Characterisation of the DNA binding domain of the transcriptional regulator encoded by VZV gene 62

by

Jessica Kate Tyler

A thesis presented for the degree of doctor of philosophy

in

The Faculty of Science at the University of Glasgow

Institute of Virology
University of Glasgow
Church Street
Glasgow
G11 5JR

March, 1994

Contents

Chapter 1: Introduction

Chapter 2: Materials and methods

List of figures and tables

Acknowledgements

I would like to thank Professor John H. Subak-Sharpe for providing research facilities in the Institute of Virology, and for his overall supervision of my studies in Glasgow.

I am indebted to my project supervisor Dr Roger D. Everett, for his unflagging enthusiasm and encouragement, which has undoubtedly been a major determinant in the accomplishment of this project. I would also like to take this opportunity to thank him for his critical and prompt proof-reading of this Thesis.

Throughout my postgraduate studies I have benefited from the help and friendship of many of my colleagues within the Institute. I am particularly grateful to Dr Joe Conner and Dr Graham Hope for advice concerning the protein purification, and to the members of Lab 104, past and present, for solving my often trivial queries on countless occasions. I would like to thank Dr Ben Luisi for stimulating discussions, and Dr Howard Marsden and Gillian McVey for kindly raising the polyclonal antisera. I would also like to give thanks for the invaluable assistance provided by my co-workers Anne Orr and Maggie Elliott, who together with Liz Allen and Michayla Meredith, succeeded in making my stay in Lab 106A an enjoyable one (in spite of the glassware rota).

I am eternally grateful to my Glasgow friends, particularly to Angela, Celia and Joe, for keeping me sane during the past three and a half years. I am especially appreciative of their support and understanding during the writing of this Thesis. Also, the delis of Glasgow's West End deserve a special mention for providing inspiration, notably in recent months.

Finally, I would like to give special thanks to my parents Valerie and Terry, for their continual encouragement throughout my time at University.

The author was a recipient of a Medical Research Council studentship. Except where specified, all of the results described in this Thesis were obtained by the author's own efforts.

Summary

The understanding of gene regulation during varicella-zoster virus (VZV) infection is still rudimentary, although insight can be gained from the more extensively studied system of the genetically related herpes simplex virus type 1 (HSV-1). Upon infection, HSV-1 expresses three main classes of genes in a coordinately regulated temporal cascade; the immediate-early (IE), early (E) and late (L) genes. The product of the HSV-1 IE3 gene, Vmw175, is considered to be the major regulatory protein of HSV-1 because functional Vmw175 is an absolute requirement for transactivation of E and L gene expression and for the down-regulation of its own IE3 promoter. The polypeptide encoded by VZV gene 62 (VZV 140k) is the sequence homologue of HSV-1 Vmw175. Transient transfection experiments have demonstrated that VZV 140k is a potent and promiscuous transactivator of gene expression, and that it can also negatively autoregulate its own gene 62 promoter. In addition, VZV 140k can complement HSV-1 viruses lacking functional Vmw175. Therefore, by analogy to its HSV-1 counterpart, VZV 140k is likely to play a critical role in the regulation of VZV gene expression. The research described in this Thesis aimed to investigate the properties and activities of the DNA binding domain of the VZV 140k protein, in order to gain a better understanding of the mechanisms of 140k-mediated gene regulation during infection.

The primary sequences of VZV 140k and HSV-1 Vmw175 have been divided into five co-linear regions on the basis of the extent of their predicted amino acid identity. The region 2 sequences are particularly highly conserved between the two proteins and the regulatory functions of Vmw175 require the integrity of region 2. Furthermore, sequences spanning HSV-1 Vmw175 region 2 constitute a physically separable, stable DNA binding domain. Many of the experiments in this study stemmed from the initial demonstration that the corresponding region of the VZV 140k protein also yields a sequence-specific DNA binding domain when expressed as a non-fusion polypeptide in bacteria.

The bacterially expressed VZV 140k DNA binding domain specifically recognised multiple DNA sequences in the vicinity of its target promoters; the 140k binding sites identified within its own VZV gene 62 promoter may have a role in the negative autoregulatory function of the 140k protein. Several of the 140k recognition sites showed sequence similarity to the Vmw175 consensus binding sequence, although gel retardation assays with a systematically mutagenised binding site found the VZV 140k DNA binding

domain peptide to be less sequence-specific than the corresponding domain of HSV-1 Vmw175. These and further differences that have been identified between the DNA binding activities of the VZV 140k and HSV-1 Vmw175 DNA binding domains may explain the previously reported variations between the regulatory functions of 140k and Vmw175.

The minimal region of the VZV 140k protein required for sequence-specific DNA binding has been defined by a deletion analysis to within 162 residues, essentially comprising the highly conserved region 2. However, additional sequences from the C-terminal end of region 1 were necessary for the full DNA binding interaction, as determined by DNase I footprinting analysis. As such, the DNA binding domain of VZV 140k is larger than those reported for other transcriptional regulators (which are often less than 80 amino acids) and therefore it may constitute a novel type of DNA binding structure. A short sequence that exhibits intriguing homology to the conserved DNA recognition helix of the homeodomain DNA binding motif is found at the centre of the VZV 140k DNA binding domain. The alterations to the DNA binding interaction that resulted from the substitution of single amino acids within this region indicate its probable involvement in the DNA recognition by VZV 140k. Of particular note, mutation of a single lysine residue drastically reduced the DNA binding activity and destroyed the transactivation function of VZV 140k; this result emphasises the importance of DNA binding for VZV 140k-mediated gene regulation.

The highly purified VZV 140k DNA binding domain was a stable dimer in solution, as demonstrated by gel filtration, glycerol gradient centrifugation and cross-linking techniques. The fact that the VZV 140k DNA binding domain binds to DNA as a dimer was indicated by gel retardation assays that showed novel protein-DNA complexes of intermediate mobility following *in vitro* co-translation of pairs of VZV 140k peptides of differing sizes. In addition, the VZV 140k and HSV-1 Vmw175 DNA binding domain peptides readily heterodimerised on co-translation, which indicated that these two related domains have similar modes of dimerisation. *In vitro* co-translation of a wide range of truncation and insertion mutation versions of the VZV 140k and HSV-1 Vmw175 DNA binding domains, followed by analysis of their dimerisation by co-immunoprecipitation and their DNA binding ability by gel retardation, has indicated that the structural requirements for dimerisation are lower than for the DNA binding interaction.

Finally, construction of a hybrid Vmw175 protein with the VZV 140k DNA binding domain replacing that of HSV-1 Vmw175 has allowed the relatedness of the two domains to be assessed directly. The hybrid protein was functional in tissue culture transient transfection assays and furthermore, supported viral growth when expressed

from a recombinant HSV-1 genome. The details of the regulatory activities of the hybrid protein in the transient assays implied that the characteristic DNA binding activities of the VZV 140k DNA binding domain play an important part in determining its specific regulatory functions.

Abbreviations

A	adenine or absorbance or Amp(s)
Ac	acetate
APS	ammonium persulphate
ATP	adenosine-5'-triphosphate
ATCC	American Type Culture Collection
BHK	baby hamster kidney
BME	β–mercaptoethanol
bp	base pair(s) (nucleic acid)
BSA	bovine serum albumin
C	cytosine or carboxy (-terminal end of protein)
^{14}C	radioisotope carbon-14
CAT	chloramphenicol acetyltransferase
CHAPS	2-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate
Ci	Curie(s)
CMI	cell mediated immunity
CNS	central nervous system
CoA	coenzyme A
cpe	cytopathic effect
CS	calf serum
CT	calf thymus
D	Dalton(s)
dATP	2'-deoxyadenosine-5'-triphosphate
dCTP	2'-deoxycytidine-5'-triphosphate
DDAB	dimethyldioctadecyl-ammonium bromide
ddATP	2'-, 3'-dideoxyadenosine-5'-triphosphate
ddCTP	2'-, 3'-dideoxycytidine-5'-triphosphate
ddGTP	2'-, 3'-dideoxyguanosine-5'-triphosphate
ddNTP	2'-, 3'-dideoxyribonucleoside-5'-triphosphate
ddTTP	2'-, 3'-dideoxythymidine-5'-triphosphate
dGTP	2'-deoxyguanosine-5'-triphosphate
dH_20	sterile distilled water
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethylsuphoxide
DNA	2'-deoxyribonucleic acid

DNase	deoxyribonuclease
dNTP	2'-deoxyribonucleoside-5'-triphosphate
DOPE	L α–phosphatidylethanolamine, dioleoyl (C18:1, [cis]-9)
DTE	dithioerythritol
DTT	dithiothrietol
dTTP	2'- deoxythymidine-5'-triphosphate
ECL	enhanced chemiluminescence
E. coli	*Escherichia coli*
EDTA	ethylenediaminotetra-acetic acid
Et	ethidium
FITC	fluorescein isothiocyanate
FPLC	fast protein liquid chromatography
G	guanine
GCG	Genetics Computer Group (University of Madison, WI)
gD	glycoprotein D
GST	glutathione-S-transferase
^{3}H	radioisotope - tritium
HEPES	N-2 hydroxyethylpiperazine-N'-2-ethanesulphonic acid
HRP	horse radish peroxidase
HSV	herpes simplex virus
I	inosine
IE	immediate-early (gene)
Ig	immunoglobulin
IPTG	isopropyl-D-thiogalactoside
MAb	monoclonal antibody
MCR	multiple cloning region
MOPS	3-(N-morpholine) propanesulphonic acid
N	amino (-terminal end of protein)
NP40	Nonidet p40
NPT	non-permissive temperature
OD	optical density
ORF	open reading frame
^{32}P	radioisotope phosphorous-32
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG (8000)	polyethylene glycol (average molecular weight 8kD)
pfu	plaque forming unit(s)
pi	post-infection
PMSF	phenylmethylsulphonyl flouride
EHV	equine herpesvirus
ICP	infected cell polypeptide

RF	replicative form (bacteriophage DNA)
RGB	resolving gel buffer
RNA	ribonucleic acid
RNase A	ribonuclease A
rpm	revolutions per minute
RT	room temperature (20-25°C)
S	cytosine or guanosine
^{35}S	radioisotope sulphur-35
SαM	sheep anti-mouse immunoglobulin
SDS	sodium dodecyl sulphate
SGB	stacking gel buffer
ss	single-stranded (nucleic acid)
SV40	simian virus 40
syn	syncytial plaque morphology locus (+/- phenotypes)
T	thymine
TAF	TATA-binding protein associated factor
TBS	tris buffered saline
TEMED	N-, N-, N'-, N'-tetramethylethylenediamine
tk	thymidine kinase
TLC	thin layer chromatography
Tris	tris (hydroxymethyl) aminomethane
tRNA	transfer ribonucleic acid
TWEEN-20	polyoxyethylene-sobitanmonolaurate
ts	temperature sensitive
u	unit
uv	ultra-violet
V	volt(s)
v/v	volume / volume
vol	volume(s)
Vmw	apparent molecular weight of virus-induced polypeptide
VZV	varicella-zoster virus
W	Watt(s)
w/v	weight / volume
w/w	weight / weight
wt	wild type
x-gal	5-bromo-4-chloro-3-indoyl-β-D-galactoside
Y	pyrimidine

Amino acid symbols

A	Ala	alanine	M	Met	methionine
C	Cys	cysteine	N	Asn	asparagine
D	Asp	aspartic acid	P	Pro	proline
E	Glu	glutamic acid	Q	Gln	glutamine
F	Phe	phenylalanine	R	Arg	arginine
G	Gly	glycine	S	Ser	serine
H	His	histidine	T	Thr	threonine
I	Ile	isoleucine	V	Val	valine
K	Lys	lysine	W	Trp	tryptophan
L	Leu	leucine	Y	Tyr	tyrosine

Chapter 1: Introduction

The product of varicella-zoster virus gene 62 (VZV 140k) is likely to be the major transcriptional regulatory protein of VZV, by homology to its counterpart HSV-1 protein, Vmw175. The research presented in this Thesis concerns the properties and functions of the DNA binding domain of the VZV 140k polypeptide. This introduction aims to overview briefly the biology and molecular biology of VZV and to review the current state of knowledge of gene regulation among the *Alphaherpesvirinae*, with particular emphasis on the HSV-1 prototype system and the fundamental role played by the Vmw175 protein. Since regulation of gene expression by viral immediate-early polypeptides requires interactions with the host cell transcriptional machinery, the later Sections of this Introduction will discuss relevant aspects of eukaryotic gene expression, with particular emphasis on the mechanisms whereby regulatory proteins influence RNA polymerase II transcription.

1A The biology of VZV

1A.1 Definition, morphology and classification of the herpesviruses

Members of the family *Herpesviridae* are characterised by the inclusion of a linear double-stranded DNA genome in an enveloped virion particle of 120-300nm in diameter (reviewed by Roizman and Furlong, 1974; Rixon, 1993). The herpesvirus virion (Fig. 1A.1) comprises four distinct layers as follows: the **core** contains the viral DNA (Epstein, 1962) and lies within the icosadeltahedral **capsid.** The capsid is 100-110nm in diameter and comprises 162 capsomeres (150 hexameric and 12 pentameric) (Wildy *et al.*, 1960). Surrounding the nucleocapsid is an amorphous proteinaceous **tegument** (Roizman and Furlong, 1974). The outer **envelope** consists of a host cell derived trilaminar membrane with viral glycoprotein spikes on the surface (Wildy *et al.*, 1960; Wildy and Watson, 1962; Spear and Roizman, 1972).

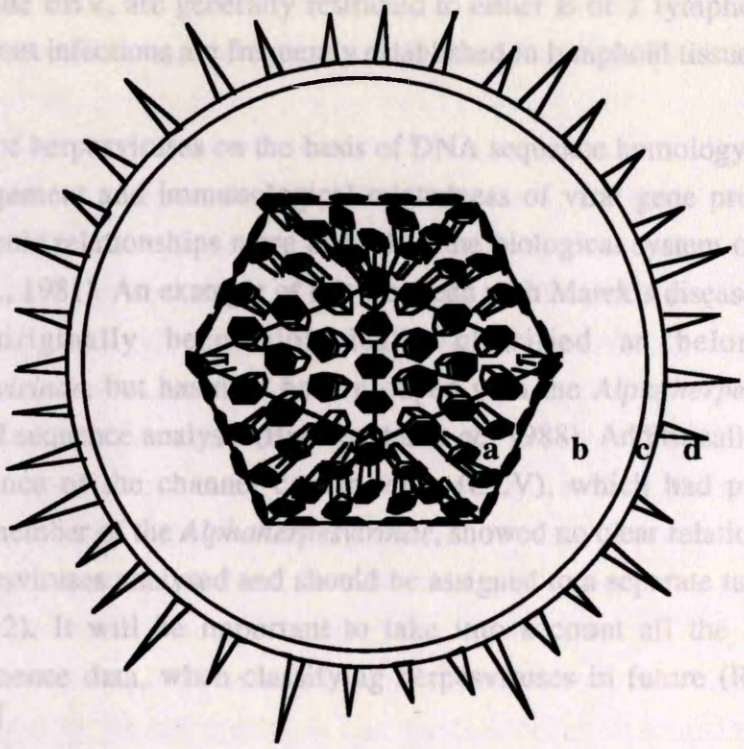


Figure 1A.1 The morphology of the herpesviruses. Schematic representation of the herpesvirion seen through a cross-section of the envelope (c), with glycoprotein spikes (d) of varying lengths projecting from its surface (Stannard *et al.*, 1987). The sides of the capsid (a) show twofold symmetry and the hexameric capsomeres are illustrated; the viral DNA is located within the capsid. The tegument (b) is located between the envelope and the nucleocapsid.

To date nearly 100 herpesviruses have been characterised, with diverse animal hosts and biologies. In common, all replicate in the nucleus, are able to remain latent in their natural hosts, and production of progeny virus is accompanied by destruction of the infected cell. Seven herpesviruses have been isolated from humans: herpes simplex virus 1 and 2 (HSV-1 and HSV-2), human cytomegalovirus (HCMV), varicella-zoster virus (VZV), Epstein-Barr virus (EBV) and human herpesvirus 6 and 7 (HHV-6 and HHV-7). These human herpesviruses are distributed in three subfamilies of the *Herpesviridae: Alpha-*, *Beta-*, and *Gammaherpesvirinae*, divided on the basis of their biological properties (Matthews, 1982; Roizman, 1982). Members of the ***Alphaherpesvirinae*** have a relatively short reproductive cycle (typically less than 24hr in tissue culture) and primarily establish latent infections in the sensory ganglia. This subfamily includes HSV-1, HSV-2, VZV, pseudorabies virus (PRV) and equine herpesvirus 1 (EHV-1) (the latter two are infections of swine and horses respectively). The ***Betaherpesvirinae*** members have a restricted host range in tissue culture and long reproductive cycles. Infection spreads slowly in tissue culture and results in enlargement of infected cells (cytomegalia). The virus can be maintained in the latent form in various tissues and members include HCMV, HHV-6 and HHV-7. The ***Gammaherpesvirinae,*** whose

members include EBV, are generally restricted to either B or T lymphocytes in tissue culture, and latent infections are frequently established in lymphoid tissues.

Classification of herpesviruses on the basis of DNA sequence homology, similarities in genome arrangement and immunological relatedness of viral gene products tends to reflect phylogenic relationships more truly than the biological system of classification (Roizman *et al.*, 1981). An example of this was seen with Marek's disease virus (MDV), which had originally been biologically classified as belonging to the *Gammaherpesvirinae*, but has now been grouped with the *Alphaherpesvirinae* on the basis of limited sequence analysis (Buckmaster *et al.*, 1988). Additionally, the complete genome sequence of the channel catfish virus (CCV), which had previously been regarded as a member of the *Alphaherpesvirinae*, showed no clear relationship to any of the other herpesviruses analysed and should be assigned to a separate taxonomic group (Davison, 1992). It will be important to take into account all the available data, especially sequence data, when classifying herpesviruses in future (Roizman *et al.*, 1992).

1A.2 Biology and pathology of VZV infections

Varicella-zoster virus is an unusual herpesvirus in causing two common distinct clinical syndromes: chickenpox (varicella) upon primary infection, and shingles (herpes zoster) upon reactivation of latent virus. Chickenpox is a highly contagious but commonly mild exanthem, usually ubiquitous in children. Shingles is a less common endemic local vesicular condition, often extremely painful, which usually affects older patients. Although neither disease is life-threatening, both have possible serious consequences in immunocompromised individuals. The pathology of each disease will be discussed in more detail.

1A.2.1 *History*

The first report of chickenpox is attributed to a ninth century Persian physician, Rhazes, and was described as a pustular skin disease that conferred no protection against smallpox. The term 'varicella' is a variation on the word *variola* meaning smallpox - a more severe disease that was often confused with chickenpox until Heberden clinically distinguished them in the late 18th century (Gordon, 1962). In 1892, Bokay suggested that chickenpox in children could be related to herpes zoster exposure and Lipschutz first described in 1921, that the histology of vesicular lesions present in chickenpox and in zoster appeared similar (Jako and Jako, 1986). Following the first isolation of VZV in tissue culture in 1953 (Weller, 1953), Weller and co-workers finally demonstrated the

identity between viruses isolated from varicella and zoster patients in 1958, leading them to suggest the name 'varicella-zoster virus' (Weller *et al.*, 1958; Weller and Witton, 1958).

1A.2.2 *Varicella*

VZV entry is usually through the mucosa of the upper respiratory tract and oropharynx. Viral replication begins at the site of primary inoculation and after 4-5 days, primary viremia is established during which virus disseminates via the bloodstream and lymphatic system. Subsequently, virus taken up by cells of the reticuloendothelial system undergoes multiple cycles of replication. The host's immune response initially limits viral replication, but is soon overwhelmed and a more extensive secondary viremia develops. High titres of virus are released into the bloodstream, followed by infection of capillary endothelial cells and subsequent viral spread to epithelial cells of the endodermis and certain mucosal surfaces. Intra- and extracellular edema develop in the epidermis, producing a characteristic fluid filled vesicle containing cell-free VZV that may be shed to the environment. The viremic phase is terminated by specific humoral and cellular responses (CMI), approximately 3 days after the appearance of lesions (Asano *et al.*, 1985; Arvin *et al.*, 1986).

Complications of varicella result from increased viral spread and failure to limit replication. Varicella pneumonia is the most common serious complication in adults, neonates and the immunocompromised. Neurological conditions, encephalitis for example, develop in approximately 0.3% of varicella cases with a 5-15% rate of mortality (Guess *et al.*, 1986).

1A.2.3 *Herpes zoster*

VZV is thought to pass from varicella cutaneous lesions to the corresponding sensory ganglia via the nerve fibers, where it remains quiescent until reactivation as herpes zoster. The observation that reactivation is often precipitated by conditions associated with immunosuppression led Hope-Simpson (1965) to suggest that deterioration of host defences below the level of containment allows reactivated VZV to replicate in the sensory ganglia. By this model (Fig. 1A.2), most such reactivations (or reversions) are contained by circulating antibody or CMI, but occasionally the host defences fail to contain viral multiplication, leading to necrosis and intense inflammation of the ganglia and often also neuralgia. Infectious virus spreads down the sensory nerve and infects the skin, producing characteristic dermatomal zoster vesicles that gird the trunk or other parts of the body (*zoster* is derived from the Greek word for 'girdle').

A common complication of herpes zoster is post-herpetic neuralgia (pain lasting more than 1 month). Additionally, spread of infection to adjacent areas of the nervous system can result in paralysis or palsy and in rare cases persistent CNS infection can lead to encephalitis. As with varicella, herpes zoster tends to be more severe with immunocompromised patients and the severity of the condition is related to the status of VZV-specific CMI (Dolin *et al.*, 1978).

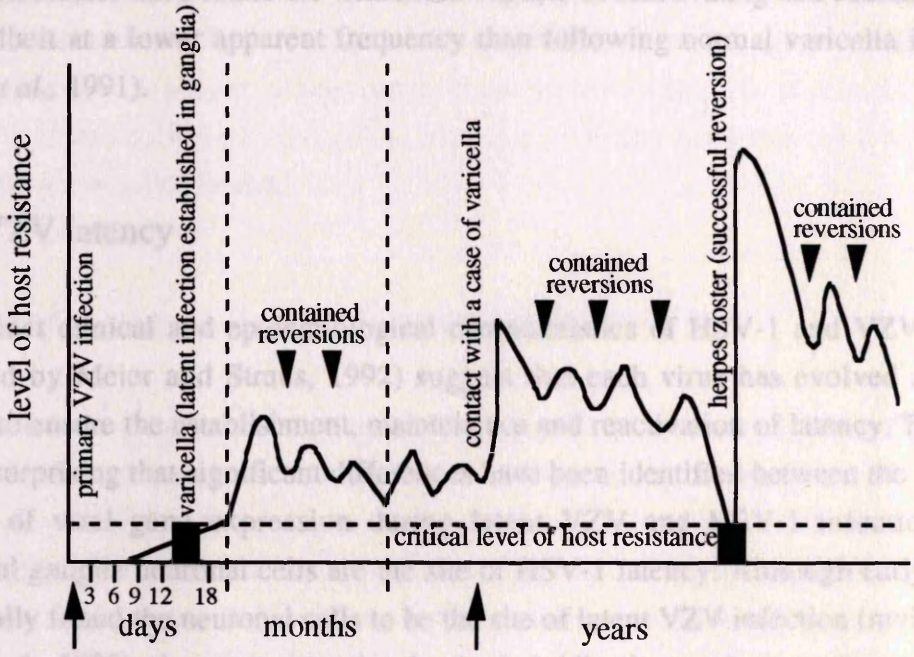


Figure 1A.2 Hope-Simpson's model of herpes zoster pathogenesis. The two arrows at the bottom indicate time of exposure to VZV; and black boxes represent clinically apparent disease; the dashed horizontal line indicates the critical level of host resistance required to contain VZV replication and reversion. Increased viral replication is marked by a corresponding increase in antibody production (increased level of host resistance); this neutralises the virus and is followed by a drop in host resistance in the case of successful reversions. Clinical disease occurs when the antibody level falls below the critical level. The Figure is an adaptation from Hope-Simpson (1965).

1A.2.4 *Treatment and prevention*

Three anti-viral compounds have proved efficacious for treatment of varicella, although specific therapy is often unnecessary in immunocompetent children; these are interferon, vidarabine (adenosine arabinoside) and acyclovir (an acyclic guanosine derivative) (reviewed by Whitley, 1992). The latter two are anti-VZV nucleoside analogues that specifically terminate DNA replication mediated by the viral DNA polymerase. These antivirals have also been reported to modestly alleviate acute herpes zoster if administered early in the course of the disease.

Preventive measures from varicella are not usually taken because infection generally provides lifelong immunity, although administration of VZIG prophylaxis (a high-titre VZV antibody preparation) greatly reduces the severity of varicella in high risk patients

exposed to VZV. A live attenuated VZV vaccine (Oka strain) has been produced and used extensively in Japan. This affords a high degree of protection against varicella and is not associated with an increased incidence of zoster (Gershon *et al.*, 1992). Since chickenpox in adulthood would have more serious consequences than in childhood, the main concern regarding the attenuated vaccine has been the duration of immunity and this will only be resolved by longer follow-up studies of vaccinated individuals. Indeed, population studies have found the Oka strain capable of reactivating and causing herpes zoster, albeit at a lower apparent frequency than following normal varicella infection (Hardy *et al.*, 1991).

1A.3 VZV latency

The distinct clinical and epidemiological characteristics of HSV-1 and VZV latency (reviewed by Meier and Straus, 1992) suggest that each virus has evolved a unique strategy to ensure the establishment, maintenance and reactivation of latency. Therefore it is not surprising that significant differences have been identified between the sites and patterns of viral gene expression during latent VZV and HSV-1 infections. The trigeminal ganglia neuronal cells are the site of HSV-1 latency. Although early studies additionally found the neuronal cells to be the site of latent VZV infection (reviewed by Cohrs *et al.*, 1992), the most extensive *in situ* hybridisation analysis performed to date found VZV latent transcripts to be specifically localised to the satellite cells immediately surrounding the neuron (Croen *et al.*, 1988). In contrast, both neuronal and non-neuronal cells support lytic VZV infections, during which all regions of the VZV genome are expressed in both types of cell (Nagishama *et al.*, 1975; Croen *et al.*, 1988).

In situ hybridisation analysis using RNA probes representing almost the entire VZV genome, identified transcription from at least three distinct and widely separated regions of the genome during latency (Croen *et al.*, 1988). The genes selectively expressed during VZV latency include some, but not all of the putative immediate-early genes (including gene 62) and certain early genes (Meier *et al.*, 1993) (Section 1B.4.1). The latent VZV transcripts are polyadenylated and appear to represent the same gene products that are transcribed during lytic infection. Thus VZV latency is characterised by a broad pattern of latency-associated transcription. This is in marked contrast to the situation with HSV-1 latency where two or three co-linear transcripts, termed the 'latency associated transcripts' (or LATs), are detected in latently infected cells at greatly enhanced levels compared to during lytic infection (reviewed by Stevens, 1989). The HSV-1 LATs are spliced and include open reading frames which span a portion of the IE1 gene in the opposite polarity (see Fig. 1B.1), although it is not clear if a protein

product is made *in vivo*. There have been no reports of VZV transcripts equivalent to the HSV-1 LATs, but if VZV latent specific transcripts exist they are unlikely to be analogous to the LATs because there is no true VZV counterpart of the HSV-1 IE1 ORF and also no splicing has been detected in the corresponding region of the VZV genome. However, the role played by the HSV-1 LATs during latent infection is far from clear.

During latency the HSV genome is extrachromosomal, circular and organised into an array of nucleosomes that resemble cellular chromatin (Fraser *et al.*, 1990; Deshmane and Fraser, 1989). It would be interesting to ascertain whether this is also the case for the VZV genome during latency, and in particular, whether the transcriptionally active regions are nucleosome bound.

1A.4 VZV growth *in vitro* and animal models

VZV has been propagated in a variety of primary cell lines, usually of human or simian origin, although continuous cell lines such as Vero or human melanoma cells have also been employed with equal success. However, even cultured cells derived from the natural host are considered to be only partially permissive for VZV growth as they do not readily support the production of cell-free infectious virus (Weller and Witton, 1958). Because of the highly cell associated nature of VZV, infected cells are routinely used as an inoculum; however this presents difficulties for tissue culture experiments requiring quantitative controlled infections. Highly optimised conditions for isolating cell-free virus produce yields at best one order of magnitude higher than the initial inoculum, and infection with the cell-free virus is very inefficient compared to using infected cells as the inoculum (Schmidt and Lennette, 1976). The various alternative explanations for the high particle/infectivity ratio of VZV replicated *in vitro* include the hypothesis that VZV particles are inactivated following targeting to lysosomal compartments during egress of the virus from the cell (Gershon *et al.*, 1973). The investigation of all aspects of VZV biology and structure have been significantly impaired by the inability to isolate stable, high titre stocks of cell-free virus. Consequently VZV growth in tissue culture is very poorly characterised.

The lack of a good animal model for infection has also hampered the understanding of VZV pathogenesis. The guinea-pig and common marmoset may be infected experimentally with VZV (Myers *et al.*, 1980; Provost *et al.*, 1987), but neither animal provides a good model for vesicular disease or latency. Simian varicella virus (SVV) infection of non-human primates is probably a closer model to the human varicella disease (Dueland *et al.*, 1992).

1A.5 The VZV Genome

The VZV DNA molecule is one of the smallest herpesvirus genomes at approximately 80 x 10^6 Daltons in size (Dumas *et al.*, 1980) and has a G+C content of 46%, which is rather less than that of most herpesviruses (Ludwig *et al.*, 1972). Restriction endonuclease mapping has demonstrated that the linear duplex genome comprises two covalently linked segments, L (long) and S (short). The VZV DNA exists predominantly in two equally abundant isomeric forms, as the short segment is capable of inversion with respect to the long segment (Dumas *et al.*, 1981) (right panel, Fig. 1A.3). The unique sequences of the long and short segments (U_L and U_S) are flanked by distinct inverted repeats R_L and R_S, and the location of each repeat is defined by I (internal) or T (terminal) (left panel, Fig. 1A.3) (Ecker and Hyman, 1982; Straus *et al.*, 1982; Gilden *et al.*, 1982; Davison, 1984). More recently it has been found that the L segment inverts with respect to the S sequence approximately 5% of the time (Kinchington *et al.*, 1985; Hayakawa and Hyman, 1987). The viral DNA therefore exists as four genome isomers, two major and two minor (right panel, Fig. 1A.3).

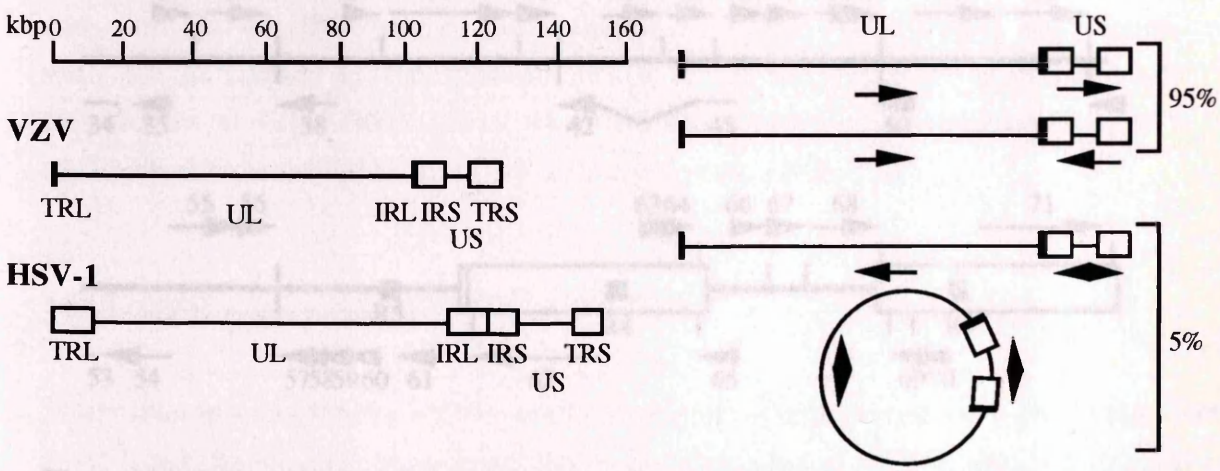


Figure 1A.3 Comparison of VZV and HSV-1 genomes, and the VZV genome isomers. The left panel shows schematics of the VZV (125kbp) and HSV-1 (152kbp) genomes and a scale in kbp is shown above. Inverted repeats (TRL/IRL and TRS/IRS) are shown as rectangles. The right panel shows the different isomers of the VZV genome; the two predominant forms contain the long segment in a single orientation, while the short segment is detected in an equimolar ratio in either of the two possible orientations, as indicated by the arrows. A small percentage of the VZV DNA molecules exist either in a circular form or with the long segment inverted.

An unpaired nucleotide is found at the 3' end of each VZV DNA strand and the genome is thought to circularise after viral entry and replicate by a rolling circle mechanism (Davison, 1984). In support of this hypothesis, circular VZV genomes have been visualised by electron microscopy (Straus et *al.*, 1981; Kinchington *et al.*, 1985). Additionally, five regions of reiterated VZV sequences (R1 to R5) were identified

initially by restriction mapping, and the number of repeats within each reiteration varies between VZV strains (see Fig. 1A.4). The role of the reiterations is unknown, but they have proved useful in investigating epidemiology of VZV infections, and genome variability of VZV isolates as determined by restriction endonuclease mapping (Straus *et al.*, 1983; Hayakawa *et al.*, 1986; Hondo *et al.*, 1987; Kinoshita *et al.*, 1988; Hondo and Yogo, 1988).

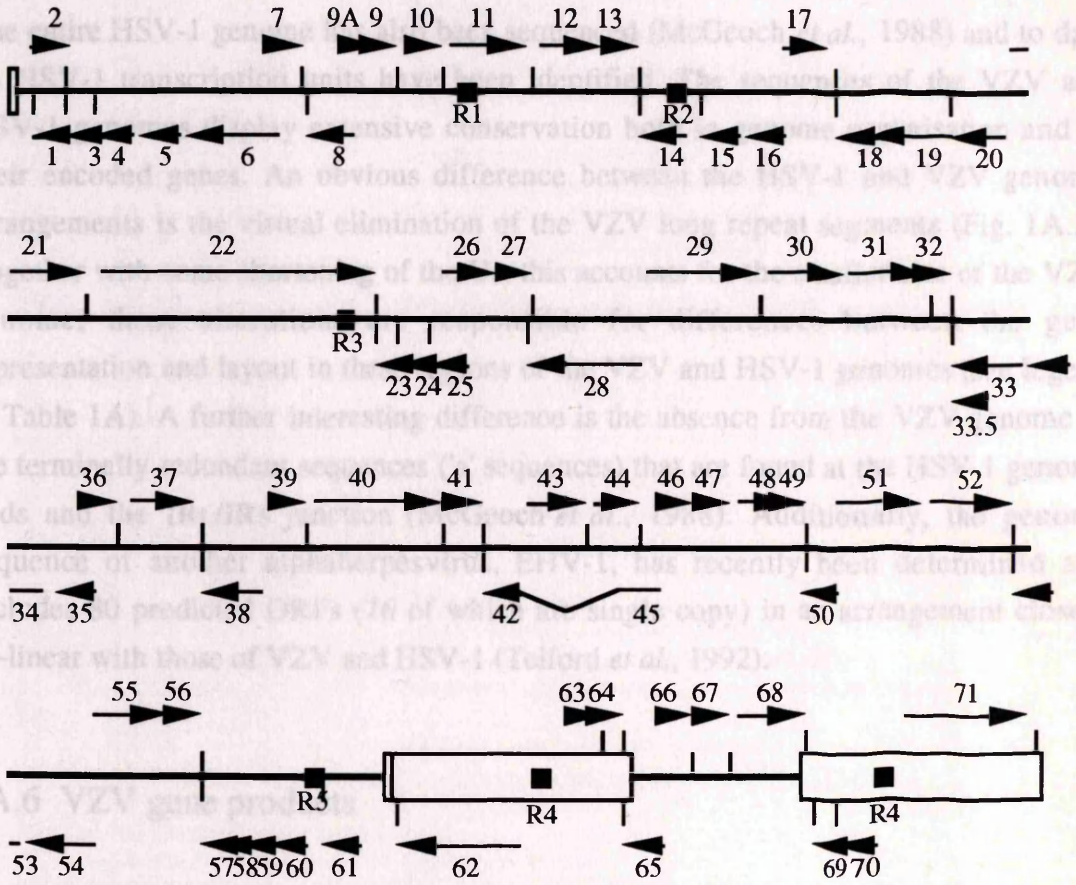


Figure 1A.4 The VZV gene arrangement. Heavy horizontal lines represent the unique regions of the genome and open rectangles depict the inverted repeats flanking the unique regions, as indicated in Figure 1A.3. ATG-initiated ORFs predicted to encode proteins are shown as arrows against the appropriate DNA strand, with the given nomenclature. The open reading frames were identified by Davison and Scott (1986), with the exceptions of ORF 9A (Barnett *et al.*, 1992) and gene 33.5 which encodes a truncated form of the ORF33 product (lacking the N-terminal portion) (Telford *et al.*, 1992). The location of the intron between ORFs 42 and 45 is indicated. Light vertical lines show positions of potential polyadenylation sites in the appropriate strand and the locations of sequence reiterations (R1 to R5) are depicted by black boxes. The Figure was adapted from Davison and Scott (1986).

Davison and Scott (1986) have sequenced the entire genome of the Dumas strain of VZV. The total genome contained 124,884bp; U_L is 104836bp; TR_L and IR_L 88.5bp each; Us 5232bp; and TRs and IRs 7319.5bp each. They originally identified a total of 71 ATG initiated ORFs, but subsequently a further two ORFs have been proposed (Fig. 1A.4). However, the gene assignment process is potentially fallible because it is based

primarily on interpretation of DNA sequences; it remains to be determined whether all the ORFs are utilised and it is possible that more will be identified in the future. Of the 73 VZV transcription units, 66 are present in single copy and three are located within the genomic repeats and are therefore diploid. ORFs 42 and 45 are probably expressed as a spliced transcript, as indicated in Figure 1A.4. The VZV genes are arranged compactly on both DNA strands but do not overlap extensively (Fig. 1A.4).

The entire HSV-1 genome has also been sequenced (McGeoch *et al.*, 1988) and to date 76 HSV-1 transcription units have been identified. The sequences of the VZV and HSV-1 genomes display extensive conservation both in genome organisation and in their encoded genes. An obvious difference between the HSV-1 and VZV genome arrangements is the virtual elimination of the VZV long repeat segments (Fig. 1A.3). Together with some shortening of the Us, this accounts for the smaller size of the VZV genome; these alterations are responsible for differences between the gene representation and layout in these regions of the VZV and HSV-1 genomes (see legend to Table 1A). A further interesting difference is the absence from the VZV genome of the terminally redundant sequences ('a' sequences) that are found at the HSV-1 genome ends and the IRL/IRS junction (McGeoch *et al.*, 1988). Additionally, the genome sequence of another alphaherpesvirus, EHV-1, has recently been determined and includes 80 predicted ORFs (76 of which are single copy) in an arrangement closely co-linear with those of VZV and HSV-1 (Telford *et al.*, 1992).

1A.6 VZV gene products

Initial attempts to identify VZV-specific polypeptides and analyse their functions were complicated by the difficulties associated with purification of VZV virions. The functions of the HSV-1 gene products have been investigated more extensively, and knowledge of the VZV genome sequence has allowed comparison of the predicted amino acid sequences of VZV ORFs with those of characterised HSV-1 proteins (McGeoch *et al.*, 1986; Davison and Scott, 1986; McGeoch *et al.*, 1988). Independent information as to the likely functions of the predicted products of certain VZV genes has also been gained from either sequence comparisons with non-herpesvirus proteins or from experimental work (reviewed by Davison, 1991). Thus, tentative functions, based mainly on the sequence comparisons, have been ascribed to many of the VZV genes (Table 1A). The VZV genome sequence predicts five membrane glycoproteins (gpI-gpV; Davison *et al.*, 1986) each having an HSV-1 counterpart (Table 1A) (Davison and Scott, 1986); each prediction has been shown to be correct experimentally (see Section 1A.7.2).

