVUL48 QNFYLGELEA REKSYATMFY GYCRALAEYI RQSAIKDLRD ARVEDKNIGA

                ... ... :::::.....: .....:.:.. .........  ..:.......:
  HSV-1VP16  GEMLRATIAD RYYRETARLA RVLFLHLYLF LTREILWAAY AEQMMRPDLF
  HSV-2VP16  REMLRTTIAD RYYRETARLA RVLFLHLYLF LSREILWAAY AEQMMRPDLF
EHV-1gene12  MHKFKQVVRD RYYREAANLA RLLYLHLYVS VTREVSWRLH ASQVINQGVF
EHV-1gene5B  MHKFKQVVRD RYYRETANLA RLLYLHLYIS VTREVSWRLH ASQVVNQGIF
  BHV-1UL48  ADRLRQLVAA RYYREASRLA RLAFAHMYVA TAREVSWRLH SQQSQAQGVF
   VZVORF10  YTQLRQSILL RYYREVASLA RLLYLHLYLT VTREFSWRLY ASQSAHPDVF
     MDVUL48 CSKARQYIAE RYYREAARFA KLLYVHLYLS TTRDVSQRLE ASQMGRQNIF

             . :.. . .. :.....:... ....:..... . .... .:.  .:..:....
  HSV-1VP16  DCLCCDLESW RQLAGLFQPF MFVNGALTVR GVPIEARRLR ELNHIREHLN
  HSV-2VP16  DGLCCDLESW RQLACLFQPL MFINGSLTVR GVPVEARRLR ELNHIREHLN
EHV-1gene12  VSLHYFWAQR RKFECLFHPV LFNHGVVILE NDPLEFHDLQ RINYRRRELG
EHV-1gene5B  VSLHYTWPQR RKFECLFHPV LFNHGVVILE NDPLEFNDLQ RINYRRRELG
  BHV-1UL48  VSLYYAWPQR RQFTCLFHPV LFNHGVVALE DGFLDAAELR RLNYRRRELG
   VZVORF10  AALKFTWTER RQFTCAFHPV LCNHGIVLLE GKPLTASALR EINYRRRELG
     MDVUL48 VYLKCEWLQE RHFHCLFQPV IFNHGVVIVE GRVLTAPELR AQNYIRSEFG
```

```
                :::.: ...: .   ::   :  ... ..:.. ...   .. : .....  ..
  HSV-1VP16  LPLVRSAATE EPGAPLTTPP TLHGNQARAS GYFMVLIRAK LDSYS.SFTT
  HSV-2VP16  LPLVRSAAAE EPGAPLTTPP VLQGNQARSS GYFMLLIRAK LDSYS.SVAT
EHV-1gene12  LPLIRAGLIE EENSPLEAEP LFSGKLPRTI GFLTHQIRTK MEAYSDAHPA
EHV-1gene5B  LPLIRAGLIE EENLPLESEP TFSGKLPRTI GFLTHQIRTK MEAYSNAHPS
  BHV-1UL48  LPLVRAGLVE VEVGPLVEEP PFSGSLPRAL GFLNYQVRAK MGA..PAEAG
   VZVORF10  LPLVRCGLVE ENKSPLVQQP SFSVHLPRSV GFLTHHIKRK LDAYAVKHPQ
    MDVUL48  LPLIRCKLVE EPDMPLISPP PFSGDAPRAS VYLLQCIRSK LEVYSLSHPP

                . . ....:  ... .     .:::.:..  .......... .....   ..
  HSV-1VP16  SPSEAVMREH AYSRAR.TKN NYGSTIEGL. LDLP.DDDAP EEAGLAAPRL
  HSV-2VP16  SEGESVMREH AYSRGR.TRN NYGSTIEGL. LDLPDDDDAP AEAGLVAPRM
EHV-1gene12  TP.LFPLAEH SYSKRIGGRL SYGTTTEAM. MDPPSPSAVL PGDPVPPLTV
EHV-1gene5B  TP.LFPLAEH SYSKRIDGRL SYGTTAEAM. MDPPSPSAVL PGDPVPPLTV
  BHV-1UL48  GR.LAPEREH SYARPRGA.I NYGTTPEAM. LRPPSPSEVL PCDPAPAATV
   VZVORF10  EP.RHVRADH PYAKVVENR. NYGSSIEAMI LAPPSPSEIL PGDPPRPPTC
    MDVUL48  NPQLHVHKEH VHVQKLESPP NYGTTVEALL MDSSDRNSIS PGDPVATTIS

                .                 .
  HSV-1VP16  SFLPAGH.TR RLS.TAPPTD VSLGDELHLD GEDVAMAHAD ALDDFDLDML
  HSV-2VP16  SFLSAGQRPR RLSTTAPITD VSLGDELRLD GEEVDMTPAD ALDDFDLEML
EHV-1gene12  GVRQTAATLA .IPSNLTLQS METDGLD... .......... ..........
EHV-1gene5B  GIRQTAETLA .LPSNLTLQS METDVLD... .......... ..........
  BHV-1UL48  RVASPATHLA QAPSAKGAAP AEFAALAGLA KPGPAPLAAA PAQAPFAAAL
   VZVORF10  GFLTR..... .......... .......... .......... ..........
    MDVUL48  TL*....... .......... .......... .......... ..........

                                                          .
  HSV-1VP16  GDGDSPGPGF TPHDSAPYGA LDMADFEFEQ MFTDALGIDE YGG*......
  HSV-2VP16  GDVESPSPGM T.HDPVSYGA LDVDDFEFEQ MFTDAMGIDD FGG*......
EHV-1gene12  .......... .......... .......... .......YSS MTGDELNQMF
EHV-1gene5B  .......... .......... .......... .......YSS ISGDELNQMF
  BHV-1UL48  ALAEPAAALA PAPLAAAPAE PAAAVAGPSP ANPFGGTYDA LLGDRLNQLL
   VZVORF10  .......... .......... .......... .......... ..........
    MDVUL48  .......... .......... .......... .......... ..........


  HSV-1VP16  ...
  HSV-2VP16  ...
EHV-1gene12  DI*
EHV-1gene5B  DI*
  BHV-1UL48  DF*
   VZVORF10  ...
    MDVUL48  ...
```

Figure 2. Linear alignment of the amino acid sequence of VP16 and its homologues. The sequence of all the proteins were analyzed using the pileup program in the GCG. Gap weight was set at 3.0 and Gap length weight of 0.1. Both N and C-termini of the homologues are not conserved in contrast to the central part of the proteins. A dot (.) indicates that the residue at that position is identical in at least four members of the homologues, a colon (:) indicates that the residue at that position is identical in all homologues.

addition to the net negative charge of the activation domain, several key aromatic and bulky hydrophobic residues play a critical role in the function of VP16. Therefore, to identify activation domains we have employed the HCA program which specifically identifies the conserved hydrophobic residues and clusters, to analyze and compare the homologues of VP16. We hypothesized that if these key aromatic and bulky hydrophobic residues are essential for activity, the HCA pattern of these residues should be conserved in VP16 homologues.

The question of whether or not these homologues of VP16 work as transcription activators still remains. Moriuchi *et al.* (1993) demonstrated that VZV ORF10 despite lacking the 80 amino acid C-terminal residues, was able to activate transcription of the viral IE promoter. This result suggests that despite the differences in the primary structure of this VP16 homologue, it is able to activate transcription. It is possible that it might also retain similar activation domains.

The sequences of HSV-1 VP16, HSV-2 VP16, BHV-1 UL48, VZV ORF10, MDV UL48, EHV-1 Gene12, and EHV-4 GeneB5 were subjected to HCA. The results are shown in Figure 3. The proteins were aligned by the conserved clusters of the homologue proteins rather than the position of amino acid residues in the primary structure. The HCA plot demonstrated that the most conserved portion among the homologues is the central portion of the protein corresponding to residues 50-390 of HSV-1 VP16. The terminal portions, both N-terminal and C-terminal, are less conserved.

The activation domains of HSV-1 VP16, N-subdomain (residues 410-456) and C-subdomain (residues 456-490) have distinct HCA patterns as shown in Figure 4 and 5. The pattern seen in the N-subdomain of VP16 from HSV-1 is a ladder like cluster of Leu residues, between residues 410-430, and a horseshoe-like shape with the critical Phe residue at the apex of the horseshoe cluster,

Figure 3. HCA of the complete sequence of VP16 and other VP16 homologues. The complete deduced sequences of HSV-1 VP16, HSV-2 VP16, ORF10, equine herpes virus I (EHV-1) Gene12, EHV-4 GeneB5, bovine herpes virus 1 (BHV-1) UL48, and Marek's disease virus (MDV) UL48 proteins were analyzed by HCA. The amino acid sequences are represented (in duplicate) along diagonals from upper left to lower right with the one-letter amino acid code except that black diamonds indicate glycine, squares indicate threonine, squares with central dots indicate serine, and stars indicate proline residues. Sets of adjacent hydrophobic residues are encompassed by contour lines. The HCA patterns were aligned by the conserved hydrophobic clusters rather than the position of the residues in their linear sequence. The vertical lines show the boundary of each conserved region.


HSV-1 VP16
HSV-2 VP16
VZV ORF10
MDV UL48
EHV-1 Gene12
EHV-4 GeneB5
BHV UL48

Figure 3

between residues 430-456 (Figure 4). Since this cluster contains the critical Phe residue and was also shown to be able to activate transcription (Emami & Carey, 1992), I have used this horseshoe shape pattern to identify the N-subdomain in other homologues. The most obvious homology was found in the BHV-1 UL48 N-terminus between residues 22-46, with the Phe residue predicted to be critical for the activity at Phe residue 33. Though the horseshoe cluster in VZV ORF10 was not as obvious, in conjunction with the linear alignment one can clearly see the conserved Phe residue, residue 28, as well as other conserved hydrophobic and acidic residues. Careful observation of the sequences of other homologues (except HSV-2) revealed that all homologues contain a fragment in the N-terminus of the protein that is homologous to the N-subdomain of HSV-1 VP16, shown in Figure 4A.

Linear sequence comparison of all the VP16 homologues in Figure 4B clearly shows the conserved critical Phe residue as well as hydrophobic and acidic residues in its vicinity. The critical Phe found at the apex of the horseshoe like shape of the cluster was conserved and used as the core of the alignment. Bulky hydrophobic residues corresponding to Leu residue 439, Met residue 446 and Leu residue 447 of HSV-1 VP16 were conserved in all the homologues except for the residue corresponding to Leu residue 444. This position in other homologues was found to be a conserved acidic residue, either Asp or Glu. Though the position of most of the acidic residues is not conserved among all the homologues, a single Asp residue following the critical Phe was found to be conserved. Therefore this Asp residue was also used, together with the critical Phe residue, to identify and align the N-subdomain in VP16 homologues.

The pattern of hydrophobic clusters in the C-subdomain is distinct from the N-subdomain and appears as a pair of ladder like clusters followed by a triangular cluster, shown in Figure 5A. Though both the N and C subdomain

Figure 4. HCA of the N-subdomain in VP16, ORF1O, and other VP16 homologues. The complete deduced sequence of HSV-1 VP16, HSV-2 VP16, ORF10, equine herpes virus I (EHV-1) Gene12, EHV-4 GeneB5, bovine herpes virus 1 (BHV-1) UL48, and Marek's disease virus (MDV) UL48 proteins were analyzed by HCA, and locations of activation domains were predicted by using VP16 as a standard. (A) HCA profiles of subdomains resembling VP16N show the characteristic horseshoe-shaped cluster and the centrally located Phe residues. (B) Linear alignment of the subdomains shown in A, with shading indicating conserved residues, boldface letters indicating conserved bulky hydrophobic residues, dark shading denoting conserved acidic residues, and open ovals representing residues conserved among homologues but not present in HSV-1 VP16.

A.
VP16-1
VP16-2
VZV ORF10
EHV-1 UL48
MDV UL48
EHV-1 GENE12
EHV-4 GENE85
B.
VP16-1 431 VAMAHADALDDFDLDMLGDGDSPGP 455
VP16-2 432 VDMTPADALDDFDLEMLGDVESPSP 456
VZV ORF10 17 RSKTDEQAVVDAFDESLFGDVASDIG 41
EHV-1 UL48 22 ATMDPYDAIEAFDDSLLGSPLAAGP 46
MDV UL48 32 PDLDILRTIEEFDETLLSETEVRTQ 56
EHV-1 GENE12 53 ELMDMDGAVASFDEGMLSALESVYS 77
EHV-4 GENE85 28 ELMDVDGVVASFDEGMLSASESIYS 52

Figure 4

seem to have ladder like clusters, they are formed by different adjacent residues in the HCA plot giving different cluster orientations. Careful observation of the VP16 homologues has also revealed similar clusters located in the C-terminal portion of HSV-2 VP16, BHV-1 UL48, EHV-1 Gene12, and EHV-4 GeneB5 shown in Figure 5A. Linear alignment of the primary sequence also shows conserved bulky hydrophobic residue corresponding to residues Met470, Phe473, Met474, Phe475, as well as acidic and polar residues in the vicinity (Figure 5B). Many of the conserved hydrophobic residues in this alignment have been shown by alanine scanning and random mutagenesis of this C-subdomain in HSV-1 VP16 to be important for activity of the C-subdomain (Susan Sullivan and Peter Horn unpublished results).

HCA predicts that all the VP16 homologues contain a conserved N-subdomain. The N-subdomain is located at the N-terminus of the VP16 homologue in BHV-1, EHV-1, EHV-4, VZV and MDV, and at the C-terminus in HSV-1 and HSV-2. The C-subdomain, however, is not conserved in all homologues. It is only present in HSV-1, HSV-2, BHV-1, EHV-1 and EHV-4 VP16 homologues, and is located at the C-terminus of all the proteins. The location of the predicted N and C subdomain of each homologue is shown in Figure 6.

Site directed oligonucleotide mutagenesis of the proposed VZV ORF10 activation domain

Results from the HCA of the VP16 homologues suggest that these proteins, except for HSV-2 VP16, contain the N-subdomain located near the N-terminus of the protein. I tested this hypothesis using VZV ORF10, already known to activate transcription from its IE promoter as well as the HSV-1 ICP0

Figure 5. HCA of the C-subdomain in VP16 and other VP16 homologues. (A) HCA profiles of subdomains resembling VP16C, drawn as in Figure 4A. (B) Linear sequence alignment of the subdomains shown in A, with shading, ovals, and boldface type as described in the legend of Figure 4B.

A.

B.

```
VP16-1          S A P Y G A L D M A D F E F E Q M F T D A L G I D E Y G G
                462                                                   490
VP16-2          P V S Y G A L D V D D F E F E Q M F T D A M G I D D F G G
                462                                                   490
EHV-1 GENE12    T D G L D Y S S M T G D E L N Q M F D I
                460                                   479
EHV-4 GENE85    T D V L D Y S S I S G D E L N Q M F D I
                435                                   454
BHV-1 UL48      P F G G T Y D A L L G D R L N Q L L D F
                486                                   505
```

Figure 5

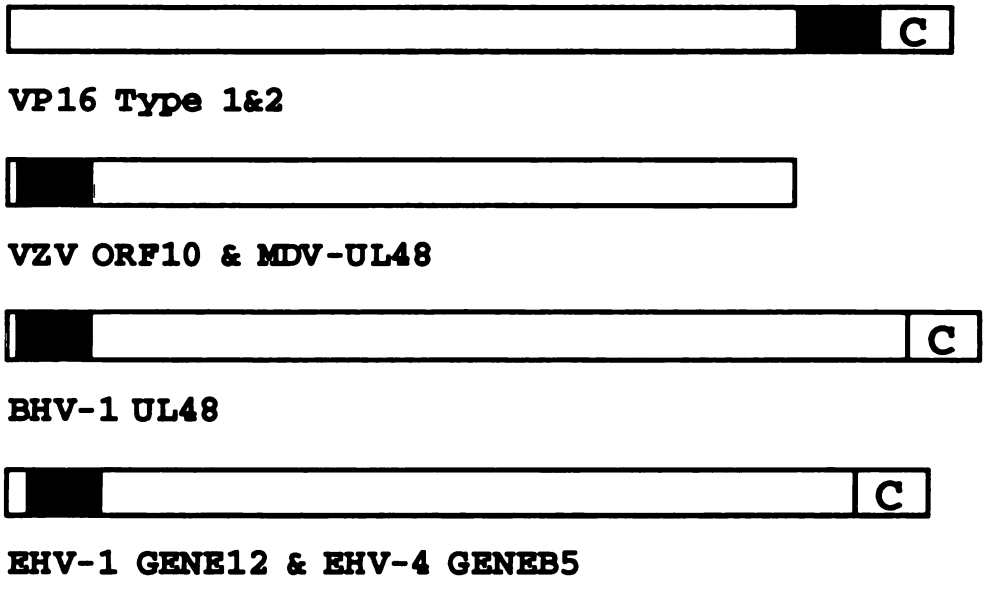


Figure 6. Locations of the N and C subdomains in VP16 and its homologues. HCA was used to analyze the deduced protein sequence of VP16 and its homologues. We predict the location of both activation subdomains of VP16 type 1 and 2 to be at the C-terminus of the protein. In other homologues the N-subdomain was predicted to be located at the N-terminus of the protein while the C-subdomain remains at the C-terminus of the protein (except MDV UL48 and VZV ORF10 where the C-subdomain is not present in the protein).

and ICP4 promoters (Moriuchi, et al., 1993). I proposed that if the activation domain of VZV ORF10 is indeed located near the N-terminus of the protein and Phe residue 28 (Phe28) is the critical residue for its activity, mutations of the Phe28 and other hydrophobic residues in the vicinity would reduce or abolish the ability of VZV ORF10 to activate transcription.

The activation domain of VZV ORF10 was initially mapped to the N-terminal portion of the protein, between residues 5-79, by H. Moriuchi. This was accomplished by fusing different deleted fragments of VZV ORF10 to the DNA binding portion of the GAL4 protein (residue 1-147) and testing these fusion proteins for the ability to activate expression of CAT from the reporter plasmid p5GAL4-E1b-CAT. VZV ORF10 fragments containing residues 267-410 and 113-351 did not show any significant activity. When a fragment containing residues 5-175 was tested, it showed approximately 100 fold activation over the control GAL4 DNA binding domain (residues 1-147). Further deletion of the VZV ORF10 fragment to residues 5-79 resulted in an increase in activity to approximately 300 fold. Dissection of residues 5-79 to residues 5-41 and 38-79 suggested that the activation domain lies within residues 38-79 (results shown in Appendix A). However, substitution of large bulky hydrophobic residues in the 38-79 region, Ile65, Leu66, Tyr67, Leu70 and Leu74 to Ala, by H. & M. Moriuchi, did not reduce the activity of GAL4-ORF10(5-79). Therefore the exact location and critical residues involved in transcription activation were unknown. This was not surprising since some random fragments from *E.coli* DNA fused to the GAL4 DNA binding domain were found to be able to activate transcription (Ma & Ptashne, 1987). Therefore interpretation of these results must be made with caution.

Although these results did not clearly identify essential residues or the exact location of the activation domain, they were sufficient to further strengthen

my hypothesis that the activation domain of VZV ORF10 is indeed located in the N-terminal portion of the protein.

I mutated the proposed critical residue for transcription function Phe28 by site directed mutagenesis. Phe28 was changed to other bulky hydrophobic residues, Tyr, Leu, Ile, smaller hydrophobic residues Val and Ala; hydrophilic residues Ser and Thr; and Pro which may disrupt secondary structures. I also designed oligonucleotides to make mutations of the flanking hydrophobic residues Val24, Val25, Leu32, and Phe33. Since a bulky hydrophobic residue at Val24 is not conserved at that position (Figure 4) I predicted that changes at Val24 should not affect the ability of GAL4-ORF10(5-79) to activate transcription.

Mutations of Phe28 generated in M13 phage were subcloned into the GAL4 fusion vector. The constructs were tested by H. Moriuchi using transient expression assays. A reporter plasmid p5GAL4-E1B-CAT was cotransfected into Vero cells with the wild type or mutated versions of the plasmid expressing GAL4-ORF10 fusion protein, and CAT activities were determined (details in Appendix A).

The results shown in Table 2 clearly demonstrated the importance of Phe28 as a critical residue in the VZV ORF10 activation domain. Mutation of Phe28 to other bulky hydrophobic residues, Tyr, Leu, and Ile reduced the activity of GAL4-ORF10 to approximately 50% of the wild type fusion protein. Substitution by the smaller hydrophobic residue Val further reduced the activity of GAL4-ORF10 to approximately 20% of wild type. Substitution by Ala and small hydrophilic residues, Ser and Thr reduce the activity to less than 15% of wild type chimeric protein. Curiously, the Pro substitution did not abolish the activity of the fusion protein, but gave similar activity to Ala and small hydrophilic residues. This shows that the structural constraint imposed by the substitution of a Pro residue at Phe28 did not have any further effect than

Table 2

Relative activity of GAL4-ORF10(5-79) bearing amino acid substitutions at position 28

Residue substituted at position 28	Relative activity (% of wild-type)
Tyr	52 ± 17
Leu	54 ± 17
Ile	50 ± 12
Val	23 ± 4
Ala	11 ± 2
Pro	13 ± 3
Ser	9 ± 4
Thr	7 ± 2

Table 2. Relative activity of GAL4-ORF10(5-79) bearing amino acid substitutions at position 28. Vero cells were cotransfected with 5 μg of plasmid expressing an amino acid substitution mutant of GAL4-ORF10(5-79) and 5 μg of a reporter plasmid containing 5 GAL4 binding sites. The relative activity is represented as the percent CAT activity of the mutant relative to that obtained from the wild type GAL4-ORF10(5-79). Means and standard deviations were calculated from at least five independent transfections.

substitution by a small hydrophobic residue such as Ala (Table 2).

I have shown that in the fusion protein GAL4-ORF10(5-79) the residue responsible for the activation activity lie within amino acid residues 5-79 from VZV ORF10 and that Phe28 is the critical residue for its activity. Mutation of this residue to small hydrophobic residues or hydrophilic residues greatly reduces the activity of this domain. These results lead to another very important question. Does this domain in the context of VZV ORF10 (rather than as a GAL4 fusion) function as an activation domain with Phe28 as a critical residue for its function.

To demonstrate that this domain is essential for VZV ORF10 activation, a mutation was constructed in which Phe28 was substituted with Ala in the context of VZV ORF10 protein. The wild type VZV ORF10 gene and VZV ORF10 bearing the substitution of Phe28 to Ala were subcloned into the pCMV expression vector and co-transfected into Vero cells with a reporter plasmid, containing the VZV ORF62 promoter (IE promoter) linked to a CAT gene. A plasmid derivative of pCMV bearing the HSV-1 VP16 gene with truncations of residues 413-490 as well as the parental pCMV vector were used as controls. The substitution of Phe28 to Ala in VZV ORF10 rendered the protein inactive in the transient expression assay. The activity of the mutant was at the same level as HSV-1 VP16 with deletions of amino acid residues 410 to 490 (Table 3). This clearly demonstrates the significance of Phe 28 in the ability of VZV ORF10 to activate transcription, and confirms our hypothesis that this domain is the VZV ORF10 activation domain.

Our laboratory has demonstrated the importance of hydrophobic residue Phe442 for the activation function of the HSV-1 VP16 N-subdomain, and has successfully used patterns of hydrophobic residues to identify the activation domain of its homologue VZV ORF10. In addition to the critical Phe residue

Table 3

Phe28 is required for transactivation of the VZV IE62 promoter by ORF10 protein

Activator	Fold induction
pCMV10	24 ± 3
PCMV10F28A	1.0 ± 0.1
pCMV16	290 ± 50
pCMV16d413-490	0.8 ± 0.1

Table 3 Phe28 is required for transactivation of the VZV IE62 promoter by ORF10 protein. Cells were cotransfected with plasmids expressing native ORF10(pCMV10), an amino acid substitution mutant of ORF10 (pCMV10F28A), full length VP16 (pCMV16) or a carboxy terminal truncation of VP16 (pMCV16d413-490) together with p62CAT (the VZV ORF62 promoter followed by a CAT gene). Fold induction is the CAT activity in transfected cell extracts relative to that obtained when using pCMV (vector control). Means and standard deviations were calculated from at least five independent transfections.

there are also other flanking hydrophobic residues present in the horseshoe like cluster (Figure 4). These residues are conserved and may play a role in the activation function. I have generated mutations of these flanking hydrophobic residues, Val25, Leu32, Phe 33 and Val24, substituting them with small hydrophobic and hydrophilic residues, Ala and Ser. Substitution at Val24 serves as a control, since bulky hydrophobic residues are not conserved at this position. A structurally conserved aromatic but slightly polar residue, Tyr, was also substituted at Phe33. These derivatives were then tested for their ability to activate transcription as GAL4-ORF10 fusion proteins, as previously described.

Substitution of Leu32 and Phe33 to Ala and Ser reduced the activity of the fusion protein to approximately 50% and 10%, respectively (Table 4). These results suggest bulky hydrophobic residues at position 32 and 33 are both important for its activity . When Tyr, another bulky hydrophobic residue, was substituted at Phe33 the activity of the GAL4-ORF10 fusion protein dropped only 25%. The reduction in activity could be a result of the slightly polar character of Tyr compared to Phe (Table 4). Curiously, substitution of Val25 to Ala had no significant effect on the activity of the fusion protein. This result was unexpected since bulky hydrophobic residues were found to be conserved at this position in the alignment shown in Figure 4. This result suggests that there is no requirement for a bulky hydrophobic residue at this position, at least in this GAL4-ORF10(5-79) chimeric construct. Substitution at Val24 to Ala did not diminished activation as expected.

These results clearly demonstrated that the activation domain of VZV ORF10 is indeed near the N-terminus of VZV ORF10 protein. The results also clearly demonstrated the importance of Phe28 for the activation function, both in the context of a chimeric GAL4-ORF10 fusion protein and the wild type VZV ORF10 protein. Flanking hydrophobic residues, at position 32 and 33, were also

Table 4

Relative activity of GAL4-ORF10(5-79) bearing amino acid substitutions at hydrophobic residues flanking Phe28

Amino acid substitution	Relative activity (% of wild-type)
Val24 to Ala	140 ± 40
Val25 to Ala	130 ± 50
Leu32 to Ala	45 ± 18
Phe33 to Ser	12 ± 4
Phe33 to Tyr	76 ± 18

Table 4. Relative activity of GAL4-ORF10(5-79) bearing amino acid substitutions at hydrophobic residues flanking Phe28. Vero cells were cotransfected with 5 μg of plasmid expressing an amino acid substitution mutant, of hydrophobic residues flanking Phe28 of GAL4-ORF10(5-79), and 5 μg of a reporter plasmid containing 5 GAL4 binding sites. The relative activity is represented as the percent CAT activity of the mutant relative to that of the wild type GAL4-ORF10(5-79). Means and standard deviations were calculated from at least five independent transfections.

shown to be important for activation by VZV ORF10, at least in the GAL4-ORF10 chimeric protein. These results support the hypothesis that the activation domain of VZV ORF10 is homologous to the activation domain of the HSV-1 VP16, N-subdomain. This strongly suggests that my predictions of the activation domain of the other VP16 homologues are likely to be correct.

Secondary structure prediction by Chou-Fasman and Garnier-Osguthorpe-Robson method

Previously, the Chou-Fasman and Garnier-Osguthorpe-Robson methods were used to predict the structure of HSV-1 VP16 alone (W. D. Cress unpublished results). In this study, the Chou-Fasman and Garnier-Osguthorpe-Robson methods were used to predict the secondary structure of the N-subdomain of all the homologues. We expected that introduction of the N-subdomain homologues to the prediction might result in some other predictions that agree with other current mutational and biophysical studies on the activation domain of VP16.

The prediction results from Chou-Fasman and Garnier-Osguthorpe-Robson method, shown in Table 5, suggest that the residues around the critical Phe residue lie within a region of the protein that is likely to form α helical structure followed by a turn. The only exception to this prediction is EHV-4 GeneB5, where the Chou-Fasman method predicted the sequence around Phe39 to be β sheet. Interestingly the critical Phe residue in all the predictions is found to be highly exposed to solvent, with a surface probability according to Emini et al. of higher than 1.0, with the exception of the EHV-1 and EHV-4 homologues for which the value is slightly less, at 0.77. This result suggests that the critical Phe residue is likely to be solvent exposed.

Table 5. Secondary structure prediction by Chou-Fasman and Garnier-Osguthorpe-Robson, and Emini surface probability prediction of the N-subdomain of HSV-1 VP16, HSV-2 VP16, VZV ORF10, BHV-1 UL48, EHV-1 Gene12, EHV-4 GeneB5 and MDV UL48. The sequences of the N-subdomain of HSV-1 VP16 (residues 341-455), HSV-2 VP16 (residues 432-456), VZV ORF10 (residues 17-41), BHV-1 (residues 22-46), EHV-1 Gene12 (residues 53-77), EHV-4 GeneB5 (residues 28-52) and MDV UL48 (residues 32-56) were subjected to secondary structure prediction by the peptidestructure program from GCG. The calculations were based on the Chou-Fasman and Garnier-Osguthorpe-Robson method. The surface probability of each residue was also calculated by the method of Emini *et al.* Abbreviations: **Position**, Position of amino acid; **Amino acid**, Amino acid residue; **Surf Pr.**, Surface probability; **CF**, Chou-Fasman prediction; **GOR**, Garnier-Osguthorpe-Robson prediction. The output of the prediction is reported as: "H", for strong α helical prediction; "h", for weak helical prediction; "B", for strong β sheet prediction; "b", for weak β sheet prediction; "T", for strong turn prediction; "t", for weak turn prediction; "." for residues that no prediction was made.

Table 5

Chou-Fasman and Garnier-Osguthorpe-Robson secondary structure prediction

HSV-1 VP16					HSV-2 VP16					VZV ORF10				
Position	Amino acid	Surf	CF	GOR	Position	Amino acid	Surf	CF	GOR	Position	Amino acid	Surf	CF	GOR
431	V	0.28	H	H	432	V	0.66	.	B	17	R	2.84	h	.
432	A	0.30	H	H	433	D	0.80	.	B	18	S	3.84	h	.
433	M	0.24	H	H	434	M	0.63	H	B	19	K	5.21	h	.
434	A	0.53	H	H	435	T	1.43	H	B	20	T	2.69	h	H
435	H	0.53	H	H	436	P	0.86	H	H	21	E	1.49	h	H
436	A	0.44	H	H	437	A	0.72	H	H	22	Q	0.55	h	H
437	D	0.73	H	H	438	D	0.83	H	H	23	A	0.64	h	H
438	A	0.9	H	H	439	A	0.90	H	H	24	V	0.37	h	H
439	L	0.77	H	H	440	L	0.77	H	H	25	V	0.19	h	H
440	D	0.77	H	H	441	D	0.77	H	H	26	D	0.31	h	H
441	D	0.63	H	H	442	D	0.63	H	H	27	A	0.72	h	H
442	F	1.27	H	H	443	F	1.32	H	H	28	F	1.30	h	H
443	D	0.75	H	H	444	D	0.78	H	H	29	D	0.64	h	H
444	L	0.37	H	H	445	L	0.39	H	H	30	E	0.55	h	H
445	D	0.43	H	H	446	E	0.44	H	H	31	S	0.63	h	H
446	M	0.42	H	H	447	M	0.44	H	H	32	L	0.63	h	H
447	L	0.51	H	T	448	L	0.40	H	H	33	F	0.27	h	H
448	G	0.51	t	T	449	G	0.40	h	H	34	G	0.20	h	H
449	D	0.69	T	T	450	D	0.54	h	H	35	D	0.33	h	H
450	G	1.30	T	T	451	V	1.01	h	.	36	V	0.64	h	H
451	D	1.30	T	.	452	E	1.37	h	.	37	A	0.45	h	H
452	S	1.20	t	.	453	S	1.27	h	.	38	S	0.27	h	.
453	P	1.55	T	.	454	P	2.18	T	.	39	D	0.46	h	T
454	G	1.19	T	.	455	S	1.61	T	.	40	I	0.58	h	T
455	P	1.13	.	.	456	P	1.53	.	.	41	G	0.56	.	T

Table 5 (cont'd)

Chou-Fasman and Garnier-Osguthorpe-Robson secondary structure prediction

BHV-1 UL48					EHV-1 GENE12					EHV-4 GENEB5				
Position	Amino acid	Surf	CF	GOR	Position	Amino acid	Surf	CF	GOR	Position	Amino acid	Surf	CF	GOR
22	A	0.90	.	.	53	E	0.88	H	H	28	E	0.88	h	H
23	T	1.10	.	.	54	L	0.68	H	H	29	L	0.51	h	H
24	M	1.34	t	.	55	M	0.89	H	H	30	M	0.67	h	H
25	D	2.21	t	.	56	D	0.51	H	H	31	D	0.38	h	H
26	P	1.55	T	.	57	M	0.63	H	H	32	V	0.35	h	H
27	Y	1.10	T	H	58	D	0.47	H	H	33	D	0.26	h	H
28	D	1.14	H	H	59	G	0.28	H	H	34	G	0.16	h	H
29	A	0.74	H	H	60	A	0.39	H	H	35	V	0.28	B	H
30	I	0.41	H	H	61	V	0.20	H	H	36	V	0.15	B	H
31	E	0.41	H	H	62	A	0.34	H	H	37	A	0.25	B	H
32	A	0.68	H	H	63	S	0.58	H	H	38	S	0.58	B	H
33	F	1.30	H	H	64	F	0.77	H	H	39	F	0.77	B	H
34	D	0.62	t	H	65	D	0.75	H	H	40	D	0.75	h	H
35	D	0.51	t	H	66	E	0.46	H	H	41	E	0.46	h	H
36	S	0.58	t	.	67	G	0.72	h	H	42	G	0.72	h	H
37	L	0.46	.	.	68	M	0.43	h	H	43	M	0.43	h	H
38	L	0.43	.	.	69	L	0.18	h	H	44	L	0.34	h	H
39	G	0.26	.	.	70	S	0.31	h	H	45	S	0.59	h	H
40	S	0.32	.	.	71	A	0.42	h	H	46	A	0.80	h	H
41	P	0.40	t	.	72	I	0.38	h	H	47	S	0.68	h	H
42	L	0.40	t	.	73	E	0.44	h	H	48	E	0.79	h	H
43	A	0.46	.	.	74	S	0.58	h	H	49	S	1.05	h	H
44	A	0.38	.	.	75	V	1.06	h	H	50	I	1.00	h	T
45	G	0.59	.	.	76	Y	0.78	.	T	51	Y	0.74	.	T
46	P	0.74	.	.	77	S	0.75	.	T	52	S	0.71	.	T

Table 5 (cont'd)

Chou-Fasman and Garnier-Osguthorpe-Robson secondary structure prediction

MDV UL48				
Position	Amino acid	Surf	CF	GOR
32	P	1.33	h	.
33	D	0.73	h	H
34	L	0.47	h	H
35	D	0.60	h	H
36	I	0.52	h	H
37	L	0.44	h	H
38	R	0.46	h	H
39	T	1.12	h	H
40	I	1.18	h	H
41	E	1.01	h	H
42	E	1.21	h	H
43	F	2.49	h	H
44	D	1.18	h	H
45	E	0.56	h	H
46	T	0.87	h	H
47	L	0.90	h	H
48	L	0.37	h	H
49	S	0.44	h	H
50	E	0.40	h	H
51	I	0.94	h	H
52	E	1.01	h	H
53	V	1.01	h	H
54	R	1.84	h	H
55	T	1.36	h	H
56	Q	2.34	h	H

Secondary structure prediction by Sequery and PHD

I have employed two new methods of secondary structure prediction which have a higher degree of accuracy to analyze the N-subdomain, expecting to find structural predictions compatible to our mutational and biophysical analysis. Recently a new method of prediction, the PHD program, which uses a computer neural network and yields a higher degree of accuracy was introduced (Rost & Sander, 1994). The prediction by the PHD program was used in this study, since it has a degree of accuracy (from 70% to 88%) for three-state prediction, α helix, β sheet and loop (as high as the overall average for homologous modeling). PHD uses the statistical information and the structural information available in the present data base in conjunction with the evolutionary information contained in multiple sequence alignments as input to the neural network.

PHD prediction results show the input amino acid sequence, the overall structural prediction for each amino acid residue and the detailed structure predictions for each of the amino acids in the given sequence. PHD results are shown in Tables 6 and 7. PHD did not suggest any strong prediction of a particular secondary structure around the critical Phe residue. Predictions of the secondary structure in the different homologues were made around the critical Phe , but the reliability was equal or less than 5 from a scale of 0-9. In most cases no predictions were made in this region. Both the N and C-terminus of the N-subdomain in all homologues were strongly predicted to be a loop structure with a reliability of 7-9 in most cases. A strong prediction of α helical structure was found in a stretch of 6-7 residues N-terminal to the critical Phe in VZV ORF10, MDV UL48, and BHV-1 UL48, but the degree of reliability drops dramatically at the critical Phe residue and following residues (Tables 6 and 7).

Table 6. PHD secondary structure predictions of the N-subdomain of HSV-1 VP16, HSV-2 VP16, BHV-1 UL48 and EHV-1 Gene12. The amino acid sequences of the N-subdomain of these homologues were subjected to the PHD secondary structure prediction program (Rost *et. al.*, 1994). Abbreviations: secondary structure; H=helix, E=extended (sheet), blank=rest (loop); AA, amino acid sequence; PHD, Profile network prediction Heidelberg; Rel, Reliability index of prediction (0-9); detail: prH, 'probability' for assigning helix; prE, 'probability' for assigning strand; prL, 'probability' for assigning loop (note: the 'probabilites' are scaled to the interval 0-9); subset: SUB, a subset of the prediction, for this subset the following symbols are used, L, is loop (for which above " " (space) is used), ".", means that no prediction is made for this residue, as the reliability is Rel < 5

Table 6

PHD secondary structure prediction of the N-subdomain of HSV-1 VP16, HSV-2 VP16, BHV-1 UL48 and EHV-1 Gene12

Protein: HSV-1 VP16, N-subdomain Length 25