VZV Gene	HSV-1 counterpart	Assigned function of VZV product	Status of HSV-1 product
1			
2			
3	UL55		
4	UL54/IE2 (Vmw63)	Regulator of gene expression	e
5	UL53 (gK)	Possible membrane protein	
6	UL52	Component of DNA helicase-primase complex	e
7	UL51		
8	UL50	Deoxyuridine triphosphatase	ne
9	UL49	Tegument protein	
9A	UL49A	Possible membrane protein	
10	UL48 (Vmw65)	Tegument protein - transactivator of IE genes	e
11	UL47	Tegument protein	ne
12	UL46		ne
13		Thymidylate synthetase	
14	UL44 (gC)	Membrane glycoprotein (gpV) - role in cell entry	ne
15	UL43	Probable integral membrane protein	ne
16	UL42	Associated with DNA polymerase	e
17	UL41	Host shutoff virion protein	ne
18	UL40	Ribonucleotide reductase small subunit	e/ne
19	UL39	Ribonucleotide reductase large subunit	e/ne
20	UL38	Capsid protein	e
21	UL37	Possible DNA binding function	
22	UL36	Tegument protein	e
23	UL35	Capsid protein	
24	UL34	Possible virion protein	
25	UL33	Role in capsid maturation/DNA packaging	e
26	UL32		
27	UL31		
28	UL30	DNA polymerase	e
29	UL29	Single-stranded DNA binding protein	e
30	UL28	Role in capsid maturation/DNA packaging	e
31	UL27 (gB)	Membrane glycoprotein (gpII) - role in cell entry	e
32			
33	UL26	Protease, acts in virion maturation	e
33.5	UL26.5	Capsid protein, processed by UL26 in HSV-1	
34	UL25	Virion protein	e
35	UL24		ne
36	UL23	Thymidine kinase	ne
37	UL22 (gH)	Membrane glycoprotein (gpIII) - role in cell entry	e
38	UL21		
39	UL20	Integral membrane protein - role in virion egress	e/ne
40	UL19	Major capsid protein	e
41	UL18	Capsid protein	e
42+45	UL17		
43	UL16		ne
44	UL15, exons 1 and 2		
46	UL14		
47	UL13	Tegument protein, probable protein kinase	ne
48	UL12	DNase - role in maturation/packaging DNA	e
49	UL11	Myristylated tegument protein - role in envelopment and transport of virions	ne
50	UL10	Probable integral membrane protein	ne
51	UL9	ORI-binding protein	e
52	UL8	Component of DNA helicase-primase complex	e
53	UL7		
54	UL6	Role in virion morphogenesis	e
55	UL5	Component of DNA helicase-primase complex	e
56	UL4		ne
57			
58	UL3		ne

VZV Gene	HSV-1 counterpart	Assigned function of VZV product	Status of HSV-1 product
59	UL2	Uracil-DNA glycosylase	ne
60	UL1(gL)		
61	RL2 or IE0 (Vmw110)	Transcriptional regulator	ne
62/71	RS1 or IE3 (Vmw175)	Transcriptional regulator	e
63/70	US1 or IE4 (Vmw68)	Transcriptional regulator	e/ne
64/69	US10	Virion protein	ne
65	US9	Tegument protein	ne
66	US3	Protein kinase	ne
67	US7 (gI)	Membrane glycoprotein (gpIV)	ne
68	US8 (gE)	Membrane glycoprotein (gpI)	ne

Table 1A Functions of the VZV proteins. The first two columns give the nomenclature of the 73 VZV genes and the corresponding HSV-1 genes, as determined by sequence comparisons (the identity of HSV-1 glycoprotein and transcriptional regulator gene products are given in brackets in the second column). The third column gives the likely functions of VZV gene products (data taken from Davison, 1991; Ostrove, 1990; McGeoch *et al.*, 1993). The references to appropriate DNA sequence data on which the assignment of VZV gene functions are based are given in Davison and Scott (1986) and functional characterisations of VZV gene products are reviewed by Davison (1991). The final column indicates (where determined) whether the HSV-1 counterpart proteins are essential (e) or non-essential (ne) for all viral growth. 'e/ne' specifies gene products that are essential for HSV-1 infection of animal models but are dispensable for viral growth in tissue culture. VZV has five genes (genes 1, 2, 13, 32 and 57) that have no HSV-1 counterparts. Likewise, three HSV-1 genes in the long segment (UL45, UL56, γ34.5) have no VZV counterparts. The complement of genes and also the gene order in the short segments differs to a greater extent between VZV and HSV-1 than the essentially co-linear long segments; 7 HSV-1 Us genes have no VZV counterparts (Davison and McGeoch, 1986).

The VZV genome also encodes homologues of four of the five HSV-1 IE proteins (VZV ORFs 4, 61, 62 and 63) and these are likely to play roles in regulation of VZV gene expression of similar importance to their HSV-1 counterparts (Sections 1B.3.4 and 1B.4.2). Like HSV-1, VZV specifies a large number of enzymes involved in DNA synthesis and this is a characteristic of herpesviruses not shared with other animal nuclear DNA viruses (see Section 1A.8).

Interestingly, VZV is unique for an alphaherpesvirus in its possession of a gene encoding a thymidylate synthetase enzyme (TS). VZV gene 13 was predicted to encode a TS enzyme by its homology to non-herpesvirus TS enzymes (Davison and Scott, 1986) and the ability of the VZV TS to complement *E. coli* TS^- strains verified its identity as a functional TS enzyme (Thompson *et al.*, 1987). Similarly, the protein kinases encoded by VZV genes 47 and 66 were identified as such by their sequence homology to known serine-threonine protein kinases (McGeoch and Davison, 1986; Smith and Smith, 1989). The thymidylate kinase (*tk*) enzyme (sometimes termed deoxypyrimidine kinase - dPyK) encoded by VZV gene 36, is probably the best characterised VZV gene product. A full decade before the *tk* gene was mapped (Sawyer *et al.*, 1986), it was recognised that VZV infection results in the induction of a *tk* activity (Dobers on *et al.*, 1976; Ogino *et al.*, 1977). Despite being sequence homologues, the VZV and HSV-1 *tk* enzymes have different enzymatic properties (Cheng *et al.*, 1979).

Thus the elucidation of the VZV sequence significantly increased the knowledge of the potential functions of VZV proteins, and the availability of clones spanning the entire VZV genome makes possible the expression and subsequent analyses of isolated VZV gene products, circumventing the problems associated with growing VZV in tissue culture. Thus the availability of the VZV and HSV-1 genome sequences provides a very favourable starting point for those studying VZV gene products, and it is likely that many interesting differences between the functional details of the VZV and HSV-1 homologue proteins will be discovered, as illustrated by the characterisation of the VZV *tk* enzyme.

1A.7 Identification and mapping of VZV transcripts

1A.7.1 *VZV transcript analysis*

Concurrent with the derivation of the VZV genome sequence, the overall transcription pattern of the genome was investigated. Northern blotting experiments identified the genomic positions that corresponded to 58 discrete VZV-specific RNA transcripts from total RNA extracts of VZV infected cells (Ostrove *et al.*, 1985). A further 20 abundant VZV transcripts were subsequently identified and the direction of transcription was discerned for each RNA species using ssRNA probes (Reinhold *et al.*, 1988). Cytoplasmic polyadenylated VZV transcripts have also been mapped (Maguire and Hyman, 1986). These studies indicated that the size and map locations of many of the VZV transcripts were consistent with the open reading frames identified within the genome sequence (Davison and Scott, 1986). In addition, the size and locations of VZV transcripts showed generalised analogy to those of the HSV genome and indicated that transcription in the VZV system grossly parallels transcription in HSV-1 infected cells.

Although much information can be gained from sequence analysis, verification of the actual situation requires experimental data; for example detailed transcript analysis has shown that predictions based on sequence data alone are not always correct. Analysis of the 5' end of the VZV thymidine kinase (*tk*) transcript has demonstrated that transcription is initiated from a specific start site that is upstream of several other putative transcriptional start sites for VZV gene 36. Additionally, the 3' end of the *tk* transcript extends beyond several predicted poly(A) sites by over 400bp (Davison and Scott, 1986). Coupled with the examination of sequences upstream and downstream of other VZV ORFs, this data led Davison and Scott (1986) to speculate that there may be considerable variation between the transcriptional and translational control of the VZV and HSV-1 genes. Regarding the VZV 140k protein, a 4.3kb transcript encoding 140k is expressed in cell lines stably transformed with gene 62 (Felser *et al.*, 1988) and 140k

was confirmed as the product of gene 62 using a 140k-specific monoclonal antibody (Forghani *et al.*, 1990). The 5' end of the gene 62 transcript has been defined by primer extension and S1 nuclease analyses (McKee *et al.*, 1990). In addition, the 5' and 3' ends of the transcripts of gpIV (gene 67) and gpV (gene 14) (Ling *et al.*, 1992; 91) and the 5' end of IE gene 61 (Nagpal and Ostrove, 1991; Stevenson *et al.*, 1992) have been mapped.

1A.7.2 ***Mapping of genes encoding specific VZV polypeptides***

Several groups have endeavoured to identify and map the transcripts encoding specific proteins of functional interest, in particular, much attention has been given to mapping the VZV glycoprotein-encoding genes. Ellis *et al.* (1985) identified the DNA encoding gpI (the major structural glycoprotein) by screening a bacterial expression vector library containing VZV DNA fragments for gpI antigens using monoclonal antibodies, followed by the mapping of the gpI transcript to VZV gene 68. *In vitro* transcription of selected mRNA species, followed by immunoprecipitation with an anti-gpII antibody was used to identify the gpII translation product and hence map the gpII transcript to gene 31 (Keller *et al.*, 1986). Yet another approach was used to identify the gene encoding the gpIII glycoprotein: the gpIII protein was purified and derivation of the N-terminal sequence allowed identification of its corresponding gene (gene 37) by hybridisation with degenerate nucleotides based upon the protein sequence (Keller *et al.*, 1987). VZV gene 67 has been identified as encoding the gpIV family of glycoproteins (Davison *et al.*, 1985). Finally, sequence analysis predicted gene 14 to encode glycoprotein gpV, and has since been shown to be the case by functional characterisation of the VZV gene 14 products (Kinchington *et al.*, 1990).

Various non-structural VZV gene products have additionally been investigated. The transcript specifying the VZV thymidine kinase (*tk*) enzyme was mapped by transfecting plasmids carrying VZV fragments into mouse cells which lacked functional cellular *tk* enzyme (tk^-), followed by screening for tk^+ transformants. The VZV transcripts from the tk^+ cells mapped to VZV gene 36 (Sawyer *et al.*, 1986). There are several further examples where genes specifying particular polypeptides have been determined experimentally, however the wealth of information that can be obtained from the VZV and HSV-1 genome sequences suggests that a simpler and more informative approach in many cases might be to identify the VZV gene with the putative properties of interest by sequence analysis, followed by characterisation of the translation products of the cloned gene. The was the approach used in the mapping of VZV gene 62, encoding 140k.

1A.8 The VZV replicative cycle

Very little is known about the VZV life cycle at the molecular level, but it is unlikely to differ greatly from that of the HSV-1 prototype system (for a review, see Grose and Ng, 1992). In particular, the biochemical processes of viral DNA replication are unlikely to differ markedly between VZV and HSV-1 because each of the seven HSV-1 genes that are essential for HSV-1 DNA replication (Wu *et al.*, 1988) have VZV sequence homologues (VZV ORFs 6, 16, 28, 29, 51, 52 and 55) (see Table 1A) (Davison and Scott, 1986). For example, the VZV DNA polymerase (ORF 28) shares more than 50% amino acid sequence identity with its HSV-1 counterpart (Becker, 1988). A VZV origin of replication (ORI) has been identified within each copy of the genome short repeat, in a position analogous to that occupied by one of the HSV-1 replication origins (ORIs) (Stow and Davison, 1986). Like the HSV-1 ORIs, the VZV ORI contains an almost perfect 45bp palindrome, with central alternating tracts of A and T residues. The VZV genome appears not to include an origin of replication in the equivalent position to the HSV-1 ORIL, although it is highly possible that one may be located elsewhere in the long unique segment of the VZV genome (Davison and Scott, 1986).

1A.9 Reconstruction of the VZV genome from sets of cosmids: a method for introduction of site-specific mutations

The poor growth properties of VZV *in vitro* have greatly hindered application of the genetic techniques that were the key to the identification of functions of many HSV-1 gene products (i.e. production and analysis of *ts* and null mutants) (reviewed in Roizman and Sears, 1993). Successful genetically engineered alteration of a VZV gene (in the viral context) has only been reported at the thymidine kinase gene locus, by recombining a plasmid carrying the altered gene into the VZV genome (Lowe *et al.*, 1987; Shiraki *et al.*, 1991). Drawbacks of this approach include the incorporation of a foreign marker required for recombinant selection, difficulties in obtaining VZV DNA for transfection, and in obtaining cell-free virus to plaque purify VZV mutants from wild type virus.

However, Van Zijl *et al.* (1988) have described a method for generating recombinant pseudorabies virus using a set of overlapping cosmid DNAs, and this has recently been applied successfully to the production of a VZV genome including a site-specific mutation (Cohen and Seidel, 1993). These experiments used four cosmids carrying

overlapping fragments of VZV DNA that together spanned the entire VZV genome and a specific mutation was introduced into one of the cosmids in the thymidylate synthetase ORF, followed by transfection of all the cosmids into tissue culture cells. The VZV genomes containing the specific mutation were produced by recombination events and a plasmid expressing the VZV 140k transactivator was also co-transfected with the overlapping cosmids to improve the infectivity of the viral DNA. This overlapping cosmid approach has also been successfully employed for regenerating the HSV-1 genome (Cunningham and Davison, 1993). However, the implications of this technology are especially important to the analysis of VZV gene functions, which has long been hampered by the difficulties discussed in the previous paragraph. As such the introduction of defined mutations into specific VZV genes using overlapping cosmids holds much promise for the future of VZV research.

1B Gene regulation in the *Alphaherpesvirinae*

The details of VZV gene regulation are only beginning to emerge and will be discussed later (Section 1B.4). However, the high degree of similarity between the VZV and HSV-1 genome arrangements and the predicted sequences of their gene products, strongly suggests that these two alphaherpesviruses are likely to employ similar transcriptional programmes. Therefore it is valuable to discuss the more extensively analysed situation regarding HSV-1 gene expression.

1B.1 Temporal regulation of HSV-1 gene expression - an outline

The transcription of viral DNA is by the unmodified host RNA polymerase II (Costanzo *et al.*, 1977; Ben-Zeev and Becker, 1977) and presumably utilises the host-cell basal transcription machinery (described in Section 1D.1.2). The HSV-1 polypeptides have been divided into three broad, coordinately regulated, sequentially produced groups of polypeptides, on the basis of their kinetics of synthesis and also their sensitivity to metabolic inhibitors during tissue culture infections (Honess and Roizman, 1974). The precise details of the period of synthesis of these groups of polypeptides varies with virus strain and host cell type and therefore makes detailed temporal description difficult.

The transcripts of the first group of HSV-1 polypeptides to be synthesised on infection, termed the **immediate-early** (IE) or alpha polypeptides (α), are expressed in the absence of *de novo* viral protein synthesis. The IE polypeptides are detected as early as 1hr post-infection (pi), although synthesis peaks between 2 and 4hr pi. At least some of the IE proteins continue to accumulate throughout infection (Ackermann *et al.*, 1984; Everett and Orr, 1991). A component of the HSV-1 virion particle is partly responsible for the induction of IE gene expression (Section 1B.2.2). As will be discussed later (Section 1B.3.4), all the HSV-1 IE polypeptides (with one exception) have been found to contribute to the regulation of synthesis of subsequent polypeptide groups.

The second group, the **early** (E) or beta polypeptides (β), are detected very soon after infection and are synthesised at high rates for most of the infectious cycle, but are less abundant at the later stages of infection. The E polypeptides are produced in advance of

viral DNA replication but have an absolute requirement for the presence of functional IE proteins for their synthesis (Honess and Roizman, 1975). The majority of polypeptides involved in viral nucleic acid metabolism are within the early group.

The remaining HSV-1 polypeptides fall into the **late** (L) or gamma (γ) group. These can be detected as early as 2hr pi, but they are produced maximally only after DNA replication has produced a large copy number of replicated templates. Normal production of late gene products requires the presence of the viral IE polypeptides and active viral DNA replication (which is itself a function of the E gene products) for maximum efficiency (Preston, 1979a; Watson and Clements, 1980; Ward and Stevens, 1975; Conley *et al.*, 1981). The late group of polypeptides includes many of the structural components of the HSV-1 virion.

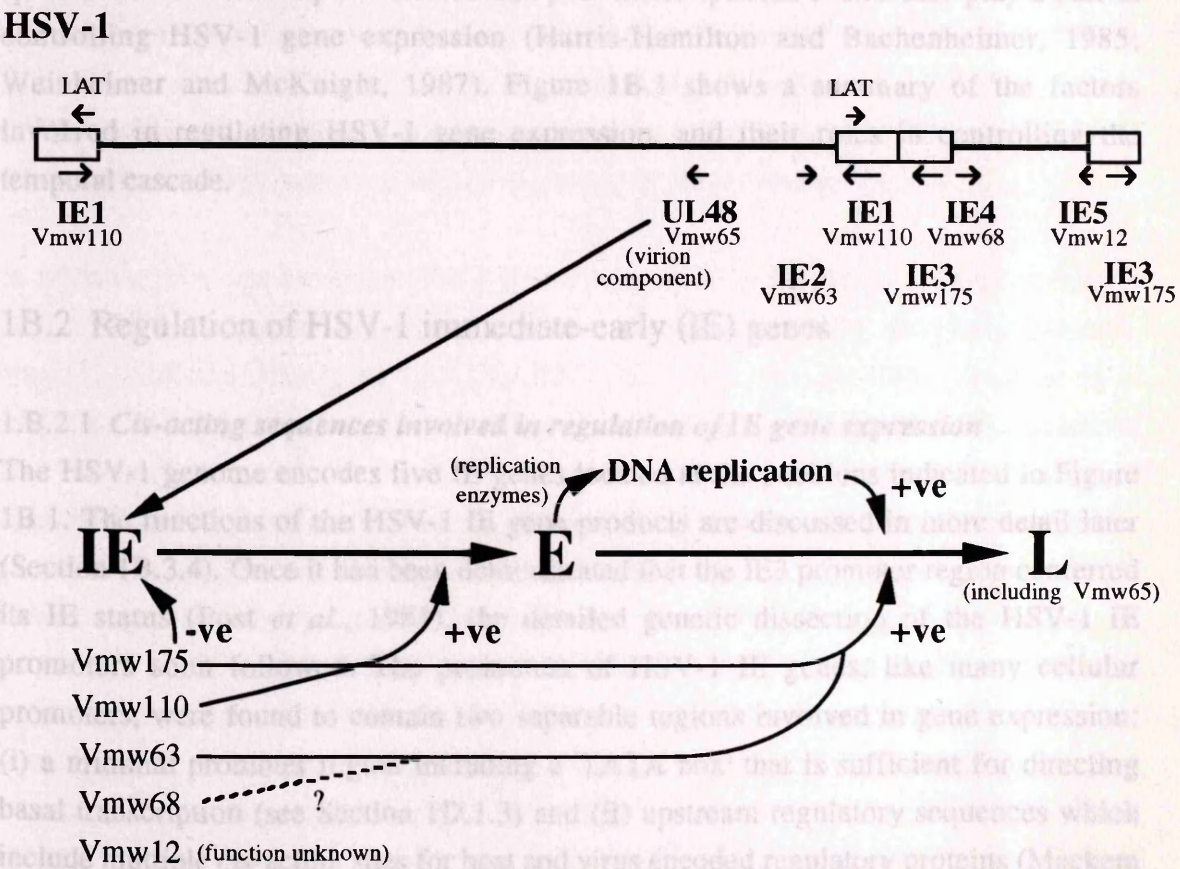


Figure 1B.1 HSV-1 gene products which regulate transcription. The positions of the genes encoding the five HSV-1 IE proteins and the IE *trans*-inducing protein, Vmw65, are shown on a map of the HSV-1 genome, below are given the gene names and corresponding products. The IE4 and IE5 genes are transcribed from a common promoter situated within the short repeat region of the genome. The positions of the LAT transcription units (Section 1A.3), antisense to a portion of the IE1 gene, are also indicated. The role of each of the transcriptional regulatory gene products in the control of HSV-1 gene expression is depicted in the lower part of the Figure, Vmw68 may have a role in activation of late gene expression but the role of Vmw12 is unknown (Sections 1B.2.2 and 1B.3.4). Early gene products, particularly those required for DNA replication, are necessary for late gene expression. The Figure is an adaptation of an original diagram by R. D. Everett.

In general there is a close correlation between the time of appearance of individual mRNA transcripts and the kinetic class of the proteins encoded by them, indicating that the major regulatory processes occur at the transcriptional level (Clements *et al.*, 1977; Jones and Roizman, 1979; Holland *et al.*, 1980; Zhang and Wagner, 1987; Smith and Sandri-Goldin, 1988). However, the situation is not as straightforward as the strict synchronous IE, E and L groupings. A distinct subclass of late genes termed 'leaky' late (or γ_1), are expressed prior to viral DNA synthesis, but require the onset of viral DNA replication for maximal expression; the remaining late genes have a stricter requirement for DNA replication and are termed 'true' late (γ_2) (Powell *et al.*, 1975; Pedersden *et al.*, 1981). In reality, genes expressed later in infection appear to form more of a continuum with respect to their time of expression and dependence on viral DNA replication (reviewed by Wagner, 1985). Detailed analysis of the synthesis and accumulation of specific mRNA transcripts indicated that post-transcriptional events also play a part in controlling HSV-1 gene expression (Harris-Hamilton and Bachenheimer, 1985; Weinheimer and McKnight, 1987). Figure 1B.1 shows a summary of the factors involved in regulating HSV-1 gene expression, and their roles in controlling the temporal cascade.

1B.2 Regulation of HSV-1 immediate-early (IE) genes

1.B.2.1 *Cis-acting sequences involved in regulation of IE gene expression*

The HSV-1 genome encodes five IE genes located at the positions indicated in Figure 1B.1. The functions of the HSV-1 IE gene products are discussed in more detail later (Section 1B.3.4). Once it had been demonstrated that the IE3 promoter region conferred its IE status (Post *et al.*, 1981), the detailed genetic dissection of the HSV-1 IE promoters soon followed. The promoters of HSV-1 IE genes, like many cellular promoters, were found to contain two separable regions involved in gene expression: (i) a minimal promoter region including a 'TATA box' that is sufficient for directing basal transcription (see Section 1D.1.3) and (ii) upstream regulatory sequences which include multiple *cis*-acting sites for host and virus encoded regulatory proteins (Mackem and Roizman, 1982a; b; Cordingley *et al.*, 1983; Preston *et al.*, 1984; Kristie and Roizman, 1984; Bzik and Preston, 1986) (Fig. 1B.3). At least in the case of the IE3 promoter, the far upstream promoter region functions as an enhancer (Lang *et al.*, 1984; Preston and Tannahill, 1984).

Upstream promoter elements for common cellular sequence-specific factors have been identified that are likely to contribute to the high constitutive activity of these IE promoters. The hexanucleotide sequence GGGCGG is present in multiple copies in the

HSV-1 IE promoters and some of these have transpired to be 'GC boxes' that bind and respond to *trans*-activation by the cellular sequence-specific factor Sp1 *in vitro* (Jones and Tjian, 1985). *Cis*-acting binding sites for other ubiquitous cellular transactivators such as the CCAAT-box binding proteins (CBP) are also found in IE distal promoter regions and contribute to the efficiency of gene expression (Spector *et al.*, 1990).

An important element for conferring IE inducibility is present in at least one copy in all the HSV-1 IE promoters and includes the sequence TAATGARAT (R = purine). Functional analyses have found this element to be essential for the response to a component of the HSV-1 virion, later identified as the protein Vmw65 (αTIF or VP16) (Mackem and Roizman, 1982b; Campbell *et al.*, 1984; Preston *et al.*, 1984; Kristie and Roizman, 1984; Gaffney *et al.*, 1985; Bzik and Preston, 1986). Overlapping the majority of TAATGARAT sequences is a consensus octamer motif that has been shown to bind the ubiquitous cellular octamer-binding transcription factor Oct-1 (NFIII, α-H1, OTF-1 or TRF) (Pruijn *et al.*, 1986; Kristie and Roizman, 1987; Gerster and Roeder, 1988; O'Hare and Goding, 1988). The interplay between Vmw65 and Oct-1 and their roles in IE-specific promoter induction will be discussed in the following Section.

In addition, GA rich elements (GA-RE) are present in the upstream regions of at least four HSV-1 IE genes and these elements are necessary for the fully efficient Vmw65-mediated activity of TAATGARAT (Bzik and Preston, 1986; Triezenberg *et al.*, 1988). Characterisation of the cellular GA-RE binding protein (GABP) identified two distinct subunits; one subunit shows sequence similarity to a family of transcriptional regulatory proteins while the other is related to a protein implicated in signal transduction (LaMarco *et al.*, 1991). It appears from DNA binding analyses that GABP cooperates with Oct-1 to enhance formation of a multi-protein complex on selective TAATGARAT sites (Bailey and Thompson, 1992); this complex mediates IE promoter induction (see below). It is interesting to speculate from the sequence homologies to the GABP subunits, that the stimulatory effect of GABP may be coupled to external stimuli or cellular events.

Finally, the IE1 and IE3 promoters also contain consensus binding sites (see Section 1C.1.3) for the HSV-1 immediate-early protein Vmw175 (Kristie and Roizman, 1986b; Muller, 1987). The significance of these binding sites for the activity of the HSV-1 IE promoters will be discussed later (Section 1B.2.3).

1B.2.2 ***Transactivation of HSV-1 IE promoters by the immediate-early complex (IEC)***

A virion component, Vmw65 (more widely known as VP16) has been shown to *trans*-induce IE gene expression selectively (Post *et al.*, 1981; Campbell *et al.*, 1984;

Dalrymple *et al.*, 1985; Pellet *et al.*, 1985). This function is important for the efficiency of lytic viral growth but it is not essential; an HSV-1 mutant defective only in the IE transactivation function of Vmw65 replicates poorly *in vitro*, has an elevated particle:pfu ratios and is avirulent *in vivo* (Ace *et al.*, 1989). Vmw65 is unusual in having two distinct functions, as it is also an essential structural component of the virion tegument (Batterson and Roizman, 1983; Weinheimer *et al.*, 1992).

The mechanism by which Vmw65 induces HSV-1 IE gene expression through the IE responsive element (TAATGARAT) has been studied intensively in recent years, not least because Vmw65 alone is unable to bind stably to the TAATGARAT element (Marsden *et al.*, 1987). As mentioned in the previous Section, a cellular protein (Oct-1) binds to a site overlapping the TAATGARAT motif. Oct-1 itself has little effect on IE promoter activity but acts to recruit Vmw65 to the 'immediate-early complex' (IEC) that forms over the TAATGARAT element (McKnight *et al.*, 1987; Preston *et al.*, 1988; Gerster and Roeder, 1988; O'Hare and Goding, 1988; O'Hare *et al.*, 1988; Stern *et al.*, 1989; apRhys *et al.*, 1989). Oct-1 is a member of the POU protein subclass of homeodomain proteins and as such has a bipartite DNA binding interface (see Section 1E.2.3). Truncation analysis of Oct-1 discerned that the homeodomain portion contacts the TAAT sequence while the POU-specific domain contacts the ATGC sequence 5' to the TAATGARAT site (Verrijzer *et al.*, 1990; 1992) (see Fig. 1B.2).

Importantly, *in vitro* experiments using purified components have found that Vmw65 and Oct-1 are not sufficient to reconstitute the IEC and that an additional cellular component is required (Gerster and Roeder, 1988; Kristie *et al.*, 1989; Xiao and Capone, 1990; Kristie and Sharp, 1990; Katan *et al.*, 1990; Stern and Herr, 1991). The association of this auxiliary protein, known as HCF (C1, CFF or VCAF), with Vmw65 is thought to 'prime' the association of Vmw65 with Oct-1 over the TAATGARAT motif. Since mutations in the GARAT portion of the TAATGARAT sequence also prevent IEC formation (O'Hare *et al.*, 1988), it is likely that the Vmw65-HCF heterocomplex recognises the GARAT sequence, once recruited by Oct-1 (Fig. 1B.2). Wilson *et al.* (1993) have recently shown that HCF comprises multiple subunits derived from a single polypeptide species by processing events. Their isolation of the cDNA clone encoding the HCF polypeptides should enable functional analysis of the complete IEC in reconstituted systems. The regions of the Oct-1 and Vmw65 proteins involved in IEC formation have been mapped by various mutational analyses. The details of these findings have been reviewed recently by O'Hare (1993) and will not be considered further here.

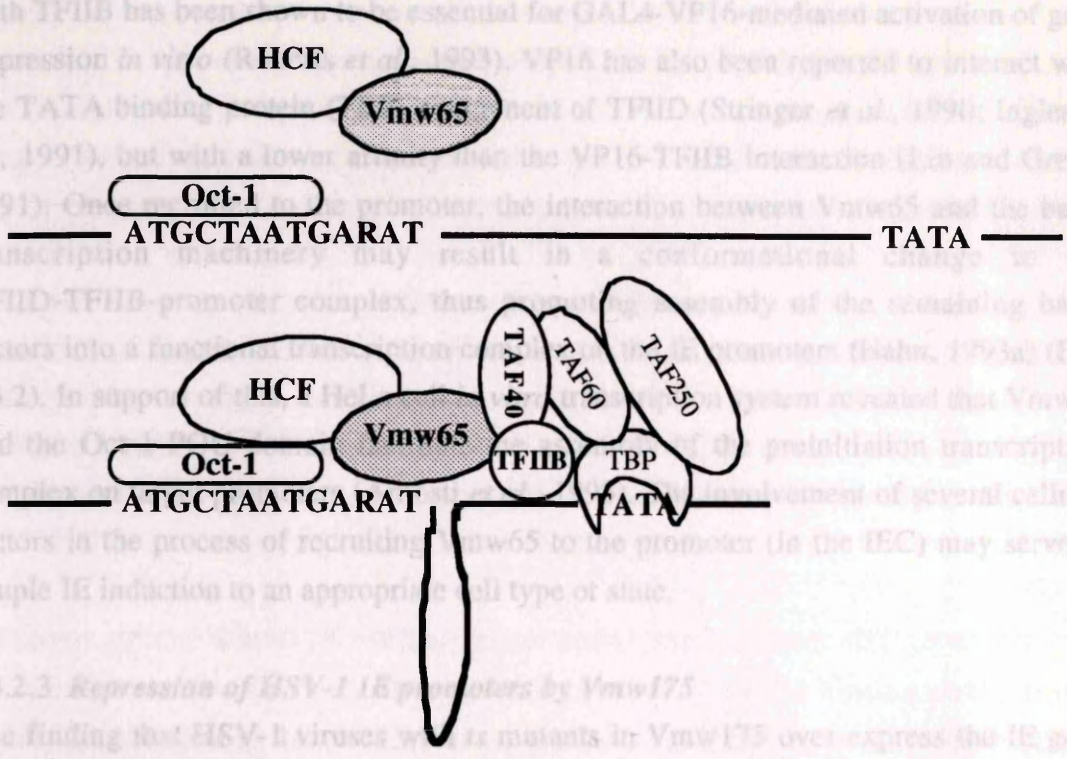


Figure 1B.2 A model for IEC formation and transcriptional activation by Vmw65. This schematic diagram depicts interactions that have been experimentally observed between the various protein components of the IEC and the basal transcription machinery. Following release of the viral DNA from the capsid, the octamer sites of the IE promoters become occupied by Oct-1. Independently, Vmw65 complexes to HCF, and this heterocomplex then recognises Oct-1 in the context of the flanking GARAT signal (specific to octamer motifs of IE genes) to form the immediate-early complex (IEC). The Vmw65 acidic activation region is now positioned to activate transcription by promoting assembly of the transcription complex at the TATA box, through interactions with the basal transcription factors TFIIB and TFIID (comprising TBP and the TAFs, see Section 1D.1.4). The *Drosophila* TAFs are shown but are thought to have human homologues, the remaining TAFs and other basal factors have been omitted for simplicity. There is no experimental data to suggest how proteins bound over the upstream TAATGARAT element and those bound at the proximal regions of the promoter are brought together to allow Vmw65 to interact with the basal factors, although looping or condensation of DNA in nucleoprotein complexes are both possibilities. The Figure is a compilation of figures by O'Hare (1993) and Goodrich *et al.* (1993).

Of particular relevance to the transcriptional regulatory activity of Vmw65 are its two extremely potent transcriptional activation domains, which reside within its highly acidic C-terminus. Each Vmw65 activation domain can function independently when fused to heterologous portions of proteins that bind directly to DNA (Sadowski *et al.*, 1988; Cousens *et al.*, 1989; Goodrich *et al.*, 1993). Investigations using 'GAL4-VP16' chimeras (containing the activation domain(s) of Vmw65 fused to the yeast GAL4 DNA binding domain) have been highly productive in elucidating the details of the Vmw65 transactivation mechanism. *In vitro* protein-protein interaction assays with GAL4-VP16 revealed that both of the Vmw65 activation domains contribute interfaces for the interaction with the basal transcription factor TFIIB, whereas the furthest C-terminal domain additionally contacts a specific TAF component of the basal factor TFIID (Lin *et al.*, 1991; Goodrich *et al.*, 1993) (see Section 1D.1.4). Furthermore, the interaction

with TFIIB has been shown to be essential for GAL4-VP16-mediated activation of gene expression *in vitro* (Roberts *et al.*, 1993). VP16 has also been reported to interact with the TATA binding protein (TBP) component of TFIID (Stringer *et al.*, 1990; Ingles *et al.*, 1991), but with a lower affinity than the VP16-TFIIB interaction (Lin and Green, 1991). Once recruited to the promoter, the interaction between Vmw65 and the basal transcription machinery may result in a conformational change in the TFIID-TFIIB-promoter complex, thus promoting assembly of the remaining basal factors into a functional transcription complex on the IE promoters (Hahn, 1993a) (Fig. 1B.2). In support of this, a HeLa cell *in vitro* transcription system revealed that Vmw65 and the Oct-1 POU domain facilitate the assembly of the preinitiation transcription complex on target promoters (Arnosti *et al.*, 1993). The involvement of several cellular factors in the process of recruiting Vmw65 to the promoter (in the IEC) may serve to couple IE induction to an appropriate cell type or state.

1B.2.3 *Repression of HSV-1 IE promoters by Vmw175*

The finding that HSV-1 viruses with *ts* mutants in Vmw175 over-express the IE gene products at the NPT was interpreted as indicating that Vmw175 is continuously required for the *trans*-repression of all HSV-1 IE genes (Preston, 1979a; Dixon and Schaffer, 1980; Watson and Clements, 1980). Transient transfection assays have been employed extensively to analyse the basis of this apparent repression of IE gene expression, using various combinations of plasmid-borne copies of HSV-1 regulatory proteins and IE promoters (reviewed by Everett, 1987a). Due to the nature of these assays the results obtained varied according to the particular system employed (Everett, 1988b), however the general consensus was that Vmw175 clearly repressed both constitutive and Vmw65-activated IE3 promoter activity, and also reduced activation of the IE1 promoter (DeLuca and Schaffer, 1985; O'Hare and Hayward, 1985b; Gelman and Silverstein, 1986; Resnick *et al.*, 1989). It was soon noted that both of these IE promoters included a 'strong' Vmw175 binding site fitting the Vmw175 consensus sequence that had been derived for several Vmw175 binding sites (Muller, 1987; Kristie and Roizman, 1986b) (see Section 1C.1.3). The Vmw175 binding site within the IE3 promoter spans the mRNA start site and has been implicated in the ability of Vmw175 to repress transcription from its own promoter (Muller 1987; DeLuca and Schaffer, 1988; Roberts *et al.*, 1988), whereas the IE1 promoter site was located further upstream (Kristie and Roizman, 1986b). The IE3 and IE1 sites are 2.6 and 3.5 helical turns away from their TATA boxes respectively, albeit on opposite sides of the TATA box, and the footprinting analyses of Didonato and Muller (1989) found that binding of Vmw175 to either the IE1 or the IE3 promoter binding site resulted in perturbations at the TATA box (less marked at the IE1 TATA box). They suggested that Vmw175 binding to the IE1 and the IE3 site may have deleterious effects on the DNA binding interactions or the

activity of the basal transcription machinery at the TATA box, leading to promoter repression. It appeared that the distance between the Vmw175 binding site and TATA box and also their stereo-specific alignment was important for determining whether Vmw175 binding results in promoter repression or transactivation, because removing the IE3 binding site to a greater distance apparently disrupted normal regulation and abolished the hypersensitivity at the TATA box. The naturally occurring Vmw175 consensus binding site in the HSV-1 gD promoter (a promoter that is activated by Vmw175) is located 8 helical turns upstream of the TATA box and Vmw175 binding to this site does not result in hypersensitivity at the TATA box (DiDonato and Muller, 1989). A further report indicates that the proximity of the Vmw175 binding site to the mRNA start site is likely to be a vital element in determining promoter repression (Koop *et al.*, 1993).

Mutation of the Vmw175 binding sites within the IE3 and IE1 promoters has demonstrated a direct functional correlation between Vmw175 binding and promoter repression in transient transfection assays (Roberts *et al.*, 1988; Resnick *et al.*, 1989). Further support of the importance of Vmw175 DNA binding for promoter repression comes from the observation that all mutations of the Vmw175 protein that result in loss of DNA binding activity also destroyed its repression phenotype (with one exception, Paterson *et al.*, 1990). Interestingly, mutation of the Vmw175 binding site in the IE1 promoter in the context of the viral genome had no apparent effect on Vmw110 polypeptide levels (Everett and Orr, 1991). In contrast, mutation of the Vmw175 binding site at the mRNA start site of the IE3 promoter linked to the *tk* gene in a recombinant viral genome, resulted in a 15 fold increase in *tk* expression (Michael and Roizman, 1993). Two additional non-consensus Vmw175 binding sites (Section 1C.1.3) in the upstream region of the IE3 promoter (Kristie and Roizman, 1986a; Michael and Roizman, 1989) additionally appeared to contribute to the down-regulation of the IE3 promoter in the viral context (Michael and Roizman, 1993). Thus the evidence suggests that during normal HSV-1 infection only the IE3 promoter is repressed effectively by Vmw175. The over-expression of the other IE mRNAs and their corresponding polypeptides observed in the absence of functional Vmw175 (above) might simply result from increased availability of limiting basal transcription factors, which might otherwise be recruited to the E and L promoters. Indeed, during normal infection, only IE3 gene expression appeared to be down-regulated at the level of transcription and significant amounts of the other IE transcripts persisted until late in infection (Weinheimer and McKnight, 1987; Harris-Hamilton and Bachenheimer, 1985). It can be envisaged that high levels of the Vmw175 transcriptional activator protein may have deleterious effects on the later stages of the viral replication cycle and this negative

autoregulatory mechanism presumably provides a homeostatic mechanism to prevent this occurring.

The existence of a temperature-sensitive HSV-1 virus that produces a Vmw175 protein capable of binding to the IE3 consensus binding site, yet fails to down-regulate IE3 gene expression (Paterson *et al.*, 1990) indicates that the autoregulatory function of Vmw175 is likely to require interactions with other proteins in addition to DNA binding. Recent insight into the mechanism by which Vmw175 autoregulates its own expression has been provided by DeLuca and co-workers. They demonstrated that Vmw175, and the basal transcription factors TFIID (or TBP) and TFIIB (see Section 1D.1.2) can simultaneously co-occupy their respective binding sites within the IE3 promoter region, and furthermore the three proteins appeared to bind in a cooperative manner (Smith *et al.*, 1993). By using a reconstituted *in vitro* transcription analysis of the IE3 promoter, backed up by analysis of viral mutants, Gu *et al.* (1993) demonstrated that Vmw175-mediated repression only occurs in the presence of the cellular transcriptional regulatory protein Sp1 and this repression required functional Vmw175 and Sp1 binding sites. Taken together these data indicate that the down-regulation of IE3 gene expression is not a consequence of disruption of the basal transcription machinery, but instead is caused by a reduction in Sp1-mediated induction (which is an important factor in the high constitutive level of IE promoter activity; Bzik and Preston, 1986). Additionally, the TAFs were implicated in repression of the IE3 promoter, as the TATA binding protein (TBP; TFIID but lacking the TAFs, see Section 1D.1.4) was not interchangeable with TFIID for repression *in vitro*. Thus Vmw175 may repress the IE3 promoter by interfering with the interactions of Sp1 or its co-activator/TAF with the basal transcription machinery (Gu *et al.*, 1993).