```
          ....,....1....,....2....,....3
AA       |VAMAHADALDDFDLDMLGDGDSPGP|
PHD      |           HHHHHH         |
Rel      |98996577662315413868999 99|
detail:
prH-     |00002211125556653021000 00|
prE-     |01000000000000000000000 00|
prL-     |98997787773342345878999 99|
subset:  SUB |LLLLLLLLLL...H...LLLLLLLL|
```

Protein: HSV-2 VP16, N-subdomain Length 25

```
          ....,....1....,....2....,....3
AA       |VDMTPADALDDFDLEMLGDVESPSP|
PHD      |                          |
Rel      |97785456654653258999999 89|
detail:
prH-     |00002211122223320000000 00|
prE-     |01100010000000000000000 00|
prL-     |98886677776776578999999 99|
subset:  SUB |LLLLL.LLLL.LL..LLLLLLLLLL|
```

Protein: BHV-1 UL48, N-subdomain Length 25

```
          ....,....1....,....2....,....3
AA       |ATMDPYDAIEAFDDSLLGSPLAAGP|
PHD      |    HHHHHHHHHH            |
Rel      |99994889999721214988889 99|
detail:
prH-     |00006889999753420011100 00|
prE-     |00000000000000100000000 00|
prL-     |99993110000144446888889 99|
subset:  SUB |LLLL.HHHHHHH.....LLLLLLLL|
```

Protein: EHV-1 Gene12, N-subdomain Length 25

```
          ....,....1....,....2....,....3
AA       |ELMDMDGAVASFDEGMLSAIESVYS|
PHD      |         EEEE HHHH HHHHH  |
Rel      |91486773134214431125651 39|
detail:
prH-     |00001112222236653346763 10|
prE-     |04201002345420013220112 30|
prL-     |95687875322342232322123 59|
subset:  SUB |L..LLLL...........HHH..L|
```

Table 7. PHD secondary structure predictions of the N-subdomain of EHV-4 GeneB5, MDV UL48 and VZV ORF10. The amino acid sequence of the N-subdomain of these homologues were subjected to the PHD secondary structure prediction program (Rost *et. al.*, 1994). Abbreviations: secondary structure; H=helix, E=extended (sheet), blank=rest (loop); AA, amino acid sequence; PHD, Profile network prediction Heidelberg; Rel, Reliability index of prediction (0-9); detail: prH, 'probability' for assigning helix; prE, 'probability' for assigning strand; prL, 'probability' for assigning loop (note: the 'probabilites' are scaled to the interval 0-9); subset: SUB, a subset of the prediction, for this subset the following symbols are used, L, is loop (for which above " " (space) is used), ".", means that no prediction is made for this residue, as the reliability is Rel < 5

Table 7

PHD secondary structure prediction of the N-subdomain of EHV-4 GeneB5, MDV UL48 and VZV ORF10

```
Protein:      EHV-4 GeneB5, N-subdomain      Length    25

                  ...,....1....,....2....,....3
         AA   |ELMDVDGVVASFDEGMLSASESIYS|
         PHD  |          EEEE     EEEEEEEE |
         Rel  |946867722222123154534466 8
 detail:
         prH- |000011122222321101222000
         prE- |021010024444311366656577 1
         prL- |967878743223445422121212 8
 subset: SUB  |L.LLLLL.........E.E...EEL|
```

```
Protein:      MDV UL48, N-subdomain      Length    25

                  ...,....1....,....2....,....3
         AA   |PDLDILRTIEEFDETLLSEIEVRTQ|
         PHD  |     HHHHHHHHHHHHHHHHEEEE  |
         Rel  |998267448975555654524334 9
 detail:
         prH- |000377668987677766632221 0
         prE- |000000120000000021146552 0
         prL- |998511210012222112111125 9
 subset: SUB  |LLL.HH..HHHHHHHHH.H.....L|
```