1B.3 Activation of HSV-1 early (E) and late (L) gene expression

1B.3.1 *Cis-acting sequences involved in regulation of E gene expression*

The finding that transcriptional induction of early promoters requires the presence of functional viral IE gene products (Section 1B.1) stimulated the analysis of *cis*-acting signals mediating virus induced gene expression. The promoter regulatory regions of two members of the HSV-1 E class of genes, namely glycoprotein D (gD) and thymidine kinase (*tk*), have been extensively investigated.

In the earliest analysis of any HSV-1 promoter, McKnight *et al.* (1981) measured the expression from *tk* promoter deletion mutant derivatives microinjected into xenopus oocytes and they later extended their findings using the now classical 'linker-scanning'

(LS) analysis (McKnight and Kingsbury, 1982). These experiments defined the minimal *tk* promoter as approximately 110 nucleotides extending upstream of the mRNA start site and identified the necessity of the TATA box for efficiency and fidelity of transcriptional initiation. A CCAAT motif and two GC boxes were identified as distal regulatory elements influencing *tk* promoter activity in oocytes. The GC boxes and CCAAT boxes were later shown to bind the well characterised cellular transactivator proteins Sp1 and CCAAT binding protein (CBP) respectively (Jones *et al.*, 1985; Graves *et al.*, 1986). Transient transfection analysis of the LS mutants followed by HSV-1 infection of tissue culture cells, failed to identify any additional promoter sequences specifically required for viral induction of *tk* gene expression (Eisenberg *et al.*, 1985; ElKareh *et al.*, 1985) (confirming the earlier results with the gD promoter described below), and this was also found to be the case in the viral context (Coen *et al.*, 1986).

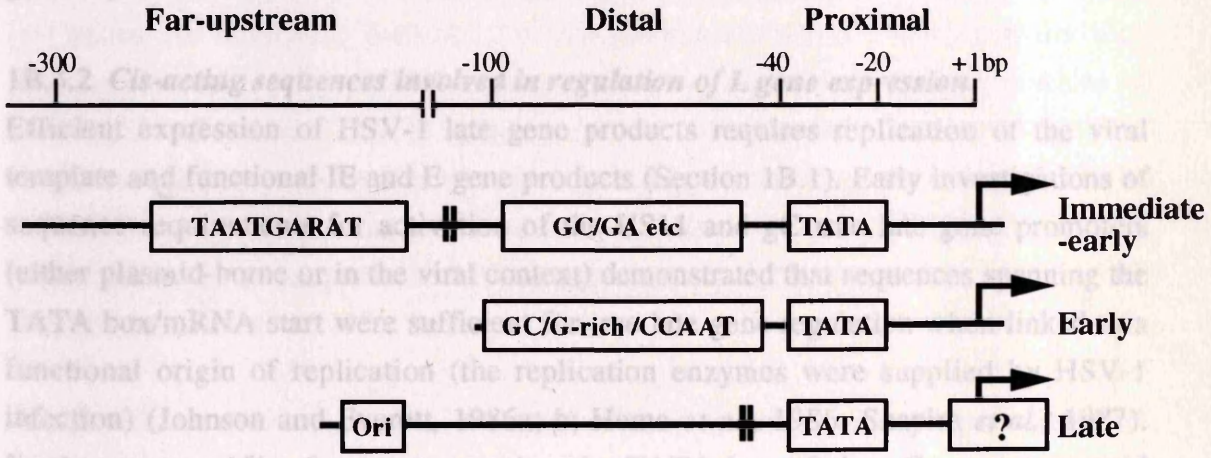


Figure 1B.3 Important sequence elements required for activity of the HSV-1 promoters. Schematic representations of promoters from each of the three temporal classes of HSV-1 genes. The top line indicates a scale in nucleotides starting from the mRNA start sites (indicated by the arrows at +1 nucleotides) as a guide to the locations of the various promoter elements. The majority of promoters in all three classes contain a typical TATA box, but vary in the upstream region between classes. The IE and E promoters include distal elements which bind a variety of host regulatory factors, such as GC boxes (which bind Sp1) and CCAAT boxes (which bind CBP) and G-rich sequences found in certain promoters bind the YY1 factor. The far-upstream regions of the IE promoters additionally include at least one TAATGARAT element, which is required for Vmw65-mediated induction of IE gene expression (Sections 1B.2.1 and 1B.2.2). Late genes require no specific upstream binding sites for regulatory proteins but have a requirement for replication of the viral template, sequences in the vicinity of the mRNA start site / leader appear to contribute to temporal expression of late promoters (Section 1B.3.2).

Using a similar approach, Everett transfected plasmids carrying fragments of the gD promoter linked to an assayable marker gene into tissue culture cells followed by HSV-1 infection, and discerned that sequences located within 83bp of the mRNA start site were sufficient for fully efficient viral-mediated induction of transcription from the gD promoter by the HSV-1 IE products (Everett, 1983). A subsequent deletion and insertion

mutational analysis identified the importance of the gD TATA box and found that induction of the gD promoter both in *trans* (by viral IE proteins) and in *cis* (by the SV40 enhancer) was governed by the TATA box and two distal G-rich elements in each case (Everett, 1984a). The distal-most G-rich element was the more important of the two for gD promoter activity (Everett, 1984a), and this sequence has recently been shown to be the recognition site for the cellular factor YY1 (Seto *et al.*, 1991; Chen *et al.*, 1992; Mills *et al.*, 1994). No virus-specific regulatory elements were identified in the gD promoter, implying that *trans*-induction by viral factors is mediated through the same sequences that are recognised by the uninfected cellular transcriptional apparatus (summarised in Fig. 1B.3). Furthermore the HSV-1 early promoters appear to most closely resemble inducible cellular RNA pol II promoters, as illustrated by the regulation of cellular globin genes in the viral context with early kinetics (Smiley *et al.*, 1987; Smiley and Duncan, 1992). The mechanism used for down-regulation of early gene expression at later times in infection remains to be determined.

1B.3.2 ***Cis-acting sequences involved in regulation of L gene expression***

Efficient expression of HSV-1 late gene products requires replication of the viral template and functional IE and E gene products (Section 1B.1). Early investigations of sequence requirements for activation of the US11 and gC true late gene promoters (either plasmid-borne or in the viral context) demonstrated that sequences spanning the TATA box/mRNA start were sufficient for true late gene regulation when linked to a functional origin of replication (the replication enzymes were supplied by HSV-1 infection) (Johnson and Everett, 1986a; b; Homa *et al.*, 1986; Shapira *et al.*, 1987). Furthermore, a 15bp fragment spanning the TATA box of the gC promoter could mediate gene expression with late kinetics (Homa *et al.*, 1988). It appeared therefore, that although replication is a strong requirement for induction of the true late promoters, no specific upstream recognition elements for virus specific or cellular transactivators are required. However, recognition sites for the cellular factor YY1 have recently been identified in the upstream promoter regions of several leaky late genes (UL19, UL37 and gB genes); deletion of this element in the UL19 promoter reduced gene induction by HSV-1 gene products approximately 9 fold, as measured by transient transfection assays (Seto *et al.*, 1991; Chen *et al.*, 1992; Mills *et al.*, 1994).

What factors then determine expression of these genes with late kinetics? Evidence from several studies suggested that true L gene regulation is defined by possession of a specific type of TATA box (Johnson and Everett, 1986b; Homa *et al.*, 1988; Flanagan *et al.*, 1991a), although conflicting evidence exists, for example substitution of the TATA box from the gC (late) promoter into the *tk* (early) locus had no effect on temporal expression of *tk* (Imbalzano and DeLuca, 1992). Further studies have found that

sequences downstream of the TATA box, around the mRNA start site / leader, contribute to L promoter activity and may even account for the requirement for replication (Blair *et al.*, 1987; Mavromara-Nazos and Roizman, 1989; Kibler *et al.*, 1991; Steffy and Weir, 1991). There are various possible means by which these 'leader' sequences could confer L gene regulation: they may bind a specific *trans*-acting factor necessary for L gene expression, or they may bind a negative factor that is dislodged by replication of the template. In support of the latter, the experiments of Mavromara-Nazos and Roizman (1987) suggested that DNA replication alleviates a negative constraint that is unique to late genes in the viral context. Alternatively, the late leader sequences may form a secondary structure (such as a stem-loop structure) that has a repressive effect on gene expression until it is disrupted by replication. The precise role played by these leader sequences awaits clarification. However, if any leader-binding factors are involved, they must have fairly degenerate DNA binding specificities as no putative *cis*-acting sequence is common to the leader region of HSV-1 late genes. An interesting possibility is that non-specific DNA binding proteins may mediate repression of late HSV-1 gene expression, such as the histone proteins of chromatin (Section 1D.2.4). However, the similarities with the cellular chromatin-mediated mechanism of reducing the basal level of gene expression will be limited, as only a small fraction of the HSV-1 DNA from lytically infected cells was found to be in a nucleosome-like arrangement (Leinbach and Summers, 1980; Muggeridge and Fraser, 1986; Seal *et al.*, 1988).

1B.3.3 ***What is the role played by Vmw175 in the activation of E and L promoters?***

From the analysis of *ts* mutant viruses, it was apparent that Vmw175 stimulates HSV-1 E and L gene expression (Section 1B.3.4) and therefore the discovery that Vmw175 exhibited sequence-specific DNA binding properties (Section 1C.1.3) led to the search for Vmw175 binding sites within its target promoters. A consensus Vmw175 binding site within the distal region of the HSV-1 gD promoter and two further non-consensus Vmw175 sites (one upstream and one downstream of the mRNA start site) have been identified (Faber and Wilcox, 1986b; Tedder *et al.*, 1989). While these sites were not specifically identified in the mutational analysis of the gD promoter (Everett, 1984a; Section 1B.3.1), they contributed to activation by Vmw175 *in vitro* (Tedder *et al.*, 1989). Multimerisation of the consensus site increased Vmw175-mediated transactivation in transient assays (Tedder and Pizer, 1988). However, removal of all three binding sites from the viral genome had little or no apparent effect on gD expression during HSV-1 infection (Smiley *et al.*, 1992). Similarly the *tk* promoter includes several non-consensus Vmw175 binding sites (Kristie and Roizman, 1986a; Imbalzano *et al.*, 1990) but a virus carrying a version of Vmw175 unable to bind specifically to the *tk* promoter binding sites *in vitro* did not show reduced levels of *tk*

gene expression during viral infection (Imbalzano *et al.*, 1990). There exists only one report of a direct involvement of a Vmw175 binding site in the transactivation of a viral promoter, but the implications of this analysis are complicated by the location of this particular site in the leader of a true late gene (Romanelli *et al.*, 1992).

Whether or not the interaction of Vmw175 with its binding sites transpires to be a requirement for promoter activation (discussed in Section 1C.1.7), it is clear that the TATA box is singly the most important promoter element for *trans*-induction of E and L genes by Vmw175. The precise sequence of the TATA box can affect the efficiency of promoter transactivation (Everett, 1988b). Additionally, HSV-1 *trans*-inducing proteins can activate many cellular genes carried on plasmids, or integrated into the cellular or viral genome (Everett and Dunlop, 1984; Everett, 1985; Smiley *et al.*, 1987; Panning and Smiley, 1989), yet they failed to activate the TATA-less rodent *aprt* gene placed in the viral context (Tackney *et al.*, 1984). These data suggest that the cellular factor that binds the TATA box (TFIID) is a critical target of Vmw175-mediated promoter activation. In support of this, Imbalzano and DeLuca (1992) found that Vmw175-mediated induction of chimeric promoters was greater when the promoter included a TATA box with a weak affinity for TBP (the TATA binding protein, a component of TFIID; see Section 1D.1.4) (i.e. the *tk* TATA box), as compared to the induction of a construct containing a TATA box having a stronger affinity for TBP (the gC TATA box). This was interpreted as indicating that Vmw175 promotes or stabilises the interaction of TBP with the promoter. Further evidence in support of the functional involvement between Vmw175 and TFIID came from results indicating that the PRV IE180 protein (the Vmw175 homologue) activates transcription *in vitro* by facilitating the rate of TFIID binding to the TATA box (Abmayr *et al.*, 1988; Section 1C.3.1). Moreover, physical interactions between HSV-1 Vmw175 and the TBP component of TFIID have recently been demonstrated and the data indicated that TBP and Vmw175 bind to DNA in a cooperative manner (Smith *et al.*, 1993).

1B.3.4 *HSV-1 gene products regulating the activity of HSV-1 E and L gene expression*

The HSV-1 IE gene products are required for the induction of E and L gene transcription (Section 1B.1), and this Section outlines the involvement of each of the IE polypeptides in the induction process (reviewed more comprehensively by Everett, 1987a). The findings of transient assays are notoriously variable (Everett, 1988b) and will only be discussed here if particularly relevant. Figure 1B.3 shows the genomic location of each of the five HSV-1 IE genes. The IE1 and IE3 genes are encoded within the genomic repeats and are therefore diploid. Although HSV-1 primary transcripts are rarely spliced, three of the IE polypeptides are translated from spliced mRNAs (Wagner,

1985). Four of the five IE proteins are phosphorylated and localise to the nucleus where they associate with chromatin (Pereira *et al.*, 1977; Hay and Hay, 1980; Fenwick *et al.*, 1980; Ackermann *et al.*, 1984); the exception is Vmw12 which is non-phosphorylated and cytoplasmic (Marsden *et al.*, 1982). In addition, Vmw12 is the only HSV-1 IE protein lacking a VZV homologue (Section 1B.4.2).

Vmw175 (ICP4): Viruses with *ts* lesions or large deletions in the IE3 gene have demonstrated that Vmw175 is an essential polypeptide that is required for the transition from IE to later classes of gene expression (Marsden *et al.*, 1976; Courtney *et al.*, 1976; Watson and Clements, 1978; Preston, 1979a; Dixon and Schaffer, 1980; Preston, 1981; DeLuca *et al.*, 1985), and furthermore is continuously required for expression of the later temporal classes of genes (Preston, 1979b; Watson and Clements, 1980). The properties and functions of Vmw175 are further discussed in Section 1C.1, but suffice to say here that Vmw175 is considered to be the major transcriptional regulator of HSV-1.

Vmw110 (ICP0): Vmw110 is not essential for virus growth in tissue culture, but insertion or deletion inactivation of Vmw110 leads to a cell type, cell cycle and multiplicity-dependent defect in viral gene expression (Stow and Stow, 1986; Sacks and Schaffer, 1987; Everett, 1989; Cai and Schaffer, 1991). HSV-1 infections *in vivo* may be at low multiplicity, therefore these defects of the Vmw110 deficient viruses observed at low moi suggest that Vmw110 may be required early in infection to up-regulate gene expression to a level that allows establishment of lytic infection. Additionally, viruses lacking Vmw110 failed to stimulate reactivation of latent viral genomes in an *in vitro* latency system (Russell *et al.*, 1987) and addition of Vmw110 is sufficient to reactivate latent HSV-2 from the same system (Harris *et al.*, 1989). The possible involvement of Vmw110 in establishment of lytic infection and reactivation of latent virus is an area of much investigation in the HSV field.

Transient transfection experiments have found Vmw110 to be a promiscuous transactivator of all classes of viral genes both by itself and in synergy with Vmw175 (Everett, 1984b; O'Hare and Hayward, 1985a; b; Quinlan and Knipe, 1985; Gelman and Silverstein, 1985; 1986) (see Section 1C.1.6). The mechanism of Vmw110-mediated activation of gene expression is not well understood, but is likely to be mediated by protein-protein interactions as no sequence-specific Vmw110 DNA binding activity has been detected (Everett *et al.*, 1991b).

The functionally important regions of Vmw110 have been mapped using various mutagenic approaches (reviewed by Everett *et al.*, 1991c). An essential cysteine-rich sequence near the N-terminus forms a C_3HC_4 zinc-finger, as demonstrated by

expression, purification and NMR spectroscopy of the equivalent domain of the EHV-1 gene 63 protein (Everett *et al.*, 1993a; Barlow *et al.*, 1994). Unlike the previously characterised types of zinc-binding domains (Section 1E.4), the C_3HC_4 motif does not appear to be involved in sequence-specific DNA binding but is more likely to mediate interactions with other polypeptides. This specific type of zinc-finger is found in a large family of proteins involved in a wide range of nuclear processes (Freemont *et al.*, 1991; Freemont, 1993). As yet, the precise function of none of these related domains has been determined. However, the zinc-binding domain of Vmw110 is essential for its roles in gene activation in transfection assays, in the onset of virus infection, and in reactivation of latent virus *in vitro* (Everett, 1988a; 1989; Harris *et al.*, 1989).

An additional region important for the transactivation function of Vmw110, as determined by transient assays, has been mapped to the C-terminus of the protein. This region forms a strong and specific interaction with a cellular 135kD protein found in all the cell types analysed (Meredith *et al.*, 1994); further characterisation of this protein is likely to shed light on the role played by Vmw110 during infection. Vmw110 localises to, and subsequently causes redistribution of discrete nuclear domains during infection (Maul *et al.*, 1993), and the pattern of Vmw110 localisation is altered by deletion of either the zinc-finger or the C-terminal domain (Maul and Everett, 1994). It can be speculated that the transactivation functions of Vmw110 may occur as a result of modification of nuclear processes, probably mediated by protein-protein interactions formed by both its zinc-finger and C-terminal domain.

Vmw63 (ICP27): Studies with viruses containing *ts* lesions or a deletion in Vmw63 have demonstrated the essential role played by Vmw63 in HSV-1 growth. These viruses overproduce certain IE and E gene products, yet fail to induce late gene expression (Sacks *et al.*, 1985; McCarthy *et al.*, 1989; Rice and Knipe, 1990). Evidence from transfection studies supported these findings, demonstrating that Vmw63, in the presence of Vmw175 and Vmw110, can repress some E promoters and enhance activation of some L promoters (Everett, 1986; Rice and Knipe, 1988; Sekulovich *et al.*, 1988; Su and Knipe, 1989). The multi-functional nature of Vmw63 is borne out by the identification of both repressor and activator domains near the C-terminal end of the protein (Hardwicke *et al.*, 1989; Rice *et al.*, 1989; McMahan and Schaffer, 1990).

The mechanisms by which Vmw63 performs its regulatory roles are poorly understood. Comparisons of the accumulation of specific transcripts and their protein products during infection with a Vmw63 *ts* virus, indicated that Vmw63 probably functions post-transcriptionally (Smith *et al.*, 1992b). McLauchlan *et al.* (1992) reported that functional Vmw63 is required for the 'late polyadenylation factor' activity, which

selectively increases the usage of a virus late gene poly(A) site in preference to that of an IE site during *in vitro* polyadenylation and also in the viral context (McLauchlan *et al.*, 1989). Although this effect was not shown to discriminate clearly between poly(A) sites of different temporal classes, it is possible that Vmw63 may contribute to the temporal regulation of HSV-1 gene expression by influencing the processing of pre-mRNA transcripts selectively. In addition, Sandri-Goldin and Mendoza (1992) demonstrated that the level of host-cell pre-mRNA splicing was reduced following transfection of Vmw63; this may be a consequence of Vmw63-mediated redistribution of the small nuclear ribonucleoprotein (snRNP) complexes that are involved in splicing events (Phelan *et al.*, 1993). Of the spliced HSV-1 transcripts, all but two are expressed from IE genes, whereas most cellular pre-mRNAs must be spliced. Therefore this redistribution of snRNP complexes at later times in infection may serve to inactivate snRNPs and reduce the relative efficiency of host-cell gene expression in favour of viral transcription.

Vmw68 (ICP22): Viruses with a deletion of the IE4 gene were viable in tissue culture (Post and Roizman, 1981) indicating that Vmw68 is non-essential. However one such deletion virus grew poorly on certain cell types, under-expressed at least one L gene and was avirulent in mice (Sears *et al.*, 1985).

Vmw12 (ICP47): The non-essential nature of the Vmw12 polypeptide is inferred from the ability of deletion mutants lacking the IE5 gene to grow efficiently in tissue culture (Longnecker and Roizman, 1986; Mavromara-Nazos *et al.*, 1986a). No function has been assigned to Vmw12 but its cytoplasmic location alone indicates that it is unlikely to be involved in transcriptional regulation.

Other viral products involved in the regulation of early and late genes: The major DNA binding protein (UL29) has been implicated in a general negative regulation of gene transcription and may play a role in maintaining the highly ordered cascade of HSV-1 gene expression, by an as yet unknown mechanism (Godowski and Knipe, 1983; 1985; 1986).

1B.4 Regulation of VZV gene expression

1B.4.1 *VZV protein synthesis in a temporal manner*

Analysis of the kinetics of VZV protein synthesis by the classic metabolic-inhibitor experiments used to define HSV-1 genes (Section 1B.1) has been complicated by the difficulty in achieving efficient, or synchronous VZV infections of tissue culture cells

(Section 1A.4). However, cycloheximide reversal experiments have identified at least four VZV immediate-early gene products (Lopetegui *et al.*, 1985; Shiraki and Hyman, 1987); phosphoproteins with sizes around 185kD and 43kD, and also a protein around 34kD were common to both studies. These experiments detect IE proteins produced in tissue culture infections by initially inhibiting *de novo* protein synthesis by cycloheximide, which allows accumulation of IE mRNA, followed by translation of the IE mRNAs after washing out the cycloheximide. Further transcription is inhibited by the addition of actinomycin D. The analysis of the kinetics of VZV protein synthesis has proved difficult, but relatively successful infections were obtained when monolayers were overlayed with VZV infected cells (showing 100% cpe), at a ratio of 4 uninfected cells per infected cell. Radiolabelling enabled the protein species which are synthesised at various times post infection (pi) to be analysed (Ruyechan *et al.*, 1991). The first viral polypeptides detected by this approach appeared at 10hr pi and had disappeared by 45hr pi. At 21hr pi a second set of polypeptides appeared, and these had largely disappeared by 52hr pi. A final set of polypeptides appeared between 21 and 28hr pi and were synthesised until late in infection. These data indicate that VZV does encode polypeptides whose expression can be divided into three broad temporal classes in a manner analogous to the HSV-1 cascade. In addition, the polypeptide synthesis pattern of another alphaherpesvirus, EHV-1, follows a coordinately regulated and sequentially ordered immediate-early / early / late cascade, as determined by use of metabolic inhibitors in tissue culture infections, although it should be noted that in contrast to the situation with HSV-1 or VZV, EHV-1 expresses a single IE transcript (Caughman *et al.*, 1985; Gray *et al.*, 1987a; b; for a review, see Harty *et al.*, 1991).

1B.4.2 *The VZV IE gene products*

The VZV genome is held to encode four immediate-early peptides. The 185kD protein identified in the cycloheximide several experiments (above) is likely to correspond to the gene 62 product (Felser *et al.*, 1988), and the 45kD protein probably corresponds to ORF63 (D. Stevenson, personal communication). However the sizes of the gene 61 product (heterologous phosphoprotein of 62-65kD) (Stevenson *et al.*, 1992) and the gene 4 product (51.5kD) do not correlate with any of the remaining VZV proteins produced during cycloheximide reversal. The similar genome locations and homology between the predicted sequences of the HSV-1 and VZV IE proteins strongly suggested that these proteins would be functionally equivalent (Davison and Scott, 1986). That VZV encoded transcriptional activators was demonstrated experimentally when VZV infection of transfected cells induced gene expression from plasmid-borne HSV-1 gD or rabbit β–globin promoters (Everett and Dunlop, 1984). Further transfection experiments demonstrated that a DNA fragment spanning VZV ORF 62 was sufficient for transactivation of HSV-1, VZV and cellular promoters (Everett, 1984b; 1985; Inchauspe

et al., 1989). Subsequently the involvement of all four putative VZV IE products in transcriptional regulation has been tested, mainly in transient transfection assays due to the difficulties that would be associated with propagation of site-specific mutated VZV in tissue culture; these transient assays have characteristically yielded some inconsistencies between different systems. In the following analyses, the *tk* (early) and gpI and gpII (late) (Vafai *et al.*, 1984) promoters were invariably used as representatives for the putative temporal classes of VZV genes. Although each VZV IE protein has an HSV-1 sequence homologue, interesting differences in the details of the functions of the VZV proteins, as compared to their HSV-1 counterparts (Section 1B.3.4), are becoming apparent.

140k (ORF62): The product of VZV gene 62 has a predicted molecular weight of 140,000D, and is therefore referred to as 140k (Davison and Scott, 1985). VZV 140k is the sequence homologue and functional counterpart of the product of HSV-1 gene IE3, Vmw175 (Section 1C.2.1). Transient transfection assays have been used to demonstrate that VZV 140k is a potent transactivator of all classes of VZV promoters and also heterologous promoters at the transcriptional level (Everett, 1984b; Inchauspe *et al.*, 1989; Cabirac *et al.*, 1990; Inchauspe and Ostrove, 1989; Perera *et al.*, 1992a; b; Section 1C.2.3). Like its HSV-1 counterpart protein (Section 1B.2.3), VZV 140k negatively autoregulates its own expression in transient transfection assays, although this appears to occur in a cell type dependent manner (Disney *et al.*, 1990; Perera *et al.*, 1992b). The high abundance of VZV 140k within the virion tegument (Kinchington *et al.*, 1992) may have implications for VZV gene regulation ; following release from the tegument, 140k may activate the other IE promoters in advance of *de novo* viral protein synthesis, playing a role analogous to HSV-1 Vmw65 (Section 1B.2.2). The functions and properties of VZV 140k are discussed in detail in Section 1C.2.

ORF61: The VZV gene 61 product shows limited sequence homology to HSV-1 Vmw110 (encoded by the HSV-1 IE1 gene) but in contrast to IE1, gene 61 is only represented once in the viral genome. The ORF61 protein has recently been characterised as a nuclear, phosphorylated heterogeneous species of 62-65 kD (Stevenson *et al.*, 1992). Even though sequence similarities between HSV-1 Vmw110 and VZV ORF61 are confined to the N-terminus, to sequences spanning the Vmw110 zinc-finger (Section 1B.3.4), a stable cell line expressing ORF61 is able to complement a Vmw110 deletion mutant HSV-1 virus (Moriuchi *et al.*, 1992). Transient assays have demonstrated VZV ORF61 to be a potent *trans*-repressor or activator of ORF4 and ORF62 mediated activation, depending on the cell type used (Nagpal and Ostrove, 1991; Perera *et al.*, 1992a). In some experiments, the ORF61 *trans*-repressor effect was specific to VZV promoters and occurred at the transcriptional level (Nagpal and

Ostrove, 1991), while in another report ORF61 was found to be a promiscuous transactivator in the absence of other viral gene products (Moriuchi *et al.*, 1993b). These differences are likely to be a consequence of the variations between the details of the experimental systems used. Interestingly, ORF61 expressing cell lines complemented an HSV-1 virus (*in* 1814) that lacks the IE *trans*-inducing function of Vmw65 (Section 1B.2.2); VZV DNA transfected into these cell lines had enhanced infectivity (approximately 10 fold) (Moriuchi *et al.*, 1993b). The evidence so far suggests that VZV ORF61 may be important for up-regulating VZV gene expression at an early stage in the viral life cycle.

ORF4: The product of VZV gene 4 is the sequence homologue of HSV-1 Vmw63 (the HSV-1 IE2 gene product). However, VZV gene 4 only partially complements an HSV-1 virus with a *ts* lesion in Vmw63 when supplied by co-infection with VZV or from a transformed cell line (Felser *et al.*, 1987; Moriuchi *et al.*, 1994). Furthermore, the cell line expressing VZV ORF4 fails to complement an HSV-1 virus carrying a Vmw63 deletion mutation (Moriuchi *et al.*, 1994). VZV ORF4 was found by transient transfection analysis to activate expression of a variety of VZV and cellular genes both alone and in synergy with VZV 140k. This transactivation by ORF4 shows a degree of selectivity, as only IE gene 62 and a selection of putative early VZV genes were activated, while the inducibility of VZV late genes was not consistent between analyses (Inchauspe *et al.*, 1989; Inchauspe and Ostrove, 1989; Perera *et al.*, 1992a; Defechereux *et al.*, 1993; Moriuchi *et al.*, 1994). The regulatory functions of ORF4 appeared to influence both transcriptional and post-transcriptional events, as the increase in reporter mRNA levels following transfection of a plasmid expressing gene 4 could not alone account for the high levels of reporter enzyme activity (Defechereux *et al.*, 1993). Therefore, in common both VZV ORF4 and its HSV-1 counterpart appear to influence pre-mRNA processing, but the two proteins show numerous other functional differences and this presumably accounts for the weak complementation seen in the experiments described above.

ORF63: The product of VZV gene 63 shows partial sequence homology to HSV-1 Vmw68 (encoded by HSV-1 gene IE4), although VZV gene 63 is diploid in contrast to HSV-1 IE4 (Table 1A). As is the case with HSV-1 Vmw68, the functions of the VZV protein are poorly understood. To date, there exists a single published transient transfection study of ORF63 function; this reports that VZV ORF63 inhibits expression of VZV gene 62, activates early but has no effect on putative late VZV promoters (Jackers *et al.*, 1992). A study of the functional interplay between ORF63 and the other VZV IE proteins is required before speculation as to its likely role in infection.

1B.4.3 *Transactivation of the VZV IE promoters*

Because transcription of VZV gene 62 can be induced following HSV-1 infection of cell lines including an integrated copy of VZV gene 62, it was suggested that the HSV-1 Vmw65 virion *trans*-inducing factor might also activate the VZV IE promoters in these experiments (Felser *et al.*, 1988). Subsequently, a plasmid-borne copy of HSV-1 Vmw65 was found to strongly transactivate the gene 62 promoter in transient assays, and a deletion analysis of the gene 62 promoter defined the Vmw65-responsive element to lie within residues -410 to -131 of the mRNA start site. Interestingly, the sequence AAATGAAAT (centred at -254bp) within this region is strikingly similar to the TAATGARAT site that mediates viral induction of the HSV-1 IE promoters (McKee *et al.*, 1990) (Fig. 1B.4; Section 1B.2.1).

*Sal*I
CACACCGTCGACGTGCATTTGATTAACTAGATGCCGGATGGGTGGAAACAACCCGTGTTA -356

octamer (TATA)
TATAAGATG**TTTTGCAT**GTGAGACAACCCCAATTGTGTTTATGTATATTATAT**ATCGTC**T -296

(mRNA) *Dra*I **octamer/TG**
GTAGACACACGATGATTGGTTGTTATTTAAACAT**ATGTAAATGAAAT**TCACATGTCTGGT -236

*Rsa*I
ATCCCTTGTTATGATGTTGTAAGGTATGCGGAAATAGACACCGGGCGTACATCGCCAACC -176

*Xho*I
AGCGGTCTCTCCTTAAACGCATACTATGGTCCATGAACTTCCCGCCTCGAGTCTCGTCCA -116

ATCACTACATCGTCTTATCATTAAGAATATTTACACGGTGACGACACGGGGAGGAAATAT -56

*Taq*I **TATA** *Rsa*I **mRNA**
GCGGTCGAGGGGGGGGCACAACACG**TTTTAA**GTACTGTTGGAACTCCCTCACCAA**C**CGCA +5

*Ava*I
ATCGCAATCCTTTGAAGGCTGCGAGAGCGTTTGGAAAACTCGGGTACGTCTAAATTCACC +65

CCACTGCG**ATG**GATACGCCGCCGATGCAGCGCTCTACACCCCAACGC +112

Figure 1B.4 The VZV gene 62 promoter. Nucleotides are numbered from the mRNA start site at +1 that was identified by transcript analysis (McKee *et al.*, 1990) and a TATA box homology is found at positions -25 to -30. Octamer and TAATGARAT-like motifs (TG) are found within a 13bp sequence centred over -255 (another octamer motif in inverse orientation is present between nucleotides -346 to -339). The first ATG codon of the gene 62 ORF is shown in bold. The mRNA start site previously proposed on the basis of sequence analysis alone is indicated at -287 nucleotides and the corresponding TATA box homology at nucleotides -301 to -312 is also indicated (Felser *et al.*, 1988). Putative *cis*-acting sequences for common cellular transcriptional regulator proteins (i.e. Sp1 and CBP) have been predicted in the upstream regulatory region for each proposed mRNA start site, but are not shown. Also shown in bold, are the hexamer homologies to the ATCGTC 5'motif of the HSV-1 Vmw175 consensus binding site (see Section 1C.2.3). Restriction sites used in preparation of radiolabelled DNA fragments derived from the gene 62 promoter (Fig. 3A.6) are also given.

Again by analogy to the HSV-1 situation, sequences similar to the octamer element have been found in the gene 62 promoter (Davison and Scott, 1986), and one of particular interest overlaps the gene 62 AAATGAAAT sequence (Fig. 1B.4). DNase I footprinting analysis has shown that in the presence of a cellular extract, HSV-1 proteins were able to specifically bind over the AAATGAAAT element (McKee *et al.*, 1990). It was possible therefore, that VZV utilises a mechanism similar to that seen in the HSV-1 system (Section 1B.2.2), for induction of its IE promoters.

The VZV genome does indeed encode a sequence homologue of HSV-1 Vmw65, VZV ORF10 and like its HSV-1 counterpart it is also incorporated into the virion tegument (Dalrymple *et al.*, 1985; Kinchington *et al.*, 1992). However, VZV ORF 10 is 80 amino acids shorter than Vmw65, lacking sequences similar to the Vmw65 C-terminal acidic transcriptional activation region (Section 1B.2.2). The absence of the C-terminal sequences from the VZV protein may account for the failure of ORF10 to direct formation of the IEC complex (Section 1B.2.2) on the gene 62 AAATGAAAT element, and also the failure of VZV ORF10 to transactivate the VZV gene 62 promoter in transient assays (McKee *et al.*, 1990). In contradiction to these earlier findings, Moriuchi *et al.* (1993a) have observed VZV ORF10-mediated activation of the VZV gene 62 promoter and also the promoters of HSV-1 genes IE1 and IE3 in transient assays, albeit at a lower level than Vmw65-mediated transactivation. Additionally an HSV-1 mutant lacking functional Vmw65 (virus *in*1814) was complemented on a stable cell line expressing ORF10. This cell line enhanced the infectivity of HSV-1 virions (Moriuchi *et al.*, 1993a), a situation very similar to that seen with cell lines expressing VZV ORF61 in the previous Section.

No obvious TAATGARAT sequence homologies have been found within the other VZV IE promoters, and it has been postulated that the VZV transcriptional programme might be more similar to that of EHV-1, with its single IE protein, than to the HSV-1 programme (Stevenson *et al.*, 1992). It is possible that VZV ORF10 functions to up-regulate expression of the gene 62 promoter during infection in a similar manner to HSV-1 Vmw65, mediated by an activation domain not conserved with its HSV-1 counterpart. The gene 62 product, 140k (either synthesised *de novo* or released from the virion tegument) could then transactivate the remaining IE promoters and initiate the temporal cascade of VZV gene expression.

Similarly EHV-1 gene 12 encodes a sequence counterpart of HSV-1 Vmw65. The upstream region of the EHV IE promoter includes a TAATGARAT element and is transactivated by the EHV-1 ORF12 protein (Purewal *et al.*, 1992). Even though the EHV-1 ORF12 protein includes some of the C-terminal sequences not present in VZV

ORF10, they lack the acidic character of the C-terminal transcriptional activation region of HSV-1 Vmw65. PRV however, apparently does not encode a Vmw65 homologue (Hampl *et al.*, 1984) although the promoter of the sole PRV IE gene (Cheung, 1989; Vlcek *et al.*, 1989) contains a TAATGARAT element and can be *trans*-induced by HSV-1 Vmw65 (Campbell and Preston, 1987). This raises the question of the function of the TAATGARAT element in the PRV promoter. It has been suggested that the IE promoter *cis*-acting sequences themselves (such as the TAATGARAT element) are sufficient to confer the IE kinetics of gene expression in the absence of Vmw65 (Ace *et al.*, 1989), since HSV-1 IE proteins were still highly expressed following infection with virus *in*1814 during cycloheximide reversal experiments (virus *in*1814 lacks functional Vmw65). In the case of PRV, it may be that binding of host factors to the TAATGARAT element is sufficient to induce IE promoter activity in the cell types infected during natural infection, although there is no evidence to indicate the existence of a cellular counterpart of HSV-1 Vmw65. It is interesting that the TAATGARAT element is conserved in the promoters of the major IE genes of each of these alphaherpesviruses, yet the sequence and even the possession of the viral *trans*-inducing factor is variable. This indicates that the mechanisms used for induction of IE gene expression are likely to differ between the alphaherpesviruses and may contribute to the differences in their biologies.

1C The major transcriptional regulatory proteins of the *Alphaherpesvirinae*

All the alphaherpesviruses analysed so far have been found to encode a sequence homologue of the VZV 140k transcriptional regulatory protein. Of these, the HSV-1 counterpart protein, Vmw175, has been most extensively analysed and therefore the majority of this Section will discuss the properties of HSV-1 Vmw175. By analogy, VZV 140k and the other alphaherpesvirus homologue proteins are likely to possess similar properties to HSV-1 Vmw175. As will be discussed in the later parts of the Section, this is transpiring to be the case.

1C.1 HSV-1 Vmw175

The sequence-specific DNA binding protein Vmw175 is essential for the expression of early and late HSV-1 genes (Section 1B.3.4) and for the down-regulation of its own promoter (Section 1B.2.3), and is therefore held to be the major transcriptional regulator of HSV-1. Vmw175 is encoded by the diploid IE3 gene, one copy of which is present in each of the short repeat regions of the HSV-1 genome (Fig. 1B.1) (Rixon *et al.*, 1982). HSV-1 Vmw175 was first identified as a viral polypeptide by Honess and Roizman (1973) and was subsequently shown to be a large phosphorylated, nuclear-localised protein, that exists as a homodimer in solution (Pereira *et al.*, 1977; Metzler and Wilcox, 1985). Migration on SDS-PAGE shows Vmw175 to be a heterogeneous species (see below) with apparent molecular weights of around 175kD on SDS-PAGE (Pereira *et al.*, 1977), although the predicted molecular weight of the 1298 amino acid product of the IE3 ORF is 133kD (McGeoch *et al.*, 1986). Vmw175 is present in approximately 100 copies in the tegument of each virus particle (Yao and Courtney, 1989), an amount which is significantly lower than the corresponding VZV 140k protein (Kinchington *et al.*, 1992). However, a recent report states that Vmw175 is predominantly located in non-infectious particles that lack a nucleocapsid, rather than in the infectious virions (McLauchlan and Rixon, 1992; Szilagyi and Cunningham, 1991).

1C.1.1 *Post-translational modifications*

Vmw175 is present in three phosphorylated forms in infected cells, having slightly differing mobilities on SDS-PAGE gels (Pereira *et al.*, 1977). The phosphates are

located on serine and threonine residues and can cycle on and off during infection (Wilcox *et al.*, 1980; Faber and Wilcox, 1986a). A serine-rich tract located between Vmw175 residues 176 and 206 is thought to be the major site of phosphorylation; this serine-rich tract is also found in the other members of this alphaherpesvirus family of proteins. Other parts of the protein are additionally phosphorylated during infection (Faber and Wilcox, 1986a; DeLuca and Schaffer, 1988), and this is likely to account for the viability of an HSV-1 virus lacking the serine-rich sequences (Paterson and Everett, 1990). The extent of phosphorylation of the Vmw175 protein appears to differentially affect its DNA binding activity, as discussed in Section 1C.1.3.

As Vmw175 includes several consensus GTP-binding elements and can be adenylated and guanylated in isolated nuclei, it has been suggested that Vmw175 may belong to a class of GTP-binding proteins which function in transcriptional transactivation (Blaho and Roizman, 1991). Additionally, Vmw175 can be poly(ADP-ribosyl)ated in isolated nuclei, and an antibody specific for poly(ADP-ribose) was used to indicate that all the forms of Vmw175 isolated from infected tissue culture cells are poly(ADP-ribosyl)ated (Preston and Notarianni, 1983; Blaho *et al.*, 1992). The functional relevance of these modifications of Vmw175, if any, are not clear.