```
Protein:      VZV ORF10, N-subdomain   Length    25

                  ...,....1....,....2....,....3
         AA   |RSKTEQAVVDAFDESLFGDVASDIG|
         PHD  |    HHHHHHHHH          EEE   |
         Rel  |995799999873112245312217 9
 detail:
         prH- |002899999886455432231110 0
         prE- |000000000000000000125441 0
         prL- |997100000113544566533347 9
 subset: SUB  |LLLHHHHHHHH......L.....LL|
```

A method termed Sequery was also used. Secondary structure prediction by Sequery was done using a sliding window of four residues from residues 439 to 447 in HSV-1 VP16. The residues allowed at each position were selected based on the conserved residues from the homologue proteins. The residues allowed at each position are shown in Figure 1. The results are shown in Table 8. The sample size indicates the number of proteins with matching sequence resulting from the search (the list of the proteins is shown in the Appendix B). The structures of the query sequence observed from matching proteins for region 439-442 shows equal probability of α helix and loop/extended structure. This suggests that the segment probably does not have strong preference for any secondary structure. This region was followed by an extended loop or strand structure, encompassing the critical Phe residue, residues 440-445, followed by a strong α helical prediction at residues 444-447. These results suggest that the residues around Phe 442 are likely to form an extended structure, either a loop or strand, since the extended structure found did not have the classical hydrogen bonding pattern of β sheet structures.

Discussion

Using HCA to compare the sequence of VP16 and its homologues, we predicted the location of the activation domains of these homologues in a way that was not obvious by the classical linear alignment method (Figure 4-5). We tested our hypothesis using the VP16 homologue VZV ORF10. VZV ORF10 was the first homologue that was found to have the ability to activate transcription (Moriuchi, et al., 1993). It functions as an activator despite the lack of the 80 amino acids at the C-terminus that were found to be necessary and sufficient for transcription activation in HSV-1 VP16 (Sadowski, et al., 1988; Triezenberg, et al.,

Table 8

Secondary structure prediction by Sequery

Search window	Amino acid residues	Sample size	Secondary structure (%)			
			α helix	β sheet	Loop/ Extend	Turn
1	439-442	9	45	0	55	0
2	440-443	7	14	0	86	0
3	441-444	12	8	17	67	8
4	442-445	9	0	33	56	11
5	443-446	12	58	0	42	0
6	444-447	20	95	0	0	5

Table 8. Secondary structure prediction by Sequery. The favorable structure of the string of sequence searched by Sequery was calculated. The sample size is the number of similar strings of sequence that Sequery found from the search. For each string, the corresponding protein structure was called up and viewed using the computer program Insight II (Version 2.3.0, Biosym Technologies, 1993). The structure of the particular sequence string was recorded and used to calculate the most favorable structure of the string of sequence. The percentage of each structure found at each window of four residues is shown.

1988a). VZV ORF10 was predicted by HCA to contain a region homologous to the N-subdomain of HSV-1 VP16. Deletion mutagenesis and transient expression assays performed by Moriuchi *et al.* confirmed our hypothesis concerning the location of the VZV ORF10 activation domain.

Linear alignment, using HCA as a guide, suggested the critical residues in VZV ORF10 that could be involved in the activation function of VZV ORF10. Phe28 as well as conserved flanking hydrophobic residues were predicted to be important for activation. Site directed oligonucleotide mutagenesis changing Phe28 into other bulky hydrophobic residues, Tyr, Leu, and Ile, medium size hydrophobic residue Val, or small hydrophilic residues Ser and Thr, decreased the activity of GAL4-ORF10(5-79) from 50%-10%, respectively (Table 2). Changing the flanking hydrophobic residues, Leu33 and Phe34 to smaller hydrophobic or hydrophilic residue also decreased the activity of the fusion protein (Table 5).

These results clearly demonstrate that Phe28 is indeed a critical residue essential for VZV ORF10 transcription activity. They also show the importance of the flanking hydrophobic residues which were predicted by HCA and linear alignment. These results confirm our hypothesis that the activation domain of VZV ORF10 lies at the N-terminus of the protein and is homologous to the N-subdomain of HSV-1 VP16.

Since HCA proved to work well to identify the activation domain of VZV ORF10, I extended my predictions further to other VP16 homologues, BHV-1 UL48, EHV-1 Gene12, EHV-4 GeneB5 and MDV UL48. I predicted the location as well as critical residues that are likely to be involved in the activation function of these proteins, shown in Figure 4-5. We predicted that the N and C activation subdomains of BHV-1 UL48 are separated, in contrast to HSV-1 VP16, as shown in Figure 6. The N-subdomain of BHV-1 UL48 is predicted to lie at the N-

terminus of the protein (residues 22-46). Phe33 is predicted to be the critical Phe residue homologous to Phe442 of HSV-1 VP16, with essential flanking hydrophobic residues Ile30, Leu37 and Leu38 (Figure 4B). The C-subdomain of BHV-1 UL48 is located at the C-terminus of the protein (residues 486-505). Leu residues at position 494, 499, 502 and 503, are predicted to be important for its activity (Figure 5B).

We also predicted the activation domains of EHV-1 Gene12 and EHV-4 GeneB5 to be separated, similar to BHV-1 UL48. The N-subdomains of EHV-1 Gene12 and EHV-4 GeneB5 are located at the N-termini of the protein. EHV-1 Gene12 N-subdomain (residues 53-77) has the critical Phe residue at position 64 and EHV-4 GeneB5 N-subdomain (residues 28-52) has the critical Phe residue at position 39 (Figure 4B). The C-subdomain of EHV-1 Gene12 and EHV-4 GeneB5 are located at the C-terminus of the protein. EHV-1 Gene12 C-subdomain (residues 460-479) have conserved bulky hydrophobic residues Met468, Leu473, Met476 and Phe477 and EHV-4 GeneB5 C-subdomain (residues 435-454) have conserved bulky hydrophobic residues Ile443, Leu448, Met451 and Phe452 (Figure 5B).

Previous studies by Peter Horn and Susan Sullivan indicated the importance of the Glu residue 474 in HSV-1 VP16 for activity, therefore, Gln residue 500 in BHV-1 UL48, residue 474 in EHV-1 Gene12 and residue 449 in EHV-4 GeneB5, could also be essential for the activity of these C-subdomains. There are also other conserved hydrophobic, acidic and polar residues within this C-subdomain that might also be important.

Recent studies in BHV-1 UL48 demonstrated that the C-terminal portion of BHV-1 UL48 was able to activate transcription in a GAL4 fusion context. However, deletion of these amino acids from the C-terminus of BHV-1 UL48 did not totally eliminate the ability of the protein to transactivate IE-1, a major BHV-1

α gene promoter (Misra, 1994). These data suggest that BHV-1 UL48 may not be functionally collinear with HSV-1 VP16, since transactivation of BHV-1 UL48 requires both the C-terminus and another region present within the N-terminal portion of the protein. Recently Misra et al. has shown that the N-terminus of BHV-1 UL48 indeed contains another activation domain (personal communication), consistent with my prediction.

Studies on the EHV-1 Gene12 homologue of VP16 were also performed. The activation domain of EHV-1 Gene12 was mapped to the C-terminal end of the protein (Elliott, 1994). A Met residue at position 476 seems to be essential for its activity.

A prediction of the MDV UL48 activation domain location and critical residues was also made. MDV UL48 is similar to VZV ORF10 in that it bears only one of the two subdomains, the N-subdomain, which is located at the N-terminus of the protein. The critical Phe residue is located at position 43 with flanking hydrophobic residues Ile40 and Leu48 and Leu49.

This hypothesis has been tested by Mekki Boussaha in Dr. Paul Coussens group (Molecular Virology Laboratory, Department of Animal Science, Michigan State University). The predicted critical Phe residue 43 was substituted with Tyr, Ala or Pro in the context of wild type MDV UL48 protein. The mutant protein was tested in a transient expression assay. The expression of wild type or derivatives of MDV UL48 was driven by a CMV promoter and the MDV ICP4 promoter linked to the CAT gene was used as a reporter plasmid. When Phe43 was substituted with Ala or Pro, the activity of MDV UL48 dropped to less than 10% of the wild type protein. The activity of the mutant carrying the Tyr substitution, on the other hand, was comparable to that of wild type MDV UL48 protein (M. Boussaha, manuscript in preparation).

Our results together with the findings from other laboratories investigating homologues of VP16 generally supported the predictions of activation subdomains made by HCA, except for the N-subdomain of EHV-1 Gene12 predicted by HCA, which did not have detectable activity in a study by Elliott and co-workers. Interestingly, the activation domain of VP16 homologues appear separated in all homologues other than HSV-1 and HSV-2. The N-subdomain is located at the N-terminus of the protein while the C-subdomain remains at the C-terminus, as summarized in Figure 6. VZV ORF10 and MDV UL48 are different from other homologues in that they lack the C-subdomain.

These findings raise interesting ideas in terms of the structure and function of VP16 and is homologues. It suggests that transcription activators may be equipped with more than a single activation domain. These domains could be located either in tandem or at separate locations in the protein; even at different termini of the protein. We may also perceive the nature of these activation domains as "activation modules" that are self-sufficient and work at different sites within the protein. The results from HCA and mutagenesis of VZV ORF10, also suggest that the N-subdomains from VP16 homologues are likely to have a similar mechanism of transcription activation. However, the mechanism by which these domains work when they are placed in tandem or when they are at opposite ends of the protein may or may not be identical. This is also a very interesting aspect for further investigation.

Knowledge of the structural aspects of a protein is always needed to fully understand the function and mechanism of how a protein works. Activation domains from transcription activator proteins have been an interesting specimen for protein chemists, because of their essential role in gene regulation of living organisms. However, most efforts to determine the basic secondary structure of an activation domain have been unsuccessful. NMR or X-ray crystallographic

analysis has not determined the structure for any activation domain. The domain is believed to be very mobile and unstructured in solution (Shen, et al., 1996a; Shen, et al., 1996b; O'Hare & Williams, 1992; Donaldson & Capone, 1992). Since biophysical means have not yielded any structural information for transcription activation domains, structural predictions based on a group of known transcription activation domains might give insight into this question.

Earlier attempts have been made to predict the secondary structure of the activation domain of VP16. The activation domain of HSV-1 VP16 between residues 421-444 was predicted to be α helical by the classical Chou-Fasman and Garnier-Osguthorpe-Robson method (W. D. Cress unpublished results). The α helical structure was predicted to be amphipathic and this structural feature was proposed to be essential for the activity of the VP16 N-subdomain. In the process of testing the hypothesis Doug Cress introduced Pro residues, via site directed oligonucleotide mutagenesis, substituting His residue 425, and Ala residues 432 and 435, simultaneously. The double substitution, Ala residue 432 and Ala residue 435, was constructed because this region was strongly predicted to presume α helical structure. Neither the single or double substitution had any detectable effect on activation by the VP16 N-subdomain. This demonstrates that α helical structure of the N-subdomain is not likely to be essential for its function. Therefore, it is very unlikely that the N-subdomain assumes α helical structure. This structural prediction for HSV-1 VP16 was contradictory to the mutational analysis results (Cress & Triezenberg, 1991b).

Taking advantage of the collection of N-subdomain homologues identified by HCA, we subjected the sequences of the N-subdomain of all the VP16 homologues to several currently available secondary structure prediction programs, including the classical Chou-Fasman and Garnier-Osguthorpe-Robson method.

The Chou-Fasman method established the statistical conformational potential of all the amino acid residues to form α helix, β sheet and turns , from a set of 15 to 29 proteins. These values are arranged in their hierarchical order and used for the prediction of the structural conformation of a primary sequence given. The structure determination by this method is made by scoring these assigned statistical values, according to a specific set of rules, on a given segment of the protein (Chou & Fasman, 1978).

The Garnier-Osguthorpe-Robson method assigns a window of 17 residues, encompassing 8 residues N-terminal (-8) and 8 residues C-terminal (+8) of the position to be predicted. The statistical conformation potential was calculated for each residue at all of the 17 position within the window (from a set of 26 proteins). The statistical values of a particular residue are different depending on the position of the residue within the window. These values are used in the Garnier-Osguthorpe-Robson predictions (Garnier, et al., 1978). Therefore the prediction by Garnier-Osguthorpe-Robson takes into account the effects of the position of the residues from -8 to +8 upon the conformational potential of each residue, where the Chou-Fasman prediction has a fixed conformational potential for each residue. However, both methods have a comparable degree of accuracy.

The Chou-Fasman and Garnier-Osguthorpe-Robson methods predicted the sequence around the critical Phe residue in most of the homologues to be α helical follow by a turn. The only exception was EHV-4 GeneB5 where the Chou-Fasman method predicted the sequence around the critical Phe residue to be β sheet. Though the Chou-Fasman and Garnier-Osguthorpe-Robinson methods are attractive in terms of their simplicity and acceptable degree of accuracy, the statistical values were based on the calculation of only 15-29 proteins and predictions were also biased by the algorithm used.

An interesting prediction of surface probability according to the Emini *et al.* method was made, suggesting the critical Phe residue in all the homologue proteins is highly exposed to solvent. This result does agree with results from fluorescence studies by Fan Shen (Shen, et al., 1996a).

The two currently available programs, PHD and Sequery, gave some interesting data but did not give a common prediction. The PHD program did not give any reliable secondary structure prediction within the region of the critical Phe residue. In most cases the program did not assign a specific structure to the residue, since it found that particular sequence to have probabilities for either α helical, β sheet or loop. Moreover, PHD could not find similar sequence from the available data base that matches with our sequence. This lowers the degree of accuracy of the prediction by PHD to less than 72% for the three-state prediction. We conclude from the results given by PHD that in most cases the region surrounding the critical Phe residue does not have a strong preference to fold into any particular secondary structure.