1C.1.2 *Nuclear localisation*

HSV-1 Vmw175 is transported to the nucleus after synthesis where it becomes associated with chromatin (Courtney and Benyesh-Melnick, 1974; Cabral *et al.*, 1980; Hay and Hay, 1980). The nuclear localisation of several other viral proteins, notably Vmw110, is impaired during infection with certain *ts* viruses which encode Vmw175 proteins that remain in the cytoplasm, indicating that Vmw175 may facilitate transport of these viral proteins to the nucleus (Knipe and Smith, 1986). The Vmw175 nuclear localisation signal has been mapped to between residues 682 and 774; this region includes a short basic sequence (PREGRKRKSP) that is related to the SV40 large T antigen and adenovirus E1a nuclear localisation signals (DeLuca and Schaffer, 1987; Paterson and Everett, 1988a; Kalderon *et al.*, 1984; Krippl *et al.*, 1985). The nuclear distribution of Vmw175 alters during the course of infection. At early times Vmw175 is diffusely located throughout the nucleus, but following the onset of viral replication, Vmw175 redistributes to discrete 'replication compartments' which also contain the major DNA binding protein, the viral DNA polymerase and viral DNA (Quinlan *et al.*, 1984; Randall and Dinwoodie, 1986; Knipe *et al.*, 1987); these replication compartments are likely to be the sites of transcription of viral genes expressed later in infection.

1C.1.3 *DNA binding*

HSV-1 Vmw175 in crude cell extracts was found to bind to non-specific DNA-cellulose columns (Powell and Purifoy, 1976; Hay and Hay, 1980), although an early purification of Vmw175 resulted in loss of DNA binding activity and the suggestion that Vmw175 bound DNA via a host polypeptide (Freeman and Powell, 1982). Affinity purified Vmw175 has subsequently been found to bind DNA directly, in a sequence-specific manner (Kattar-Cooley and Wilcox, 1989), and the intrinsic DNA binding ability of Vmw175 was also indicated by south-western blotting experiments (Michael *et al.*, 1988). An equilibrium dissociation constant of the Vmw175-DNA complex was reported to be approximately 1.1nM (at 4°C in the presence of 100mM NaCl using affinity purified Vmw175 on a consensus binding site) (Kattar-Cooley and Wilcox, 1989); this value is in the same range as those of the SV40 T antigen and the *E. coli* trytophan (*trp*) repressor (Muller *et al.*, 1987; Klig *et al.*, 1987).

Vmw175 is the only HSV-1 immediate-early protein that possesses sequence-specific DNA binding activity and various techniques including filter binding, gel retardation, Exonuclease III digestion and nuclease and chemical footprinting analyses have been employed to investigate its DNA targets. Specific sequences in the HSV-1 IE1, IE3 and gD promoters, and two sites within the *tet* gene of pBR322 were among the first Vmw175 binding sites to be identified (Beard *et al.*, 1986; Faber and Wilcox, 1986b; 88; Kristie and Roizman, 1986a; b; Muller, 1987). These binding sites displayed a degree of sequence similarity and Faber and Wilcox (1986b) put forward the consensus sequence ATCGTCnnnnYCGRC (n = any, Y = pyrimidine, R = purine) based on comparisons of three of these sites. Although highly conserved, the 5' ATCGTC motif is not sufficient for DNA recognition by Vmw175 (Roberts *et al.*, 1988); methylation interference footprinting analyses and mutations within the consensus sites have demonstrated the importance of both the 5' and 3' motifs of the sequence for DNA recognition by Vmw175 (Muller, 1987; Roberts *et al.*, 1988; Michael and Roizman, 1989; Pizer *et al.*, 1991; Everett *et al.*, 1991a). The consensus ATCGTnnnnnYSG was derived from the findings of an extensive mutational analysis that quantified the effect of single substitutions throughout the region spanning the HSV-1 IE3 consensus binding site (Everett *et al.*, 1991a); this experimentally defined consensus sequence closely corresponds to the original consensus of Faber and Wilcox (1986b). Importantly, this mutagenesis found no single base to be essential for DNA recognition by Vmw175 and the DNA binding specificity of Vmw175 appeared to be fairly relaxed (Everett *et al.*, 1991a).

However, the majority of Vmw175-responsive promoters lack such consensus binding sites. Sequences recognised by Vmw175 that do not obviously fit the consensus have

been described in HSV-1 IE, early and late promoter regions and the Vmw175 interaction with these 'non-consensus' Vmw175 binding sites may be of a lower affinity (Michael *et al.*, 1988; Tedder *et al.*, 1989; Michael and Roizman, 1989; Imbalzano *et al.*, 1990; DiDonato *et al.*, 1991; Flanagan *et al.*, 1991a). The findings of missing contact analyses led DiDonato *et al.*, (1991) to suggest that all Vmw175 binding sites might be related and they proposed a more degenerate version of the Vmw175 binding consensus that could describe both the consensus and non-consensus sites. This is consistent with the findings of the mutational analysis of the IE3 binding site (above) which indicated that the inherent DNA binding activity of Vmw175 is of a fairly low specificity. In addition, a statistical matrix was compiled from their binding site data and was used to derive a matrix mean model that described the probability of any one base occurring at each position of a Vmw175 binding site; this was found to be an accurate predictor of known Vmw175 binding sites (DiDonato *et al.*, 1991). Analysis of the HSV-1 genome with this matrix mean model demonstrated that potential Vmw175 binding sites occur frequently throughout the entire genome, in both coding and non-coding regions of genes from all three temporal classes.

The interaction of Vmw175 with non-consensus sites in HSV-1 E and L promoter regions has been found to require the presence of cellular or infected-cell factors, indicating that additional proteins may be involved in the recognition of these sequences by Vmw175 (Papavassiliou and Silverstein, 1990a; Flanagan *et al.*, 1991a; Papavassiliou *et al.*, 1991). The complexes formed appeared to be less stable with infected cell extracts than with non-infected cell extracts (Papavassiliou and Silverstein, 1990b). Also the extent of phosphorylation of Vmw175 differentially affected its ability to bind sequences within the different temporal classes of promoters (Papavassiliou *et al.*, 1991; Michael *et al.*, 1988). It is possible that the phosphorylation state of Vmw175 determines the interaction with these additional factors and hence acts to differentially control binding site recognition and promoter activity. It is interesting that the bacterially expressed isolated Vmw175 DNA binding domain (which is not phosphorylated) is capable of interacting with both consensus and non-consensus binding sites (Wu and Wilcox, 1990); it may be that the phosphorylation of sites elsewhere regulates the DNA binding activity of the intact Vmw175 protein.

In addition, the binding of Vmw175 to DNA is facilitated by the basal transcription factors TFIID and TFIIB (Smith *et al.*, 1993) (Section 1D.1.2), and Vmw175 amino acids between 142 and 210 are required for the cooperative interactions of Vmw175, TFIID (or TBP) and TFIIB to form the tripartite complex on the DNA. Interestingly, this region spans the serine-rich tract (Section 1C.1.1) that may serve as a transcriptional activation region (Shepard *et al.*, 1989; Section 1C.1.9). Finally, the DNA binding

interaction of Vmw175 induces a significant DNA bend at its occupied binding sites in the HSV-1 gD and IE3 promoters *in vitro* (Everett *et al.*, 1992). The functional relevance of this distortion to the DNA conformation is unknown, but it is a feature of the DNA binding interaction of many sequence-specific DNA binding proteins. The portions of the Vmw175 protein involved in DNA binding have been defined, as discussed later (Sections 1C.1.9 and 1C.1.10).

1C.1.4 ***Dimerisation***

The native Vmw175 protein has been purified as a homodimer which has physical parameters to indicate that it is a highly elongated molecule (Metzler and Wilcox, 1985). Lower abundance tetrameric forms of Vmw175 have additionally been observed, but it is not known whether these represent a distinct population present during infection, or merely a consequence of dimer aggregation during the concentration steps of the purification procedure (Kattar-Cooley and Wilcox, 1989). That Vmw175 binds to DNA as a dimer was indicated by the production of heterodimer bands upon gel retardation analysis of pairs of differently sized Vmw175 partial polypeptides; one of each pair was expressed from the viral genome and the other from a transfected plasmid or transformed cell line (Shepard and DeLuca, 1989; Shepard *et al.*, 1990; Shepard and DeLuca, 1991a). In accordance, Michael and Roizman (1989) reported that two molecules of Vmw175 bind to each Vmw175 binding site, as indicated by the two sequential supershifts of the Vmw175-DNA complex observed upon addition of a Fab fragment of a Vmw175 MAb to a gel retardation experiment. It is noteworthy that the Vmw175 consensus binding site ATCGTCnnnnYCGRC (see above) is unusual for a target site of a dimeric DNA binding protein (Section 1E), in that it does not show any sequence symmetry or palindromic nature. Section 3D.10 discusses the alternative ways in which the Vmw175 dimer could interact with its non-palindromic binding site.

Further evidence of the dimeric nature of the Vmw175 protein during infection was demonstrated by the ability of selected pairs of functionally impaired Vmw175 proteins to complement each other by forming a functional heterodimer molecule (Shepard and DeLuca, 1989; Shepard and DeLuca, 1991a). In one pair of partial Vmw175 peptides of particular note, a peptide lacking sequences necessary for DNA binding (deleted for residues 185 to 309) could supply in *trans* the sequences required for the transactivation function that were lacking from a truncated Vmw175 protein, comprising residues 1-774 (Shepard and DeLuca, 1989). In another case, heterodimerisation with a truncated peptide (X25) inhibited the regulatory functions of wild type Vmw175 in a *trans*-dominant manner, apparently by altering the conformation of the wild type polypeptide (Shepard *et al.*, 1990). Furthermore, expression of the X25 peptide in transgenic mice conferred a significant antiviral effect *in vivo* (Smith and DeLuca, 1992).

Regarding the sequences mediating dimerisation, one report proposed that the N-terminal 90 residues were involved in dimerisation of Vmw175 (DeLuca and Schaffer, 1989) although these sequences are inessential for the functions of Vmw175 (DeLuca and Schaffer, 1987; 1988). The dimeric nature of the isolated Vmw175 DNA binding domain indicated that sequences between residues 275 and 523 are required for dimerisation (Everett *et al.*, 1991a) (Section 1C.1.10). The specific sequences mediating dimerisation have not been further refined, although several mutations have been reported within this region that do not affect dimerisation: insertion mutations at residues 320 and 338 do not disrupt heterodimer formation with a truncated Vmw175 peptide (Shepard and DeLuca, 1991a), and Vmw175 DNA binding domain peptides including single amino acid substitutions within the highly conserved tract of amino acids between residues 445 and 487 (see Section 1C.1.9) can still be purified as dimers (Allen and Everett, 1994). The sequences within the DNA binding domain that are required for dimerisation have been investigated extensively in this study (Section 3D).

1C.1.5 ***Temperature-sensitive mutants of Vmw175***

The pivotal role played by Vmw175 in the HSV-1 transcriptional programme was demonstrated by the analysis of HSV-1 viruses with temperature-sensitive (*ts*) lesions of Vmw175 (Section 1B.3.4). Various *ts* lesions in Vmw175 have been described and these fall into two classes, and the mapping of the lesion sites indicated that Vmw175 may comprise distinct functional domains. The more stringent class of *ts* mutants over-expressed IE gene products and exhibited no early or late gene expression at the NPT (Courtney *et al.*, 1976; Marsden *et al.*, 1976; Watson and Clements, 1978; Preston, 1979a; Dixon and Schaffer, 1980). These studies highlighted the crucial role of Vmw175 during infection. Later temperature shift experiments demonstrated a continuous requirement for Vmw175 for E and L gene expression (Watson and Clements, 1980). Where mapped, these mutations have been found to lie within the N-terminal portion of Vmw175. For example, the *ts*K mutation is caused by a single alanine to valine substitution at residue 475 within the DNA binding domain of Vmw175 (Preston, 1979a; Preston, 1981; Davison *et al.*, 1984) (Fig. 1C.3). The *ts*K protein fails to bind DNA at the NPT *in vitro*, suggesting that DNA binding is a requirement for the regulatory functions of Vmw175 (Paterson *et al.*, 1990). The second class of Vmw175 *ts* viruses are capable of low levels of early gene expression, undergo limited DNA replication but fail to express late polypeptides. The lesions of these partially active mutants map to the C-terminal third of Vmw175, suggesting that a distinct function required for late gene expression may require the integrity of this part of Vmw175 (DeLuca *et al.*, 1984; Paterson *et al.*, 1990). It has been shown that plasmids expressing *ts* Vmw175 polypeptides were also *ts* in their ability to activate transcription in transfection assays and in their DNA binding abilities (DeLuca and

Schaffer, 1985; Paterson *et al.*, 1990). This is important, as it supports the functional relevance of the findings of transient transfection analyses (below).

1C.1.6 *Transactivation*

Following the genetic analyses with the *ts* mutants, the regulatory functions of Vmw175 were investigated exhaustively by various different experimental approaches. An extension of the *ts* virus studies indicated that Vmw175 activated gene expression in *trans* from transfected plasmids bearing early promoter-reporter constructs (Everett, 1983; O'Hare and Hayward, 1984). Other early studies utilised transformed cell lines expressing either Vmw175 or the target promoter or sometimes both, to investigate the effects of Vmw175 on expression of HSV-1 E and L genes (Dennis and Smiley, 1984; Silver and Roizman, 1985; Persson *et al.*, 1985; Arskenasis *et al.*, 1988). The findings of these experiments were highly variable and the ill-defined nature of the stably transformed cell lines cast doubt on their relevance. However, the cell lines expressing Vmw175 have been invaluable in the construction and propagation of viruses with engineered mutations within the IE3 ORF (Section 1C.1.9) and have additionally been used to demonstrate that Vmw175 can complement an E1a-deficient mutant adenovirus (Tremblay *et al.*, 1985).

The approach of transiently co-transfecting tissue culture cells with plasmids carrying cloned effector and reporter constructs was considered more reliable and has been employed widely. The ability of the Vmw175 product of the cloned IE3 gene to transactivate plasmid-borne reporter genes was first described by Everett (1984b). Numerous studies measuring accumulation of mRNA or assaying reporter activity have since demonstrated the ability of Vmw175 to transactivate several HSV-1 early and late promoters (Quinlan and Knipe, 1985; O'Hare and Hayward, 1985a; b; Gelman and Silverstein, 1985; 1986). Transactivation by Vmw175 is at the level of transcriptional initiation, as indicated by *in vitro* nuclear run-off assays and *in vitro* transcription experiments (Godowski and Knipe, 1986; Beard *et al.*, 1986; Pizer *et al.*, 1986).

In addition, Vmw175 has been reported to activate gene expression synergistically with the HSV-1 IE1 gene product, Vmw110; that is, the effect of Vmw175 and Vmw110 together was greater than their combined individual effects (Everett, 1984b; Quinlan and Knipe, 1985). It should be noted that the activities of Vmw175, Vmw110 and the combined activities of the two proteins were not always consistent between different transient transfection studies and this was probably a reflection of differences in the cell-types used and experimental detail (Everett, 1988b). Even though Vmw110 has also been reported to be a promiscuous transactivator in transient assays (Section 1B.3.4), it is clear that Vmw110 cannot functionally substitute for Vmw175, as viruses lacking

functional Vmw175 do not grow on normal cell lines (DeLuca *et al.*, 1985; Russell *et al.*, 1987; DeLuca and Schaffer, 1988; Paterson and Everett, 1990). The synergistic phenotype could be explained by Vmw175 and Vmw110 influencing two different rate-limiting steps in the transcription process and the data is consistent with the regulatory function of Vmw110 acting at a later stage to Vmw175. However, the functional relevance of the synergy observed between Vmw175 and Vmw110 in transfection assays is unclear, as viruses which lack a functional IE1 gene activate viral gene expression as efficiently as wild type virus (at high moi) (Everett, 1989).

Experiments have shown that HSV-1 Vmw175 is able to activate a variety of cellular promoters located in unnatural contexts (Section 1B.3.3). Perhaps of greater relevance to the situation during infection, it has been found that various endogenous cellular genes are induced by HSV-1 infection in a Vmw175-dependent manner, as determined by differential RNA analysis (Kemp and Latchman, 1988), although the transcription of most cellular genes is shut-off upon infection (reviewed by Fenwick, 1984). The identity of these *trans*-induced cellular proteins or the significance of their increased gene expression is not known.

1C.1.7 ***Is DNA binding a requirement for transactivation by Vmw175?***

Whether the DNA binding function of HSV-1 Vmw175 is essential for transactivation of its target promoters has yet to be firmly established. In favour of the requirement for DNA binding is the highly conserved nature of the DNA binding domain between Vmw175 and the other *Alphaherpesvirinae* family members (Fig. 1C.3). Furthermore, virtually all mutations within the defined Vmw175 DNA binding domain that disrupt the ability of Vmw175 to bind to DNA, also disrupt the transactivation function, the *ts*K mutation being the classical example (Preston, 1979a; Paterson *et al.*, 1990) (Section 1C.1.5). The single exception to this was provided by a pseudorevertant Vmw175 polypeptide (ri12) that exhibited an impaired DNA binding activity *in vitro* but was able to support viral growth in tissue culture (Shepard and DeLuca, 1991b). The existence of the single ri12 mutation is sufficient to question the requirement for DNA binding, although it appears possible that the DNA binding experiments of DeLuca and colleagues did not reflect the situation during infection, because in our hands the isolated DNA binding domain of the Vmw175 ri12 protein has a normal DNA binding phenotype at room temperature (Allen, 1993).

Certain mutations in the C-terminal third of HSV-1 Vmw175, outside the DNA binding domain, also influence the DNA binding activity of Vmw175 (see Section 1C.1.9). Four Vmw175 proteins with temperature sensitive lesions in this portion of the protein fail to bind DNA, yet activate certain early genes during infection at the NPT (Paterson *et al.*,

1990). It is possible that the interaction of these *ts* proteins with additional viral or cellular factors (such as TFIIB and TFIID, see below) may enable sufficient DNA binding for the activation of early gene expression, and as such the results of DNA binding assays *in vitro* may not be an accurate reflection of the situation within the cell.

As discussed in Sections 1B.3.1, 1B.3.2 and 1B.3.3, sites that uniquely specify induction by Vmw175 have not been identified as requirements for HSV-1 promoter activity during infection, even though multimerisation of the Vmw175 binding sites in the gD promoter increased promoter activation by Vmw175 during transient transfection assays (Tedder and Pizer, 1988) and their presence contributed to activation *in vitro* (Tedder *et al.*, 1989). Given that the highly degenerate nature of Vmw175 binding is likely to result in the presence of Vmw175 binding sites scattered throughout the HSV-1 genome (Section 1C.1.3), the data is consistent with the idea that the removal of specific Vmw175 binding sites will be functionally compensated for by other Vmw175 binding sites nearby.

The alternative argument might be that Vmw175 binding to DNA is not required for the transactivation of early and late viral promoters, but instead is mediated by modification of the activities of the basal transcription machinery. The interaction with basal transcription factors is indeed likely to be an important part of the transactivation mechanism because Vmw175 forms a tripartite complex with the basal transcription factors TFIIB and TFIID (or TBP) on promoters *in vitro,* in a cooperative manner. However, the interaction of Vmw175 with the basal transcription factors still requires the presence of a Vmw175 binding site (Smith *et al.*, 1993) and if Vmw175 is acting as a nucleating agent by binding DNA and stabilising the DNA association of the proteins that it contacts, then its DNA binding function would be predicted to play a critical role in transactivation. It is interesting that mutant Vmw175 proteins that are unable to form a tripartite complex in a cooperative manner, are only one-half to one-third as active as wild type Vmw175 in transient assays (Smith *et al.*, 1993). This indicates that in addition to stimulating TFIID/TFIIB recruitment to the promoter, transactivation by Vmw175 also involves other mechanisms or interactions with other proteins, perhaps mediated by sequences in the C-terminal third of the protein that appear to be important for the transactivation function but not for the interaction with the basal factors (see Fig. 1C.2).

The available data suggest that DNA binding is required for Vmw175-mediated transactivation of early and late HSV-1 promoters. It can be imagined from the large size of the Vmw175 protein and the wide sequence variation in its target binding sites, that Vmw175 utilises multiple mechanisms and interactions with various cellular

proteins to mediate transactivation of its wide range of target promoters. Further details of these possible interactions and the transactivation mechanism are likely to be discerned in the near future using the reconstituted *in vitro* transcription system of DeLuca and colleagues (Gu *et al.*, 1993). A model for the transactivation mechanism of the alphaherpesvirus major regulatory proteins is proposed in Section 4B.

1C.1.8 ***Structure***

The predicted amino acid sequence of Vmw175 shares an overall identity of 29% with its VZV 140k counterpart, and the two proteins have been divided up into five regions on the basis of the extent of amino acid conservation between them (McGeoch *et al.*, 1986) (Fig. 1C.1 and Table 1C). This pattern of amino acid conservation between VZV 140k and HSV-1 Vmw175 is additionally found in other members of this family of alphaherpesvirus proteins (see later).

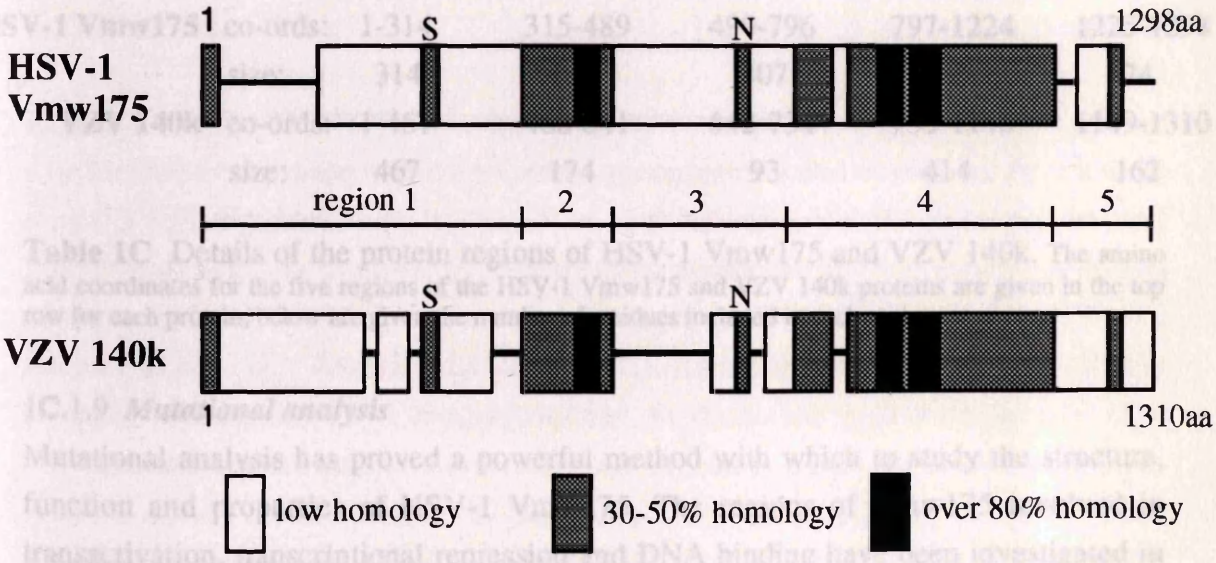


Figure 1C.1 Structure of the HSV-1 Vmw175 and the VZV 140k proteins. The scale demarks the regions 1 to 5 that were defined on the basis of amino acid similarity between the VZV 140k and HSV-1 Vmw175 proteins (McGeoch *et al.*, 1986). Protein sequences are indicated by blocks, where the shading represents the corresponding amino acid identity between the two sequences, as discussed in the text. The heavy horizontal lines are inserted to allow alignment of sequences, and do not themselves represent protein sequences. The letter S indicates the position of the conserved serine-rich tract (Section 1C.1.1) and the letter N indicates the putative nuclear localisation signal (Section 1C.1.2).

Region 1 is poorly conserved between the HSV-1 Vmw175 and VZV 140k proteins in size and amino acid composition, except for the conservation of a tract of serine residues (Section 1C.1.1; marked 'S' in Fig. 1C.1) and limited sequence similarity at the N-terminus. **Region 2** shares 46% identity between the two proteins. Towards the C-terminus of region 2 is a stretch of residues that are more than 80% conserved and the HSV-1 *ts*K mutation maps within this highly conserved area (Section 1C.1.5; Fig. 1C.3). The conservation of **region 3** sequences is restricted to the nuclear localisation signal

(marked 'N' in Fig. 1C.1) and region 3 of Vmw175 is 214 residues longer than its VZV counterpart (Table 1C). **Region 4** is highly conserved (56% identity) between HSV-1 Vmw175 and VZV 140k and includes stretches of near identity between the two proteins. Several *ts* lesions have been mapped to these highly conserved regions (Preston, 1981; Paterson *et al.*, 1990). **Region 5** is largely non-conserved in sequence and size.

The mapping of HSV-1 Vmw175 *ts* lesions to regions 2 and 4 (Section 1C.1.5), together with the high degree of conservation between VZV 140k and HSV-1 Vmw175 within these regions, indicated that regions 2 and 4 were likely to be functionally important; this was borne-out by the mutational analyses described in the following Section.

		region 1	*region 2*	*region 3*	*region 4*	*region 5*
HSV-1 Vmw175	co-ords:	1-314	315-489	490-796	797-1224	1225-1298
	size:	314	175	307	428	74
VZV 140k	co-ords:	1-467	468-641	642-734	735-1148	1149-1310
	size:	467	174	93	414	162

Table 1C Details of the protein regions of HSV-1 Vmw175 and VZV 140k. The amino acid coordinates for the five regions of the HSV-1 Vmw175 and VZV 140k proteins are given in the top row for each protein, below are given the number of residues included in each region.

1C.1.9 *Mutational analysis*

Mutational analysis has proved a powerful method with which to study the structure, function and properties of HSV-1 Vmw175. The regions of Vmw175 involved in transactivation, transcriptional repression and DNA binding have been investigated in several laboratories. Deletion, insertion and truncation mutations have been constructed in plasmid-borne copies of the IE3 gene and the activities of the resulting mutant Vmw175 proteins analysed in transient transfection assays (Mavromara-Nazos *et al.*, 1986b; DeLuca and Schaffer, 1987; Paterson and Everett, 1988a; b; Shepard *et al.*, 1989). Additionally, some of these mutations have been recombined into the normal IE3 gene loci of the viral genome (DeLuca and Schaffer, 1988; Shepard *et al.*, 1989; Paterson and Everett, 1990; Allen and Everett, 1994). Cell lines expressing integrated copies of the IE3 gene (DeLuca *et al.*, 1985; Davison and Stow, 1985) were used for propagation of many of these viruses and their parent null mutations. The results of these mutational analyses are summarised below, sub-divided into the five regions of the Vmw175 coding region for clarity:

Region 1 (codons 1-314) Residues 1-90 are dispensable for the regulatory functions of Vmw175 in transient assays and for virus growth in tissue culture (DeLuca and Schaffer, 1987; 1988). The region between codons 162 and 229, including the serine rich tract, is also inessential during infection (Paterson and Everett, 1990). A viable virus which includes a spontaneous deletion of Vmw175 codons 209 to 236 verifies the inessential nature of sequences in this region (Schroder *et al.*, 1985). However, deletions of the serine-rich sequences from a polypeptide already lacking the C-terminal third of Vmw175 eliminates the transactivation function in transient assays and indicates that the serine-rich tract and the C-terminal domain partially complement each other for transactivation (Shepard *et al.*, 1989). The serine-rich sequences have been predicted to form a loop structure, whose negative charge* may serve as a transcriptional activation region (Shepard *et al.*, 1989). The majority of insertion mutations within region 1 have little effect on the functions of Vmw175, although the C-terminal 60 or so amino acids of region 1 are functionally important, as discussed below (Paterson and Everett, 1988a; b; Shepard *et al.*, 1989).

* (when phosphorylated)

Region 2 (codons 315-489) The debilitating effects of insertion mutations (insertion of 4 or 2 amino acids) made within the C-terminus of region 1 and throughout the whole of region 2 demonstrated the critical nature of these sequences for the autoregulation and transactivation functions of Vmw175 in transient assays and DNA binding *in vitro*, and additionally for Vmw175 functions during virus growth (Paterson and Everett, 1988a; b; Shepard *et al.*, 1989; Allen and Everett, 1994b). These sequences of Vmw175 actually comprise a separable DNA binding domain, as discussed in Section 1C.1.10. The insertion mutational analyses indicated that the very highly conserved sequences between amino acids 445 and 487 were critical for DNA binding and it has been proposed from secondary structure predictions that this region of Vmw175 may contain a helix-turn-helix (HTH) DNA binding motif (Shepard *et al.*, 1989) (Section 1E.2). Residues 445-457 are predicted to form an α-helix, a predicted β–turn follows (458-464) and then comes another α-helical domain (briefly interrupted by a proline residue at amino acid 476; see Fig. 1C.3) (Allen and Everett, 1994); however neither of these α-helical regions of Vmw175 exhibit the amphipathic residue distribution that is characteristic of the helices within the classical HTH motif (Section 1E.2). A thorough single amino acid substitution analysis of this putative HTH region in the isolated Vmw175 DNA binding domain has defined several point mutations that result in a temperature-sensitive DNA binding phenotype, and those at Vmw175 residues 456 and 459 and a double substitution of residues 463/464 prevented DNA binding (see Fig. 1C.3); Vmw175 proteins including these substitutions were unable to support virus growth, indicating the correlation between DNA binding as measured *in vitro*, and the

ability of the protein to support the viral lytic cycle in tissue culture (Allen and Everett, 1994).

An insertion mutation at Vmw175 residue 338 selectively reduced the DNA binding activity to a consensus binding site and eliminated binding to a non-consensus binding site, suggesting that sequences at the N-terminus of region 2 may be involved indirectly in DNA binding (Shepard *et al.*, 1989). Therefore sequences at the C-terminus of region 2 and also sequences spanning the boundary of regions 1 and 2, are implicated in the DNA binding function of Vmw175. An insertion between these two segments, at residue 368, produced a Vmw175 polypeptide that was wild type for DNA binding (Shepard *et al.*, 1989); however this is in contrast to the mutations in this central portion of region 2 that eliminated DNA binding in the analysis of Paterson and Everett (1988b). It seems that this is a complex situation, with insertion mutations throughout the region causing disruption of DNA binding function. However, the effects of the insertions may be indirect, as conformational distortions are inevitable.

Region 3 (codons 490-796) The majority of mutations within region 3 have very minor effects on the functions of Vmw175, although some significantly reduce DNA binding presumably through steric effects (Paterson *et al.*, 1988a; b). However sequences within region 3 do have an important role to play, as truncated peptides including residues 1-590 possessed no regulatory activity while a peptide 1-774 was functional (DeLuca and Schaffer, 1987; 1988). This may be a consequence of removal of the nuclear localisation signal that maps to this region (Section 1C.1.2).

Region 4 (codons 797-1224) and **Region 5** (codons 1225-1298) Despite being the most highly conserved protein region between HSV-1 Vmw175 and VZV 140k, the functional role of region 4 remains unclear. Interestingly, Vmw175 polypeptides including residues 1-825 or 1-1060 had impaired regulatory activities (Mavromara-Nazos, 1986; DeLuca and Schaffer, 1987; 1988), while a shorter polypeptide including residues 1-774 was fully functional for normal autoregulation and E gene activation (DeLuca and Schaffer, 1987; 1988). Additionally, several very minor insertions or *ts* mutations within regions 4 and the N-terminal portion of region 5 severely disrupted the DNA binding and regulatory activities of Vmw175 (Paterson and Everett, 1988a; b; Paterson *et al.*, 1990). The fact that minor mutations had a much greater effect than deletion of the complete region suggests that although the C-terminal sequences are not essential for the functions of Vmw175, they do play a role, as yet unknown, in regulating the activities of Vmw175. The precise structure adopted by these regions may be important for Vmw175 activity as the mutations within region 4 that disrupted the regulatory activities of Vmw175 also appeared to affect the

conformation of the mutant Vmw175 polypeptides (Paterson and Everett, 1988b; Paterson *et al.*, 1990). However, when the 1-774 residue Vmw175 polypeptide was recombined back into the viral genome, it was found that DNA replication and late gene expression were inefficient; thus the C-terminal regions are implicated in facilitating Vmw175 function at later times in infection (DeLuca and Schaffer, 1988).

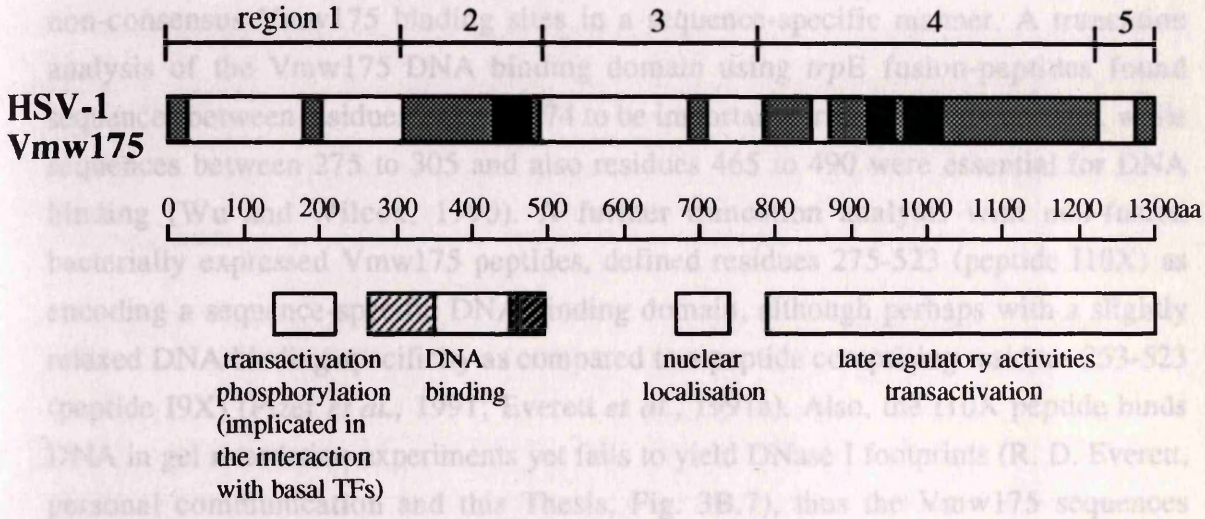


Figure 1C.2 The important functional regions of HSV-1 Vmw175. The scale at the top indicates the protein regions 1 to 5 that correspond to the representation of HSV-1 Vmw175 shown below, the intensity of shading of the portions of the protein indicate the degree of amino acid homology to VZV 140k (as defined in Fig. 1C.1). The lower scale indicates amino acid residues in the 1298 residue protein, and at the bottom are shown the functional domains of Vmw175 as defined by the mutational analyses of DeLuca and Schaffer (1988); Paterson and Everett (1988a; b), Shepard *et al.* (1989). The light shading in the region required for DNA binding indicates sequences that appeared to be indirectly involved in DNA recognition, while the dark shading indicates those sequences that appeared to be more directly involved. The sequences within region 1 are also required for the formation of a tripartite complex with the basal transcription factors TFIID/TFIIB in a cooperative manner (Smith *et al.*, 1993).

In summary, it is apparent from these mutagenic analyses that the regions that are most highly conserved between HSV-1 Vmw175 and VZV 140k (regions 2 and 4) are the most important for the functions of Vmw175 (see Fig. 1C.2). The integrity of sequences spanning region 2 is critical for the DNA binding, transactivation, and repression functions of Vmw175, and these sequences are actually sufficient for DNA binding (Section 1C.1.10). Also, the structural requirements for the negative autoregulation function appear to be lower than those for transactivation, as demonstrated by a Vmw175 partial peptide that spanned the DNA binding domain and retained autoregulatory activities in transient transfection assays (Shepard *et al.*, 1989). Transactivation of early promoter constructs additionally required either sequences spanning the serine-rich tract in region 1, or the C-terminal sequences in regions 4 and 5 (Shepard *et al.*, 1989). The evidence suggests that both the C-terminal region and the serine-rich tract of Vmw175 contribute to transactivation during infection, perhaps by contacting distinct cellular transcription factors.

1C.1.10 *The isolated Vmw175 DNA binding domain*

The mutational analyses of the intact Vmw175 protein (above) illustrated the importance of the C-terminus of region 1 and the whole of region 2 for DNA binding and this was borne-out by the ability of peptides including only these sequences to bind to DNA. Wu and Wilcox (1990) expressed in *E. coli* residues 262 to 490 of Vmw175, fused with a portion of *trpE* sequence; this fusion peptide bound to both consensus and non-consensus Vmw175 binding sites in a sequence-specific manner. A truncation analysis of the Vmw175 DNA binding domain using *trp*E fusion-peptides found sequences between residues 262 and 274 to be important for high affinity binding, while sequences between 275 to 305 and also residues 465 to 490 were essential for DNA binding (Wu and Wilcox, 1990). A further truncation analysis with non-fusion bacterially expressed Vmw175 peptides, defined residues 275-523 (peptide I10X) as encoding a sequence-specific DNA binding domain, although perhaps with a slightly relaxed DNA binding specificity as compared to a peptide comprising residues 253-523 (peptide I9X) (Pizer *et al.*, 1991; Everett *et al.*, 1991a). Also, the I10X peptide binds DNA in gel retardation experiments yet fails to yield DNase I footprints (R. D. Everett, personal communication and this Thesis, Fig. 3B.7), thus the Vmw175 sequences between 253 and 275 probably play an indirect role in DNA binding, as suggested by Wu and Wilcox (above). It is interesting that, in addition to the conserved region 2 sequences, the Vmw175 DNA binding domain also includes the less highly conserved C-terminus of region 1 (see Fig. 1C.3). The less conserved nature of these region 1 sequences is perhaps consistent with the findings of mutagenic and truncation analyses which have suggested that these region 1 sequences are likely to stabilise the structure of the domain, rather than being directly involved in sequence-specific DNA binding (Shepard *et al.*, 1989; Wu and Wilcox, 1990). At around the same time it was shown that proteolysis of full length Vmw175 could release a similar DNA binding peptide, having N- and C-terminal boundaries between residues 229 and 292, and 495 and 518 respectively. This indicates that the DNA binding sequences of Vmw175 comprise a stable and physically separable domain (Everett *et al.*, 1990).

The finding that the I9X Vmw175 DNA binding domain had a DNA binding specificity very similar to that of the intact Vmw175 protein (Everett *et al.*, 1991a) was of great importance, because it indicated that the DNA binding properties of the isolated DNA binding domain are likely to reflect the situation with the intact protein, without the complicating influences of the putative regulatory sequences found within the C-terminus of the Vmw175 protein. The convenience of the bacterial system for expressing large amounts of the DNA binding domain peptides has made this a very favourable approach for the analysis of the DNA binding activity of this family of proteins, and has been the basis for several recent studies.

```
HSV-1 Vmw175   253  AAARTPAASA GRIERRRARA AVAGRDATGR FTAGQPRRVE LDADATSGAF
VZV 140k       406  GSTKQNIQPG YRSISGPDPR IRKTKRLAGE PGRQRQKSFS LPRSRTPIIP
SVV 140k       396  DPTQKKPRPG FPSIHDPDPR IRRTRRSVGG QRRRRHRSFS LPRSITPVIP
PRV IE180      433  APAAAPRPSA SSASSSAAAS PAPAPEPARP PRRKRRSTNN HLSLMADGPP
EHV-1 IE1      367  PRPQTQTQAP AEEAPQTAVW DLLDMNSSQA TGAAAAAASA PAAASCAPGV
BHV-1 ICP4     252  GAGPGAPRGR APGPPAPSPP ETAPPPPRRQ RGRATGARAG YAAAPRDGPP
MDV ICP4       223  ASPRTRKLED EDYLPQETAN RRGGGRPRGR PPKSGRAVQR NDIQVTSSSG
Identical amino acids       ---------- ---------- ---------- ---------- ----------

                                region 1 →← region 2
HSV-1 Vmw175   303  YARYRDGYVS GEPWPGAGPP PPGRVLYGGL GDSRPGLWGA PEAEEARRRF
VZV 140k       456  PVSGPLMMPD GSPWPGSAPL PSNRVRFGPS GETREGHWED EAVRAARARY
SVV 140k       446  PLEGPLNMPD GSPWPGSGDI PADKVRFGPS GECRDGLWDD ELVRAVKARY
PRV IE180      483  PTDGPLLTPL GEPWPGSDPP ADGRVRYGGA GDSREGLWDE DDVRQAAARY
EHV-1 IE1      407  YQREPLLTPS GDPWPGSDPP PMGRVRYGGT GDSRDGLWDD PEIVLAASRY
BHV-1 ICP4     302  PLEGPLLTPS GEAWPGSAPP PPGRVRFGGA GDTREGLWDC PEIREAAARY
MDV ICP4       273  LADTSPYDLC GSVWWEVPLP PPGRCWFGGL GGHRQALTDS PEIVEAIHRF
Identical amino acids       ----PL--P- G-PWPGS-PP P-GRVR-GG- G--R-GLWD-- ----A--RY

HSV-1 Vmw175   353  EASGAPAAVW APELGDAAQQ YALITR.LLY .TPDAEAMGW LQNPRVVPGD
VZV 140k       506  EASTEPVPLY VPELGDPARQ YRALIN.LIY .CPDRDPIAW LQNPKLTGVN
SVV 140k       496  EMGSEPAPFY IPELGDPSKQ YHALIN.LIY .CPDRDPIAW LQNPKLTGDN
PRV IE180      533  RAAAGPVPVF IPEMGDSRKQ HEALVR.LIY SGAAGEAMSW LQNPRMQAPD
EHV-1 IE1      457  AEAQAPVPVF VPEMGDSTKQ YNALVR.MVF ES..REAMSW LQNSKLSGQD
BHV-1 ICP4     352  AAAAGPAAVF VPEMGDAGKQ YAALVD.LVY AR..RDAMAW LQSAKLAGPD
MDV ICP4       323  NTSHGPVPVY VEEMKDYAKQ YDALVNSLFH KSMKVNPLNW MHHGKLSPAD
Identical amino acids       -----P-PV- -PE-GD--KQ Y-AL---L-Y ---------W LQN-KL----

                                                                    hhhhh
HSV-1 Vmw175   401  VALDQAC.FR ISGAARNSSS FITGSVARAV PHLGYAMAAG RFGWGLAHAA
VZV 140k       554  SALNQFY.QK LLPPGRA.GT AVTGSVASPV PHVGEAMATG EALWALPHAA
SVV 140k       544  LALSHFC.QK LLPPGRA.GT AVTGSVACPV PHVGEAMARG EALWALPHAA
PRV IE180      582  QRFNQFC.QR RVHAPHGHGS FITGSVTPPL PHIGDAMAAQ DPLWALPHAV
EHV-1 IE1      504  QNLAQFC.QK FIHAPRGHGS FITGSVANPL PHIGDAMAAG NALWALPHAA
BHV-1 ICP4     399  LQLARLL.QR RVQGCRGHSS FITGSVTAPL PPVGDAMAAQ NALWALPHVA
MDV ICP4       373  AALNHIYVQK FQSSYDSPGA AVTGTVNRCI PHIAGAMKER KLLWAFPHIA
Identical amino acids       --L--F--Q- -----R--G- --TGSV---- PH-G-AMA-- --LWALPHAA
                                                                region 2 →
                            hhhhhhhhtt ttttthhhhh hhhhhh hhh hhhhhhhh
HSV-1 Vmw175   450  AAVAMSRRYD RAQKGFLLTS LRRAYAPLLA RENAALTGAA
VZV 140k       602  AAVAMSRRYD RAQKHFILQS LRRAFASMAY PEATGSSPAA
SVV 140k       592  AAVAMSRRYD RAQKHFILQS LRRAYATMAY PHLAQQVAGQ
PRV IE180      631  SAVAMSRRYD RTQKTFILQS LRRAYADMAY PGRAADPRAG
EHV-1 IE1      553  ASVAMSRRYD RTQKSFILQS LRRAYADMAY PRDEAGRPDS
BHV-1 ICP4     448  ACVAMSRRYD CDQKLFLLQS LRRAYAPMAY PEAGAGGSGA
MDV ICP4       423  ASIAMTRRYC KDQKTFLFRS LKKAYASMAF PDNSSETEKG
Identical amino acids       A-VAMSRRYD R-QK-FILQS LRRAYA-MAY P---------
```

Figure 1C.3 The sequences of the alphaherpesvirus homologues of HSV-1 Vmw175 spanning the DNA binding domain. The amino acid sequences are given for the C-terminal end of region 1 and region 2 of HSV-1 Vmw175 and its VZV, SVV, PRV, EHV-1, BHV-1 and MDV counterparts (McGeoch *et al.*, 1986; Davison and Scott, 1985; Clarke *et al.*, 1993; Cheung, 1989; Vlcek *et al.*, 1989; Grundy *et al.*, 1989; Schwyzer *et al.*, 1993 and Anderson *et al.*, 1992). The alignment was made using the GCG PileUp programme, although region 1 sequences are shown without gaps as negligible homology was present in the alignment. Amino acids present in five or more of the seven proteins are given below the alignment and those present in all seven sequences are shown in bold; many of the remaining residues are semi-conserved between the sequences. Shown in bold in the HSV-1 Vmw175 sequence are the *ts*K mutation (residue 475) and substitutions that disrupt the DNA binding activity of the Vmw175 DNA binding domain (residues 456, 459 and the 463/464 double mutation) (Section 1C.1.9). In the VZV 140k sequence, substitution of the residue shown in bold (residue 548) disrupted the DNA binding and regulatory functions of 140k (Section 3E). The residues that comprise the putative HSV-1 Vmw175 helix-turn-helix DNA binding motif (Section 1C.1.9) are indicated by the letters h (helix) and t (turn) (Allen and Everett, 1994).

These have included the purification of the Vmw175 DNA binding domain to homogeneity and subsequent demonstration of its dimeric nature in solution (Everett *et al.*, 1991a), and mutagenesis of the putative helix-turn-helix motif of Vmw175 (Section 1C.1.9) (Allen and Everett, 1994). The high yields of DNA binding domain peptides that can be produced by this approach will hopefully facilitate the production of crystals suitable for elucidation of the structure of the DNA binding domain.

1C.2 VZV 140k

The 1310 residue IE protein encoded by VZV gene 62 has a predicted molecular weight of 139,989D and is referred to here as VZV 140k (alternative names: IE175, ORF62, p140), although it has a mobility on SDS PAGE closer to 175,000 (Davison and Scott, 1985; Felser *et al.*, 1988). As discussed previously (Section 1C.1.8), the predicted amino acid sequence of VZV 140k shows extensive homology to the HSV-1 Vmw175 sequence and the level of conservation of regions 2 and 4 is particularly high between the two proteins (McGeoch *et al.*, 1986; Fig. 1C.1). The two copies of the gene encoding VZV 140k occupy equivalent genomic positions to those of the HSV-1 IE3 genes (Fig. 1A.4) (Davison and Scott, 1986). VZV 140k is a phosphoprotein and can be phosphorylated *in vitro* by the VZV serine protein kinase associated with VZV ORF47 (Ng *et al.*, 1994). 140k is present in abundance in the virus tegument (Kinchington *et al.*, 1992) and this is consistent with it being a major VZV antigen in the human immune response (Arvin *et al.*, 1991). Furthermore, VZV 140k has potential value as a vaccine component as guinea pigs immunised with a vector expressing VZV gene 62 are protected against subsequent VZV challenge (Sabella *et al.*, 1993).

1C.2.1 *VZV 140k, a functional counterpart of HSV-1 Vmw175*

The functional similarity of VZV 140k to its HSV-1 sequence homologue was demonstrated by the ability of either a plasmid-borne copy of VZV gene 62 or a cell line stably transformed with gene 62, to complement growth of HSV-1 viruses with deletion or *ts* lesions within the HSV-1 IE3 gene (Felser *et al.*, 1987; 1988). In addition, an HSV-1 virus containing copies of VZV gene 62 in place of the HSV-1 IE3 genes (virus HSV-140) is viable in tissue culture (Disney and Everett, 1990). Characterisation of the HSV-140 virus found that VZV 140k can perform many, but not all of the functions of HSV-1 Vmw175 in tissue culture infections; the HSV-140 virus failed to give obvious plaques on normal cell lines and produced reduced amounts of certain late polypeptides at high moi, and at low moi the polypeptide expression from HSV-140 was very inefficient. These data indicate that the VZV 140k and HSV-1 Vmw175 proteins differ in functional detail.

1C.2.2 *The 140k DNA binding domain*

VZV 140k sequences equivalent to those comprising the HSV-1 Vmw175 DNA binding domain (Section 1C.1.10) have been expressed in *E. coli* fused to 323 codons of the bacterial *trp*E gene. This fusion peptide, FP617, included 140k residues 417 to 647 that span the N-terminus of region 1 and the whole of region 2 (Wu and Wilcox, 1991) (Fig. 3A.1). The FP617 peptide exhibited sequence-specific DNA binding activity, although with a lower apparent DNA binding affinity than the equivalent domains of the corresponding HSV-1 and PRV proteins. The FP617 peptide bound various sequences in the PRV IE, HSV-1 IE3 and VZV gene 62 promoters (Wu and Wilcox, 1991), as well as binding the Vmw175 consensus binding site that lies within the HSV-1 gD promoter (Betz and Wydoski, 1993). Also a chimeric *trp*E fusion DNA binding domain which included the C-terminus of the VZV 140k domain (residues 507 to 647) linked to the N-terminus of the HSV-1 Vmw175 domain, possessed the same DNA binding specificity as the FP617 140k domain. This result suggested that the C-terminal portion of region 2 determines the DNA binding specificity, but unfortunately the complementary domain swap peptide (the N-terminus of the 140k domain fused to the C-terminus of the Vmw175 domain) was non-informative (Wu and Wilcox, 1991). Similarity between the VZV 140k and HSV-1 Vmw175 DNA binding domains has been further indicated by the identification of a common antigenic epitope spanning the boundary of regions 1 and 2 of these proteins (Everett *et al.*, 1993c).

1C.2.3 *Regulatory functions of VZV 140k*

Transactivation: Transient transfection assays have demonstrated that VZV 140k can activate VZV promoters from all three putative kinetic classes including IE (ORF4 and ORF61) (Perera *et al.*, 1992a), E (tk) and L (gpI, gpIV, gpV) (Inchauspe *et al.*, 1989; Perera *et al.*, 1992a; Nagpal and Ostrove, 1991). VZV 140k also activates heterologous viral and cellular promoters: these include HSV-1 *tk* (Inchauspe and Ostrove, 1989; Cabirac *et al.*, 1990) HSV-1 gD (Everett, 1984b; Disney, 1990) SV40, c-myc and the HIV LTR (Inchauspe and Ostrove, 1989). That VZV 140k is a more promiscuous transactivator than HSV-1 Vmw175 was indicated by the inability of HSV-1 Vmw175 to activate the VZV *tk* promoter or the enhancer-less SV40 promoter (Inchauspe and Ostrove, 1989; Sekulovich *et al.*, 1988).

In addition, VZV 140k is also a more potent transactivator than HSV-1 Vmw175 (Cabirac *et al.*, 1990; Disney, 1990). Recently a potent transcriptional activation domain has been mapped to the N-terminal sequences of VZV 140k between residues 9 and 86, by testing the ability of fragments of VZV 140k fused to the yeast GAL4 DNA binding domain to activate a reporter construct in transient transfection experiments (Cohen *et al.*, 1993; Perera *et al.*, 1993). A 140k protein with these putative activation sequences

deleted was non-functional and interfered with the regulatory function of the full-length protein, perhaps through competition for a cellular transcription factor or promoter binding site (Perera e*t al.*, 1993). The 140k activation domain has been reported to be of a comparable strength (Cohen *et al.*, 1993), or about 6-fold less potent (Perera *et al.*, 1993), than the powerful HSV-1 Vmw65 transcriptional activation region. As is the case with many eukaryotic regulatory factors (Section 1D.2.2), the VZV 140k activation sequences have an acidic nature. As yet no transcriptional activation domain has been reported for the HSV-1 Vmw175 protein, although it will not be equivalent to that of VZV 140k as the N-terminal 90 residues of Vmw175 are non-essential (DeLuca and Schaffer, 1987; 1988).

Like its HSV-1 counterpart protein (Section 1C.1.6), VZV 140k can activate gene expression in synergy with other viral IE gene products. Synergistic activation with VZV ORF4 has been observed in several studies whereas synergy with VZV ORF61 depended on the cell type used (Inchauspe *et al.*, 1989; Perera *et al.*, 1992a). One study found that VZV 140k could also activate gene expression in synergy with the HSV-1 Vmw110 protein, the partial sequence homologue of VZV ORF61 (Cabirac *et al.*, 1990); but again this activity depended on the particular system employed (Disney, 1990). It appears that the transactivation function of VZV 140k is mediated by sequence-specific DNA binding, as multiplication of the HSV-1 gD Vmw175 consensus binding site increased the efficiency of VZV 140k-mediated transactivation from the gD promoter in transient transfection experiments (Betz and Wydoski, 1993) (the FP617 fusion DNA binding domain peptide binds to this site *in vitro*).

Negative autoregulation: Treatment of cell lines expressing VZV 140k with cycloheximide increased accumulation of VZV gene 62 mRNA and indicated that VZV 140k, like its HSV-1 counterpart (Section 1B.2.3), may negatively autoregulate its own expression (Felser *et al.*, 1988). This was shown to be the case in transient transfection experiments in Vero and BHK cells (Disney *et al.*, 1990). However, contradictory results were obtained using T lymphocytes and neural cell lines, in which VZV 140k positively stimulated its own expression (Perera *et al.*, 1992b). Both studies were in agreement on the fact that VZV 140k slightly activated the HSV-1 IE3 promoter rather than down-regulating it, and this was supported by the over-expression of the VZV 140k protein observed during infection with a virus that includes the VZV ORF62 under the control of the HSV-1 IE3 promoter sequences (virus HSV-140; Disney and Everett, 1990). The fact that HSV-1 Vmw175 down-regulates the VZV gene 62 promoter, whereas VZV 140k activates the HSV-1 IE3 promoter (Disney *et al.*, 1990) indicates that even though HSV-1 Vmw175 and VZV 140k appear functionally similar in many aspects, the two proteins are not interchangeable for autoregulation. Regarding the

failure to observe either VZV 140k-mediated or HSV-1 Vmw175-mediated down-regulation of the VZV gene 62 promoter in the T lymphocyte and neural cell lines (Perera *et al.*, 1992b), it should be noted that in these cell lines HSV-1 Vmw175 also had no demonstrable affect on its cognate promoter even though it has been established that Vmw175 down-regulates its own expression during infection (Section 1B.2.3).

DNase I footprinting analyses found that the VZV 140k DNA binding domain fusion peptide FP617 recognises a sequence in the immediate vicinity of the VZV gene 62 mRNA start site defined by McKee *et al.* (1990) (Wu and Wilcox, 1991). This sequence includes a divergent form of the ATCGTC motif of the HSV-1 Vmw175 consensus binding site and by analogy to the situation regarding the autoregulation function of HSV-1 Vmw175 (Section 1B.2.3), this interaction is implicated in the down-regulation of the gene 62 promoter by VZV 140k. They proposed that binding to a specific sequence of conserved nucleotides at or near the mRNA start site may be a common autoregulatory mechanism used by the major transcriptional regulatory proteins of the alphaherpesviruses (Wu and Wilcox, 1991).

However, Wu and Wilcox (1991) found the FP617 peptide to have an even stronger affinity for another sequence adjacent to the previously proposed mRNA start site further upstream (Felser *et al.*, 1988); this sequence also included an ATCGTC motif (see Fig. 1B.4). They suggested that 140k differentially influences gene 62 transcription by binding with different affinities to the two alternative transcription start sites and to support their hypothesis they cite an unpublished observation that longer gene 62 mRNA transcripts have been detected that may have originated from the upstream start site. However, this would appear highly unlikely from the presence of translational start and stop codons located between this putative upstream mRNA start site and the ATG codon at the start of the major open reading frame.

1C.3 PRV IE180

Pseudorabies virus expresses a single IE RNA, which like its HSV and VZV counterparts is expressed from the inverted repeat sequences bounding Us, and encodes a 1446 residue polypeptide of 180kD (Ihara *et al.*, 1983). The predicted sequence of the product of the PRV IE gene, PRV IE180, exhibits the similar pattern of sequence conservation seen between HSV-1 Vmw175 and VZV 140k (Cheung, 1989; Vlcek *et al.*, 1989) (Section 1C.1.8). Experiments using the *ts*G_1 virus with a *ts* lesion in PRV IE180, have indicated that only IE180 RNA is transcribed at the NPT (Ben-Porat *et al.*, 1982; Ihara *et al.*, 1983). Therefore the functional IE180 protein is required for

expression of the later classes of PRV genes, also these experiments suggested that IE180 represses its own expression, and as such the PRV transcriptional programme shows similarities to that of HSV-1 (Section 1B.1). The promiscuous transactivating phenotype of IE180 has been demonstrated using transient expression experiments (Green *et al.*, 1983; Imperiale *et al.*, 1983; Everett, 1984b) and Thali *et al.*, (1990) have proposed that PRV IE180 functions to activate gene expression from suboptimally activated RNA polymerase II promoters. Suboptimally activated promoters are those that are not already fully activated by multiple *cis*-acting or enhancer elements, and those that are not bound in repressive chromatin structures; hence during infection the viral promoters are the most abundant targets for PRV IE180-mediated activation.

Experiments have shown that PRV IE180 binds to multiple sequences within the adenovirus major late promoter (used in the *in vitro* transcriptional analyses described below) and the human *hsp* 70 promoter, and that DNA binding to these sequences is required for promoter transactivation *in vitro* (Cromlish *et al.*, 1989). The binding of PRV IE180 to numerous sites within its target promoters is a feature conserved with the VZV 140k protein but not with the HSV-1 Vmw175 protein (this Thesis, Sections 3A.6 and 3A.8). The sequences involved in 'promoter targeting' were found to lie within the N-terminal half of IE180 (Martin *et al.*, 1990b) and a fusion protein consisting of sequences spanning the conserved region 2 of IE180 binds DNA in a sequence-specific manner (Wu and Wilcox, 1991). Its DNA targets included a site fitting the HSV-1 Vmw175 consensus binding sequence immediately adjacent to the PRV IE mRNA start site, and this may be involved in its autoregulatory function.

In addition, a potent transcriptional activation domain has been mapped to the 34 N-terminal residues of PRV IE180, by employing the same approach used to identify the activation sequences of VZV 140k (Section 1C.2.3) (Martin *et al.*, 1990b). Importantly, a PRV IE180 protein lacking these putative activation sequences is stable and fails to activate gene expression. Furthermore, it inhibited activation by the GAL4-IE180(1-34) construct. Like the N-terminal VZV 140k activation domain (Section 1C.2.3), the PRV IE180 N-terminal activation domain has an acidic nature but the sequences are not particularly well conserved with 140k or the other members of this family of proteins.

1C.3.1 *The PRV IE180 protein facilitates transcription by recruiting the basal transcription factor TFIID to the promoter*

Although the role played by IE180 during PRV infection has not been well characterised, biochemical analyses of the partially purified IE180 protein using *in vitro* transcription systems have provided important insight into the likely mechanism of

action of this family of alphaherpesvirus proteins. The finding that the transactivator proteins PRV IE180, HSV-1 Vmw175 and adenovirus E1a overlapped in function (Feldman *et al.*, 1982; Tremblay *et al.*, 1985) was the first indication that these promiscuous viral activators might function through the cellular basal transcription machinery. The ability of PRV IE180 (but not the *ts*G1 IE180 protein) to activate transcription in a cell-free system indicated that it does not function by simply increasing the concentration of a transcription factor *in vivo* (Abmayr *et al.*, 1985). In a follow up study, it was found that the effect of adding IE180 to the *in vitro* transcription reaction was very similar to either (i) titration of non-specific factors from the promoter template by the addition of excess competitor DNA, or (ii) preincubation of the promoter template with the basal transcription factor TFIID (Abmayr *et al.*, 1988). These data imply that non-specific proteins normally compete with TFIID for binding to the promoter, thus making binding of TFIID rate-limiting for transcription. PRV IE180 enables TFIID to overcome this rate-limiting step, presumably by facilitating the TFIID-promoter interaction. These analyses are consistent with the more recently observed functional interplay between the HSV-1 Vmw175 protein and TFIID (Smith *et al.*, 1993; Gu *et al.*, 1993).

Further experiments by Workman *et al.* (1988) employed an *in vitro* system where transcriptional preinitiation complex formation was in direct competition with nucleosome assembly over the target promoter region. Under these conditions, the presence of PRV IE180 facilitated the TFIID-promoter interaction in preference to the assembly of nucleosomes. Although these findings obviously have great significance for the general mechanism of promoter activation from chromatin templates, their relevance to the situation during viral infection is not clear because at least for the HSV-1 and BHV-1 genomes, only a minor fraction of the lytic DNA appears to be bound in nucleosomes (Leinbach and Summers, 1980; Muggeridge and Fraser, 1986; Seal *et al.*, 1988).

1C.4 EHV IE1

The single IE transcript of equine herpesvirus 1 again maps to the short repeats of the genome (Gray *et al.*, 1987a; b), and has a spliced leader sequence (Harty *et al.*, 1989), a feature shared with its BHV-1 counterpart (Wirth *et al.*, 1991). The largest open reading frame in this region has 1,487 codons and predicts a polypeptide of 155kD with the same pattern of amino acid conservation seen with the other alphaherpesvirus proteins (Grundy *et al.*, 1989). However, four antigenically cross-reactive EHV-1 IE polypeptides (ranging in size from 200 to 125kD) are translated from the single IE

mRNA *in vitro* (Robertson *et al.*, 1988) and are likely to arise from post-translational modifications of a larger protein precursor. The largest polypeptide, IE1, is the most abundant and is the major phosphorylated form of the protein. Cycloheximide blocking of IE mRNA translation has indicated the key role played by the IE gene products in regulating the expression of all other EHV-1 genes (Gray *et al.*, 1987a). The regulatory properties of EHV-1 IE1 have been examined in transient transfection assays on various promoters (Smith *et al.*, 1992a); IE1 activated early and late EHV promoters as well as the HSV-1 *tk* promoter, the SV40 promoter-enhancer region and the enhancer-less SV40 promoter. In these assays, transactivation of the late EHV-1 promoters also required the presence of another EHV-1 gene product that shows sequence homology to HSV-1 Vmw63 (Section 1B.3.4). Like its VZV counterpart, EHV IE1 appears to be a more promiscuous activator than HSV-1 Vmw175.

In addition, EHV-1 IE negatively autoregulated its own promoter in the transient assays and the *cis*-acting autoregulatory elements have been mapped to between -288 and +73 of the IE gene (Smith *et al.*, 1992a). By analogy to the apparent mechanism whereby HSV-1 Vmw175 down-regulates its own production (Section 1B.2.3), negative autoregulation of EHV-1 IE1 production is likely to be mediated through a sequence homologous to the Vmw175 consensus binding site that is found overlapping the EHV-1 IE mRNA start site (Smith *et al.*, 1992a). The negative regulatory function of EHV-1 IE1 appeared to be more specific than that of its HSV-1 Vmw175 counterpart, as it fails to down-regulate the HSV-1 IE3 promoter (Smith *et al.*, 1992a). This is reminiscent of the findings of similar experiments with the VZV 140k protein (Disney *et al.*, 1990). In fact, the evidence suggests that the VZV and EHV-1 protein counterparts are more closely related to each other than to the HSV-1 Vmw175 protein, because as well as having similarly promiscuous transactivating phenotypes and similar specificities of transcriptional repression, the sequences of the conserved EHV-1 and VZV protein regions 2 and 4 show a higher level of homology than to HSV-1 Vmw175. (Grundy *et al.*, 1989). This is consistent with the proposed similar nature of the VZV and EHV-1 transcriptional programmes, as discussed in Section 1B.4.3.

1C.5 Other alphaherpesvirus homologues of HSV-1 Vmw175

The sequences of other members of this family of alphaherpesvirus regulatory proteins have recently been derived and the conservation of the predicted amino acid sequences of regions 2 and 4 in all these proteins further indicates the functional importance of these particular portions of sequence (see Fig. 1C.3 for a comparison of the sequences

spanning region 2). In addition, all the proteins contain the serine-rich sequence and the basic putative nuclear localisation signal (Sections 1C.1.1 and 1C.1.2).

BHV-1 The sequence of the ICP4 protein of bovine herpesvirus type 1 is divided into 5 regions, but it is unique in having the polyserine tract displaced to region 5, rather than being in region 1. Also BHV-1 ICP4 has two striking stretches of polyglutamic acid residues within region 3 that are not found in the other counterpart proteins but may have implications for its transactivation function (Section 1D.2.2). BHV-1 ICP4 shares functional similarities with other members of the family; it repressed its own promoter and activated other E and L herpesvirus promoters in transient transfection assays (Schwyzer *et al.*, 1993).

SVV The simian varicella virus homologue of HSV-1 Vmw175 has recently been sequenced and found to share 53% amino acid identity with VZV 140k but only 24% identity with HSV-1 Vmw175 (Clarke *et al.*, 1993). This is a reflection of the similar nature of the VZV and SVV diseases and also their genome structures (Gray *et al.*, 1992).

MDV Marek's disease virus induces T-cell lymphomas in susceptible chickens and is one of the more divergent alphaherpesviruses in both characteristics of the disease and DNA sequence, as is exemplified in the predicted sequence of the MDV ICP4 protein (Anderson *et al.*, 1992).

The **HSV-2** counterpart of HSV-1 Vmw175 has a sequence very similar to that of the HSV-1 protein, and this is not too surprising from the high degree of overall homology between the two genome sequences (B. Barnett, personal communication). In agreement, the HSV-2 Vmw175 protein is functionally interchangeable with HSV-1 Vmw175 during HSV-1 infection (Smith and Schaffer, 1987).

In summary, the sequence homologues of the major transcriptional regulatory protein of HSV-1, Vmw175, show strong conservation of the functionally important regions (regions 2 and 4) and where analysed they possess similar regulatory functions. The thoroughly characterised nature of the HSV-1 Vmw175 protein provides a basis for comparative studies with the other alphaherpesvirus counterparts, and interesting differences in the details of their functions have been identified. Particularly noteworthy are the analyses of the PRV IE180 protein, which have yielded important information regarding the probable mechanism of action of the whole family of proteins, and have complemented and extended the mechanistic studies of HSV-1 Vmw175. In addition, the information that can be gained from the sequences of these evolutionarily closely

related proteins complements the functional studies on the individual proteins. The sequences of the more divergent members, such as MDV ICP4 (see Fig. 1C.3), have special value for indicating conserved portions of sequences that may mediate functions common to all the proteins, as well as identifying differences that may account for the properties specific to a particular protein.

1D Eukaryotic RNA polymerase II transcription and its regulation

Our knowledge of the details of the transcription process has advanced extremely rapidly over the last decade and this Section will review only the initiation of transcription of inducible protein-coding genes by RNA polymerase II. Emphasis will be given to the human system, although the mechanism of transcriptional initiation is remarkably conserved in eukaryotes ranging from yeast to man. As viral transactivators appear to function through mechanisms analogous to those of cellular transcriptional activator proteins, the later Sections will discuss the mechanisms used by the host cell to regulate transcriptional initiation.

1D.1 Basal RNA polymerase II transcription

1D.1.1 *RNA polymerase II*

Eukaryotic cells contain three RNA polymerases which are large multi-subunit complexes, and each is responsible for transcribing a distinct set of genes: RNA polymerase I (pol I) transcribes ribosomal RNA genes; pol II is the RNA polymerase of protein-coding genes; pol III transcribes 5S rRNA and tRNA genes (for a comprehensive review, see Young, 1991). RNA polymerase II isolated from HeLa cells contains 10 subunits ranging from 10 to 240kD. The function and subunit structure of RNA polymerase II has been well conserved among the eukaryotes and by analogy to the essential nature of the yeast pol II subunits, all the HeLa pol II subunits are likely to be functionally important. In addition, the two largest subunits of eukaryotic RNA polymerases show clear sequence homology to the β and β' subunits of the single RNA polymerase of *E. coli* (Allison *et al.*, 1985). The structure of the yeast pol II enzyme, determined at 16 Å resolution, provides a structural basis for functional analyses of the transcription process (Darst *et al.*, 1991).

Of particular interest is the unusual C-terminal domain (CTD) of the largest subunit of RNA polymerase II which comprises multiple tandem repeats of the consensus heptapeptide YSPTSPS; the number of these repeats is important for viability in yeast (reviewed by Corden, 1990). The CTD can be extensively phosphorylated at the serine and threonine residues and pol II is consequently present in phosphorylated and non-phosphorylated forms within the cell; the catalytically active enzyme is

phosphorylated (Cadena and Dahmus, 1987). The evidence so far supports a model where the non-phosphorylated form of pol II preferentially associates with the promoter-bound preinitiation complex, and the subsequent phosphorylation of the CTD potentiates the transition from transcriptional initiation to elongation (Fig. 1D.1) (Payne *et al.*, 1989; Laybourn and Dahmus, 1989; 1990; Lu *et al.*, 1991). In agreement, a CTD-specific kinase is associated with the TFIIH component of the basal transcription machinery; this only phosphorylates the CTD when pol II is contained within a complete transcriptional preinitiation complex (Lu *et al.*, 1992). A complementary CTD-phosphatase activity has been identified in HeLa cell extracts (Chesnut *et al.*, 1992), which may play a role in the recycling of RNA pol II between rounds of transcription. Affinity chromatography has demonstrated that the non-phosphorylated CTD of pol II (but not the phosphorylated form) binds to the basal transcription factor TBP (see next Section) (Usheva *et al.*, 1992). In addition, the CTD has been reported to interact non-specifically with DNA *in vitro* (Suzuki, 1990); the possible functional implications of these interactions for the transition from transcriptional initiation to elongation are discussed further in Section 1D.1.3. Genetic analyses in yeast have indicated that the CTD of pol II also mediates the activity of selected transcriptional activators and repressors, and is therefore implicated in the regulation of transcriptional initiation (Allison and Ingles, 1989; Scafe *et al.*, 1990; Peterson *et al.*, 1991).

1D.1.2 *The basal transcription factors*

In contrast to the prokaryotic RNA polymerase, the functional activities of the eukaryotic enzymes require multiple accessory proteins. Seven auxiliary factors are involved in the accurate initiation of transcription from class II promoters by RNA polymerase II; these are commonly known as the basal or general transcription factors and together with pol II they constitute the basal transcription machinery. These basal transcription factors, TFIIA, B, D, E, F, H and J, were isolated by fractionation of HeLa cell nuclear extracts and the composition and role played by each factor is given in Table 1D (reviewed by Zawel and Reinberg, 1993). The cloning of the cDNAs encoding many of the HeLa basal factors has facilitated their functional analysis *in vitro*, and has enabled sequence comparisons with components of other transcription systems (Zawel and Reinberg, 1993). The basal factors are required for transcription from all class II promoters and are therefore favourable targets of cellular transcriptional regulatory proteins (see Section 1D.2.3). Furthermore, viral regulatory proteins, for example the HSV-1 proteins Vmw175 and Vmw65 (VP16), appear to influence the level of transcription of their target promoters through interactions with components of the basal transcription machinery (Sections 1B.3.3 and 1B.2.2). [Note that HSV-1 Vmw65 will be referred to as VP16 for the remainder of Section 1D, as this is the nomenclature appropriate to the transcription field].

Several recent studies using *in vitro* reconstituted transcription systems have managed to reproduce transcription with only a subset of the basal factors: TBP, TFIID and pol II were sufficient to transcribe the *IgH* promoter when the template was negatively supercoiled, although transcription from the relaxed template required the full set of factors (Parvin and Sharp, 1993). This result inferred that the free energy of supercoiling functionally replaces the requirement for the remaining basal factors, by promoting formation of an open complex for transcriptional initiation (Section 1D.1.3). Another study found that transcription of certain promoters required only *Drosophila* TBP, TFIIB, pol II and the RAP30 subunit of TFIIF (Tyree *et al.*, 1993). Transcription with these basal factor subsets is fairly inefficient (Drapkin *et al.*, 1993), but this provides a favourable basis for the functional analysis of the remaining factors upon their addition to the reconstituted minimal *in vitro* transcription system.

Factor	Polypeptide composition	Native Factor Mass	Function
TBP	38kD*	38kD	Initiates complex assembly by binding to the TATA box Interacts with the CTD of pol II and also many TAFs Contacts activators (i.e. VP16)
TFIIA	34kD* 19kD* 14kD	Unknown	Stabilises the TFIID-DNA interaction Removes repressors of basal transcription from the TFIID complex
TFIIB	33kD*	33kD	Contacts acidic activators (i.e. VP16) Binds to DA complex Accurately positions pol II
TFIIF	30kD* 74kD*	220kD	Commonly referred to as RAP30/RAP74 Recruits pol II to the promoter Affects efficiency of transcriptional elongation
TFIIE	34kD* 56kD*	200kD	Binds to DABpolF complex
TFIIH	90kD 62kD* 43kD 41kD 35kD	230kD	Binds to DABpolFE complex Associated with CTD kinase
TFIIJ	Unknown	Unknown	Binds to DABpolFEH complex

Table 1D The polypeptide composition and known functions of the HeLa basal transcription factors. The composition and native size are given for each of the RNA polymerase II basal transcription factors: TFIIA, B, F, E, H, J and the TBP component of TFIID (see Section 1D.1.4), * factors whose cDNA has been isolated.

1D.1.3 ***Ordered assembly of the transcription complex and transcriptional initiation***

With the exception of 'housekeeping' genes which have poorly defined promoter elements, class II promoters contain at least one of two minimal promoter elements: the TATA box (typically located 30 bp upstream of the mRNA start site) and the initiator element (Inr) (which encompasses the mRNA start site) (see Fig. 1D.1, panel A). The TATA box and / or the Inr can support basal transcription in the absence of other promoter elements. As all the alphaherpesvirus promoters so far analysed contain a TATA box homology, transcriptional initiation from promoters lacking a TATA box will not be considered here. For a review of initiation from TATA-less promoters, see Zawel and Reinberg (1993).

DNA binding analyses indicated that the assembly of the basal transcription machinery over the minimal promoter elements proceeded in a highly ordered manner to form a stable transcriptional initiation complex (Van Dyke *et al.*, 1988; Buratowski et *al.*, 1989). An updated version of the assembly scheme is described below and illustrated in Figure 1D.1 (reviewed by Conaway and Conaway, 1993). TFIID is the only basal transcription factor with sequence-specific DNA binding activity, and binding of TFIID to the TATA box constitutes the initial step in the formation of a transcription-competent complex (Fig. 1D.1, panel B) (Sawadogo and Roeder, 1985; Van Dyke *et al.*, 1988). Thereafter, the remaining factors enter the complex via protein-protein interactions. As discussed in the next Section, TFIID is a multi-subunit complex composed of the TATA-binding protein (TBP) and TBP-associated factors (TAFs). In the next step of the assembly process, TFIIA binds to the TBP component of TFIID and stabilises the TFIID-DNA interaction. TFIIA can remain associated or leave at any subsequent step in the pathway, and the sole role of TFIIA may be to displace a repressor of basal transcription that binds to TFIID (see Section 1D.2.4). Next, TFIIB enters the complex to yield the TFIID/TFIIA/TFIIB or DAB complex (panel C). TFIIB is responsible for the precise positioning of pol II at the promoter, a fact cleverly demonstrated by swapping components between two yeast *in vitro* transcription systems (Li *et al.*, 1994). The recruitment of TFIID to the promoter is known to be slow (Reinberg *et al.*, 1987), and affinity purification of preinitiation complexes at various stages of assembly has indicated that the binding of TFIIB is also a rate-limiting step (Lin and Green, 1991). Thus, these two steps are likely targets for enhancement by transcriptional activators (Section 1D.2.3). This has been shown to be the case with the PRV homologue of VZV 140k (Section 1C.3.1). In the next stage of the assembly of the preinitiation complex, Pol II is escorted to the DAB complex by TFIIF, to yield the DABpolF complex (panel D) (Flores *et al.*, 1991). Then follows the ordered association of TFIIE, TFIIH and TFIIJ to produce the complete preinitiation complex (panel E).

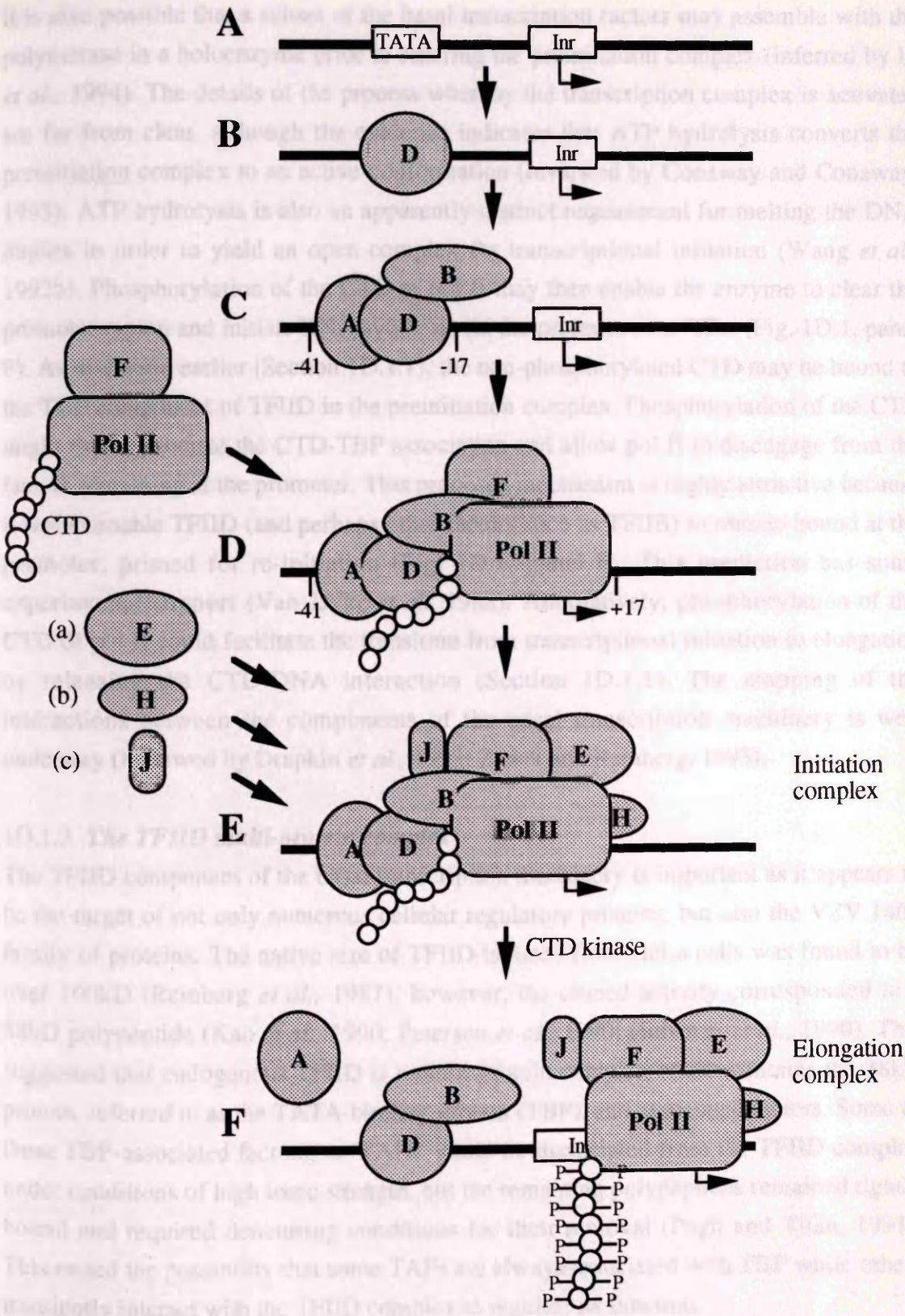


Figure 1D.1 RNA polymerase II mediated transcriptional initiation. This schematic diagram presents a model for assembly of the transcriptional initiation complex and its conversion to an elongation complex, see text for a description of the stages A to F. The basal transcription factors are indicated by letters A, B, D, E, F, H and J. Only eight of the 52 repeats found in the C-terminal domain (CTD) of mammalian RNA pol II are shown (Corden *et al.*, 1985). Nucleotide coordinates indicate the region of the promoter spanned by the complexes, as determined by footprinting analyses. The Figure is after Zawel and Reinberg (1992).