Sequery gave a unique and interesting prediction of the domain in the region of the critical Phe residue. It suggested an extended structure or loop encompassing the critical Phe residue followed by a strong α helical prediction (Table 8). This data fits well with the biophysical data currently available. Instead of converting the conformational information of the protein into statistical values using a limited set of proteins, Sequery searches for the exact match of the query sequence from the current data base of protein of know structure (more than 1300 proteins). This method would eliminate the bias imposed on the predictions caused by the limited set of proteins used to generate the statistical values as well as potential flaws in the prediction methods.

These predictions imply that the primary structure of the N-subdomain might not have a strong preference for any specific secondary structure. In this

respect these data agree with the previous biophysical studies (Shen, et al., 1996a) as well as mutational analysis (Cress & Triezenberg, 1991b). It is very plausible that the N-subdomain of VP16 and its homologues assumes a loop or extended structure.

From all the information available currently, it is most likely that the N-subdomain of VP16 and its homologues are flexible in their structure. The critical Phe residues are likely to be exposed to solvent. The structure of this domain might be expected to be induced upon interaction with its target proteins. Fan Shen has shown by anisotropy decay experiments that upon addition of TBP to the VP16 activation domain, a more ordered structure of both the N and C subdomain could be induced. Addition of TFIIB, another target of the VP16 activation domain, could also induce a slight change in the C-subdomain (Shen, et al., 1996b).

Future studies

Though mutational analysis and preliminary work from other laboratories strongly suggest that my predictions are correct, thorough mutational analysis of the proposed activation domains both in the context of fusion proteins and wild type protein should be investigated for the other homologues of VP16. In particular the homologues in EHV-1 and EHV-4 where initial results have not identified the N-subdomain of the protein need further investigation. The mechanism by which these homologues activate transcription could be elucidated, for instance, by identification of their target proteins, structural determination of the activation domain by itself or in a complex with its target protein in biophysical studies, or by crystallographic studies. The results from one homologue might help explain or raise a testable hypothesis for the others.

Swapping of the activation domains or subdomains of the homologues, with careful construction, would also be interesting to show that these domains are homologous and could be exchanged.

It would also be very interesting to expand the search for similar activation domains in other transcription activators using patterns of hydrophobic residues from HCA. My preliminary attempts have identified portions of the activation domain of GCN4, residues 107-136, and the p53 activation domain, residues 48-66, which have striking similarity to the C-subdomain of VP16 (data not shown). Recent mutational analysis of this GCN4 activation domain has shown that all residues within the hydrophobic clusters predicted by HCA to be the core of this activation domain are indeed important for its activation function (Drysdale, et al., 1995). Some of the critical hydrophobic residues in the region between residues 90-120 of GCN4 have also been correctly predicted by Doug Cress (Cress & Triezenberg, 1991b). However for p53, the hydrophobic residues that I have predicted to be important for activation have not yet been studied, although some hydrophobic residues in the N-terminus , Leu residue 22 and Trp residue 23, were shown by Lin *et al.* to be important for the activation function of p53 (Lin, Chen, Elenbaas, & Levine, 1994; Lin, Teresky, & Levine, 1995).

I have shown that HCA is a powerful tool for recognition and differentiation of the two different groups of activation domains in herpes homologues of VP16. The pattern of hydrophobic residues of the activation domain might also prove to be a very useful criterion to characterize and group the activation domains of different transcription activator proteins discovered in the future.

APPENDICES

APPENDIX A

APPENDIX A

Hydrophobic cluster analysis predicts an amino-terminal domain of varicella-zoster virus open reading frame 10 required for transcriptional activation

(herpesvirus/VP16/transcription)

HIROYUKI MORIUCHI*, MASAKO MORIUCHI*, RATH PICHYANGKURA†, STEVEN J. TRIEZENBERG†, STEPHEN E. STRAUS*, AND JEFFREY I. COHEN*

*Medical Virology Section, Laboratory of Clinical Investigation, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892; and †Department of Biochemistry, Michigan State University, East Lansing, MI 48824-1319

Communicated by Steven L. McKnight, Tularik, Inc., South San Francisco, CA, July 5, 1995

ABSTRACT **Varicella-zoster virus open reading frame 10 (ORF10) protein, the homolog of the herpes simplex virus protein VP16, can transactivate immediate-early promoters from both viruses. A protein sequence comparison procedure termed hydrophobic cluster analysis was used to identify a motif centered at Phe-28, near the amino terminus of ORF10, that strongly resembles the sequence of the activating domain surrounding Phe-442 or VP16. With a series of GAL4-ORF10 fusion proteins, we mapped the ORF10 transcriptional. activation domain to the amino-terminal region (aa 5-79). Extensive mutagenesis of Phe-28 in GAL4-ORF10 fusion proteins demonstrated the importance of an aromatic or bulky hydrophobic amino acid at this position, as shown previously for Phe-442 of VP16. Transactivation by the native ORF10 protein was abolished when Phe-28 was replaced by Ala. Similar amino-terminal domains were identified in the VP16 homologs of other alphaherpesviruses. Hydrophobic cluster analysis correctly predicted activation domains of ORF10 and VP16 that share critical characteristics of a distinctive subclass of acidic activation domains.**

Transcriptional activators play an important role in the regulation of eukaryotic genes and typically contain two domains: one conferring specific association with promoter sequences, usually a DNA-binding domain, and a second domain governing transcriptional activation. Several different types of activation domains have been identified and classified on structural grounds as acidic, glutamine-rich, or proline-rich. The functional significance of these classifications, however, is unclear. Indeed, recent studies (1, 2) demonstrated that not all activation domains of a given class interact with the same target; therefore, each of these classes may include several functionally and structurally distinct subclasses.

Genetic studies demonstrated that a net negative charge is not sufficient for transcriptional activation by acidic activation domains of proteins like VP16 (3) and GAL4 (4). Alignment of amino acid sequences of several transcriptional-activation domains also suggested the importance of bulky hydrophobic residues (3). Detailed mutagenesis studies demonstrated that the presence of bulky hydrophobic amino acids interspersed among acidic amino acids is critical for the transcriptional-activation ability of VP16 (5, 6).

VP16, a prototypical acidic activator, is a component of the tegument of herpes simplex virus type 1 (HSV-1) and type 2 (HSV-2) particles and stimulates expression of viral immediate-early (IE) genes (7). Varicella-zoster virus (VZV) open reading frame 10 (ORF10) encodes a homolog of VP16. The ORF10 product is also a tegument protein (8) and transactivates the VZV and HSV IE promoter in transient expression assays (9). These proteins share considerable amino acid sequence homology; however, the ORF10 protein is 80 amino acids shorter, lacking the carboxyl-terminal acidic tail known to be critical for transactivation by VP16 (10, 11).

To define the structural basis of activation by ORF10, we used a computerized protein sequence comparison procedure termed hydrophobic cluster analysis (HCA) (12, 13). Rather than selecting sequences according to optimal amino acid similarities, the HCA method compares clusters of hydrophobic residues presumed to constitute hydrophobic cores in globular proteins. Using HCA, we identified a motif centered at Phe-28 near the amino terminus of ORF10 protein that strongly resembles the motif surrounding Phe-442 of VP16, a residue known to be critical for the transactivating ability of the carboxyl-terminal region of VP16. Using fusion proteins bearing the DNA-binding domain of GAL4, we mapped the ORF10 transcriptional-activation domain to the amino terminal region (aa 5-79). Detailed mutagenesis studies of ORF10 support the HCA-based hypothesis that the activation domain of ORF10, like that of VP16, depends upon a specific pattern of bulky hydrophobic residues for activation of transcription. The HCA method also predicts the presence of similar domains in the amino termini of the VP16 homologs (including VP16 itself) from a number of other alphaherpesviruses. Our results demonstrate that the activation domains of VZV ORF10 and HSV-1 VP16 share critical structural features characteristic of a distinctive subclass of acidic activation domains.

MATERIALS AND METHODS

Plasmids. Plasmids expressing GAL4 ORF10 fusion proteins were constructed by ligating fragments from pMTPORFlOR (9) to plasmid pGAL4 (14) or pGAL4E (15) to maintain an open reading frame. Single amino acid substitution mutants in the activation domain of ORF10 were constructed by oligonucleotide-directed mutagenesis (as in ref. 3) of the EcoRI restriction fragment of GAL4-ORF10(5-79). Wild-type and amino acid substitution mutants of VP16 residues 1-25 were constructed by ligating double-stranded synthetic oligonucleotides into plasmid pGAL4.

An ORF10 amino acid substitution mutant (pCMV1OF28A) was constructed by using PCR-oriented mutagenesis (16). The VP16 deletion mutant was constructed by digesting pCMV16 (17) completely with *Bam*HI and partially with *Sal*I, isolating a 5.9-kb fragment, blunt-ending the DNA fragment with DNA polymerase I (Klenow fragment), and ligating it to a double-stranded oligonucleotide (CTAGTCTAGACTAG). An expression vector for the GAL4-VP16 fusion protein (18) and reporter plasmids p62CAT (19), p5GAL4-E1b-CAT, and p1GAL4-E1b-CAT (20) were described previously.

Preparation of Antiserum to ORF10 Protein. Plasmid pGEX1O(41 410) was constructed to express a glutathione S-transferase (GST) fusion protein bearing amino acids 41410 of ORF10. A *Bam*HI-*Eco*RV restriction fragment from pGORF10 [ref. 21; nt 12,280-13,636 of VZV (22)] was inserted into pGEX-2T (Pharmacia) that had been previously digested with *Bam*HI and *Sma*I. The GST-ORF10 fusion protein, expressed in and purified from *Escherichia coli* strain DH5 cells, was used to raise anti-ORF10 antisera in rabbits.

Transfections, Immunoprecipitations, and Immunoblotting. To assess the transcriptional function of various activators, Vero cells were transfected, and chloramphenicol acetyltransferase (CAT) assays were performed, as described (23). To test expressed levels of the activating proteins, COS cells were transfected with plasmids encoding GAL4 fusion proteins and cell extracts were immunoprecipitated with anti-GAL4(1-147) antibody, as described (24). For immunoblotting, extracts of MeWo cells transfected with pCMV, pCMV10, or pCMV1OF28A were electrophoresed, blotted, and then incubated with anti-GST10(41-410) antibody.

HCA. The sequences of VP16 and its homologs were processed by using an HCA program provided by B. Henrissat (Centre National de la Recherche Scientifique, Université Joseph Fourier, Grenoble, France).

RESULTS

The Amino-Terminal Region of ORF10 Protein Contains the Transcriptional-Activation Domain. To identify amino acid sequences of ORF10 that can function as transcriptional-activation domains, portions of the ORF10 protein were fused to the GAL4 DNA-binding domain (aa 1-147). These constructs (activator plasmids) were assessed in transient expression assays for

their ability to activate transcription from a reporter plasmid containing five tandem GAL4-binding sites upstream of the adenovirus E1B TATA box and the CAT gene.

Analysis of a series of GAL4-ORF10 fusion constructs showed that VZV ORF10 protein contains a potent transcriptional-activation domain near its amino terminus (Fig. 1).

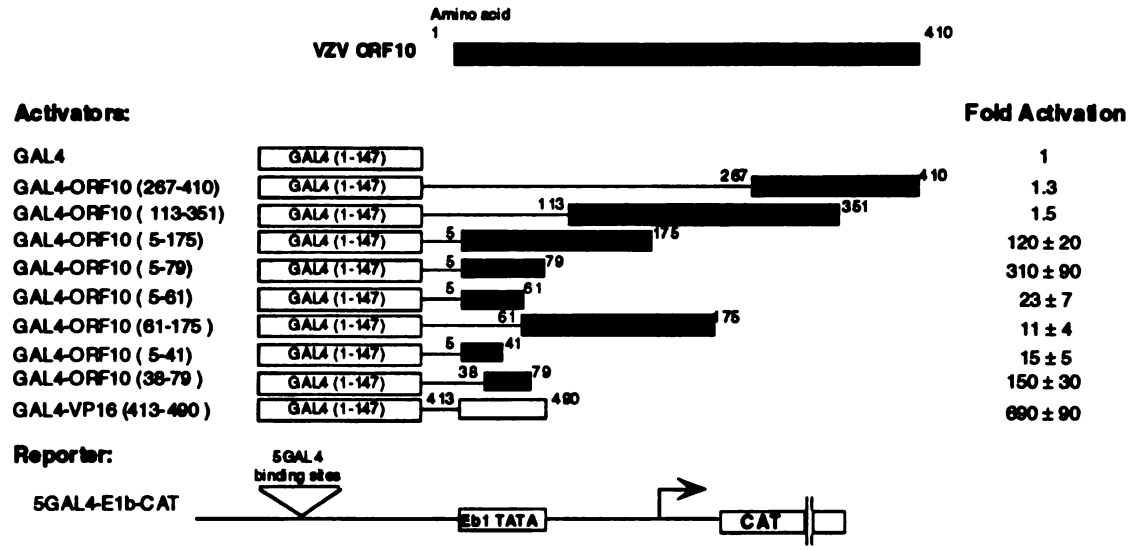


FIG. 1. Localization of the ORF10 transcriptional-activation domain. Vero cells were cotransfected with 5 μg of plasmid expressing a GAL4-ORF10 fusion protein and 5 μg of reporter plasmid containing five GAL4-binding sites, and cell lysates were assayed for CAT activity. Fold activation is the CAT activity relative to that obtained for GAL4 alone. Means and standard deviations were calculated from at least five independent transfections.

When fused to the GAL4 DNA-binding domain, an amino-terminal fragment (aa 5-175) of the VZV ORF10 protein activated transcription, while the middle or carboxyl-terminal portions of the protein failed to activate transcription. Successive carboxyl-terminal deletions within the activating fragment of ORF10 indicated that maximal activity was conferred by amino acids 5-79. Further deletion from the carboxyl terminus (aa 42-79) of GAL4-ORF10(5-79) abolished most of its activity, while deletion from the amino terminus (aa 5 37) reduced activity by 50%. While the carboxyl portion (aa 38-79) of GAL4-ORF10(5-79) had substantial activity, a single amino acid substitution within the amino-terminal portion (aa 28) abolished activity of the GAL4 construct and the native ORF10 protein (see below). The level of transcriptional activation seen with GAL4-ORF10(5-79) was about 50% of that obtained with the potent activator GAL4-VP16. GAL4ORF1O(5-79) did not transactivate a reporter plasmid lacking GAL4-binding sites, indicating that the activation was specific (data not shown). All of the fusion proteins were expressed at comparable levels in transfected cells

when assayed by immunoprecipitation with anti-GAL4 antibody (data not shown).

The HCA Method Detects a Motif Centered at Phe-28 of ORF10 That Resembles the Motif Surrounding Phe-442 of VP16. A previous alignment of the amino acid sequences of ORF10 and VP16 noted a number of highly similar regions, but failed to identify a domain similar to the VP16 activation domain (25). Given the recent evidence that the ORF10 protein is indeed a transcriptional activator (9), the complete amino acid sequences of ORF10 and VP16 were analyzed by using the HCA method in which the size, shape, and orientation of clusters of hydrophobic residues were compared. We identified a motif surrounding Phe-28 near the amino terminus of ORF10 that strongly resembles the motif centered at Phe-442 within the VP16 carboxyl-terminal activation domain (Fig. 2,4). The HCA plot (Fig. 2B) indicates that both of these motifs have horseshoe-shaped arrays of hydrophobic amino acids with Phe residues at their apices. As shown clearly in the linear sequence alignment (Fig. 2C), in both proteins, the hydrophobic residues are interspersed with numerous acidic (Asp and Glu) and potentially phosphorylated (Ser and Thr) residues.