It is also possible that a subset of the basal transcription factors may assemble with the polymerase in a holoenzyme prior to entering the preinitiation complex (inferred by Li *et al.*, 1994). The details of the process whereby the transcription complex is activated are far from clear, although the evidence indicates that ATP hydrolysis converts the preinitiation complex to an active conformation (reviewed by Conaway and Conaway, 1993). ATP hydrolysis is also an apparently distinct requirement for melting the DNA duplex in order to yield an open complex for transcriptional initiation (Wang *et al.*, 1992b). Phosphorylation of the CTD of pol II may then enable the enzyme to clear the promoter region and initiate RNA synthesis (in the presence of rNTPs) (Fig. 1D.1, panel F). As discussed earlier (Section 1D.1.1), the non-phosphorylated CTD may be bound to the TBP component of TFIID in the preinitiation complex. Phosphorylation of the CTD might then dissociate the CTD-TBP association and allow pol II to disengage from the factors remaining at the promoter. This proposed mechanism is highly attractive because it would enable TFIID (and perhaps other factors such as TFIIB) to remain bound at the promoter, primed for re-initiation (Fig. 1D.1, panel F). This prediction has some experimental support (Van Dyke *et al.*, 1988). Alternatively, phosphorylation of the CTD of pol II could facilitate the transition from transcriptional initiation to elongation by releasing the CTD-DNA interaction (Section 1D.1.1). The mapping of the interactions between the components of the basal transcription machinery is well underway (reviewed by Drapkin *et al.*, 1993; Zawel and Reinberg, 1993).

1D.1.3 ***The TFIID multi-protein complex***

The TFIID component of the basal transcription machinery is important as it appears to be the target of not only numerous cellular regulatory proteins, but also the VZV 140k family of proteins. The native size of TFIID isolated from HeLa cells was found to be over 100kD (Reinberg *et al.*, 1987); however, the cloned activity corresponded to a 38kD polypeptide (Kao *et al.*, 1990; Peterson *et al.*, 1990; Hoffman *et al.*, 1990). This suggested that endogenous TFIID is a multi-protein complex which contains the 38kD protein, referred to as the TATA-binding protein (TBP), and associated factors. Some of these TBP-associated factors, or 'TAFs', could be dissociated from the TFIID complex under conditions of high ionic strength, but the remaining polypeptides remained tightly bound and required denaturing conditions for their removal (Pugh and Tjian, 1991). This raised the possibility that some TAFs are always associated with TBP while others transiently interact with the TFIID complex to regulate its function.

In vitro, both TBP and TFIID were found to be equally capable of mediating basal transcription. However, TFIID, but not TBP, could support transcription activated by a wide variety of sequence-specific activators including VP16, Sp1 and CTF. Based on these results it was postulated that native TFIID contains 'co-activators' which physically

link transactivators to components of the basal transcription machinery (Pugh and Tjian, 1990; Peterson *et al.*, 1990; Dynlacht *et al.*, 1991; White *et al.*, 1991; Zhou *et al.*, 1992). In agreement, at least some of the purified *Drosophila* TAFs function as co-activators *in vitro* (Dynlacht *et al.*, 1991). The co-activator function of these TAFs also exhibited functional specificity, as those that mediated activation by proline-rich activators were unable to mediate activation by acidic activators (Tanese *et al.*, 1991) (see Section 1D.2.3).

Affinity purification of TFIID from HeLa cells has indicated the presence of numerous TAFs, and much recent effort has concentrated on their characterisation (Zhou *et al.*, 1992; 1993, and references therein) (see Fig. 1D.2). The largest human TAF, $TAF_{II}250$, transpired to be identical to the cell cycle regulator CCG1; this indicated a link between cell cycle regulation and the basal transcription machinery (Ruppert *et al.*, 1993). A temperature sensitive version of $TAF_{II}250$ provided the first direct evidence that TAFs mediate transcriptional activation *in vivo* (Wang and Tjian, 1994). Furthermore, the effects of this mutation on transcriptional activation were gene specific, supporting the functional specificity of the TAFs observed *in vitro*.

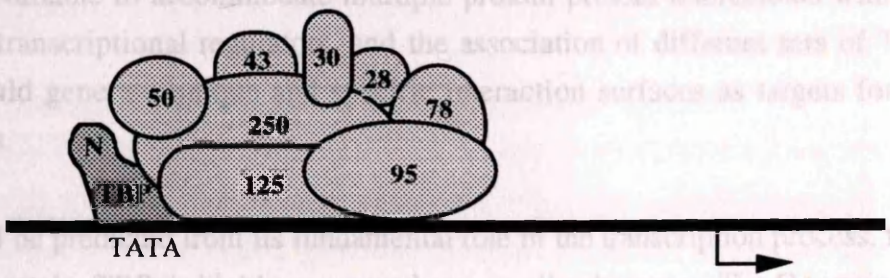


Figure 1D.2 The Human TFIID complex. Diagrammatic representation of the TBP and TAF components of HeLa TFIID. Only those TAFs that have clearly been established to be a part of the HeLa TFIID complex are shown (Zhou *et al.*, 1992, 1993). The 700kD holoTFIID shown here includes TAFs ranging from 28-250kD, and the TAFs are named according to their apparent size on SDS-PAGE (kD). The positions on the DNA and relative positions of the TAFs and TBP are shown on the basis of the available information from footprinting and far-western analyses (Zhou *et al.*, 1993, and references therein). TBP is drawn schematically to indicate the lack of interaction of the non-conserved N-terminal portion with either DNA or TAFs.

Important advances have also been made in the analysis of the *Drosophila* TAFs, including the identification of the distinct TAF components of TFIID that interact with the activation domains of the transcriptional regulators Sp1 and VP16 (Hoey *et al.*, 1993; Goodrich *et al.*, 1993; Gill *et al.*, 1994) (Fig. 1B.2). These analyses are relevant to the human situation because of the high degree of sequence identity between at least some *Drosophila* and human TAFs. Within the TFIID complex, some of the interactions between TAFs are thought to hold the subunits together, while others may be subject to dynamic change induced by regulatory factors. Indeed, an inhibitory domain in one of

the *Drosophila* TAFs controls the promoter binding activity of TFIID and may be regulated by other TAFs or regulatory factors (Kobubo *et al.*, 1994). Two of the seven cloned *Drosophila* TAFs showed significant sequence homology to histones H3 and H4 and this raises intriguing functional implications; these particular TAFs may be involved in wrapping DNA into a nucleosome-like structure or displacing the H3/H4 components of chromatin to allow transcriptional initiation (Kobubo *et al.*, 1994).

TBP has recently been found to be a requirement for transcription from promoters which lack a TATA box and also for transcription from class I and III promoters, processes which presumably require TBP to be associated in distinct multi-protein complexes (reviewed by Sharp, 1992; White and Jackson, 1992; Rigby, 1993). Indeed, multiple chromatographically separable TBP-containing complexes with different functional properties exist in HeLa cells (Timmers and Sharp, 1991; Timmers *et al.*, 1992; Brou *et al.*, 1993). Some of the TAFs were common to all these TFIID complexes, while others were unique (Brou *et al.*, 1993). Therefore, it appears that the factors associated with TBP are responsible for modulating the response to transcriptional activators and also for determining promoter specificity. The TAF components of TBP increase the protein surface available to accommodate multiple protein-protein interactions with a diverse array of transcriptional regulators, and the association of different sets of TAFs with TBP would generate unique and specific interaction surfaces as targets for different activators.

As would be predicted from its fundamental role in the transcription process, the TATA binding protein, TBP, is highly conserved among all eukaryotes. The C-terminal domain of TBP is particularly well conserved, whereas the N-terminal domain is variable in length and sequence. The C-terminal 180 amino acid domain is sufficient to fulfil the majority of the functions and interactions mediated by TBP (for reviews, see Pugh and Tjian, 1992; Hernandez, 1993). In addition to its interactions with DNA, the TAFs, TFIIB and TFIIA, the TATA binding protein (TBP) also associates with transcriptional regulatory factors, co-activators and repressors of basal transcription (see Sections 1D.2.3 and 1D.2.4). It seems highly surprising that a protein region the size of the C-terminal domain of TBP can participate in so many interactions, even given the fact that many are likely to be dynamic. A possible explanation has been provided by the recent 3-D structures (both in the presence and absence of DNA) of the plant *Arabidopsis thaliana* TBP and the conserved C-terminal domain of yeast TBP (Nikolov *et al.*, 1992; Kim *et al.*, 1993a; Chasman *et al.*, 1993; Kim *et al.*, 1993b). These show that the C-terminal domain of TBP adopts a novel protein fold for DNA binding, which gives the protein a saddle-like shape straddling the DNA and presents a large convex surface available for interactions with other proteins (Fig. 1D.3) (see Section 1E.5.2).

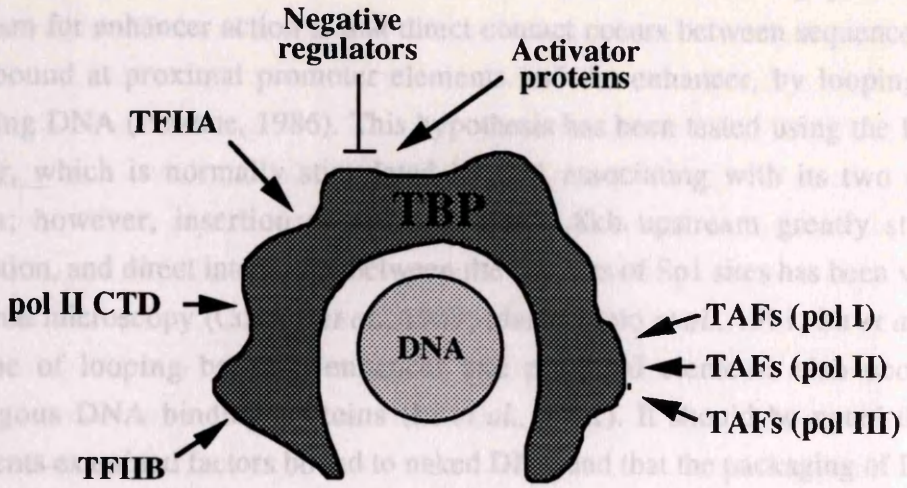


Figure 1D.3 Transcription factors that interact with the convex outer surface of TBP. The crystal structures of the C-terminal domain of TBP demonstrate that it forms a molecular saddle over the DNA, as shown here in cross-section. Many different types of factors have been shown to bind TBP as discussed in the text (also see Struhl, 1994) and the binding site for TFIIA on TBP's surface has been well-defined (Lee *et al.*, 1992).

1D.2 Regulation of RNA polymerase II transcription

1D.2.1 *Promoter elements involved in transcriptional regulation*

The activity of the basal transcription machinery is modulated by sequence-specific regulatory factors that bind to short consensus sequences usually located within 100bp of the minimal promoter region (proximal promoter elements), and to enhancer elements which are often several kbp further upstream (reviewed by Saltzman and Weinmann, 1989). Where present, enhancers characteristically function independently of distance or orientation and comprise multiple elements for different transcriptional activators; some recognition elements are shared between enhancer and proximal promoter sequences.

Cis-acting promoter elements function either separately or in various combinations, to confer responses to a wide range of signals by influencing the activity of the basal transcription machinery (see Section 1D.2.3). Some of the regulatory factors bound by these elements are ubiquitous and contribute to the constitutive activity of the promoter (for example Sp1 and CBP), while other elements bind factors which are tissue specific or are regulated by specific stimuli (for example heat shock, or steroids). The particular complement and arrangement of the *cis*-acting elements within a promoter defines its specific pattern of gene expression (reviewed by Saltzman and Weinmann, 1989). How enhancers work is not known, although they clearly promote the formation of stable

transcription complexes (Mattaj *et al.*, 1985; Weintraub, 1988). A popular proposed mechanism for enhancer action is that direct contact occurs between sequence-specific factors bound at proximal promoter elements and the enhancer, by looping-out the intervening DNA (Ptashne, 1986). This hypothesis has been tested using the HSV-1 *tk* promoter, which is normally stimulated by Sp1 associating with its two proximal elements; however, insertion of six Sp1 sites 1.8kb upstream greatly stimulated transcription, and direct interaction between the two sets of Sp1 sites has been visualised by electron microscopy (Courey *et al.*, 1989; Mastrangelo *et al.*, 1991; Su *et al.*, 1991). This type of looping between enhancer and proximal elements also occurs with heterologous DNA binding proteins (Li *et al.*, 1991). It should be noted that these experiments examined factors bound to naked DNA and that the packaging of DNA into nucleosomes *in vivo* may help bring the proximal and enhancer elements into proximity, thus allowing contacts between regulatory factors to be made. Such interactions might be either direct or indirect through non-DNA binding mediator proteins.

1D.2.2 *Domain structure of transcriptional activators*

An important discovery for the study of the regulation of gene expression was that transcriptional activator proteins are modular in structure, having a DNA binding domain that confers specificity and a separate and functionally independent activation domain(s) that enhances the activity of the basal transcription machinery (reviewed by Ptashne, 1988; Frankel and Kim, 1991). As such, these domains can be interchanged using molecular genetics to create chimeric proteins with novel activities. It is interesting that certain viral activator proteins rely on host cell proteins to supply the DNA binding domain in *trans*, as is the case with VP16 (Section 1B.2.2) and adenovirus E1a (Liu and Green, 1990). Transcriptional activation regions have been roughly classified on the basis of their primary sequence: acidic (found in proteins such as VP16 and yeast GCN4, GAL4), proline-rich (i.e. CTF), and glutamine-rich (i.e. Sp1) (reviewed by Mitchell and Tjian, 1989). In contrast to the wealth of knowledge concerning the different types of DNA binding domains (Section 1E), the structure and details of the molecular interactions of the transcriptional activation regions remain more of a mystery.

Acidic activation regions have been most extensively analysed. They appear to have no sequences or structures in common except for a net negative charge; the sequence requirements of specific acidic activation regions have been investigated by mutational analyses (reviewed by Hahn, 1993b). One study found that the net negative charge of the VP16 activation region contributed to, but was not sufficient for transcriptional activation. Instead, it appeared that bulky hydrophobic residues were more important, and one particular phenylalanine residue was critical, for the VP16 activation function

(Cress and Triezenberg, 1991). No essential residues were identified in the GAL4 acidic activation region (Gill and Ptashne, 1987) and even elimination of all of the negative charges caused only a modest reduction in GAL4 activator function (Leuther *et al.*, 1993). Contrary to earlier belief, acidic activation regions fail to adopt an α-helical structure, as indicated by circular dichroism spectroscopy (Van Hoy *et al.*, 1993). The fact that the GCN4 acidic activation domain could be extensively deleted without reducing its activity *in vivo* strongly suggested that tertiary structure is unimportant for the activation function (Hope *et al.*, 1988). Therefore, the interaction of acidic activators with their target(s) is unlikely to depend upon precise rigid, structural complementarity and this is consistent with the observation that certain acidic activation domains can function in all the eukaryotic systems tested from yeast to humans regardless of their origin. Possible mechanisms whereby the activation domains stimulate transcription will be considered below.

1D.2.3 *Mechanisms of transcriptional activation*

All transcriptional activation domains presumably stimulate the activity of the basal transcription machinery through either direct or indirect interactions. The majority of studies on the mechanism of transcriptional activation have employed the acidic activation region of HSV-1 VP16, due to its potency (Sadowski *et al.*, 1988). As discussed in Section 1B.2.2, the VP16 activation region forms direct and specific contacts with the basal transcription factor TFIIB and a TAF component of the TFIID complex *in vitro*; VP16-TBP interactions have also been reported. The TFIIB-VP16 interaction has been shown to be vital for VP16-mediated transcriptional activation *in vitro* (Roberts *et al.*, 1993). In support of the relevance of the TFIID-VP16 interaction, the transcriptional activity of mutant VP16 activation domain derivatives correlated with their affinity for TFIID, as measured by *in vitro* binding assays (Ingles *et al.*, 1991). The basal factors TFIIB and TFIID have also been found to be the targets of other activation regions. A synthetic acidic activation region increased the efficiency of TFIIB binding to the TFIID-promoter complex *in vitro* (Lin and Green, 1991) and viral activators bearing non-acidic activation regions, such as E1a and EBV Zta, have also been found to contact TFIID and apparently stabilise its association with DNA (Lee *et al.*, 1991; Lieberman and Berk, 1991).

An obvious interpretation of these data is that these transcriptional activation regions recruit TFIID and / or TFIIB to the promoter during preinitiation complex formation, thus stimulating transcription by overcoming a rate-limiting step. However, there are several other equally plausible explanations. It has been suggested that the basal transcription factors frequently assemble into non-productive complexes in the absence of an activator; this was based on the observation that a very small fraction of promoters

act as transcription templates during *in vitro* transcription experiments, even though footprinting analyses detect complexes on the majority of templates (Herschlag and Johnson, 1993). Transcriptional activators may favour assembly of productive transcription complexes, and there is evidence to support this (Wang *et al.*, 1992a). The binding of an activator to TFIID and / or TFIIB may block the assembly pathway for non-productive complexes, perhaps by inducing conformational changes in the basal factors. Alternatively, non-productive complexes may be a consequence of inhibitory factors binding to the basal transcription factors. In this scenario, the activator might compete with the repressor for binding sites on the basal transcription factor. In support of this mechanism, several activators have been observed to stimulate transcription by displacing repressors of basal transcription from TBP (see next Section). It is possible that similar types of transcriptional repressors also bind to other components of the basal transcription machinery, such as TFIIB. Therefore, by preventing negative factor binding by direct competition, activators would stimulate the rate of transcriptional initiation.

In addition to their apparent importance for the function of acidic activation domains, the TAF co-activator components of TFIID are also an essential requirement for activation mediated by the activators Sp1 and CTF *in vitro* (Section 1D.1.4). This suggests that the mechanisms of transcriptional activation by the glutamine- and proline-rich activation domains may also involve contacts with the TAFs during the initiation process *in vivo*. Activators bearing different classes of activation domains presumably contact distinct targets, and the fact that the activation domains of Sp1 and VP16 specifically interact with different TAFs *in vitro* supports this (Section 1D.1.3). That TBP can be associated with distinct factors or sets of factors which specifically mediate the activity of one or several different activation domains (Brou *et al.*, 1993), may explain the results of 'squelching' experiments which suggest that distinct transcriptional co-activators can be titrated by acidic and non-acidic activation domains (Martin *et al.*, 1990a; Tasset *et al.*, 1990). These squelching experiments test the ability of activators to interfere with each other by competing for the same targets.

Experiments in yeast also indicated that basal factors are not sufficient to promote transcriptional stimulation by the acidic activation domains of VP16 and GCN4, and that auxiliary factors, termed mediators or adaptors, are required (Berger *et al.*, 1990; Kelleher *et al.*, 1990; Flanagan *et al.*, 1991b). These mediators may represent a class of transcription factors distinct from the TAFs of higher eukaryotes, as yeast TBP is not part of a stable multi-subunit complex in the same way as the TFIID of higher eukaryotes (Schmidt *et al.*, 1989). These yeast mediators may either connect activator proteins to the basal transcription machinery in a manner analogous to the TAFs, or they

may facilitate the interaction of activators with chromatin proteins to bring about antirepression of chromatin-bound templates (see Section 1D.2.4). Furthermore, the results of squelching experiments in yeast have suggested that not all acidic activators target the same mediator, indicating that there are several different classes of acidic activators that function through distinct pathways (Berger *et al.*, 1992).

Another interesting observation is that transcriptional activator proteins bound to the same promoter can stimulate gene expression cooperatively or synergistically, with activator proteins bearing the same class or different classes of activation domains. This enhanced activity may be through their stimulation of different rate-limiting steps in the transcriptional initiation process or through multiple favourable contacts with the basal transcription machinery. The suggestion that certain activators contact several basal factors (as seen with VP16) may explain the greatly enhanced transcriptional stimulation which results from certain combinations of activators.

1D.2.4 *Transcriptional antirepression*

Many of the functional studies of activation domains described in the previous Section used reconstituted *in vitro* transcription systems in which the magnitude of activator-mediated transcription observed did not approach the physiological levels of activation. It was apparent that additional factors repressed the basal level of gene expression *in vivo*, but were absent from these purified *in vitro* systems. The presence of general repression mechanisms within the cell would act to lower the resting level of gene expression, and therefore enable a much greater response to transcriptional activators. Thus activators could function by counteracting transcriptional repression ('antirepression') as well as stimulating the inherent basal transcription process ('true activation'). Two distinct types of generalised repression mechanisms have been identified that have wide ranging implications for transcriptional regulation, as described below.

A large body of data from genetic and biochemical experiments indicates that chromatin structure plays an important role in the regulation of gene expression, and *in vitro* systems have demonstrated that pre-assembly of nucleosomes and / or histone H1 onto DNA templates drastically inhibits transcriptional initiation, presumably by sterically blocking the assembly of the transcription complexes (reviewed by Workman and Buchman, 1993; Croston and Kadonaga, 1993). Thus, activators must function at least in part to counteract chromatin-mediated repression. This was observed *in vitro* when Sp1 or GAL4-VP16 was bound to a naked DNA template, prior to chromatin assembly and transcriptional analysis (Croston *et al.*, 1992; Kamakaka *et al.*, 1993). However, these experiments would only reflect the situation immediately following DNA

replication *in vivo* (i.e. when the DNA may be transiently naked), but in the majority of cases the rapid induction of gene expression observed *in vivo* would require the activator to bind and transactivate a promoter that is already assembled into chromatin. Therefore, a major question in the regulation of gene expression is how do the transcription factors gain access to the DNA bound in nucleosomes? Chromatin disruption is known to accompany transcriptional activation *in vivo* (reviewed by Croston and Kadonaga, 1993); significantly these alterations include the depletion of histone H1, an observation which supports the identification of H1 as a general repressor *in vitro* (Croston *et al.*, 1991). Activators may be responsible for directing this reconfiguration of the chromatin to allow the basal factors access to the promoter and to assemble into the transcriptional initiation complex. Recent *in vitro* experiments with reconstituted chromatin assembly-transcription systems have demonstrated that this is the case, but additional proteins are required to facilitate this activator-mediated chromatin reconfiguration and antirepression process (Tsukiyama *et al.*, 1994; J. T. Kadonaga, personal communication).

The other general repression mechanism that has been identified involves the binding of small proteins to the basal transcription factor TFIID; these include the HeLa cell proteins Dr1 and Dr2, and the yeast proteins NC1, NC2 and AD1 (Meisterernst and Roeder, 1991; Meisterernst *et al.*, 1991; Inostroza *et al.*, 1992; Zawel and Reinberg, 1993; Auble and Hahn, 1993). The Dr1 repressor functions in two ways: (i) it binds to TFIID and precludes further assembly of the preinitiation complex (i.e. the binding of TFIIA and TFIIB) and (ii) it disrupts preinitiation complexes such as DABpolF (see Fig. 1D.1). In addition, the Dr1-TFIID complexes can be disrupted by the viral transactivators E1a and SV40 (Zawel and Reinberg, 1993). Another inhibitory factor NC1 (that may be a counterpart of Dr2), binds to TFIID and down-regulates the basal promoter activity (Meisterernst *et al.*, 1991). Both NC1 and Dr2 are displaced when excess TFIIA is added, or by the presence of an acidic activator (Zawel and Reinberg, 1993). The yeast repressor AD1 can also be displaced by TFIIA, but in contrast to the other inhibitory factors, AD1 binding to TFIID dissociates the TFIID-DNA interaction (Auble and Hahn, 1993).

In addition to these general mechanisms that are used to repress basal transcription, other repressors affect transcription by recognising specific DNA sequences in selected promoters and then reducing the activity of the basal transcription machinery or blocking the stimulatory effects of activators (reviewed by Levine and Manley, 1989; Renkawitz, 1990). Passive repressors compete with activators or basal transcription factors for occupancy of promoter elements; for example the autoregulation function of SV40 large T antigen is due to its binding to a site overlapping the minimal elements of

its own promoter (Tjian, 1981). Active repressors either disrupt the function of specific activators, or the basal transcription machinery, through protein-protein interactions. For example, *Drosophila engrailed* represses transcription from its proximal promoter elements via a repression domain that is functionally transferable to other DNA binding domains, in a manner analogous to activation domains (Jaynes and O'Farrell, 1991). Because some transcriptional repression domains, for example the *Drosophila* even-skipped domain (Han and Manley, 1993), can function on minimal promoters comprising only a TATA box, it is likely that at least some repression domains modify the activities of the basal machinery directly, or indirectly through co-repressors (as has been suggested by squelching experiments). Interestingly, many repressors can also act as activators in different situations, and they have distinct repression and activation domains that are presumably subject to differential regulation. Additionally, differential splicing can produce related activators and repressors from the same gene (reviewed by Foulkes and Sassone-Corsi, 1992).

In summary, the initiation of transcription is a complex multi-step process, and as such provides numerous stages for the regulation of transcription. Repression and activation mechanisms both play crucial roles in determining the activity of the basal transcription machinery. The level of transcription of a eukaryotic gene at any one time is likely to be determined by the concerted actions of several activators and repressors bound to multiple promoter elements. The TFIID complexes present large and varied surfaces for specific protein-protein interactions with these regulatory factors to allow specific effects on the activity of the basal machinery.

1E DNA binding domains of sequence-specific transcription factors

1E.1 Overview of DNA binding strategies

Sequence-specific DNA binding proteins possess a discrete DNA binding domain that is folded in a way that presents a protruding surface, a flexibly extended structure, or both in order to contact the DNA bases of the target binding site. These domains are usually stabilised by a hydrophobic core and are usually physically separable from the rest of the protein. Many DNA binding proteins bind to DNA as symmetrical dimers and their DNA binding domains often include a specific type of dimerisation motif closely associated with the DNA binding motif. On the basis of sequence conservation of these motifs, DNA binding domains are classed into structural families that employ distinct modes of DNA recognition. However, there are examples of DNA binding domains that show very little sequence homology yet adopt similar structures, as seen with the variant HTH motifs (see Section 1E.2.3).

The DNA recognition element is often an α-helix which fits snugly into the major groove, although DNA binding strategies for several proteins using β-sheets for DNA recognition have also been elucidated (Section 1E.5). The different classes of DNA binding domains characteristically contact bases in either the major or minor, or both grooves of DNA. In addition, multiple hydrogen bonds with the DNA phosphate backbone are critical for correctly positioning the recognition element, and also for fixing the DNA in a specific conformation to allow the optimal protein-DNA interaction (discussed further by Luisi, 1994). The salient features of each class of DNA binding domain, as elucidated by nuclear magnetic resonance and X-ray crystallographic analyses, will be discussed in the following Sections (for reviews, see Harrison, 1991; Pabo and Sauer, 1992; Burley, 1994).

1E.2 The helix-turn-helix (HTH) DNA binding motif

1E.2.1 *The prokaryotic HTH motif*

The helix-turn-helix (HTH) was the first DNA recognition motif to be characterised structurally and it has been found in the majority of analysed bacterial and bacteriophage transcriptional regulatory proteins (for reviews, see Brennan and Matthews, 1989; Harrison and Aggarwal, 1990). The HTH motif is defined as a 20 residue segment comprising two α-helices separated by a four residue turn; the helices lie at 120° to each other. The second amphipathic helix is the DNA recognition element that lies in the major groove of DNA and contacts the bases. Additional residues outside the HTH motif are often important for DNA recognition, as is the case with the λ repressor-DNA complex where a flexible N-terminal arm makes critical contacts in the major groove (Jordan and Pabo, 1990). The HTH motif cannot fold or function by itself, but is found in diverse structural environments as part of a larger DNA binding domain. The prokaryotic HTH proteins bind to DNA as dimers and each monomer recognises a symmetrical half-site.

1E.2.2 *Homeodomains*

As discussed in Section 3E.1, the homeodomain is a conserved 60 residue DNA binding domain, found in a large number of eukaryotic transcription factors that play important roles in development (reviewed by Kornberg, 1993; Wright, 1994). The homeodomain is structurally related to the HTH motif and consists of three α-helices and an N-terminal arm; however, the prokaryotic and homeodomain HTH units dock against the DNA in significantly different ways (Kissinger *et al.*, 1990). The DNA co-crystal structures for homeodomains from yeast MATα*2* and *Drosophila Antennapedia* and *engrailed*, demonstrate that the homeodomain binds to DNA as a monomer by inserting the long third α-helix into the major groove of the DNA and its N-terminal arm into the adjacent minor groove (Otting *et al.*, 1990; Kissinger *et al.*, 1990; Wolberger *et al.*, 1991).

1E.2.3 *POU domains and variant eukaryotic HTH motifs*

Certain subfamilies of homeodomain proteins have an additional DNA binding motif, flanking the homeodomain, that is required for high affinity and sequence-specific DNA binding (Scott *et al.*, 1989; Laughon, 1991). A well characterised example of this is provided by the divergent homeodomain of the POU proteins, referred to as the POU-specific domain (see Section 3E.7) (reviewed by Verrijzer and van der Vliet, 1993; Wright, 1994). The POU-specific domain lies on the N-terminal side of the homeodomain and was first identified in the Pit-1, Oct-1, Oct-2, and Unc-86 proteins

(hence the name POU) (Herr *et al.*, 1988). The structures of the Oct-1 POU-specific domain show it to comprise four α-helices arranged in a manner surprisingly similar to the HTH-containing DNA binding domains of the bacteriophage λ and 434 repressors (Assa-Munt *et al.*, 1993; Dekker *et al.*, 1993). Recent additional structures of variant HTH motifs have been reported: the third tandem repeat of the proto-oncogene c-Myb, the HNF-3γ */fork head* DNA binding motif, the globular domain of histone H5, rat LFB1/HNF1 and the yeast heat shock transcription factor (HSF), and all show extended versions of the HTH motif with the inserted amino acids accommodated in the turn region (Ogata *et al.*, 1992; Clark *et al.*, 1993; Ramakrishnan *et al.*, 1993; Ceska *et al.*, 1993; Harrison *et al.*, 1994). The DNA binding domain of HSF includes a four-stranded antiparallel β-sheet capping the HTH motif, while the DNA binding domains of HNF-3γ and GH5 include loops flanking the HTH which make additional contacts with the minor groove (for a review of this 'winged-helix' motif see Brennan, 1993). The LFB1/HNF1 protein is unusual for this type of domain, because it binds to DNA as a dimer.

1E.3 Basic-leucine zipper (bZIP) and basic-helix-loop-helix (bHLH) DNA binding domains

The bZIP and bHLH DNA binding domains are found in proteins with important roles in differentiation and development, and have been grouped together here because they share many similar features (for reviews, see Baxevanis and Vinson, 1993; Ellenberger, 1994). Transcription factors within each family can associate as homodimers or heterodimers with distinct DNA binding and transcriptional properties. The bZIP and bHLH domains both comprise a C-terminal amphipathic α–helix which mediates dimerisation (the HLH or ZIP region, see below), and a basic N-terminal DNA binding region. The basic regions of the bHLH and bZIP proteins are unstructured in solution but become α-helical upon insertion into the major grove for recognition of the DNA target sequence (see below) (Weiss *et al.*, 1990; Anthony-Cahill *et al.*, 1992). The leucine zipper and helix-loop-helix dimerisation segments contribute to DNA binding specificity by appropriately positioning the basic region helices and determining which subunits form stable dimers.

1E.3.1 *The bZIP domain*

The bZIP domain usually comprises 60-80 residues and is found in a wide variety of transcription factors (Hope and Struhl, 1987; Landschulz *et al.*, 1988). The DNA co-crystal structures of the bZIP domain of yeast GCN4 (Ellenberger *et al.*, 1992; Konig and Richmond, 1993) show that the bZIP dimer is composed of two smoothly curving

α-helices, associated in a coiled coil at their C-terminal ends (the leucine zipper). The helices diverge at their N-terminal ends and extend through the major groove of DNA where the basic regions contact the DNA bases and phosphate backbone. Dimerisation is mediated by the leucine zipper: an apolar residue (usually leucine) is found at every 4th position of each heptad repeat within the C-terminal 30-40 residues of the amphipathic helix, and the hydrophobic faces of a helix from each monomer pack together in a two-stranded parallel coiled coil with the leucines interdigitating (O'Shea *et al.*, 1991; reviewed by Alber, 1992).

1E.3.2 *The bHLH domain*

Like the leucine zipper proteins, the activity of the bHLH proteins (Murre *et al.*, 1989) is modulated by heterodimer formation. The bHLH motif bears resemblance to the bZIP motif: the structures of homodimer-DNA complexes for the oncoprotein Max and the transcription factor USF demonstrate that the HLH is a parallel four-helix bundle with an extended strand (the loop) connecting helices 1 and 2; a pair of protruding basic region helices contact the DNA in a similar manner to the basic region helices of the bZIP motif (Ferre-D'Amare *et al.*, 1993; 1994). Dimerisation is mediated by coiled coil interactions between the amphipathic C-terminal helix of each HLH monomer. Some bHLH proteins, including Max, have an additional zipper segment C-terminal to the HLH that adds extra stability and specificity to the dimer interaction.

1E.4 Zinc-binding domains

Zinc-binding domains are found in proteins involved in very diverse aspects of eukaryotic gene regulation and contain zinc as a structural element; they are classified according to the mode of zinc coordination, secondary structure elements, and modularity (reviewed by Kaptein, 1992; Schmiedeskamp and Klevit, 1994).

1E.4.1 *Cys_2His_2 zinc fingers*

The original 'zinc finger' is an approximately 30-residue module with one zinc ion coordinated by two cysteines and two histidines, TFIIIA being the classical example. (Miller *et al.*, 1985). The modules are usually present in at least three copies and bind adjacent triplets of bases in the major groove of DNA. Numerous structural studies have shown that each module is a peptide loop containing a two strand β-hairpin and a single α-helix (αββ motif); this fold is stabilised by the coordination of the zinc ion via the side-chains of two cysteine and histidine amino acids.

1E.4.2 *Steroid / nuclear receptor DNA binding domains*

Typified by the glucocorticoid and oestrogen receptors, these DNA binding domains contain about 70 residues and have two loop-helix modules, each with four cysteines coordinating a zinc ion. The two Cys_4 modules coalesce to form a single globular domain with the two α-helices packed against each other at nearly right angles (reviewed by Freedman and Luisi, 1993). The α-helix of the N-terminal module contacts bases of the target site within the major groove of DNA, whereas the helix of the C-terminal module presents a dimerisation interface for coiled coil interactions. Proteins containing this type of domain bind to palindromic recognition elements as dimers (Mader *et al.*, 1989).

1E.4.3 *Others zinc-binding domains*

Another class of zinc motif, containing two zinc ions coordinated by six cysteine residues, has been found only in yeasts. The X-ray structure of the GAL4 DNA binding domain-DNA complex revealed the globular nature of this α–helical domain, in which two of the cysteines are shared between the zincs (Marmorstein *et al.*, 1992). GAL4 binds to DNA as a dimer, with each monomer inserting its metal-binding domain into the major groove on opposite faces of the DNA; dimerisation is mediated by formation of a coiled coil between α-helical regions of each domain. Other recent structures demonstrate the striking diversity of zinc-containing DNA binding domains: the TFIIS elongation factor contains a novel Cys_4 domain comprising a compact three-stranded antiparallel β-sheet, and can bind polynucleic acids in a sequence independent manner (Qian *et al.*, 1993). The GAGA-1-DNA complex illustrated a novel Cys_4 α/β motif which makes both major and minor groove base contacts (Omichinski *et al.*, 1993).

1E.5 DNA binding domains with β-sheet recognition elements or scaffolds

1E.5.1 *Prokaryotic ribbon-helix-helix proteins*

The *met* repressor of *E. coli* and the *arc* repressor of phage P22 are members of a family of proteins that use a β-ribbon recognition motif to probe the major groove of DNA (Breg *et al.*, 1990; Somers and Phillips, 1992; reviewed by Raumann *et al.*, 1994). Dimers of these proteins have a core with four α-helices, and a single antiparallel β-ribbon (formed from a segment of each monomer) protrudes from the core and lies in the major groove of DNA. Another similar family of proteins with β-sheet recognition motifs includes the prokaryotic HU DNA-packaging protein and the *E. coli* integration host factor (IHF). The HU dimer also has a core of four α-helices, but it has a separate β-ribbon projecting from each monomer (Tanaka *et al.*, 1984); at least for the IHF

protein, the two β-ribbons are believed to contact the minor groove (Yang and Nash, 1989).

1E.5.2 *Novel eukaryotic DNA binding motifs*

The **TATA binding protein** (TBP) plays a fundamental role in all eukaryotic transcription (see Section 1D.1.4) and its structure demonstrates a novel mode of DNA recognition (reviewed by Klug, 1993; Struhl, 1994). The conserved C-terminal domain of TBP folds into a symmetrical ten-stranded, antiparallel β-sheet, which provides a large concave surface (referred to as the 'molecular saddle') to follow the minor groove of DNA. Strikingly, the DNA is kinked in two places under the 'saddle' of TBP to give a severe DNA bend (about 100° overall) towards the major groove; this results in an opening and shallowing of the minor groove which complements the concave surface of TBP (Kim *et al.*, 1993a; b).

The X-ray structure of the dimeric DNA binding domain of the **papilloma virus E2 transactivator** has revealed an unusual protein scaffold comprising a barrel-like structure formed by eight antiparallel β-strands (the β-barrel) (Hedge *et al.*, 1992). A long α-helix connects β-strands 1 and 2 of each half-barrel and serves as the DNA recognition helix; the DNA bends smoothly around the β-barrel allowing the α-helix of each monomer to penetrate the major groove and interact with the bases of its target site.

The HMG box occurs in a variety of chromosomal proteins and transcription factors, several of which are known to interact with the minor groove of DNA. The three α-helices of the DNA binding domain of HMG-1 adopt a novel 'arrow-head' structure (Weir *et al.*, 1993).

In summary, several common themes have emerged from the structures of DNA binding domains and the strategies used for DNA recognition, yet there is also significant diversity within each family of DNA binding domains. No doubt more novel DNA binding domain structures will be determined in the future which may explain the mode of binding of domains, such as the VZV 140k DNA binding domain, that do not obviously conform to the established families.

1F The aims of the research presented in this Thesis

The initial aims of this study were to investigate the DNA binding properties of a bacterially expressed fragment of the VZV 140k ORF encompassing its DNA binding domain and to define its limits. Purification of the domain would allow the analysis of its biochemical properties. By analogy to the situation with the HSV-1 Vmw175 DNA binding domain, it seemed likely that the VZV 140k domain would be dimeric. Therefore, truncation and insertion mutant DNA binding domains (expressed in *E. coli* or *in vitro*) could be constructed and assayed by a wide range of techniques with the aim of identifying sequences mediating dimerisation. Combined with DNA binding analyses of the mutant domains, these experiments would aim to define the sequences that were responsible for DNA binding and dimerisation interactions.

The subsequent intention was to compare the characteristics of the VZV 140k DNA binding domain with those of its HSV-1 Vmw175 counterpart, in order to investigate the basis for the differences in the regulatory properties of the 140k and Vmw175 proteins. Construction of a hybrid protein with the VZV 140k DNA binding domain in the background of the HSV-1 Vmw175 protein, and characterisation of its regulatory properties in transient transfection assays, would enable the functional relevance of possible differences between the DNA binding properties of the 140k and Vmw175 DNA binding domains to be assessed. Recombining the hybrid ORF back into the HSV-1 genome would indicate whether the DNA binding domains of 140k and HSV-1 Vmw175 were functionally interchangeable in the viral context.

Chapter 2: Materials and methods

2A Materials

2A.1 Enzymes

Restriction enzymes, proteinase K and calf intestinal phosphatase were supplied by Boehringer Mannheim. RNase A, bovine pancreas DNase I, lysozyme and thrombin were obtained from Sigma. RQ1 placental RNasin ribonuclease inhibitor was supplied by Promega. T4 polynucleotide kinase and T4 DNA polymerase were from New England Biolabs, T4 DNA ligase was supplied by Gibco. *Taq* DNA polymerase was obtained from Cetus. Klenow fragment DNA polymerase was purified in house by Miss M. Watson from recombinant *E. coli* cells.

2A.2 Radiochemicals

With the exception of ^{14}C chloramphenicol (1mM) which was supplied by Dupont, radiochemicals were supplied by Amersham at the following specific activities:

$\alpha^{32}P$-dNTPs	3,000Ci/mmol (10μCi/μl)
^{14}C chloramphenicol	50.13mCi/mmol (0.1μCi/μl)
^{35}S L-methionine	approximately 800Ci/mmol (15μCi/μl)
^{3}H Leucine	140Ci/mmol (5μCi/μl)
^{14}C-labelled high molecular weight protein markers	10-50μCi/mg protein

2A.3 Synthetic oligonucleotides

The M13 universal sequencing primer was obtained from New England Biolabs. All other synthetic oligonucleotides used in this study were either produced in house using a

Biosearch model 8600 DNA synthesiser by Dr J. McLauchlan, or on a Cruachem PS250 automated synthesiser by Mr F. R. VanDeursen and Miss S. Fitzpatrick.