Importance of Aromatic or Bulky Hydrophobic Amino Acids at Position 28 of ORF10 Protein. Previous studies demonstrated that mutations of Phe-442 of VP16 markedly influence transactivation level (3, 5). Substitution of aromatic (Tyr or Trp) or bulky hydrophobic (Leu, Ile, or Met) amino acids for Phe-442 decreased but did not abolish function, while substitution with most other amino acids reduced the transactivating activity of a truncated form of VP16 by ≥90% (5). To assess whether an aromatic or bulky hydrophobic amino acid at position 28 of ORF10 (which aligns with VP16 Phe-442 by HCA) is also important for transactivating activity, Phe-28 was replaced with other amino acids in GAL4-ORF10(5-79) fusion proteins, and these proteins were assessed for their ability to transactivate p5GAL4-E1b-CAT. All constructs tested were shown to be expressed in transfected cells by immunoprecipitation with anti-GAL4 antibody (data not shown).

Substitution of Phe-28 with Tyr, another aromatic amino acid, or with either Leu or Ile, two bulky hydrophobic amino acids, reduced the activity of the wild-type construct by 50% (Table 1). Substitution with smaller hydrophobic amino acids (Val, Ala, or Pro) reduced the activity by 80 90%, while substitution with hydrophilic amino acids (Ser or Thr) reduced the activity by >90%. These findings implicate an aromatic or bulky hydrophobic amino acid at position 28 as being important for the activation function of ORF10.

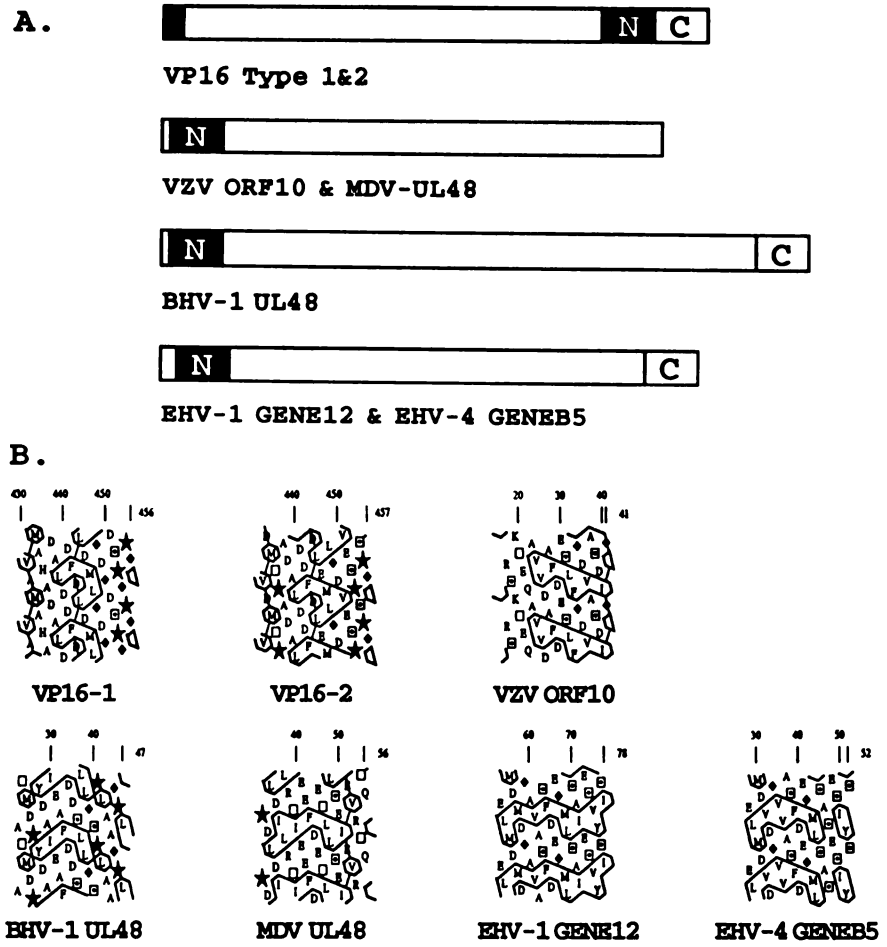


FIG. 2. HCA of VP16, ORF1O, and other VP16 homologs. The complete sequence of HSV-1 VP16 (25), HSV-2 VP16 (26), ORF10 (22), equine herpesvirus I (EHV-1) gene 12 (27), EHV-4 gene B5 (28), bovine herpesvirus 1 (BHV-1) UL48 (29), and Marek disease virus (MDV) UL48 (30) proteins were analyzed by HCA, and locations of activation domains were predicted by using VP16 as a standard, (A) Bars represent the approximate relative sizes of VP16 and its homologs. The VP16 activation domain comprises two subdomains designated N (black box) and C (open box). The amino-terminal black box of VP16 represents a cryptic activation domain. (B) HCA profiles (I 2) of subdomains resembling VP16N show the characteristic horseshoe-shaped cluster and the centrally located Phe residues. The amino acid sequences are represented (in duplicate) along diagonals from upper left to lower right with the one-letter amino acid code except that black diamonds indicate glycine, squares indicate threonine, squares with central dots indicate

serine, and stars indicate proline residues. Sets of adjacent hydrophobic residues are encompassed by contour lines. (C) Linear alignment of the subdomains shown in B, with shading indicating conserved residues, boldface letters indicating conserved bulky hydrophobic residues, dark shading denoting con served acidic residues, and open ovals representing residues conserved among homologs but not present in HSV-1 VP16.

Table 1. Relative activity of GAL4-ORF10(5-79) bearing amino acid substitutions at position 28

Residue at position 28	Relative activity, %
Tyr	52 ± 17
Leu	54 ± 17
Ile	50 ± 12
Val	23 ± 4
Ala	11 ± 2
Pro	13 ± 3
Ser	9 ± 4
Thr	7 ± 2

Vero cells were cotransfected with 5 mg of plasmid expressing an amino acid substitution mutant of GAL4 ORF10(5-79) and 5 mg of a reporter plasmid containing five GAL4-binding sites. Data are presented as the percent CAT activity of the mutant relative to that obtained for wild-type GAL4-ORF10(5-79). Means and standard deviations were calculated from at least five independent transfections.

Residues Flanking Phe-28 Also Influence Activity. Several hydrophobic residues (Val-25, Leu-32, and Phe-33) flank Phe-28 in the horseshoe-like hydrophobic cluster (Fig. 2). Some of the corresponding amino acids in VP16 (Leu-439 and Leu-444) have previously been shown to influence its function (5). To determine whether these residues also contribute to the activity of ORF10, site-specific mutations were made in ORF10(5-79). While substitution of Val-24 (not in the horseshoe-like cluster) by Ala did not reduce the activity, substitution of Leu-32 (in the horseshoe-like cluster) by Ala reduced activity to 45% of wild type (Table 2). Substitution of Phe-33 (in the horseshoe-like cluster) by Ser markedly reduced activity, while substitution of Phe-33 by Tyr, another aromatic amino acid, retained 80% of the wild-type activity. Contrary to our expectation, substitution of Val-25 (in the horseshoe-like cluster) by Ala did not reduce activity. Each of the constructs was expressed at comparable levels in transfected cells when assayed by immunoprecipitation (data not shown). Thus, while some of the hydrophobic residues flanking Phe-28 contribute to the transactivating ability of GAL4-ORF10(5-79), other residues may not be essential for activity.

The carboxyl-terminal portion (aa 62-79) of the ORF10 transcriptional-activation domain (aa 5-79) also contains any hydrophobic residues but does not form a characteristic horseshoe-like cluster. In contrast to the hydrophobic residues that flank Phe-28, substitution of any of the hydrophobic residues in the carboxyl-terminal portion of the activation domain (Ile-65, Leu-66, Tyr-67, Leu-70, Ile-71, or Leu-74) by Ala did not reduce the activity of GAL4-ORF10(5-79) (data Dot shown). These results suggest that none of these hydrophobic residues is critical for the transactivating function of the domain.

Table 2. Relative activity of GAL4-ORF10(5-79) bearing amino acid substitutions at hydrophobic residues flanking Phe-28

Amino acid substitution	Relative activity, %
Val-24 --> Ala	140 ± 40
Val-25 --> Ala	130 ± 50
Leu-32 --> Ala	45 ± 18
Phe-33 --> Ser	12 ± 4
Phe-33 --> Tyr	76 ± 18

Mutant activation domains were tested as described in the legend to Table 1.

A Motif Surrounding Phe-28 Is Essential for Transactivation by Native ORF10 Protein. To determine if the region surrounding Phe-28 is required for transactivation by native ORF10 protein, we mutated the ORF10 gene at this site and assessed its ability to transactivate the VZV ORF62 target promoter. Substitution of Phe-28 with Ala abolished transactivation by ORF10 (Table 3). An immunoblot analysis indicated that the ORF10 substitution mutant construct (pCMV1OF28A) was expressed in transfected cells at levels comparable to that of native ORF10 (pCMV1O) (data not shown). As previously shown (10, 11), deletion of the carboxyl-terminal region of VP16 abolished transactivation. These results indicate that a motif centered at Phe-28 is critical for transactivation by native ORF10 protein.

Table 3. Phe-28 is required for transactivation of the VZV IE62 promoter by ORF10 protein

Activator	Fold induction
pCMV1O	24 ± 3
pCMV1OF28A	1.0 ± 0.1
pCMV16	290 ± 50
pCMV16d4l3-490	0.8 ± 0.1

Cells were cotransfected with plasmids expressing native Of (pCMV10), an amino acid substitution mutant of ORF10 (pCMV1OF28A), full-length VP16 (pCMV16), or a carboxyl-terminal truncation of VP16 (pCMV16d4l3-490) together with p62CAT (the VZV ORF62 promoter followed by the CAT gene). Fold induction is the CAT activity in transfected cell extracts relative to that obtained when using pCMV (vector control). Means and standard deviations were calculated from at least five independent transfections.

The HCA Method Predicts That VP16 Homologs of Other Alphaherpesviruses have Transcriptional-Activation Do. mains in Their Amino Termini. Using HCA, we analyzed the structures of diverse VP16 homologs from other alphaherpes-viruses. VP16 has two distinct transcriptional-activation subdomains near its carboxyl end (Fig. 2A). The first of these subdomains surrounds Phe-442 and has been designated HI (6) or VP16N (2). The HCA method predicts the presence of similar subdomains in the amino-terminal region of each of the alphaherpesvirus VP16 homologs examined (Fig.

2A and B). Each amino-terminal subdomain has the characteristic horseshoe-like shape of hydrophobic amino acids with a Phe residue at the apex and interspersed with acidic amino acids (Fig. 2B). The most conserved features of this subdomain are the Phe, a subsequent Asp, and bulky hydrophobic residues three positions to the amino-terminal side and four and five positions to the carboxyl-terminal side (Fig. 2C).

At the extreme carboxyl-terminal region (aa 452-490), VP16 has a second activation subdomain, previously designated H2 (6) or VP16C (2). The HCA method predicts the presence of similar carboxyl-terminal domains in BHV-1 UL48 and EHV-1 gene 12 proteins but not in VZV ORF10 or MDV UL48 proteins (Figs. 2,4 and 3 A and B). The five similar domains include a pattern of hydrophobic residues interspersed with acidic or polar amino acids. Both alanine-scanning and random mutagenesis of the VP16C subdomain indicate that the positions in this pattern that are most conserved among the homologs are also the most critical residues in this region (P. Horn, S. Sullivan, V. Olson, and S. T., unpublished data). The extreme carboxyl-terminal domains of BHV-1 UL48 and EHV-1 gene 12 proteins are important for transactivation (31, 32).

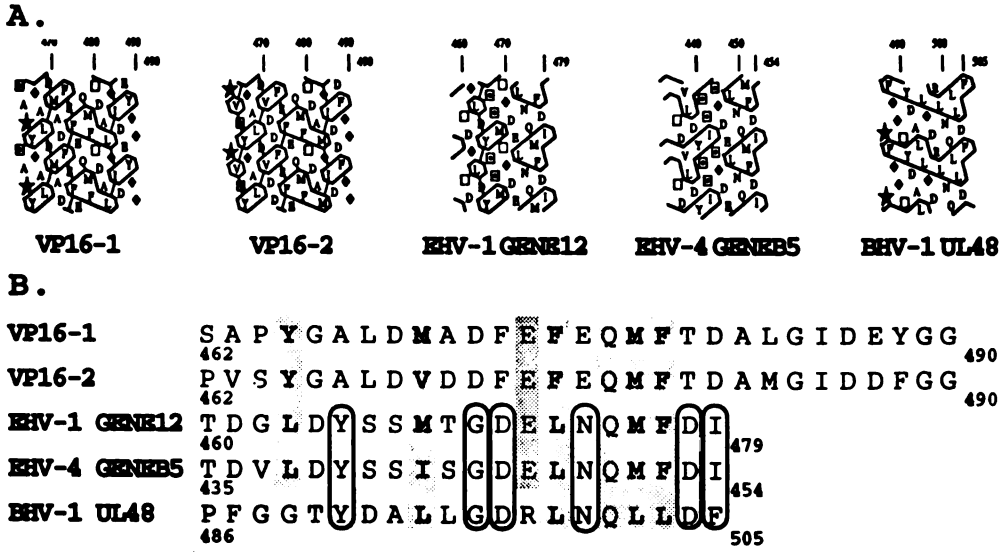


FIG 3. (A) HCA profiles of subdomains resembling VP16C, drawn as in Fig. 2B. (B) Linear sequence alignment of the subdomains shown in A, with shading, ovals, and boldface type as described in the legend of Fig. 2C.

The VP16 Amino-Terminal Region Contains a Weak Transcriptional-Activation Domain. The previous analysis suggests that among the alphaherpesviruses, the HSV VP16 proteins are distinct in having a VP16N-like subdomain as part of the larger carboxyl-terminal activation domain, whereas in all other VP16 homologs, this subdomain is found near the amino terminus. Deletion mutations removing the carboxyl-terminal 80 amino acids of VP16

abolished transcriptional activation, suggesting that amino-terminal sequences of VP16 are not strong activators in the native protein (10, 11). While HCA profiles of residues 1-40 of VP16 show little obvious similarity to the amino-terminal regions of ORF10, direct alignment of the sequences indicated some potential similarities (Fig. 4). To test the hypothesis that residues 1-25 of VP16 make up a cryptic activation domain, this region was fused to the GAL4 DNA-binding domain, and transcriptional activation was assessed in transient expression assays. GAL4-VP16(1-25) activated transcription up to 90-fold, compared with a 690-fold activation by GAL4-VP16(413-490) (data not shown).

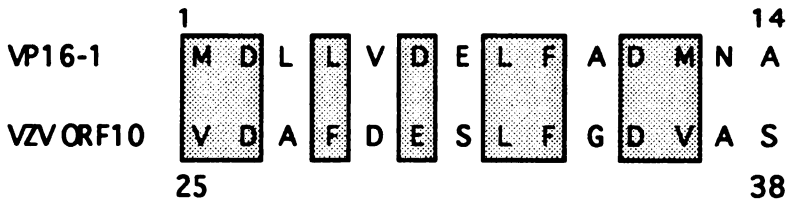


FIG. 4. Comparison of amino terminal sequences of VP16 and ORF10. Shading indicates identical or conservatively changed amino acids.

To determine whether the VP16 amino-terminal domain has structural properties similar to those seen with the VP16 carboxyl-terminal or ORF10 amino-terminal domains, amino acid substitutions were made at Leu-4. Substitution of Leu-4 by the aromatic amino acids Trp or Phe augmented activity, while substitution by the smaller hydrophobic amino acid Ala or the hydrophilic amino acid Thr virtually abolished activity (Table 4). Replacement of Met-1, Leu-8, and Phe-9, which flank Leu-4 and make up a characteristic horseshoe-like cluster in this domain, with Ala nearly abolished activity. Thus, the properties of the cryptic VP16 amino-terminal domain somewhat resemble those of the VP16 carboxyl-terminal and ORF10 amino-terminal domains.

Table 4. Relative activity of GAL4-VP16(1-25) mutants bearing various amino acid substitutions

Mutation in VP16 (1-25)	Relative activity, %
Leu-4 --> Trp	210 ± 20
Leu-4 --> Phe	150 ± 40
Leu-4 --> Ala	5.5 ± 1.7
Leu-4 --> Thr	6.8 ± 1.0
Met-1 --> Ala	8.4 ± 2.8
Leu-8 --> Ala	1.6 ± 0.4
Phe-9 --> Ala	1.3 ± 0.4

Relative activity is the CAT activity in transfected cell extracts relative to that obtained when using wild-type GAL4-VP16(1-25). Means and standard deviations were calculated from at least five independent transfections.

DISCUSSION

Using a series of GAL4-ORF10 fusion proteins, we mapped the ORF10 transcriptional-activation domain to the amino terminus of the protein. Using HCA, we detected a motif surrounding Phe-28 near the amino terminus of ORF10 that strongly resembles the motif centered at Phe-442 near the carboxyl terminus of VP16. These two motifs are composed of characteristic horseshoe-like clusters of hydrophobic residues surrounded by numerous acidic residues. Amino acid substitutions for Phe-28 in GAL4-ORF10 fusion proteins confirmed the importance of aromatic or bulky hydrophobic residues at this position, as shown previously for Phe-442 of VP16 (5), although VP16 shows a more marked preference for aromatic amino acids at this position. The critical role of Phe-28 was verified by the complete absence of activating ability by the intact ORF10 protein when Ala was substituted for Phe-28. Certain hydrophobic residues flanking Phe-28 of ORF10 also contributed to the transactivating activity, similar to the results seen with mutagenesis of residues flanking Phe-442 of VP16 (5). These data conclusively support the HCA-based prediction that the ORF10 activation domain shares critical structural features with the VP16 carboxyl-terminal activation domain.

Other members of the alphaherpesvirus subfamily, BHV-1 and EHV-1, have VP16 homologs that transactivate IE promoters of their respective viruses (31, 32). Using HCA, we located activation domains similar to VP16 and ORF10 near the amino termini of BHV-1 UL48, MDV UL48, EHV-1 gene 12, and EHV-4 gene B5 proteins. Each of these domains is characterized by a horseshoe-like hydrophobic cluster interspersed with numerous acidic residues. Although the activation domains of these proteins have not yet been thoroughly characterized, our results predict that these amino-terminal domains are important; moreover, key roles for specific amino acids can be predicted.

We identified a domain at the amino terminus of VP16 that shares some features of the VP16 carboxyl-terminal and ORF10 amino-terminal domains. This VP16 amino-terminal domain proved to be a weak transcriptional activator when fused to a heterologous DNA-binding domain; however, it is not noticeably active in the context of the native VP16 molecule. A VP16 carboxyl-terminal truncation mutant that retains the amino-terminal domain does not have transactivating activity (10, 11). We hypothesize that the amino-terminal domains in VP16 and its homologs evolved from common ancestral domains. Unlike its homologs, however, the VP16 amino-terminal domain lost or failed to develop transcriptional-activation function during evolution because VP16 acquired a potent carboxyl-terminal activation domain.

VP16 has a second domain at its extreme carboxyl terminus that is important for transcriptional activation (2, 5, 6) and that consists of two clusters of hydrophobic amino acids with adjacent acidic residues (Fig. 3A and B). Similar domains that participate in transcriptional activation are located at the extreme carboxyl termini of BHV-1 UL48 and EHV-1 gene 12 proteins (31, 32). The VZV ORF10 and MDV UL48 proteins apparently lack this domain.

HCA has previously been used to compare the sequences of numerous enzyme families, as well as DNA-binding domains (33) and dimerization domains (34) of eukaryotic transcription factors. The results reported here

demonstrate the value of HCA for predicting and comparing transcriptional-activation domains among groups of related proteins.

We thank B. Henrissat for discussions regarding HCA, D. Last for providing anti-GAL4 antibody, L. P. Perera for plasmids pCMV and p62CAT, and J. Lillie for plasmids pE1b-CAT, p1GAL4-E1b-CAT, and p5GAL4-E1b-CAT. This work was supported in part by grants from the National Institutes of Health (RO5-AI27323, K04-AI01284) and the American Cancer Society (JFRA-328) to S.J.T.

---

1. Gill, G., Pascal, E., Tseng, Z. H. & Tjian, R. (1994) *Proc, Natl. Acad. Sci. USA* **91,** 192 196.
2. Goodrich, J. A., Hoey, T., Thut, J., Admon, A. & Tjian, R. (1993) *Cell* **75,** 519 530.
3. Cress, W. D. & Triezenberg, S. J. (1991) *Science* **251,** 87-90.
4. Leuther, K. K, Salmeron, J. M. & Johnston, S. A. (1993) *Cell* **72,** 575-585.
5. Regier, J. L., Shen, F. & Triezenberg, S. J. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 883-887.
6. Walker, S., Greaves, R. & O'Hare, P. (1993) *Mol. Cell. Biol.* **13,** 5233-5244.
7. Batterson, W. & Roizman, B. (1983) *J. Virol.* **46,** 371-377.
8. Kinchington, P. R., Hougland, J. K., Arvin, A. M., Ruyechan, W. T. & Hay, J. (1992) *J. Virol.* **66,** 359-366.
9. Moriuchi, H., Moriuchi, M., Straus, S. E. & Cohen, J. I. (1993) J. *Virol.* **64,** 2739-2746.
10. Triezenberg, S. J., Kingsbury, R. C. & McKnight, S. L. (1988) *Genes Dev.* **2,** 718-729.
11. Cousens, D. J., Greaves, R., Goding, C. R. & O'Hare, P. (1989) *EMBO J.* **8,** 2337-2342.
12. Lemesle-Varloot, L., Henrissat, B., Gaborlaud, C., Bissery, V., Morgat, A. & Mornon, J.-P. (1990) *Biochimie* **72,** 555-574.
13. Woodcock, S., Mornon, J.-P. & Henrissat, B. (1992) *Protein Eng.* **5,** 629-635.
14. Keegan, L., Gill, G. & Ptashne, M. (1986) *Science* **231,** 699-704,
15. Cohen, J. I., Heffel, D. & Seidel, K. (1993) J *Virol.* **67,** 4246-4251.
16. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, R. L. (1989) *Gene* **77,** 51-59.
17. Moriuchi, M., Moriuchi, H., Straus, S, E. & Cohen, J. I. (1994) *Virology* **200,** 297-300.
18. Sadowski, I., Ma, J., Triezenberg, S. J. & Ptashne, M. (1988) *Nature (London)* **335,** 563-564.
19. Perera, L. P., Mosca, J. D., Sadeghi-Zadeb, M., Ruyechan, W. T. & Hay, J. (1992) *Virology* **191,** 346-354,
20. Lillie, J. W. & Green, M. R. (1989) *Nature (London)* **335,** 39-44.

21. Meier, J. L. & Straus, S. E. (1993) *J. Virol.* **67,** 7573-7581.
22. Davison, A. J. & Scott, J. E. (1986) *J. Gen. Virol.* **67,**1759-1816.
23. Moriuchi, H., Moriuchi, M., Straus, S. E. & Cohen, J. I., (1993) *J. Virol.* **67,** 4290-4295.
24. Cohen, J. 1. & Kieff, E. (1991) *J. Virol.* **65,** 5880-5885.
25. Dalrymple, M. A., McGeoch, D. J., Davison, A. J. & Preston, C. M. (1985) *Nucleic Acids Res.* **13,** 7865-7979.
26. Cress, W. D. & Triezenberg, S. J. (1991) *Gene* **103,** 235-238.
27. Telford, E. A. R., Watson, M. S., McBride, K, & Davison, A. J. (1992) *Virology* **189,** 304-316.
28. Purewal, A. S., Allsopp, R., Riggio, M., Telford, E. A. R., Azam, S., Davison, A. J. & Edington, N. (1994) *Virology* **198,** 385-389.
29. Carpenter, D. E. & Misra, V. (1991) *Gene* **119,** 259-263.
30. Yanagida, N., Yoshida, S., Nazerian, K. & Lee, L. F. (1993) J. *Gen. Virol.* **74,**1837-1845.
31. Misra, V., Bratanich, A. C., Carpenter, D. & O'Hare, P. (1994) *J. Virol.* **68,** 4898-4909.
32. Elliot, G. D. (1994) *J. Virol.* **68,** 4890-4897.
33. Laget, M., Callebaut, I., delaunoit, Y., Stehelin, D., & Mornon, J.-P. (1993) *Nucleic Acids Res,* **21,** 5987-5996.
34. Mornon, J.-P., Thoreau, E., Rowlands, D., Callebaut, I. & Moreau, G. (1994) *C. R. Acad. Sci. Ser. III* **317,** 597-606.

APPENDIX B

APPENDIX B

LIST OF PROTEINS FROM SEQUERY SEARCH

Protein	Matching residues	Protein	Matching residues
1st window		4th window (cont'd)	
1pdg -	390-393	4enl -	258-261
2pmg A	253-256	1bov A	14-17
2scp A	147-150	1pcd A	84-87
3sc2 A	179-182	5th window	
1vsg A	105-108	1wsy A	159-162
2pmg A	153-156	7xia -	55-58
1cpc A	107-110	3cla -	80-83
1gla G	120-123	1gsg P	16-19
1phs -	49-52	1ace -	259-262
2nd window		1etu -	142-145
1aso A	318-321	3pgk -	10-13
1aso A	399-402	1nip B	69-72
7xia -	294-297	1gky -	168-171
7xia -	358-361	1nrd -	327-330
1min A	385-388	1dri -	190-193
2cyp -	34-37	1fha -	25-28
5p21 -	30-33	6th window	
3rd window		1btc -	423-426
2hip A	52-55	1wsy A	160-163
1ofv -	126-129	4fxn -	41-44
1gpb -	251-254	1pya B	208-211
1caj -	69-72	3grs -	441-444
1mdc -	68-71	2ccy A	6-9
2cyp -	76-79	1plc -	60-63
1hdd C	19-22	1wsy B	78-81
1hc6 -	579-582	1etu -	143-146
1tlk -	91-94	1lpe -	49-52
1lap -	129-132	1nip B	265-268
2lbp -	48-51	1pda -	292-295
1atn A	361-364	1phg -	286-289
4th window		1tro A	59-62
1pcd A	147-150	2zta A	10-13
1plc -	41-44	3adk -	73-76
2pmg A	174-177	8atc A	199-202
2dnj A	97-100	3chy -	66-69
1bbp A	121-124	1dpi -	402-405
1tlk -	92-95	3pgk -	397-400

APPENDIX C

APPENDIX C

LIST OF PUBLICATIONS

1. Koelle, D. M., Corey, L., Burke, R. L., Eisenberg, R. J., Cohen G. H., Pichyangkura, R., Triezenberg, S. J. (1994). Antigenic specificities of human $CD4^+$ T-cell clones recovered from recurrent genital herpes simplex virus type 2 lesions. Journal of Virology, Volume 68, pages 2803-2810

2. Moriuchi, H., Moriuchi, M., Pichyangkura, R., Triezenberg, S. J., Straus, S. E., & Cohen, J. I. (1995). Hydrophobic cluster analysis predicts an amino-terminal domain of varicella-zoster ORF10 required for transcriptional activation. Proceedings of the National Academy of Science USA, Volume 92, pages 9333-9337

LIST OF REFERENCES

LIST OF REFERENCES

Ace, C. I., Dalrymple, M. A., Ramsay, F. H., Preston, V. G., & Preston, C. M. (1988). Mutational analysis of the herpes simplex virus type 1 trans-inducing factor VP16. J. Gen. Virol. 69, 2595-2605.

Ace, C. I., McKee, T. A., Ryan, J. M., Cameron, J. M., & Preston, C. M. (1989). Construction and characterization of a herpes simplex virus type 1 mutant unable to transinduce immediate-early gene expression. J. Virol. 63, 2260-2269.

Axelrod, J. D., Reagan, M. S., & Majors, J. (1993). GAL4 disrupts a repressing nucleosome during activation of GAL1 transcription in vivo. Genes Dev. 7, 857-869.

Berger, S. L., Pina, B., Silverman, N., Marcus, G. A., Agapite, J., Regier, J. L., Triezenberg, S. J., & Guarente, L. (1992). Genetic isolation of ADA2 - a potential transcriptional adaptor required for function of certain acidic activation domains. Cell 70, 251-265.

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

Bzik, D. J., & Preston, C. M. (1986). Analysis of DNA sequences which regulate the transcription of herpes simplex virus immediate early gene 3: DNA sequences required for enhancer-like activity and response to trans-activation by a virion polypeptide. Nucleic Acids Res. 14, 929-943.

Cadena, D. L., & Dahmus, M. E. (1987). Messenger RNA synthesis in mammalian cells is catalyzed by the phosphorylated form of RNA polymerase II. J. Biol. Chem. 262, 12468-12474.

Campbell, E. N., Palfreyman, J. W., & Preston, C. M. (1984). Identification of herpes simplex virus DNA sequences which encode a trans-acting polypeptide responsible for stimulation of immediate early transcription. J. Mol. Biol. 180, 1-19.

Carey, M., Leatherwood, J., & Ptashne, M. (1990). A potent GAL4 derivative activates transcription at a distance in vitro. Science 247, 710-711.

Carpenter, D. E., & Misra, V. (1991). The most abundant protein in bovine herpes 1 virions is a homologue of herpes simplex virus type 1 UL47. J. Gen. Virol. 72, 3077-3084.

Carpenter, D. E., & Misra, V. (1992). Sequences of the bovine herpesvirus 1 homologue of herpes simplex virus type-1 alpha-trans-inducing factor (UL48). Gene 119, 259-263.

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

Chao, D. M., Gadbois, E. L., Murray, P. J., Anderson, S. F., Sonu, M. S., Parvin, J. D., & A., Y. R. (1996). A mammalian SRB protein associated with an RNA polymerase II holoenzyme. Nature 380, 82-85.

Chen, J., Attardi, L. D., Verrijzer, C. P., Yokomori, K., & Tjian, R. (1994). Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. Cell 79, 93-105.

Chou, P. Y., & Fasman, G. D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47, 45-148.

Cleary, M. A., & Herr, W. (1995). Mechanisms for flexibility in DNA sequence recognition and VP16-induced complex formation by the Oct-1 POU domain. Mol. Cell. Biol. 15, 2090-2100.

Collawn, J. F., Kuhn, M. A., Liu, L.-F. S., Tainer, J. A., & Trowbridge, I. S. (1991). Transplanted LDL and mannose-6-phosphate receptor internalization signal promote high -efficiency endocytosis of the transferrin receptor. EMBO J. 10, 3247-3253.

Collawn, J. F., Stangel, M., Kuhn, L. A., Esekogwu, V., Jing, S., Trowbridge, I. S., & Tainer, J. A. (1990). Transferrin receptor internalization sequence YXRF implicates a tight turn as a structural recognition motif for endocytosis. Cell 63, 1061-1072.

Cordingley, M. G., Campbell, M. E. M., & Preston, C. M. (1983). Functional analysis of a herpes simplex virus type I promoter: Identification of far-upstream regulatory sequences. Nucleic Acids Res. 11, 2347-2365.

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

Cress, A., & Triezenberg, S. J. (1991a). Nucleotide and deduced amino acid sequences of the gene encoding virion protein 16 from herpes simplex virus type 2. Gene 103, 235-238.

Cress, W. D., & Triezenberg, S. J. (1991b). Critical structural elements of the VP16 transcriptional activation domain. Science 251, 87-90.

Croston, G. E., Kerrigan, L. A., Lira, L. M., Marshak, D. R., & Kadonaga, J. T. (1991). Sequence-specific antirepression of histone H1-mediated inhibition of basal RNA polymerase II transcription. Science 251, 643-649.

Davison, A. J. (1991). Varicella-zoster virus. J. Gen. Virol. 72, 475-486.

Doerfler, W. (1983). DNA methylation and gene activity. Ann. Rev. Biochem. 52, 93-124.

Doerfler, W., Toth, M., Kochanek, S., Achten, S., Freisemrabien, U., Behnkrappa, A., & Orend, G. (1990). Eukaryotic DNA methylation - facts and problems. FEBS Lett. 268, 329-333.

Donaldson, L., & Capone, J. P. (1992). Purification and characterization of the carboxyl-terminal transactivation domain of Vmw65 from herpes simplex virus type 1. J. Biol. Chem. 267, 1411-1414.

Drysdale, C. M., Duenas, E., Jackson, B. M., Reusser, U., Braus, G. H., & Hinnebusch, A. G. (1995). The transcriptional activator GCN4 contains multiple activation domains that are dependent on hydrophobic amino acids. Mol. Cell. Biol. 15, 1220-1233.

Dynan, W. S. (1989). Understanding the molecular mechanism by which methylation influences gene expression. Trends Genetics 5, 35-36.

Elliott, G., Mouzakitis, G., & O'hare, P. (1995). VP16 interacts via its activation domain with VP22, a tegument protein of herpes simplex virus, and is relocated to a novel macromolecular assembly in coexpressing cells. J. Virol. 69, 7932-7941.

Elliott, G. D. (1994). The extreme carboxyl terminus of the equine herpesvirus 1 homolog of herpes simplex virus VP16 is essential for immediate-early gene activation. J. Virol. 68, 4890-4897.

Emami, K. H., & Carey, M. (1992). A synergistic increase in potency of a multimerized VP16 transcriptional activation domain. EMBO J. 11, 5005-5012.

Emini, E. A., Hughes, J. V., Perlow, D. S., & Boger, J. (1985). Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J. Virol. 55(3), 836-839.

Fang, S. M., & Burton, Z. F. (1995). RNA polymerase II-assosiated protein (RAP) 74 binds transcription factor (TF) IIB and blocks TFIIB-RAP30 binding. J. Biol. Chem. 271, 11703-11709.

Feng, D.-F., & Doolittle, R. F. (1987). Progressive sequence alignment as a prerequisite to correct phylogenitic tree. J. Mol. Evol. 25, 351-360.
Friedman, A. D., Triezenberg, S. J., & McKnight, S. L. (1988). Expression of a truncated viral trans-activator selectively impedes lytic infection by its cognate virus. Nature 335, 452-454.

Garnier, J., Osguthorpe, D. J., & Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. 120, 97-120.

Gibson, W., & Roizman, B. (1974). Proteins specified by herpes simplex virus. X. Staining and radiolabelling properties of B capsid and virion proteins in polyacrylamide gels. J. Virol. 13, 155-165.

Gill, G., Pascal, E., Tseng, Z. H., & Tjian, R. (1994). A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII110 component of the Drosophila TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. USA 91, 192-196.

Goodrich, J. A., Hoey, T., Thut, C. J., Admon, A., & Tjian, R. (1993). Drosophila TAF(II)40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. Cell 75, 519-530.

Graessmann, A., Sandberg, G., Guhl, E., & Graessmann, M. (1994). Methylation of single sites within the herpes simplex virus tk coding region and the simian virus 40 T-antigen intron causes gene inactivation. Mol. Cell. Biol. 14, 2004-2010.

Graham, F. L., & Eb., A. J. V. D. (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology 52, 456-467.

Greaves, R., & O'Hare, P. (1989). Separation of requirements for protein-DNA complex assembly from those for functional activity in the herpes simplex virus regulatory protein Vmw65. J. Virol. 63, 1641-1650.

Guentte, D. K., Ritzenthaler, J. D., Foley, J., Jackson, J. D., & Smith, B. D. (1992). DNA methylation inhibits transcription of procollagen alpha 2(I) promoters. Biochem. J. 283, 699-703.

Ha, I., Roberts, S., Maldonado, E., Sun, X. Q., Kim, L. U., Green, M., & Reinberg, D. (1993). Multiple functional domains of human transcription factor-IIB - distinct interactions with 2 general transcription factors and RNA polymerase-II. Genes Dev. 7, 1021-1032.

Harrington, M. A., Jones, P. A., Imagawa, M., & Karin, M. (1988). Cytosine methylation does not affect binding of transcription factor Sp1. Proc. Natl. Acad. Sci. 85, 2066-2070.

Heine, J. W., Honess, R. W., Cassain, E. N., & Roizman, B. (1974). Proteins specified by herpes simplex virus. XII. The virion polypeptides of type 1 strains. J. Virol. 14, 640-651.

Hengartner, C. J., Thompson, C. M., Zhang, J., Chao, D. M., Liao, S. M., Koleske, A. J., Okamura, S., & Young, R. A. (1995). Association of an activator with an RNA polymerase II holoenzyme. Genes Dev. 9, 897-910.

Hoey, T., Weinzierl, R. O. J., Gill, G., Chen, J. L., Dynlacht, B. D., & Tjian, R. (1993). Molecular cloning and functional analysis of drosophila TAF110 reveal properties expected of coactivators. Cell 72, 247-260.

Homa, F. L., Krikos, A., Glorioso, J. C., & Levine, M. (1991). Functional analysis of regulatory regions controlling strict late HSV gene expression. In E. K. Wagner (Eds.), Herpesvirus transcription and its regulation (pp. 207-222). CRC Press, Inc.

Horiuchi, J., Silverman, N., Marcus, G. A., & Guarente, L. (1995). ADA3, a putative transcriptional adaptor, consists of two separable domains and interacts with ADA2 and GCN5 in a trimeric complex. Mol. Cell. Biol. 15, 1203-1209.

Ingles, C. J., Shales, M., Cress, W. D., Triezenberg, S. J., & Greenblatt, J. (1991). Reduced binding of TFIID to transcriptionally compromised mutants of VP16. Nature 351, 588-590.

Johnson, C. A., Goddard, J. P., & Adams, R. L. (1995). The effect of histone H1 and DNA methylation on transcription. Biochem. J. 305, 791-798.

Johnson, P. F., & McKnight, S. L. (1989). Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. 58, 799-839.

Johnson, P. F., Sterneck, E., & Williams, S. C. (1993). Activation domains of transcriptional regulatory proteins. J. Nutr. Biochem. 4, 386-398.

Katan, M., Haigh, A., Verrijzer, C. P., Vandervliet, P. C., & O'Hare, P. (1990). Characterization of a cellular factor which interacts functionally with Oct-1 in the assembly of a multicomponent transcription complex. Nucleic Acids Res. 18, 6871-6880.