2A.4 Plasmids and bacteriophage

The well characterised universal cloning vectors pUC9, pUC19, pACYC184 and M13mp18 were used in this study, but are not described here. The other plasmids utilised are described briefly below, and were kindly provided by the acknowledged authors (coordinates given are relative to the transcription start site (+1) of the gene in question):

p585.4 (C. M. Preston, unpublished) is a derivative of the pET-8c T7 expression vector (Studier *et al.*, 1990) with minor alterations including removal of the original *Hind*III site and creation of a novel unique *Hind*III site between the *Nco*I and *Bam*HI sites.

pUCIE3 (R. D. Everett, unpublished) includes HSV-1 IE3 gene sequences spanning the IE3 consensus binding site, on an *Ava*I-*Bam*HI fragment (-18 to +27), cloned into vector pUC19.

pBEND2 (Zweib *et al.*, 1989) includes unique *Xba*I and *Sal*I restriction sites flanked by direct 236bp repeats bearing 17 duplicated restriction sites, in a pBR322 background.

pGEX2TN3 (R. D. Everett, unpublished) was derived from the commercial vector pGEX2T (Pharmacia) by inserting a DNA fragment (carrying *Nco*I, *Eco*RI and *Hind*III sites) between the *Bam*HI and *Eco*RI sites downstream of the GST coding sequence (this cloning removed the original *Bam*HI and *Eco*RI sites).

pT7I9X (Pizer *et al.*, 1991) includes sequences encoding HSV-1 Vmw175 amino acid residues 252 to 523, on an *Eco*RI-*Bam*HI fragment (+752 to +1566) of plasmid pI9 (Paterson and Everett, 1988a), cloned into plasmid p585.4.

pT7I10X (Pizer *et al.*, 1991) includes sequences encoding HSV-1 Vmw175 amino acid residues 275 to 523, on an *Eco*RI-*Bam*HI fragment (+823 to +1566) of plasmid pI10 (Paterson and Everett, 1988a), cloned into vector p585.4.

pJI10X (R. D. Everett, unpublished) includes the identical HSV-1 Vmw175 sequences carried by plasmid pT7I10X, cloned into vector p585T7a (Section 3A.2).

pT7I10i12X-i20X series of plasmids (R. D. Everett, unpublished) derivatives of pT7I10X, with single 12bp insertions in the IE3 gene sequences that have been described previously (Paterson and Everett, 1988a).

pT7110 (Everett *et al.*, 1991b) contains the entire HSV-1 IE1 cDNA cloned on an *Nco*I *Hpa*I fragment cloned into the *Nco*I-*Hind*III sites of plasmid p585.4.

p140 (McKee *et al.*, 1990) contains VZV gene 71 (gene 62) coding regions and its promoter-leader sequences (-1146 to the genomic terminus), on a *Cla*I-*Pst*I fragment of plasmid pVZVSstf (Davison and Scott, 1983) cloned into vector pUC9.

pC33 (or **pV17**) (Disney, 1990) a version of p140 containing a 12bp linker insertion at nucleotide +1253, disrupting amino acid Ser 418.

pC34 (or **pV18**) (Disney *et al.*, 1990) a version of p140 containing a 12bp linker insertion at nucleotide +1418, disrupting amino acid Ser 472.

pVZV*ts*K (R. D. Everett, unpublished), carries a version of VZV gene 62 with alterations to the coding sequence which result in the substitution Ala 628 Val, equivalent to the HSV-1 Vmw175 *ts*K mutation (Ala 475 Val).

p140BT (McKee *et al.*, 1990), a derivative of p140, with a unique *Bgl*II site inserted at position +57.

p140SV (Disney *et al.*, 1990) carries VZV gene 62 sequences from nucleotide +57, under the control of the SV40 early promoter-enhancer regions.

p175 (Everett, 1987b) contains HSV-1 IE3 gene sequences from +27 (on a *Bam*HI-*Sst*I fragment) under the control of the SV40 early promoter-enhancer regions.

pI8 (Paterson and Everett, 1988a) a derivative of plasmid p175 created by insertion of a 12bp *Eco*RI linker at nucleotide position 684 of the HSV-1 IE3 gene.

pI9 (Paterson and Everett, 1988a) a derivative of plasmid p175 created by insertion of a 12bp *Eco*RI linker at nucleotide position 752 of the HSV-1 IE3 gene.

pLI10/11 (Allen and Everett, 1994) a derivative of plasmid pI10 (Paterson and Everett, 1988a) created by removal of the second *Bam*HI site in region 3 of the HSV-1 IE3 coding sequence, by oligonucleotide-directed mutagenesis.

p111 (Everett, 1987b) contains the entire HSV-1 IE1 coding region and promoter sequences cloned on a *Sst*I-*Hpa*I fragment into vector pUC9.

pMC1 (Campbell *et al.*, 1984) includes the complete HSV-1 gene UL48, encoding Vmw65, cloned into vector pUC9.

pSVEB (Paterson and Everett, 1988a) contains the SV40 early promoter-enhancer regions (not linked to any coding sequences) cloned into vector pBR322.

pgDCAT (Everett, 1986) contains the promoter of the HSV-1 gD gene (-392 to +11) on a *Sst*I-*Hind*III fragment fused to the CAT gene in vector pBWL2 (Gaffney *et al.*, 1985).

pIE3CAT (Stow *et al.*, 1986) includes the HSV-1 IE3 promoter (-331 to +27), on a *Hind*III-*Bam*HI fragment fused to the CAT gene in a vector derived from plasmid pBLW2 (Gaffney *et al.*, 1985).

p140CAT (McKee *et al.*, 1990) includes the VZV gene 62 promoter-leader sequences from -1146 to +57 on a *Cla*I-*Bgl*II fragment of plasmid p140BT fused to the CAT gene in vector pCAT, derived from plasmid pBLW2 (Gaffney *et al.*, 1985).

pSVECAT (R. D. Everett, unpublished) includes the SV40 early promoter-enhancer sequences linked to the CAT gene in a vector derived from plasmid pBLW2 (Gaffney *et al.*, 1985).

ptkCAT (Everett, 1986) contains the promoter sequences of the HSV-1 tk gene on a *Bam*HI-*Bgl*II fragment (approximately -650 to +53) linked to the CAT gene in a vector derived from plasmid pBLW2 (Gaffney *et al.*, 1985).

pIE1CAT (Everett and Orr, 1991) contains the HSV-1 IE1 gene promoter on a *Sst*1-*Nco*I fragment linked to the CAT gene in the vector pCAT, derived from plasmid pBLW2 (Gaffney *et al.*, 1985).

M13IE3EB series of plasmids (Everett *et al.*, 1991a), each includes a single nucleotide substitution in the IE3 consensus binding site on the *Eco*RI-*Bam*HI cap site fragment (-108 to +27) of the HSV-1 IE3 gene, cloned into M13mp18.

2A.5 Bacterial strains

E. coli strain DH5α (F'/*endA1 hsdR17* ($r_k^-m_k^+$) *supE44 thi-1 recA1 gyrA* (Nal^r) *relA1 Δ(lacZYA-argF)U169 (ϕ80dlacΔ(lacZ)M15*) was the usual host for construction, maintenance and propagation of recombinant plasmids. *E. coli* strain JM101 (F' *traD36 lacI^{q}Δ(lacZ)M15 proA$^+$B$^+$/supE thi Δ (lac-proAB*) (Messing, 1979) or strain TG1 (K12 *Δ (lac-pro) supE thi hsdD5/F'traD36 proA$^+$B$^+$ lacI^q lac ZΔM15*) were used for growth and maintenance of M13 bacteriophage. *E. coli* strain BL21 (DE3) pLysS (F^- *ompT* r^-_B m^-_B) (Studier *et al.*, 1990) was used for the expression experiments; bacteriophage DE3 encodes an inducible T7 RNA polymerase and plasmid pLysS encodes T7 lysozyme and carries a chloramphenicol resistance marker. *E. coli* strain BL21 (DE3), lacking the pACYC184 based pLysS plasmid, was used for the *in vivo* co-translation experiments described in Section 3D.3.1.

2A.6 Bacterial culture media

The following bacterial culture media were used:

L-broth:	10g NaCl, 10g Bactopeptone, 5g yeast extract per litre.
LB agar:	L-Broth plus 1.5% (w/v) agar.
2YT broth:	5g NaCl, 16g Bactopeptone, 10g yeast extract per litre.
H top agar:	8g NaCl, 10g Bactotryptone, 8g agar per litre.
SOC broth:	10mM NaCl, 2.5mM KCl, 2% Bactotryptone, 0.5% yeast extract.

Bacterial culture media for growth and maintenance of pUC or pBR322 based plasmids was supplemented with ampicillin at 100μg/ml, media for growing pACYC184 based plasmids was supplemented with chloramphenicol at 25μg/ml. Bacterial media for *E. coli* strain BL21 (DE3) pLysS was additionally supplemented with chloramphenicol at 25μg/ml to maintain the pLysS plasmid.

2A.7 Cells and tissue culture media

WS HeLa cells (obtained from Dr W. Schaffner, Zurich) were grown in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco) supplemented with 100units/ml penicillin, 100µg/ml streptomycin, 2.5% foetal calf serum (Gibco) and 2.5% newborn calf serum (Gibco).

BHK-21 C-13 cells, a fibroblastic cell line derived from baby hamster kidneys (MacPherson and Stoker, 1962), were grown in Glasgow Modified Eagle's Medium (GMEM) (Gibco) with 10% newborn calf serum supplemented with 100 unit/ml penicillin, 100µg/ml streptomycin and 10% tryptose phosphate broth.

Vero cells (obtained from the ATCC via Flow Laboratories) were grown in GMEM (Gibco) with 10% foetal calf serum supplemented with 100units/ml penicillin, 100µg/ml streptomycin and 10% tryptose phosphate broth.

2A.8 Viruses

Virus D30EBA was derived from the HSV-1 Glasgow strain 17 syn^+, which forms non-syncytial plaques on BHK cells (Brown *et al.*, 1973). D30EBA (Paterson and Everett, 1990) has a large deletion created within both copies of the IE3 gene which removes codons 83 to 1236.

2A.9 Antisera

The following monoclonal antibodies (MAbs) were obtained from Dr A. Cross of this Institute:

MAb 10,084 (Everett *et al.*, 1993c) was raised against the purified bacterially expressed HSV-1 Vmw175 DNA binding domain peptide, I9X, and recognises an epitope between residues 310 and 329 of HSV-1 Vmw175. MAb 10,084 cross-reacts with an epitope spanning residues 462 and 472 of VZV 140k.

MAbs 11,060 and 10,503 (Everett *et al.*, 1993b) were raised against the purified Baculovirus expressed HSV-1 Vmw110 polypeptide. MAb 11,060 recognises an epitope within the 105 N-terminal residues of Vmw110. MAb 10,503 recognises an epitope within the 140 C-terminal residues of Vmw110.

Normal mouse serum was produced from non-immunised mice.

Rabbit polyclonal antisera 108 and 109 were raised against the purified bacterially expressed VZV 140k DNA binding domain peptide, VT2 (Section 3C.5).

Protein A horse radish peroxidase (HRP) conjugate, anti-mouse IgG whole molecule peroxidase conjugate, sheep anti-mouse (SαM) and fluorescein isothiocyanate (FITC) conjugated sheep anti-rabbit immunoglobulin were purchased from Sigma.

2A.10 Solutions

Blue juice (10 X agarose gel loading buffer): 1 X TBE, 1% SDS, 50% Glycerol, 1mg/ml bromophenol blue.

Coomassie stain: 50:50:7, Methanol:dH_2O:acetic acid, 0.2% R250 Coomassie brilliant blue.

20 X dNTPs: 1mM dATP, dCTP, dGTP, dTTP.

Fix (protein gel fixative): 4:1:27, Acetic acid:ethanol:dH_2O.

Formamide dyes mix: 10mM EDTA, 1mg/ml xylene cyanol FF, 1mg/ml bromophenol blue in formamide.

Giemsa stain: 1.5% Giemsa in glycerol, heated to 50°C for 2hr and diluted with an equal volume of methanol.

Glycine SDS PAGE buffers:

3 X Boiling mix: 1ml SGB, 1ml 20% SDS, 500µl BME, 1ml glycerol, 1mg/ml bromophenol blue.

RGB (resolving gel buffer): 181.5g Tris, 4g SDS, 1l dH_2O, pH 8.9 (HCl).

SGB (stacking gel buffer): 59g Tris, 4g SDS, 1l dH_2O, pH 6.7 (HCl).

Tank buffer: 3.16g Tris, 2.0g glycine, 0.5g SDS, 1l dH_2O.

2 X HBS: 280mM NaCl, 50mM HEPES, 1.5mM Na_2HPO_4, pH to 7.12 with NaOH.

IP buffer: 0.05M Tris.HCl (pH 7.5), 0.3M $NaCl_2$, 1mM EDTA, 0.2% NP40, pH to 7.5.

5 X Kinase buffer: 350mM Tris.HCl (pH 7.5), 50mM $MgCl_2$, 25mM DTT.

5 X Ligase buffer: 250mM Tris.HCl (pH 7.6), 50mM $MgCl_2$, 5mM DTT, 5mM ATP, 25% (w/v) PEG 8000.

M buffer (1 X): 10mM Tris.HCl, 10mM $MgCl_2$, 50mM NaCl, 1mM DTE, pH 7.5.

5 X non-SDS loading buffer: 10mM Tris.HCl (pH8.0), 1mM EDTA, 10mM BME, 0.1% CHAPS, 50% glycerol and 0.25mg/ml bromophenol blue.

PBSA: 170mM NaCl, 3.4mM KCl, 10mM HPO_4, 1.8mM KH_2PO_4, pH 7.2.

PBSc: PBSA plus $CaCl_2 2H_2O$ and $MgCl_2 6H_2O$ at 1g/l.

10 X seq buffer: 100mM Tris-HCl, 100mM $MgCl_2$, pH 8.5.

STET: 8% Sucrose, 5% triton X-100, 50mM EDTA (pH 8.0), 50mM Tris (pH 8.0).

TBE: 90mM Tris.HCl (pH 8.0), 90mM boric acid, 1mM EDTA.

TBS: 25mM Tris.HCl (pH 7.5), 137mM NaCl, 5mM KCl, 0.7mM $CaCl_2$, 0.5mM $MgCl_2$, 0.6mM Na_2HPO_4, pH to 7.4 with HCl.

TE: 10mM Tris.HCl, 1mM EDTA, pH 8.0.

TEN: 150mM NaCl, 40mM Tris.HCl, 1mM EDTA, pH 7.5.

TEP: 1mM Tris.HCl, 0.5mM EDTA, pH 7.5.

Towbin (blotting buffer): 25mM Tris.HCl, 192mM glycine, 20% (v/v) methanol, pH 8.3.

Trypsin: 0.25% (w/v) trypsin dissolved in Tris.saline.

Tricine SDS PAGE buffers:

3 X Boiling mix: 150mM Tris, 12% SDS, 36% Glycerol, 6% BME, 1mg/ml bromophenol blue, pH 6.8.

Gel buffer: 3.0M Tris, 0.3% SDS, pH 8.45.

Lower tank buffer: 200mM Tris, pH 8.9.

Upper tank buffer: 0.1M Tris, 0.1M tricine, 0.1% SDS.

TTBS: TBS + 0.05% TWEEN-20.

Versene: 0.6mM EDTA in PBSA, 0.002% (w/v) phenol red.

2A.11 Chemicals and reagents

Unless specifically stated below, all chemicals and reagents were obtained from BDH Chemicals UK or Sigma Chemical Co. and were at least of analytical grade. The sources of the reagents and materials given in the Methods Sections are not repeated here.

Ammonium hydroxide (Fisons).
Ammonium persulphate, Bio-Rad protein assay reagent, Coomassie brilliant blue, gelatin, TEMED and tricine (Bio-Rad).
Ampicillin (Beecham Research).
Boric acid, butan-1-ol, butan-2-ol, chloroform, ethanol, glacial acetic acid, glycerol, hydrochloric acid, isopropanol and methanol (Prolabo).
Caesium chloride (Melford Laboratories Ltd.).
Diethylether (Rathburn).
Ecoscint A (National Diagnostics).
En^{3}hance (Dupont).
Formamide, formaldehyde and piperidine (Fluka).
Formic acid (May and Baker Ltd.).
Glutaraldehyde (Agar Scientific Ltd.).
polydI.polydC and 7-deaza-dGTP, dNTPs and ddNTPs (Pharmacia).
Protease inhibitors; leupeptin, aprotinin and bestatin (Boeringer Mannheim).

2B DNA manipulation and cloning procedures

2B.1 Large scale plasmid DNA preparation

Single *E. coli* colonies, from agar plates selecting for a plasmid-borne resistance marker, were inoculated into 10ml of 2YT broth (plus appropriate antibiotics, see Section 2A.6), and shaken overnight at 37°C. The overnight culture was inoculated into 350ml of L-broth (plus antibiotics) and shaken for 20hr at 37°C.

Plasmid DNA or RF M13 DNA was prepared by the 'maxi-boiling' technique as described by Holmes and Quigley (1981), and supercoiled plasmid DNA was purified on

caesium chloride / ethidium bromide gradients. Bacterial cultures were centrifuged at 5krpm for 6min at 4°C (Sorvall GS3 rotor) and the pellet resuspended in 20ml of STET buffer. 2ml of a fresh 10mg/ml solution of lysozyme in STET was added and the suspension was brought to the boil and placed in a boiling water bath for 45sec. The lysate was cleared by centrifugation at 18krpm for 45min at 4°C (Sorvall SS34 rotor). DNA was precipitated from the supernatant with an equal volume of isopropanol and pelleted by centrifugation at 3krmp for 10min at RT (Beckman CPR centrifuge). After draining, the pellets were resuspended in 5.5ml TE and the total volume determined (Xml). X+0.7g of CsCl was added to the DNA solution to give a final density of around 1.6g/ml, and 200µl of 10mg/ml EtBr was added to give a final concentration of around 0.25mg/ml. After several hours at 4°C, the solution was cleared by centrifugation at 3krpm for 10min at 4°C. The supernatant was transferred to a sealable centrifuge tube (Dupont 03945 tubes) and spun at 40krpm (Beckman TV865 vertical rotor) for ≤16hr at 14°C. After removal of the upper chromosomal DNA band from the gradient, the lower supercoiled plasmid band was collected (using syringes fitted with wide-bore needles through the side of the tube).

EtBr was removed from the plasmid DNA solution by at least two extractions in butan-1-ol (TE saturated) followed by dialysis against TE for 3hr at RT. Next, the solution was treated with 100µg/ml RNase A for 1hr at 65°C, followed by a 1hr incubation with 100µg/ml proteinase K and 0.1% SDS at 37°C. The DNA was then purified by extractions with phenol (saturated in TE) and chloroform, then by ethanol precipitation. The DNA pellet was resuspended in 400µl 0.3M NaAc and ethanol precipitated again. The lyophilised pellet was washed with 80% ethanol, resuspended in TE and stored at 4°C. DNA concentration was determined by measuring the absorbance at 260nm, assuming $1A_{260} = 50\mu g/ml$.

A more rapid procedure for large scale preparation of plasmid DNA, by alkali lysis of bacteria and polyethylene glycol precipitation of DNA, was occasionally used following the method of Ausubel *et al.* (1993a).

2B.2 Miniprep plasmid DNA preparation

Plasmid DNA or RF M13 DNA was prepared on a small scale as follows: single colonies of transformed bacteria were picked and inoculated into 3ml aliquots of L-broth containing antibiotics as appropriate (Section 2A.6), cultures were grown overnight at 37°C with shaking. The bacterial cells from 1.5ml of culture were pelleted for 30sec (benchtop microcentrifuge) and resuspended in 200µl STET. To this was added 5µl of a fresh 10mg/ml solution of lysozyme in STET and the tube immediately placed in a boiling water

bath for 45sec. The suspension was centrifuged for 10min and the pellet removed. The supernatant was transferred to a fresh tube and 0.9vol isopropanol added to precipitate the DNA. After centrifugation for 10min, the plasmid DNA pellet was washed with 80% ethanol, lyophilised, resuspended in 20µl TE and stored at -20°C. Prior to sequencing (Section 2B.12), miniprep DNA was treated with RNase A to 20µg/ml for 20min at 37°C, followed by phenol, chloroform extractions and ethanol precipitation.

2B.3 Restriction enzyme digestion of DNA

Digests with appropriate restriction enzymes were generally carried out in 20µl volumes containing 1 unit enzyme per 0.5µg DNA per 1hr at the appropriate temperature, using the commercial buffers specified and supplied by the manufacturer. In general, the Boeringer Mannheim A, B, L, M, H buffer system was used. Usually, 0.5µg plasmid DNA was digested for diagnostic analysis on agarose gels, whereas 2–10µg plasmid DNA was digested for preparative isolations of DNA fragments. Restriction digest analysis of miniprep DNA preparations used 5µl of DNA, digested in a 20µl reaction volume including 50µg/ml RNase A.

2B.4 Electrophoretic separation and purification of DNA fragments

2B.4.1 *Non-denaturing agarose gels*

DNA fragments (over 200bp) produced by restriction enzyme digests and PCR reactions were resolved by non-denaturing agarose gel electrophoresis. Samples in 0.1vol Blue juice were loaded onto horizontal slab gels of 0.5-1.5% (w/v) agarose in 1 X TBE, and electrophoresed in 1 X TBE at up to 12V/cm. Appropriate DNA fragment size markers (i.e. *Hpa*II digested pBR322, *Hind*III digested phage λDNA or *Hae*III digested phage ϕX174 DNA) were also ran on the gel. After the dye front had migrated the required distance, the gel was stained in 1µg/ml EtBr for 10min, rinsed thoroughly in H_2O and photographed with short-wave uv transillumination (long-wave for preparative gels) using a Polaroid instant camera or photographed onto heat-sensitive paper using The Imager (Appligene).

2B.4.2 *DNA purification from agarose gels by silica matrix adsorbtion ('genecleaning')*

A commercial kit, GENECLEANII (BIO 101 Inc., La Jolla, CA) was used for recovery of DNA from agarose blocks which had been excised from agarose gels during visualisation under long-wave uv transillumination. The kit contains a silica matrix that binds DNA in the presence of high concentrations of sodium iodide (Vogelstein and

Gillespie, 1979). Blocks containing the required DNA fragments were mixed with 4.5vol of saturated NaI and 0.5vol TBE modifier and incubated for 5min at 55°C until the agarose had disolved. Silica matrix was added (5µl for up to 5µg of DNA) and the mixture incubated for 5min at RT. Following a 5sec centrifugation, the pellet was washed 3 times using 'NEW' wash (a Tris-buffered mixture of NaCl, ethanol and water). The DNA was eluted from the silica matrix by resuspending the pellet in 20µl TE and incubating for 2min at 55°C. After a 30sec centrifugation, the 'genecleaned' DNA solution was collected and stored at -20°C. The GENECLEANII procedure was also applied to situations where electrophoresis was not required, for example, to purify DNA away from enzymes, nucleotides or unwanted buffer components. In these later cases, 3vol of NaI was used and the TBE modifier omitted.

2B.4.3 ***Non-denaturing polyacrylamide gels***

DNA fragments (40-400bp) were resolved on 5-8% polyacrylamide (a+b:b; 20:1, where a = acrylamide and b = NN'-methylenebisacrylamide) vertical gels run in 1 X TBE. The gels (1.5mm thick) contained 1 X TBE and were polymerised by addition of 0.001vol TEMED and 0.01vol 10% APS. Samples in 0.2vol 5 X non-SDS loading buffer were loaded and electrophoresed at up to 16V/cm. The position of ^{32}P-radiolabelled DNA fragment bands were identified by brief autoradiography (for approximately 5min), otherwise EtBr staining was used (Section 2B.4.1).

2B.4.4 ***Denaturing polyacrylamide gels***

Denaturing polyacrylamide (a+b:b; 20:1) vertical gels were used to resolve the products of DNA sequencing reactions (Section 2B.12), DNase I footprinting reactions (Section 2D.4) and also to purify oligonucleotides (Section 2B.7). The gels contained 1 X TBE, 8M urea, and were polymerised by addition of 0.001vol TEMED and 0.01vol 10% APS. The 8% or 12% gels used for sequencing and DNase I footprinting analyses were 0.35mm thick X 35cm long; samples in formamide dyes mix were denatured by boiling for 2min prior to loading and electrophoresis in 1 X TBE at 40W for 2-5hr. The 16% gels used for purifying synthetic oligonucleotides were 1.5mm thick, 25cm long and were ran in 1 X TBE at 250V for 5hr.

2B.4.5 ***DNA purification from polyacrylamide gels by elution***

DNA fragments that had been resolved on non-denaturing polyacrylamide gels and also oligonucleotides ran on denaturing polyacrylamide gels were purified by passive elution from gel slices. Gel slices containing the relevant DNA fragments (excised during visualisation by long-wave uv transillumination or after detection by autoradiography) were fragmented, and the DNA eluted into 500µl TE by shaking overnight at 37°C. The suspension was filtered through siliconised glass wool and the DNA was ethanol precipitated from the filtrate, lyophilised, resuspended in TE and stored at -20°C.

2B.5 End repair and DNA ligation

Before ligating DNA fragments together, it was occasionally necessary to enzymatically modify the ends of the fragments to allow the ligation reaction to proceed. After any of the end repair procedures described below, enzymes were heat inactivated by incubation for 10min at 65°C and the DNA isolated by 'genecleaning' (Section 2B.4.2), before further restriction enzyme digests and prior to setting up ligation reactions.

2B.5.1 *Filling-in of 5' overhangs*

To enable blunt end ligation, 5' overhangs were filled in to produce DNA fragments with flush ends. After heat inactivation of the restriction enzyme, the buffer conditions were adjusted to resemble those of M buffer (by dilution or salt addition) and this DNA solution was incubated with 1µl 20 X dNTPs and 2u Klenow fragment for 30min at 37°C. To obtain DNA fragments with a single blunt end, the end-filling reaction was included prior to digestion with the second restriction enzyme.

2B.5.2 *Removal of 3' overhangs*

T4 DNA polymerase is more efficient than the Klenow fragment at digesting back the 3' overhangs formed by some restriction enzymes, to yield flush fragment ends for blunt end ligations. After heat inactivation of the restriction enzyme, 8u T4 polymerase was added to the DNA solution and incubated for 30min at 37°C.

2B.5.3 *Phosphate removal from 5' ends*

The 5' phosphate was removed from DNA fragments in order to prevent re-ligation of vector fragments with complementary sticky ends that were produced by digestion with a single restriction enzyme. After heat inactivation of the restriction enzyme, the digested vector DNA was incubated with 1u calf intestinal alkaline phosphatase for 30min at 37°C in a buffer giving final concentrations of 1mol/l diethanolamine, 10mmol/l 4-nitrophenyl phosphate, 0.5mmol/l $MgCl_2$, pH 9.8.

2B.5.4 *Phosphate addition to 5' ends*

It was necessary to add phosphate to the 5' ends of synthetic oligonucleotides or PCR products, to facilitate their ligation into plasmid vectors. Kinasing reactions included 10u T4 polynucleotide kinase, 2µl 10mM ATP and 5 X Kinase buffer, and were incubated for 30min at 37°C.

2B.5.5 *Ligation reactions*

DNA fragments and cut vector DNA (in a molar ratio of 10:1), which had been previously digested with the appropriate restriction enzyme(s), end-repaired if necessary, isolated and purified from agarose gels (Sections 2B.4.1 and 2B.4.2), were ligated together in

10-20µl reactions containing 2u T4 DNA ligase and 5 X ligase buffer for 4-16hr at RT. After the incubation period, the ligase enzyme was heat inactivated at 65°C for 10min, followed by purification of the DNA by phenol and chloroform extractions or 'genecleaning' (Section 2B.4.2), prior to transformation of half the ligation mix into competent *E. coli* cells as described below.

2B.6 Preparation and transformation of competent bacteria for plasmid growth and maintenance

A 1ml overnight culture of *E. coli* strain DH5α, inoculated from a single fresh colony, was diluted into 100ml of L-broth and shaken for about 3hr at 37°C until the culture OD_{630} reached about 0.3. The culture was cooled on ice for 20min prior to harvesting the bacteria by centrifugation at 5krpm for 10min at 4°C. The pellet was resuspended in 25ml cold 0.1M $MgCl_2$ and immediately re-pelleted at 5krpm for 10min at 4°C. Resuspension in 5ml cold 0.1M $CaCl_2$ and incubation on ice for 45min, was followed by centrifugation at 5krpm for 10min at 4°C. The cells were resuspended in 5ml cold 0.1M MOPS, 50mM $CaCl_2$, 20% glycerol. After 20min on ice, cells were aliquoted into 100-500µl amounts, flash frozen, and stored at -70°C.

About 10ng plasmid DNA, or half a ligation reaction, was incubated with 100µl of competent bacterial cells for 30min on ice, followed by a 45sec 42°C heat shock. The transformation reaction was incubated on ice for a further 2min before addition of 0.5ml SOC broth supplemented with $MgCl_2$, $MgSO_4$ and glucose, each to 20mM. The transformation mix was shaken for 1hr at 37°C, before plating cells onto LB agar containing the appropriate antibiotics (Section 2A.6) for selection of transformants. Plates were incubated at 37°C overnight to allow single colonies to form.

2B.7 Purification of synthetic oligonucleotides

Synthetic oligonucleotides were produced by the phosphoramidite method and supplied on 200µl synthesis columns. Oligonucleotides were eluted by pushing 1.5ml ammonium hydroxide through the column, 200µl at a time with 20min incubations at RT between each fresh addition. The ammonium hydroxide solution was collected from the bottom of the column. After the final incubation period, the ammonium solution was pushed back and forth through the column several times (using syringes fitted to both ends of the column). The oligonucleotide-ammonium solution was incubated for 5hr at 55°C to remove the base protecting groups, followed by lyophilisation. The DNA pellet was dissolved in 200µl TE, boiled for 2min in an equal vol of formamide, and electrophoresed on a 16% polyacrylamide gel containing 8M urea (Section 2B.4.4). Oligonucleotide bands

were visualised under uv light, as dark shadows against a fluorescing TLC plate. Oligonucleotides were eluted and precipitated from slices of the gel as described previously (Section 2B.4.5), resuspended in 100µl TE and stored at -20°C. The concentration of the oligonucleotide solution was determined by spectrophotometry ($OD_{260}1.0 = 50\mu g/ml$).

2B.8 Polymerase chain reaction (PCR) amplification of DNA

Conditions used for the PCR amplification of sequences encoding the VZV 140k DNA binding domain were determined empirically and were as follows: PCR reactions were performed in 25µl reaction volumes containing 10ng plasmid DNA template, 16pmoles of each primer (described in Fig. 3F.1), 50mM KCl, 20mM Tris.HCl (pH 8.4), 15mM $MgCl_2$, 800µM dNTPs and 2u of *Taq* thermostable DNA polymerase. Reactions were overlaid with 50µl mineral oil and cycled for 25 rounds of PCR in a Perkin-Elmer DNA Thermal cycler 460. PCR cycling parameters were as follows: (a) template duplex melting at 94°C for 3min for the first cycle, subsequent cycles had a 30sec incubation at 94°C to melt the primer-template duplex; (b) primer annealing for 30sec at 60°C for 3 cycles, 58°C for 3 cycles, 56°C for 3 cycles and the remaining cycles at 55°C; (c) primer elongation at 72°C for 1.5min. After 25 cycles, the remaining DNA ends were elongated by incubation at 72°C for 7min, before cooling to 0°C. PCR products were analysed by agarose gel electrophoresis (Section 2B.4.1), then purified by 'genecleaning' (Section 2B.4.2) prior to subsequent cloning steps (Section 3F.2).

2B.9 Preparation of single-stranded template M13 DNA

Small scale preparations of single-stranded template M13 DNA were required for oligonucleotide-directed mutagenesis reactions (Section 2B.10), for sequencing of M13 vectors (Section 2B.12), and for preparation of 'prime cut' probes (Section 2D.1.2). A single colony of M13 host bacterium strain JM101 (strain TG1 was used for all procedures related to the oligonucleotide-directed mutagenesis) was inoculated into 10ml 2YT broth and shaken overnight at 37°C. 200µl of the overnight culture was diluted into 100ml 2YT broth, this was then dispensed into universals in 3ml amounts. Single plaques harbouring bacteriophage M13 were picked from fresh transformation plates, inoculated into universals, and shaken for 3hr at 37°C. Approximately 1.5ml of each bacterial culture was centrifuged for 30sec at RT, rotated through 180° and re-spun. Next 800µl of the supernatant was removed to a fresh tube containing 200µl 20% PEG / 2.5M NaCl, mixed and incubated for 30min at RT. The bacteriophage precipitate was spun down for 5min and resuspended in 100µl TE. The bacteriophage DNA was purified by phenol and

chloroform extractions, then 10µl 4M NaAc was added to the solution prior to an ethanol precipitation. The pellet was washed with 80% ethanol, lyophilised, resuspended in 20µl 0.2 X TE, and stored at -20°C.

2B.10 Oligonucleotide-directed mutagenesis of M13 vectors

Specific amino acid substitutions were introduced into the VZV 140k sequence using the Oligonucleotide-Directed *In Vitro* Mutagenesis System Version 2 (Amersham), which is based on the protocol of Eckstein and co-workers (Sayers *et al.*, 1988). The target sequence for mutagenesis had previously been cloned into RF M13mp18 and single-stranded template DNA prepared (Section 2B.9). To check that the mutagenic oligonucleotide anneals with the intended sequence of the template DNA, the oligonucleotide was used as a primer in a sequencing reaction before being used to direct mutagenesis. The manufacturer's mutagenesis protocol was followed and a control mutagenesis reaction was performed in parallel, using the mutant oligonucleotide and template supplied with the system. An outline of the stages in the oligonucleotide-directed mutagenesis system is given in Figure 2B. Briefly, the mutagenesis was as follows:

First of all, 4pmoles of phosphorylated oligonucleotide (Section 2B.5.4) was annealed to 5µg of template M13 DNA, by placing in a 70°C water bath for 3min, followed by incubation for 30min at 37°C. To this was added 3u Klenow fragment, 3u T4 ligase and a nucleotide mix including an analogue of dCTP to allow primer extension and ligation of the mutant DNA strand, during an incubation period of 16hr at 15°C. After this time, single-stranded (non-mutant) DNA was removed using centrifugal filter units. The heteroduplex RF DNA molecules were ethanol precipitated from the filtrate and the non-mutant strand was nicked with the enzyme *Nci*I (Section 2B.3); the nucleotide analogue protects *Nci*I sites in the mutant DNA strand from cleavage. Selective degradation of the nicked non-mutant DNA strand was achieved by incubation with 50u Exonuclease III for 15min at 37°C, followed by heat inactivation of the enzymes. Incubation with 1.25u DNA polymerase I, nucleotides and 1.5u T4 ligase for 3hr at 15°C ensured synthesis of closed circular homoduplex mutant RF M13 DNA. Finally, 1vol 4M ammonium acetate (pH 5.4) was added to facilitate ethanol precipitation, and the precipitated DNA pellet was washed with 80% ethanol, lyophilised and resuspended in 100µl TE. 20µl of the mutagenesis reaction products were transformed into competent TG1 bacterial cells, plated out and screened for mutants (Section 2B.11).

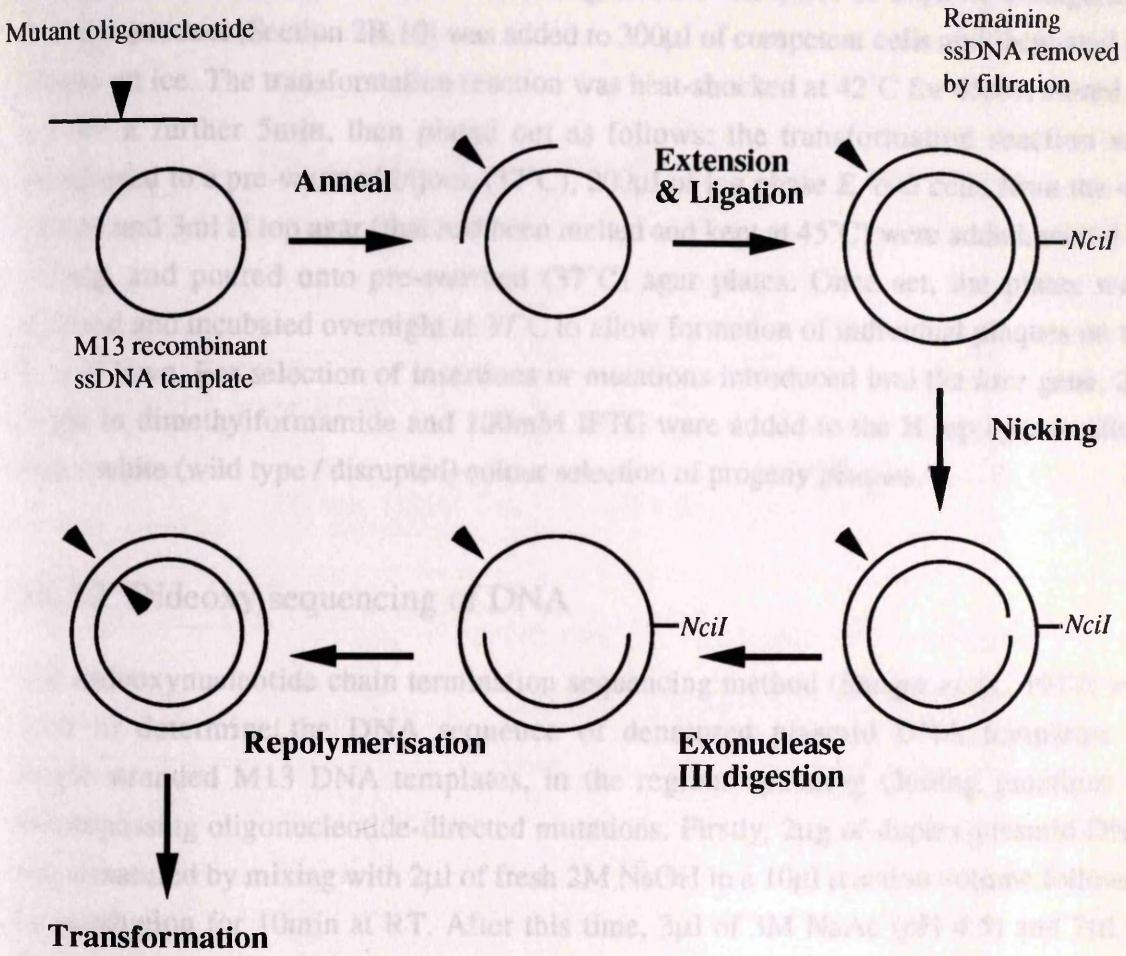


Figure 2B Plan of the strategy used by the Amersham oligonucleotide-directed *in vitro* mutagenesis system. The arrow head indicates the position of a specific mutation in the synthetic oligonucleotide. The mutation is initially introduced into a single strand of the bacteriophage DNA, and then copied into the complementary DNA strand to produce a homoduplex mutant bacteriophage, as described in the text.

2B.11 Preparation and transformation of competent bacteria for bacteriophage growth and maintenance

A single colony of JM101 or TG1 cells was inoculated into 10ml 2YT broth and shaken overnight at 37°C. Next, 1ml of the overnight culture was diluted into 100ml 2YT broth, and shaken for around 3hr until OD_{630} of 0.3 was reached. Competent JM101 cells were prepared using the procedure previously described in Section 2B.6. TG1 cells were pelleted at 3krpm for 5min at 4°C, resuspended in 0.5vol pre-chilled 50mM $CaCl_2$ and left on ice for 20min. Cells were pelleted again at 3krpm for 2min at 4°C, before resuspending in 0.1vol cold 50mM $CaCl_2$, and competent cells were stored on ice until required.