Klemm, R., Goodrich, J., Zhou, S., & Tjian, R. (1995). Molecular cloning and expression of the 32-kDa subunit of human TFIID reveals interactions with VP16 and TFIIB that mediate transcriptional activation. Proc. Natl. Acad. Sci. USA. 92, 5788-5792.

Koleske, A. J., & Young, R. A. (1994). An RNA polymerase II holoenzyme responsive to activators. Nature 368, 466-469.

Koleske, A. J., & Young, R. A. (1995). The RNA polymerase II holoenzyme and its implications of gene regulation. Trends Biochem. Sci. 20, 113-116.

Koptidesova, D., Kopacek, J., Zelnik, V., Ross, N. L. J., Pastorekova, S., & Pastorek, J. (1995). Identification and characterization of a cDNA clone derived from the Marek's disease tumour cell line RPL1 encoding a homologue of alpha-transinducing factor (VP16) of HSV-1. Arch. Virol. 140, 355-362.

Kraus, V. B., Inostroza, J. A., Yeung, K., Reinberg, D., & Nevins, J. R. (1994). Interaction of the Dr1 inhibitory factor with the TATA binding protein is disrupted by adenovirus E1A. Proc. Natl. Acad. Sci. USA 91, 6279-6282.

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

Kristie, T. M., & Roizman, B. (1984). Separation of sequences defining basal expression from those conferring alpha gene recognition within the regulatory domains of herpes simplex virus 1 alpha genes. Proc. Natl. Acad. Sci. USA 81, 4065-4069.

Kristie, T. M., & Sharp, P. A. (1990). Interactions of the Oct-1 POU subdomains with specific DNA sequences and with the HSV alpha-trans-activator protein. Genes Devel. 4, 2383-2396.

Laybourn, P. J., & Kadonaga, J. T. (1991). Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. Science 254, 238-245.

Lemaster, S., & Roizman, B. (1980). Herpes simplex virus phosphoproteins. II. Characterization of the virion protein kinase and of the polypeptides phosphorylated in the virion. J. Virol. 35, 798-811.

Lemesle-Varloot, L., Henrissat, B., Gaboriaud, C., Bissery, V., Morgat, A., & Mornon, J.-P. (1990). Hydrophobic cluster analysis: procedures to derive structural and functional information from 2-D-representation of protein sequences. Biochimie 72, 555-574.

Lewis, J., & Bird, A. (1991). DNA methylation and chromatin structure. FEBS Lett. 285, 155-159.

Lieberman, P. (1994). Identification of functional targets of the Zta transcriptional activator by formation of stable preinition complex intermediates. Mol. Cell. Biol. 14, 8365-8375.

Lieberman, P. M., & Berk, A. J. (1994). A mechanism for TAFs in transcriptional activation: activation domain enhancement of TFIID-TFIIA-promoter DNA complex formation. Genes Dev. 8, 995-1006.

Lillie, J. W., & Green, M. R. (1989). Transcriptional activation by the adenovirus E1A protein. Nature 338, 39-44.

Lin, J., Chen, J., Elenbaas, B., & Levine, A. J. (1994). Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein. Genes Dev. 8, 1235-1246.

Lin, J., Teresky, A., & Levine, A. (1995). Two critical hydrophobic amino acids in the N-terminal domain of the p53 protein are required for the gain of function phenotypes of human p53 mutants. Oncogene 10, 2387-2390.

Lin, Y. S., Ha, I., Maldonado, E., Reinberg, D., & Green, M. R. (1991). Binding of general transcription factor TFIIB to an acidic activating region. Nature 353, 569-571.

Ma, J., & Ptashne, M. (1987). A new class of yeast transcriptional activators. Cell 51, 113-119.

Mackem, S., & Roizman, B. (1982a). Differentiation between alpha promoter and regulator regions of herpes simplex virus type 1: The functional domains and sequence of a movable alpha regulator. Proc. Natl. Acad. Sci. USA 79, 4917-4921.

Mackem, S., & Roizman, B. (1982b). Structural features of the herpes simplex virus alpha gene 4, 0, and 27 promoter-regulatory sequences which confer alpha regulation on chimeric thymidine kinase genes. J. Virol. 44, 939-949.

Marcus, G. A., Silverman, N., Berger, S. L., Horiuchi, J., & Guarente, L. (1994). Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors. EMBO J. 13, 4807-4815.

McGeoch, D. J., Dalrymple, M. A., & Davison, A. J. (1988). The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. 69, 1531-1547.

McGeoch, D. J., Dolan, A., Donald, S., & Brauer, D. H. K. (1986). The complete sequence of the short repeat region in the genome of herpes simplex virus type 1. Nucleic Acids Res. 14, 1727-1745.

McKee, T. A., Disney, G. H., Everett, R. D., & Preston, C. M. (1990). Control of expression of the varicella-zoster virus major immediate early gene. J. Gen. Virol. 71, 897-906.

Merino, A., Madden, K. R., Lane, W. S., Champoux, J. J., & Reinberg, D. (1993). DNA topoisomerase I is involved in both repression and activation of transcription. Nature 365, 227-234.

Misra, V., Bratanich, A.C., Carpenter, D., O'Hare, P. (1994). Protein and DNA elements involved in transactivation of the promoter of the bovine herpesvirus (BHV) 1 IE-1 transcription unit by the BHV α gene *trans*-inducing factor. J. Virol. 68, 4898-4909.

Moriuchi, H., Moriuchi, M., Pichyangkura, R., Triezenberg, S. J., Straus, S. E., & Cohen, J. I. (1995). Hydrophobic cluster analysis predicts an amino-terminal domain of varicella-zoster ORF10 required for transcriptional activation. Proc. Natl. Acad. Sci. USA 92, 9333-9337.

Moriuchi, H., Moriuchi, M., Straus, S. E., & Cohen, J. I. (1993). Varicella-zoster virus open reading frame 10 protein, the herpes simplex virus VP16 homolog, transactivates herpesvirus immediate-early gene promoters. J. Virol. 67, 2739-2746.

O'Hare, P., & Williams, G. (1992). Structural studies of the acidic transactivation domain of the Vmw65 protein of herpes simplex virus using 1H NMR. Biochemistry 31, 4150-4156.

Ossipow, V., Tassan, J. P., Nigg, E. A., & Schibler, U. (1995). A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation. Cell 83, 137-146.

Paranjape, S. M., Kamakaka, R. T., & Kadonaga, J. T. (1994). Role of chromatin structure in the regulation of transcription by RNA polymerase II. Ann. Rev. Biochem. 63, 265-297.

Paranjape, S. M., Krumm, A., & Kadonaga, J. T. (1995). HMG 17 is a chromatin-specific transcriptional coactivator that increases the efficency of transcription initiation. Genes Dev. 9, 1987-1991.
Pazin, M. J., Kamakaka, R. T., & Kadonaga, J. T. (1995). ATP-dependent nuclepsome reconfiguration and transcriptional activation from preassembled chromatin templates. Science 266, 2007-2011.

Peek, R., Niessen, R. W. L. M., Schoenmakers, J. G. G., & Lubsen, N. H. (1991). DNA methylation as a regulatory mechanism in rat gamma-crystallin gene expression. Nucleic Acids Res. 19, 77-83.

Poon, A. P. W., & Roizman, B. (1995). The phenotype in vitro and in infected cells of herpes simplex virus 1 alpha-trans-inducing factor (VP16) carrying temperature sensitive mutations introduced by substitutions of cysteines. J. Virol. 69, 7658-7667.

Post, L. E., Mackem, S., & Roizman, B. (1981). Regulation of alpha genes of herpes simplex virus: Expression of chimeric genes produced by fusion of thymidine kinase with alpha gene promoters. Cell 24, 555-565.

Preston, C. M., Frame, M. C., & Campbell, M. E. M. (1988). A complex formed between cell components and an HSV structural polypeptide binds to a viral immediate early gene regulatory DNA sequence. Cell 52, 425-434.

Purewal, A. S., Allsop, R., Riggio, M., Telford, E. A. R., Azam, S., Davison, A. J., & Edington, N. (1994). Equid herpesvirus 1 and 4 encode functional homologs of the herpes simplex virus type 1 virion transactivator protein, VP16. Virology 198, 385-389.

Purewal, A. S., Smallwood, A. V., Kaushal, A., Adegboye, D., & Edington, N. (1992). Identification and control of the cis-acting elements of the immediate-early genes of equid herpesvirus type 1. J. Gen. Virol. 73, 513-519.

Regier, J. L., Shen, F., & Triezenberg, S. J. (1993). Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. Proc. Natl. Acad. Sci. USA 90, 883-887.

Roberts, S. G. E., Green, M. R. (1994). Activator-induced conformational change in general transcription factor TFIIB. Nature 371, 717-720.

Roberts, S. G. E., Ha, I., Maldonado, E., Reinberg, D., & Green, M. R. (1993). Interaction between an acidic activator and transcription factor IIB is required for transcriptional activation. Nature 363, 741-744.

Roizman, B. (1990). Herpesviridae: A brief introduction. In Virology (ed. B. Fields) (pp. 1787-1794).

Roizman, B., & Sears, A. E. (1990). Herpes simplex viruses and their replication. In Virology (ed. B. Fields) (pp. 1795-1842).

Rost, B., & Sander, C. (1994). Combining evolutionary information and neural networks to predict protein secondary structure. Proteins 15, 55-72.

Roy, A. L., Malik, S., Meisterernst, M., & Roeder, R. G. (1993). An alternative pathway for transcription initiation involving TFII-I. Nature 365, 355-359.

Sadowski, I., Ma, J., Triezenberg, S., & Ptashne, M. (1988). GAL4-VP16 is an unusually potent transcriptional activator. Nature 335, 563-564.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). In Molecular Cloning. A Laboratory Manual (2nd ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S., & Coulson, S. A. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Shen, F., Triezenberg, S. J., Hensley, P., Porter, D., & Knutson, J. R. (1996a). Critical amino acids in the transcription activation domain of the herpesvirus protein VP16 are solvent-exposed in highly mobile protein segments. J. Biol. Chem. 271 , 4819-4826.

Shen, F., Triezenberg, S. J., Hensley, P., Porter, D., & Knutson, J. R. (1996b). Transcriptional activation domain of the herpesvirus protein VP16 becomes conformationally constrained upon interaction with basal transcription factors. J. Biol. Chem. 271 , 4827-2837.

Smale, S. T., & Baltimore, D. (1989). The "initiator" as a transcription control element. Cell 57, 103-113.

Smibert, C. A., Popova, B., Xiao, P., Capone, J. P., & Smiley, J. R. (1994). Herpes simplex virus VP16 forms a complex with the virion host shutoff protein vhs. J. Virol. 68, 2339-2346.

Spear, P. G., & Roizman, B. (1972). Proteins specified by herpes simplex virus. V. Purification and structural proteins of the herpesvirion. J. Virol. 9, 431.

Stern, S., & Herr, W. (1991). The herpes simplex virus trans-activator VP16 recognizes the Oct-1 homeo domain: evidence for a homeo domain recognition subdomain. Genes Devel. 5, 2555-2566.

Stringer, K. F., Ingles, C. J., & Greenblatt, J. (1990). Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. Nature 345, 783-786.

Tanaka, M., & Herr, W. (1994). Reconstitution of transcriptional activation domains by reiteration of short peptide segments reveals the modular organization of a glutamine-rich activation domain. Mol. Cell. Biol. 14, 6056-6067.

Thanos, D., & Maniatis, T. (1992). The high mobility group protein HMG I(Y) is required for NF-kappa B-dependent virus induction of the human IFN-beta gene. Cell 71, 777-7789.

Thut, C. J., Chen, J., Klemm, R., & Tjian, R. (1995). p53 transcriptional activation mediated by coactivators $TAF_{II}40$ and $TAF_{II}60$. Science 267, 100-104.

Tiley, L. S., Madore, S. J., Malim, M. H., & Cullen, B. R. (1992). The VP16 transcription activation domain is functional when targeted to a promoter-proximal RNA sequence. Genes Dev. 6, 2077-2087.

Triezenberg, S. J. (1995). Structure and function of transcriptional activation domains. Curr. Opin. Genetics Dev. 5, 190-196.

Triezenberg, S. J., Kingsbury, R. C., & McKnight, S. L. (1988a). Functional dissection of VP16, the *trans*-activator of herpes simplex virus immediate early gene expression. Genes Dev. 2, 718-729.

Triezenberg, S. J., LaMarco, K. L., & McKnight, S. L. (1988b). Evidence of DNA:protein interactions that mediate HSV-1 immediate early gene activation by VP16. Genes Dev. 2, 730-743.

Usheva, A., & Shenk, T. (1994). TATA-binding protein-independent initiation: YY1, TFIIB, and RNA polymerase II direct basal transcription on supercoiled template DNA. Cell 76, 1115-1121.

Van Hoy, M., Leuther, K. K., Kodadek, T., & Johnston, S. A. (1993). The acidic activation domains of the GCN4 and GAL4 proteins are not α helical but form β sheets. Cell 72, 587-594.

Verrijzer, C. P., Chen, J.-L., Yokomori, K., & Tjian, R. (1995). Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. Cell 81, 1115-1125.

Verrijzer, C. P., Yokomori, K., Chen, J. L., & Tjian, R. (1994). Drosophila TAFII 150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA. Science 264, 933-941.

Walker, S., Greaves, R., & O'Hare, P. (1993). Transcriptional activation by the acidic domain of Vmw65 requires the integrity of the domain and involves additional determinants distinct from those necessary for TFIIB binding. Mol. Cell. Biol. 13, 5233-5244.

Watt, F., & Molly, P. L. (1988). High mobility group proteins 1 and 2 stimulate binding of a specific transcription factor to the adenovirus major late promoter. Nucleic Acid Research 16, 1471-1486.

Weinheimer, S. P., Boyd, B. A., Durham, S. K., Resnick, J. L., & O'Boyle, D. R. (1992). Deletion of the VP16 open reading frame of herpes simplex virus type 1. J. Virol. 66, 258-269.

Weis, L., & Reinberg, D. (1992). Transcription by RNA polymerase-II - initiator-directed formation of transcription-competent complexes. FASEB J. 6, 3300-3309.

Werstuck, G., & Capone, J. P. (1989a). Mutational analysis of the herpes simplex virus trans-inducing factor Vmw65. Gene 75, 213-224.

Werstuck, G., & Capone, J. P. (1989b). Identification of a domain of the herpes simplex virus trans-activator Vmw65 required for protein-DNA complex formation through the use of protein A fusion proteins. J. Virol. 63, 5509-5513.

Whitley, R. J. (1990). Herpes simplex viruses. In Virology (ed. B. Fields) (pp. 1843-1888).

Wilson, C. J., Chao, D. M., Imbalzano, A. N., Schnitzler, G. R., Kingston, R. E., & Young, R. A. (1996). RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling. Cell 84, 235-244.

Wolffe, A. P. (1994). Transcription: in tune with the histones. Cell 77, 13-16.

Wölfl, S., Schräder, M., & Wittig, B. (1991). Lack of correlation between DNA methylation and transcriptional inactivation: The chicken lysozyme gene. Proc. Natl. Acad. Sci. USA 88, 271-275.

Woodcock, S., Mornon, J.-P., & Henrissat, B. (1992). Detection of secondary structure elements in proteins by hydrophobic cluster analysis. Protein Eng. 5, 629-635.

Xiao, H., Pearson, A., Coulombe, B., Truant, R., Zhang, S., Regier, J. L., Triezenberg, S. J., Reinberg, D., Flores, O., Ingles, C. J., & Greenblatt, J. (1994). Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53. Mol. Cell. Biol. 14, 7013-7024.

Xiao, P., & Capone, J. P. (1990). A cellular factor binds to the herpes simplex virus type-1 transactivator Vmw65 and is required for Vmw65-dependent protein-DNA complex assembly with Oct-1. Mol. Cell. Biol. 10, 4974-4977.

Yanagida, N., Yoshida, S., Nazerian, K., & Lee, L. F. (1993). Nucleotide and predicted amino acid sequences of Marek's disease virus homologues of herpes simplex virus major tegument proteins. J. Gen. Virol. 74, 1837-1845.

Yankulov, K., Blau, J., Purton, T., Roberts, S., & Bentley, D. L. (1994). Transcriptional elongation by RNA polymerase II is stimulated by transactivators. Cell 77, 749-759.

Young, R. A. (1991). RNA polymerase-II. Annu. Rev. Biochem. 60, 689-715.

Zawel, L., & Reinberg, D. (1993). Initiation of transcription by RNA polymerase II: a multi-step process. Prog. Nucleic Acids Res. Mol. Biol. 44, 67-108.

Zawel, L., & Reinberg, D. (1995). Common themes in assembly of eukaryotic transcription complexes. Ann. Rev. Biochem. 64, 533-561.