Four hours before transformation of competent cells with bacteriophage DNA, a 10ml culture of JM101 or TG1 cells was set up in 2YT broth, using an overnight culture

inoculum, and shaken at 37°C. Around 1ng of RF M13 DNA or 20µl of mutagenesis reaction product (Section 2B.10) was added to 300µl of competent cells and incubated for 40min on ice. The transformation reaction was heat-shocked at 42°C for 45sec, stored on ice for a further 5min, then plated out as follows: the transformation reaction was transferred to a pre-warmed bijoux (37°C), 200µl of log phase *E. coli* cells from the 4hr culture and 3ml H top agar (that had been melted and kept at 45°C) were added, mixed by rolling, and poured onto pre-warmed (37°C) agar plates. Once set, the plates were inverted and incubated overnight at 37°C to allow formation of individual plaques on the *E. coli* lawn. For selection of insertions or mutations introduced into the *lacz* gene, 2% X-gal in dimethylformamide and 100mM IPTG were added to the H top agar to allow blue / white (wild type / disrupted) colour selection of progeny plaques.

2B.12 Dideoxy sequencing of DNA

The dideoxynucleotide chain termination sequencing method (Sanger *et al.*, 1977) was used to determine the DNA sequence of denatured plasmid DNA templates or single-stranded M13 DNA templates, in the regions spanning cloning junctions or encompassing oligonucleotide-directed mutations. Firstly, 2µg of duplex plasmid DNA was denatured by mixing with 2µl of fresh 2M NaOH in a 10µl reaction volume followed by incubation for 10min at RT. After this time, 3µl of 3M NaAc (pH 4.5) and 7µl of dH_2O were added, followed by 60µl of ethanol, incubation on dry ice for 10min and centrifugation for 10min. The pellet of denatured DNA was washed with 80% ethanol, lyophilised and resuspended in 10µl dH_2O.

To anneal the sequencing primer to the DNA template: 8µl of denatured plasmid DNA or single stranded bacteriophage M13 template (2µg) (Section 2B.9) was mixed with 1µl of 5ng/µl oligonucleotide primer and 1µl 10 X seq buffer and incubated for 20min at 37°C. 2u Klenow fragment buffered in 10 X seq was added to the annealed DNA/primer reaction. Each annealing reaction was divided into four wells of a 96 well reaction plate (Nunclon), for the four chain termination reactions. 2µl of a termination mix (comprising 12µl of a sequencing mix containing one ddNTP + dNTPs (see below), plus 0.57µM dATP and 5µCi α^{32} P-dATP in a final volume of 14µl) was added to the relevant well and mixed by centrifugation at 1krpm for 10sec. The plate was incubated for 15min at 37°C to allow primed DNA synthesis until terminated by the incorporation of a ddNTP. Next, 2µl of a 'chase' mix containing all 4 dNTPs, each at 0.25mM, was added to the wells, mixed as before and incubated for 30min at 37°C; this elongates non-terminated chains into higher molecular weight DNA. 2µl of formamide dyes mix was added to each well, and the plate boiled for 2min immediately prior to loading samples into the spaces formed between the teeth of a sharks-tooth comb. Electrophoretic separation used 8% denaturing

polyacrylamide gels as described previously (Section 2B.4.4). The dried gel was exposed to Kodak X-omat S film overnight.

Nucleotide components of sequencing mixes

Sequencing solutions:

	dA-O	dT-O	dC-O	dG-O
5mM dTTP	20 (µl)	1	20	20
5mM dCTP	20	20	1	20
5mM 7-deaza dGTP	20	20	20	1
10 X TE seq	50	50	50	50
dH_2O	540	370	370	370

(TE seq: 10mM Tris.HCl (pH 8.0), 0.1mM EDTA)

Sequencing mixes:

	dN-O mix	ddNTP	dH_2O
T seq	500 (µl)	500 (600µM ddTTP)	0
C seq	500	105 (140µM ddCTP)	395
G seq	500	155 (200µM ddGTP)	345
A seq	500	250 (140µM ddATP)	250

2C Expression and purification of DNA binding domain peptides

2C.1 Preparation and transformation of competent bacteria for protein expression

A single fresh colony of *E. coli* strain BL21 (DE3) pLysS was inoculated into 10ml 2YT broth (supplemented with chloramphenicol, see Section 2A.6) and shaken at 37°C for around 3hr until the OD_{630} reached about 0.3. Cells were pelleted at 3krpm for 5min at 4°C, resuspended in 2.5ml ice cold 0.1M $MgCl_2$ and re-centrifuged at 3krpm for 5min at 4°C. The pellet was resuspended in 300µl ice cold 0.1M $CaCl_2$ and stored on ice until used. Approximately 10ng of plasmid DNA was transformed into 100µl competent bacteria as described previously (Section 2B.6) and plated onto agar containing antibiotics as appropriate (Section 2A.6).

2C.2 Bacterial expression and partial purification of DNA binding domain peptides

Several BL21 (DE3) pLysS colonies, freshly transformed with a T7 expression plasmid, were inoculated into 10ml 2YT broth (supplemented with appropriate antibiotics, Section 2A.6) and shaken overnight at 37°C. 100µl of the overnight culture was diluted into 100ml 2YT broth (plus antibiotics as above), and shaken at 37°C until the OD_{630} was about 0.5. IPTG was added to 0.1mM and peptide expression induced at 26°C, for a further 2hr. Bacteria were harvested by centrifugation at 5krpm for 5min at 4°C (Sorvall GS3 rotor) and resuspended in 0.1vol resuspension buffer (50mM HEPES (pH 7.2), 100mM NaCl, 6mM CHAPS, 1mM PMSF and 0.1mM DTT). After freeze-thawing, the bacteria were lysed by sonication for 4 X 10sec at 80-90W (using a Dawe soniprobe; Type 7532A). The solution was incubated with DNase I to 24µg/ml; RNase A to 100µg/ml and $MgCl_2$ to 6mM for 30min on ice. NaCl was added to 1M, followed by incubation for 10min on ice. Polymin-P was added to 0.2% and the mixture incubated for a further 10min. The resulting suspension was centrifuged at 15krpm for 15min at 4°C (Sorvall SS34 rotor). The supernatant was adjusted to 35% ammonium sulphate (by dropwise addition of a saturated solution) and incubated on ice for 40min. The precipitated proteins were collected by centrifugation at 10krmp for 10min at 4°C and dissolved in 300µl resuspension buffer. The extract was clarified by centrifugation at 10krpm for 5min at 4°C and the partially purified soluble peptides were stored at -70°C. Samples taken throughout the purification procedure were analysed by SDS PAGE as described below.

2C.3 SDS polyacrylamide gel electrophoresis (SDS PAGE) of proteins

Denatured proteins were resolved by electrophoresis through polyacrylamide minigels containing SDS (Laemmli, 1970) using the Bio-Rad miniprotean II apparatus and the glycine buffer system. Polyacrylamide resolving gels (a+b:b; 40:1; final acrylamide concentration between 10% and 16% as appropriate) contained 0.25 X RGB; 0.01vol 10% APS and 0.001vol TEMED were added to polymerise the gel. Butan-2-ol was layered over the resolving gel mix to give a smooth interface. Once set, the butanol was removed and the gel rinsed, prior to overlaying the stacking gel (5% polyacrylamide, 0.25 X SGB, 0.01vol 10% APS and 0.001vol TEMED). Samples and molecular weight markers (Sigma) (as required), were boiled for 2min in appropriate amounts of 3 X SDS boiling mix before loading and electrophoresis in tank buffer at ≥200V until the dye front reached the bottom of the gel. The tricine buffer system was used to give better resolution of low molecular weight proteins. Tricine polyacrylamide (a+b:b, 35:2) gels use a single

gel buffer for resolving and stacking gels, but the tricine system has separate tank buffers for lower and upper tank chambers.

Following electrophoresis, gels were either stained with Coomassie stain for 5min followed by destaining in methanol:acetic acid:water, (50:7:50), before photographing the gel on heat-sensitive paper using The Imager (Appligene). Alternatively, to locate radiolabelled proteins, gels were washed for 10min in Fix, followed by a 10min wash in En^{3}hance, rinsed in H_2O, dried under vacuum and placed in contact with Kodak X-omat S film for autoradiography (fluorography for ^3H) or placed against a phosphor screen (Molecular Dynamics) for PhosphorImager scanning analysis.

2C.4 Western blotting analysis of proteins

2C.4.1 *Electroblotting to nitrocellulose*

Proteins resolved on SDS PAGE minigels were transferred to nitrocellulose by the method of Towbin *et al.* (1979). The western blotting gel sandwich described below was prepared under Towbin buffer, and all components of the sandwich had been pre-soaked in this buffer. The minigel was placed on a square of Whatman 3MM paper, and a sheet of nitrocellulose paper (Schleicher and Schuell Inc.), cut to size, was placed on top of the gel, ensuring that no air bubbles were trapped between the layers. Another square of Whatman paper was placed on top and the sandwich arranged in the Bio-Rad mini transblot cell according to the manufacturer's instructions. Proteins were electroblotted to nitrocellulose at 0.28A for 3hr.

2C.4.2 *Immunodetection*

Nitrocellulose membranes were blocked in 200ml 3% gelatin in TTBS by shaking gently for 30min at 37°C. This was repeated with a fresh 200ml 3% gelatin in TTBS. The membrane was washed twice in 200ml TTBS for 5min before adding the first antibody. Polyclonal antiserum (rabbit) or specific monoclonal antibody (mouse) was diluted as appropriate (1:1000 to 1:10,000) in 1% gelatin in TTBS, and the membrane was incubated with 20ml of this first antibody dilution overnight at RT, with constant agitation. The membrane was washed 6 times with 100ml TTBS for 5min each, before adding the second antibody. Protein A horse radish peroxidase (HRP) conjugate or anti-mouse IgG whole molecule peroxidase conjugate (as appropriate) was diluted 1:1000 in 1% gelatin in TBS; the membrane was incubated with 10ml of this second antibody dilution for 1hr at RT, with agitation. After washing twice with 200ml TTBS, proteins were either detected using colour reagents or enhanced chemiluminescence (ECL). For colour detection of proteins, 60mg HRP colour development reagent (Bio-Rad) was mixed with 20ml ice cold methanol. To this was added a solution of 60µl H_2O_2 (30%

stock) in 100ml TBS. The membrane was washed twice with 200ml TBS, prior to incubation with the colour detection solution in the dark until protein bands appeared (about 10min). ECL detection of proteins used ECL western blotting detection reagents (Amersham). 1ml of each of the two ECL reagents were combined and poured onto the drained membrane, and after 1min the membrane was placed between two sheets of saran wrap and briefly exposed to Kodak X-omat S film.

2C.5 FPLC purification of VZV 140k DNA binding domain peptides

Partially purified extracts of VZV 140k DNA binding domain peptides were prepared from 1l bacterial cultures (scaled up from Section 2C.2) and purified by FPLC column chromatography. Purification used three preparative column steps; two sequential ion exchange chromatography steps, followed by a gel filtration step. The following FPLC buffer was used for all column elutions at the salt concentrations specified in Section 3C.2: 50mM HEPES (pH 7.5), 1mM PMSF, 0.1mM DTT, 0.01% CHAPS. PMSF and DTT were added immediately before use and all buffers were filtered through 22µm filters (Milllipore) and de-gassed before use. Columns were equilibrated with at least 5 column volumes of elution buffer prior to loading samples. Aliquots of column fractions were analysed by SDS PAGE (Section 2C.3) to identify peptide-containing fractions for subsequent chromatography steps.

2C.5.1 *Mono Q anion exchange column*

The first step in the FPLC purification procedure used a 20ml Mono Q anion exchange column (HR 16/10, Pharmacia) equilibrated in 0.1M NaCl FPLC buffer. 1ml of a pre-cleared partially purified peptide extract (from 330ml bacterial culture), diluted to 5ml in 0.1M NaCl FPLC buffer, was loaded onto the Mono Q column and eluted with a linear 0.1 to 1.0M NaCl gradient.

2C.5.2 *Mono S cation exchange column*

Pooled Mono Q flow-through fractions were diluted to 0.05M NaCl if necessary (Section 3C.2.4) or loaded directly onto an equilibrated 20ml Mono S cation exchange column (HR 16/10, Pharmacia) and a linear A to B NaCl gradient eluted the 140k peptides (where A is low salt and B is high salt FPLC buffer, as specified in Section 3C.2). The gradient was temporarily levelled off as appropriate, to obtain maximum separation of the VZV 140k peptides from bacterial contaminant proteins. Peak fractions were concentrated using Centricon 10 centrifugation filtration units (Amersham).

2C.5.3 ***Superose 12 gel filtration column***

The peptide sample from the concentrated Mono S peak fractions (in a volume of 250µl) was loaded onto a Superose 12 column (Pharmacia) equilibrated and eluted in an FPLC buffer of the NaCl concentration at which the 140k peptide eluted from the Mono S column. Peak fractions were pooled and concentrated by centrifugation filtration to yield >90% homogeneous peptide solutions, as determined by SDS PAGE analysis. Purified peptides were stored at -70°C.

2C.6 Purification of bacterially expressed GST-fusion peptides

Portions of the VZV 140k ORF fused to the glutathione-S-transferase (GST) gene of the parasitic helminth *Schistosoma japonicum* were bacterially expressed and purified following the method of Ausubel *et al.* (1993b). pGEX-fusion vectors were transformed into competent DH5α bacteria as previously described (Section 2B.6). 1l cultures of transformed bacteria were grown in 2YT broth (plus antibiotics, see Section 2A.6) at 37°C until the OD_{630} reached about 0.4. Peptide expression was induced for a further 2hr at 37°C by addition of IPTG and bacteria harvested (as described in Section 2C.2). The pellet was resuspended in 10ml PBSA and placed at -20°C. The thawed cells were lysed by incubation with 100µl 40mg/ml lysozyme and 1ml 5% Na-deoxycholate for 40min on ice. Triton X-100 was added to a final concentration of 1%, after a further 5min the solution was centrifuged at 20krpm for 10min at 4°C (Sorvall SS34 rotor). 1ml of a 50% slurry of glutathione agarose beads was incubated with the supernatant for about 1hr at RT. The beads were spun down briefly and washed twice with 20ml 1% Triton X-100 in PBSA. The beads were then washed in 20ml 'wash' buffer (50mM Tris.HCl (pH 7.5), 150mM NaCl). After a final wash in 20ml 'thrombin cleavage' buffer (2.5mM $CaCl_2$ in 'wash' buffer) the beads were resuspended in 1ml 'thrombin cleavage' buffer containing 16µl of 0.4µg/ml thrombin (1% (w/w) fusion protein). Thrombin cleaved the fusion peptides at the thrombin cleavage site, C-terminal to the GST sequences, during a 1hr incubation at RT. The beads were spun out to yield a supernatant containing the purified VZV 140k peptide that was released by the thrombin cleavage. Samples were taken throughout this procedure for SDS PAGE analysis.

2C.7 *In vitro* transcription-translation in rabbit reticulocyte lysates

DNA binding domain peptides were expressed *in vitro* using the TNT T7 Coupled Reticulocyte Lysate System (Promega), in the presence of ^{35}S-methionine (or ^{3}H-leucine), following the manufacturer's instructions. Briefly, T7 expression plasmids (0.001-1µg) were transcribed by T7 RNA polymerase and translated by components of

the lysate for 90min at 31°C, in 50µl reaction mixes containing the following: 25µl TNT rabbit reticulocyte lysate, 2µl TNT reaction buffer, 1µl TNT T7 RNA polymerase, 1µl 1mM amino acid mix minus methionine (or minus leucine), 40µCi ^{35}S-methionine (or 25µCi ^{3}H-leucine) and 40u RQ1 RNasin ribonuclease inhibitor. Low plasmid amounts were used (about 0.001µg) when extracts were prepared for DNA binding analyses, to prevent the plasmid DNA from sequestering the DNA binding peptides. The amount of each plasmid required to produce equivalent peptide yields was determined empirically for co-translation experiments, due to variations in the efficiency of expression from different T7 expression constructs. For co-immunoprecipitation experiments, 1µg of each plasmid DNA was included in a 50µl translation reaction. The translation products (1-5µl) were analysed by SDS PAGE followed by autoradiography (Section 2C.3), before use in gel retardation (Section 2D.2) or co-immunoprecipitation analyses (Section 2E.4).

2D DNA binding analyses

2D.1 Radiolabelling of DNA

2D.1.1 *End-labelling of DNA fragments and synthetic oligonucleotides*

End-labelled DNA probes were produced by filling-in the overhangs produced by restriction enzyme digestion or by annealing overlapping oligonucleotides, in the presence of a radiolabelled dNTP. To construct radiolabelled restriction fragments, the plasmid was cut with the first restriction enzyme (Section 2B.3), followed by heat inactivation of the enzyme for 10min at 65°C. Then 20µCi of a suitable α^{32} P-dNTP (choice depending on the template nucleotides), 1µl of 1mM of the remaining three dNTPs and 2u Klenow fragment were added and buffer conditions adjusted to those of buffer M (by dilution or addition of salt). After incubation for 30min at 37°C, 1µl 20 X dNTPs was added and incubation continued for a further 10min. Heat inactivation of the Klenow fragment was followed by digestion with the appropriate second restriction enzyme. Labelling of both restriction fragment ends was achieved by cutting with both enzymes prior to the labelling reaction. To radiolabel complementary overlapping oligonucleotides, equal amounts of each single stranded oligonucleotide were annealed together for 5min at 80°C and slowly cooled to RT. The labelling reaction contained 20ng duplex oligonucleotide and was otherwise exactly as above. Radiolabelled DNA probes were purified on either 5% (100-400bp) or 8% (20-100bp) non-denaturing polyacrylamide gels (Section 2B.4.3) and eluted as described previously (Section 2B.4.5).

2D.1.2 ***Preparation of 'prime cut' radiolabelled DNA fragments***

DNA fragment probes with equivalent specific activities were prepared by the 'prime cut' method (Biggin *et al.*, 1984), for comparative gel retardation analyses using DNA fragments with point mutations in the HSV-1 IE3 consensus binding site (Section 3A.9). Single-stranded template DNAs (3µl) (Section 2B.9) from the M13IE3EB series of bacteriophage (including the mutagenised IE3 binding site), were annealed with 2µl M13 universal primer in the presence of 1µl 10 X seq buffer, for 90min at 60°C. 5µl of a label mix (2µl ^{32}P dATP, 4µl 0.5mM three remaining dNTPs, 1µl 11.8µM ATP, 2µl 10 X seq buffer, 4u Klenow fragment, 3µl dH_2O) was dispensed into each tube and incubated for 30min at RT. After this time, 1µl of 20 X dNTPs was added and incubated for a further 10min. Heat inactivation of the enzyme was followed by restriction enzyme digests of the duplex bacteriophage and the resulting 'prime cut' radiolabelled DNA fragments were resolved on 5% non-denaturing polyacrylamide gels (Section 2B.4.3) and the required fragment isolated as described previously (Section 2B.4.5). Radioactivity concentrations of probes were equalised by Cerenkov counting and making suitable dilutions in TE. 'Prime cut' probes used for any series of experiments were prepared using the same label mix, and a wild type probe was prepared in each series of probe preparations, to allow standardisation of the DNA binding results.

2D.2 Gel retardation assays

The abilities of peptides to bind specifically to radiolabelled DNA molecules was assayed by gel retardation. Optimal binding conditions were determined empirically, depending on the particular type of assay, and were as follows:

Appropriate amounts of bacterially expressed DNA binding domain peptides or HeLa nuclear extracts were mixed with buffer giving final concentrations of 10mM Tris.HCl (pH8.0), 1mM EDTA, 100mM NaCl, 0.1% NP40, 50µg/ml BSA. Non-specific competitor (1µg polydI.polydC or 0.5µg pUC9) was also included. The mixture was incubated for 20min on ice with 1ng of end-labelled DNA fragment probe or freshly made 'prime cut' probe. *In vitro* translated extracts were incubated in buffer containing 10mM Tris.HCl (pH8.0), 1mM EDTA, 100mM NaCl, 0.2% CHAPS, 100mM DTT.

For analyses of the stability of protein-DNA complexes, 1µg of unlabelled oligonucleotide specific competitor was added to incubation reactions for increasing lengths of time before electrophoresis. For analyses of the specificity of DNA binding, the protein was incubated with increasing amounts of unlabelled specific competitor for 20min, prior to incubation with the radiolabelled probe.

After incubation, 0.2vol of 5 X non-SDS loading buffer was added and the samples applied to non-denaturing 4% polyacrylamide gels (a+b:b; 30:1) (1.5mm wide X 20cm long) containing 0.5 X TBE (diluted from 10 X TBE stock), 0.01vol 10% APS and 0.001vol TEMED. Gels were electrophoresed in 0.5 X TBE at 150V for about 3hr at 4°C. When required, 35cm gels were used to increase the resolution of complexes. After electrophoresis, gels were soaked in Fix for 10min and dried down under vacuum. Complexed and unbound probe were detected by autoradiography of dried gels using Kodak X-omat S film and intensifying screens at -20°C for 1-14 days.

Where appropriate, the extent of binding was determined by scintillation counting (Beckman LS 5000CE) of dried gel slices (in 5ml ecoscint A) corresponding to the positions of free and bound probe, previously located by autoradiography. To allow for background values contributed by complexes with *E. coli* proteins, dried gel slices from control lanes were also cut out and counted. Control values were subtracted from the sample counts, to allow calculation of the percentage of total probe bound. For more recent experiments, quantification was by PhosphorImager analysis of gels exposed to phosphor screens, using Image-Quant software (Molecular Dynamics).

2D.3 Pore exclusion electrophoresis

The sizes of protein-DNA complexes were estimated by electrophoresis through polyacrylamide gels of graded porosity, using the method of Andersson *et al.* (1972). Partially purified bacterially expressed peptides were incubated with radiolabelled DNA as described above. Samples were loaded onto linear 4% to 24% non-denaturing polyacrylamide gradient gels (a+b:b, 30:1) containing 0.5 X TBE, 0.01vol 10% APS and 0.001vol TEMED. An appropriate selection of protein standards of known molecular weights (Sigma) were run on the same gel. Gels were run in 0.5 X TBE at 150V at 4°C until the macromolecules had reached their equilibrium position in the polyacrylamide molecular sieve (after about 40hr). The migration distance of protein-DNA complexes, visualised by autoradiography, were compared to those of the Coomassie-stained protein standards, to give an estimate of the size of complexes.

2D.4 DNase I footprinting assay

Site specific protection of DNA fragments from partial DNase I digestion, by DNA binding domain peptides, was determined by modification of the method of Galas and Schmitz (1978). DNA binding reactions contained 4ng of target DNA fragment (radiolabelled at a single end) and suitable amounts of partially purified bacterially expressed peptides in buffer giving final concentrations of 10mM Tris.HCl (pH 7.5),

1mM EDTA, 10mM BME, 0.1% CHAPS and 100mM NaCl. After a 20min incubation at RT, $MgCl_2$ to 10mM and $CaCl_2$ to 2mM were added. Then 0.0625u of DNase I was added to the protein-DNA solution, and after 60sec at RT, digestion was halted by addition of 20µg proteinase K in 100µl of a buffer containing 1% SDS, 100µg/ml tRNA, 200mM NaCl, 20mM EDTA, and incubating at 50°C for 20min. The DNA fragments were purified by phenol and chloroform extractions, ethanol precipitation and resuspended in formamide dyes mix. A G+A Maxam and Gilbert chemical sequencing reaction (Maxam and Gilbert, 1980) was run on the same gel as the footprinting reactions, to provide a sequencing ladder for orientation of the footprint. The G+A reaction was as follows: 4ng of the lyophilised DNA probe fragment was resuspended in 2% diphenylamine, 1mM EDTA, 66% formic acid and incubated for 5min at RT. The mix was diluted in 45µl dH_2O and extracted three times with 400µl diethylether, and then lyophilised. The DNA pellet was resuspended in 100µl of freshly diluted 1M piperidine and incubated at 90°C for 20min to cleave modified purine bases. The products were purified by two ethanol precipitations before resuspension in formamide dyes mix. Samples were applied to 8% (probes ≤100bp) or 12% (probes >100bp) denaturing polyacrylamide sequencing gels and electrophoresed as described previously (Section 2B.4.4). The dried gel was exposed to Kodak X-omat S film for 1-14 days with an intensifying screen at -20°C.

2E Physical analyses of DNA binding domain peptides

2E.1 Glutaraldehyde cross-linking

The ability of DNA binding domain peptides to form multimers was analysed by incubation with glutaraldehyde, which irreversibly and covalently cross-links α- and ε-amino groups of lysine residues of proteins in close proximity (Peters and Richards, 1977). Amounts of partially purified or FPLC purified peptides that yield clearly visible bands upon Coomassie-staining of SDS PAGE gels (approximately 4µg), were incubated in a buffer containing 50mM HEPES (pH 7.5), 0.1M NaCl, 1mM PMSF, either with or without 0.01% glutaraldehyde, for 30min at 31°C. Parallel incubations, including 4µg of monomeric protein standard (bovine carbonic anhydrase; Sigma) were performed to indicate the background level of non-specific cross-linking by glutaraldehyde. The products of the cross-linking were resolved by SDS PAGE (Section 2C.3) and visualised

by Coomassie blue staining, or transferred to nitrocellulose and immunodetected (Section 2C.4).

2E.2 Superdex 75 analytical gel filtration chromatography

The native molecular mass of DNA binding domain peptides in solution was determined using the high resolution FPLC Superdex 75 gel filtration column (Pharmacia). About 50µg of partially purified, FPLC purified DNA binding domain peptide, or gel filtration protein standard (Sigma) was loaded in a volume of 200µl onto the Superdex 75 column equilibrated in a buffer containing 50mM HEPES (pH 7.5), 0.3M NaCl, 1mM PMSF, 0.1mM DTT, 0.01% CHAPS. The proteins were eluted from the column in this buffer and detected by the uv monitor ($0D_{280}$ setting of 0.05) upon exit from the column. The identity of proteins in the peaks were verified by SDS PAGE analysis (Section 2C.3) of column fractions. The accurate retention volume of each protein peak was provided by evaluation of the stored peak data for each column run using the software in the FPLC equipment. The approximate molecular mass of a DNA binding domain peptide in solution was derived from a standard curve produced from the retention volumes of a range of protein standards (Fig. 3C.6).

2E.3 Glycerol gradient centrifugation sedimentation analysis

Sedimentation on 10-30% glycerol gradients was used to estimate the true molecular sizes of DNA binding domain peptides, as compared to protein standards. 5ml linear gradients of 10-30% glycerol in 50mM HEPES (pH 7.5), 0.5M NaCl were prepared in AH650 tubes (Dupont). Air bubbles were carefully removed and the gradients allowed to settle for 30min at RT. Approximately 10µg of FPLC purified peptide and appropriate protein standards were layered on top of the gradient in a 100µl volume containing 50mM HEPES (pH7.5), 0.5M NaCl, 0.1% CHAPS and 1mM PMSF. Gradients were centrifuged at 40krpm for 46hr at 14°C (Sorvall AH650 rotor), with low acceleration and deceleration. 200µl fractions were collected from the top of gradients and analysed by SDS PAGE (Section 2C.3).

2E.4 Co-immunoprecipitation of *in vitro* translated peptides

Interactions between DNA binding domain peptides in solution were detected by co-immunoprecipitation of peptides co-translated *in vitro* (Section 2C.7) with an epitope tagged peptide (see Fig. 3D.10). Extracts from rabbit reticulocyte lysate translation reactions (50µl) were pre-cleared by incubation for 30min at 4°C with 10vol IP buffer

containing 5µl normal mouse serum, 12.5µl sheep anti-mouse Ig (SαM) and 60µl 50% slurry of protein A sepharose beads. After a 2min spin, 20µl of the pre-cleared supernatant was incubated for 2hr at 4°C with 2µl MAb, 5µl (SαM) and 60µl 50% slurry protein A sepharose in a volume made up to 400µl with IP buffer. Immune complexes were pelleted by centrifugation for 2min. The pellet was washed 3 times with IP buffer, and then washed with a buffer containing 0.6M LiCl, 0.1M Tris.HCl (pH 7.5), 1% BME. The pellet was finally washed in PBSA and resuspended in 20µl 1 X SDS boiling mix and boiled for 5min. The beads were pelleted and the supernatant applied to an SDS PAGE gel (Section 2C.3), with radiolabelled protein markers as appropriate. After electrophoresis, the proteins were electroblotted to nitrocellulose membranes (Section 2C.4.1). Co-immunoprecipitated proteins were detected by autoradiography or phosphorimaging of the dried membranes. Inclusion of a cocktail of protease inhibitors throughout the immunoprecipitation procedure reduced proteolytic degradation (bestatin, leupeptin, aprotinin, PMSF, at concentrations recommended by the manufacturer).

2E.5 Urea denaturation - renaturation

Equal amounts (approximately 50µg) of partially purified or FPLC purified bacterially expressed peptides were diluted to 400µl in buffer R; where R buffer contained 50mM HEPES (pH 7.5), 0.3M NaCl, 0.01% CHAPS, 1mM PMSF, 0.1mM DTT. Peptide dilutions containing either a single peptide or a mix of two peptides were denatured by dialysis overnight at 4°C against 400ml buffer R containing 8M urea (previously deionised with Amberlite resin). The dialysis sacks were transferred to 400ml buffer R for 4hr at 4°C to renature the peptides. Samples (around 1µl) of the peptide dilutions, both before and after this treatment, were analysed by gel retardation assay (Section 2D.2).

2F Tissue culture

2F.1 Growth of tissue culture cells

Eukaryotic cells for transfection experiments and virus work were grown in tissue culture disposable plasticware (Nunc, Sterilin Ltd.). BHK, Vero and WS HeLa cells were passaged in 175cm^2 flasks in the relevant tissue culture media (described in Section 2A.7) and were grown at 37°C under 5% CO_2 in a humidified incubator. Confluent monolayers

were harvested in 10ml medium after a 20ml versene wash and brief trypsinisation with 20ml trypsin:versene (1:4). Cells were resuspended by pipetting and used directly, although BHK and Vero cells remained viable for several days at 4°C.

2F.2 Calcium phosphate mediated transfection for CAT assay

WS HeLa, BHK or Vero cells were seeded at 1 X 10^6 cells per 50mm petri dish and were transfected by the calcium phosphate precipitation method 24hr later (Corsalo and Pearson, 1981). Transfections used 4µg of CAT reporter constructs, which carry a heterologous promoter-leader sequence fused to the chloramphenicol acetyltransferase (CAT) open reading frame. Varying amounts of effector plasmids (0-10µg) were used as appropriate and pUC9 was added to equalise the amount of DNA within experiments. All transfection buffers were freshly filtered using 0.2µm sterile Acrodiscs (Gelman Sciences). TEP buffer was added to the DNA to bring the volume to 140µl, followed by dropwise addition of 20µl 2M $CaCl_2$. The DNA solution was in turn added dropwise to 160µl of 2 X HBS buffer. After standing for 30min at RT, the precipitate was sprinkled onto the cell monolayer. After 24hr at 37°C the cells were washed and re-fed with fresh medium. 24hr later, cells were scraped into 3ml TEN, pelleted at 2krpm for 5min at 4°C, resuspended in 0.25M Tris, pH 7.8 and sonicated. The sonicates were cleared by a 2krpm spin for 5min, and stored at -20°C.

2F.3 Chloramphenicol acetyltransferase (CAT) assay

Activation of gene expression from CAT-reporter plasmids by transiently transfected effector plasmids, was assayed by determining the CAT enzyme activity in cell sonicates (Section 2F.2), using a method adapted from Gorman *et al.* (1982). Assay mixtures contained 1µl of 50mM acetyl CoA, 1µl of ^{14}C-chloramphenicol, 13µl dH_2O and 25µl of cell extract. Typically, 5µl cell sonicates from transactivation experiments were diluted to 25µl in 0.25M Tris, pH 7.8, and assayed for 30min at 37°C. Sonicates from repression experiment were assayed undiluted for 4hr at 37°C. After incubation, the reactions were extracted with 200µl of ethyl acetate, lyophilised, resuspended in 20µl ethyl acetate and spotted onto 0.25mm Silica gel TLC plates (Camlab). Ascending chromatography was performed using 100ml of 19:1, chloroform:methanol tank buffer until the solvent front had reached the top of the plate. The dried TLC plate was placed against a phosphor screen. The percentage conversion to product was calculated by quantitating the radioactivity in the substrate and product spots using a PhosphorImager scanner with Image-Quant software (Molecular Dynamics).

2F.4 Determination of protein concentration of cell extracts

The total concentration of proteins in each cell extract used in the CAT assay was determined using the method of Bradford (1976). 195µl of a 1:4 dilution of protein assay reagent (Bio-Rad) was aliquoted into the wells of a flat bottomed microtitre plate (all dilutions were in 0.25M Tris, pH 7.8). To the 'sample' wells were added 5µl of a 1:9 dilution of each extract, to the 'standard' wells were added 5µl of stock dilutions of BSA containing 0.2, 0.4, 0.6, 0.8, 1.0, 1.2 or 1.4mg/ml BSA. 5µl of 0.25M Tris (pH 7.8) was added to the 'blank' well. After around 20min at RT, the absorbances at 610nm were measured using an Anthos Labtec HT2 plate reader, which subtracted the 'blank' reading, derived a standard curve from the BSA 'standard' readings and calculated the protein content of sample extracts. For each extract, the percentage conversion from substrate to product per microgram protein was derived by dividing the percentage conversion to product, calculated from the CAT assay (Section 2F.3), by the average protein concentration from duplicate protein determinations.

2F.5 Liposome mediated transfection (Lipofection)

BHK cells were seeded at 5 X 10^5 cells per 50mm petri dish and 24hrs later were transfected with plasmid DNA, using a liposome-mediated DNA transfection procedure based on the method of Felgner *et al.* (1987) as described below. Liposomes were kindly prepared by Mrs M. Elliott as follows: 4µg DDAB was added to 1ml DOPE (supplied as 10mg/ml in chloroform) and the chloroform evaporated using a N_2 supply. The lyophilised pellet was resuspended in 10ml sterile dH_20 and sonicated until clear, to yield liposomes ready for use. 6µg plasmid DNA was made up to a volume of 170µl with a buffer containing 150mM NaCl, 20mM HEPES.NaOH pH 7.4. The DNA solution was mixed with 2ml Opti-MEM (low serum medium, Gibco), before addition of 30µl liposomes, mixing and standing for 10min at RT. Meanwhile, the sub-confluent monolayers were washed once with Eagle's medium (Gibco) (plus penicillin/streptomycin) (Section 2A.7) and drained thoroughly. The liposome-DNA mix was transferred to the monolayer and incubated at 37°C. After 4 hours, 2ml Eagle's medium plus 5% calf serum (CS) (and penicillin/streptomycin as previously) was added to each plate and incubation continued for 24hr. Monolayers were washed, re-fed with BHK medium and incubated for a further 24hr before immunofluorescence analysis or viral infection (see below).

2F.6 Immunofluorescent staining of VZV 140k in transfected cells

BHK cells prepared as described above were washed with PBSc and the monolayer fixed with 1.8% formaldehyde, 2% sucrose in PBSc for 10min at RT. The monolayer was washed 3 times with PBSc, and permeabilised with 0.5% NP40, 10% sucrose, 1% CS in PBSc for 5min at 20°C. Cells were washed 3 times with PBSc plus 1% CS, and dried briefly at 37°C. Polyclonal antiserum 109 was diluted 1:500 in PBSc plus 1% CS, then one drop of this first antibody dilution was placed on a marked spot on the monolayer for 1hr at RT (care was taken to prevent cells from drying out during incubations). The monolayer was washed 3 times with PBSc plus 1% CS, dried, then incubated for 1hr in the dark at RT with 10µl of a 1:30 dilution of FITC-conjugated sheep anti-rabbit Ig in PBSc, placed on the same spot as before. The monolayer was washed 3 times with PBSc, and dried. 10µl of 90% glycerol in PBSc was placed on the spot, covered with a coverslip and the cells observed by uv fluorescence microscopy. Suitable fields of view were photographed using Kodak ASA400 black and white film.

2F.7 Production of recombinant virus

Lipofection (Section 2F.5) was used to introduce linearised plasmid DNA into BHK cells as the initial step in the production of recombinant HSV-1 based viruses. BHK cells were seeded at 4 X 10^5 per 35mm petri dish and grown for 24hr at 37°C. Transfection mixes included 2µg of *Pst*I linearised plasmid DNA (previously 'genecleaned' and checked by agarose gel electrophoresis for complete linearisation), the total DNA in each transfection was made upto 3µg with calf thymus (CT) DNA. A negative control plate was transfected with 3µg CT DNA to indicate the background level of wild type HSV-1 in the D30EBA viral stock. After lipofection (using half the volumes given in Section 2F.5), cells were grown for 24hr at 37°C before infecting with virus. A D30EBA viral stock, (made as described below) was diluted 1:1000 in BHK medium, cell monolayers were drained and 100µl of diluted virus sprinkled onto each plate. The plates were incubated at 37°C, with agitation every 10min. After 1hr the monolayers were re-fed with BHK medium and cells grown for a further 24hr, before washing and replacing fresh medium. Cells were grown for 2-3 days at 37°C until cpe became visible. Cells were harvested into the medium, sonicated for 10sec and stored at -70°C. These sonicates were used as viral stocks without further treatment.

2F.8 Titration of virus stocks

Virus stocks produced as described above were titrated by plaque assay. 10-fold serial dilutions of stocks were prepared in BHK serum. The growth medium was removed from

confluent BHK monolayers (seeded at 1 X 10^6 cells per 35mm petri dishes 24hr previously). 100µl of diluted virus was added to each dish and was adsorbed to the cells by incubation for 1hr at 37°C (with gentle agitation every 10min). The monolayers were overlaid with Eagle's medium (Gibco) plus 5% human serum (Flow) to prevent secondary plaque formation. After incubation for 3–5 days at 37°C, cells were fixed with Cidex (Johnson and Johnson) for around 30min and stained with Giemsa stain for a further 30min. Plates were washed and plaques counted using a dissection microscope.

2G Computing

Computer analysis of DNA and protein sequences were performed on the MicroVAX II computer in this Institute using the software package form the University of Wisconsin Genetics Computer Group (GCG Inc., Madison, Wis.). For details refer to Genetics Computer Group (1991) Program Manual for the GCG Package, Version 7, April 1991, 575 Science Drive, Madison, Wisconsin, USA 53711.

Chapter 3: Results

3A Expression and DNA binding characteristics of the VZV 140k DNA binding domain

3A.1 Strategy

The majority of studies described in this thesis utilise peptide fragments that encompass the DNA binding domain of the product of VZV gene 62, 140k. The approach of analysing the DNA binding domain in isolation from other viral sequences was adopted because of the difficulties encountered in a previous attempt to detect sequence-specific DNA binding by the intact VZV 140k polypeptide (Disney, 1990). As discussed in the Introduction (Section 1C.1.10), the bacterially expressed DNA binding domain of HSV-1 Vmw175 is a stable separable domain that binds to DNA with a very similar affinity and specificity to the intact polypeptide. Therefore it is reasonable to assume that the binding properties of the VZV 140k DNA binding domain would reflect those of the full 140k polypeptide. The DNA binding domain of HSV-1 Vmw175 comprises the highly conserved region 2 of the protein, although the C-terminal portion of the less highly conserved region 1 is additionally required for DNA binding (Figs. 1C.2 and 1C.3). Due to the functional and sequence similarities between HSV-1 Vmw175 and VZV 140k (see Section 1C.2.1 and Fig. 1C.1), it seemed likely that the DNA binding activity of 140k would reside in comparable regions of the protein to that of Vmw175. Immediately prior to the commencement of these PhD studies, Wu and Wilcox communicated before publication (1991), that this was indeed the case. They expressed in *E. coli*, 140k codons 417 to 647, preceded by 323 codons derived from the bacterial *trp*E ORF. They found that this fusion peptide, FP617 (Fig. 3A.1), bound DNA in a sequence-specific manner (Wu and Wilcox, 1991; Section 1C.2.2). The experiments in this Section use the same regions of 140k (amino acid residues 417 to 647) expressed as a non-fusion polypeptide, VT2 (Fig. 3A.1), to eliminate the possible influence of a large stretch of heterologous *E. coli* coding region on the properties of the fusion protein. By

homology to the HSV-1 Vmw175 DNA binding domain (Wu and Wilcox, 1990; Pizer *et al.*, 1991), the VT2 peptide includes sequences from the C-terminus of region 1 and the whole of region 2 of the 140k protein. The intention was to confirm and extend the findings of Wu and Wilcox (1991) using the non-fusion VZV 140k DNA binding domain.

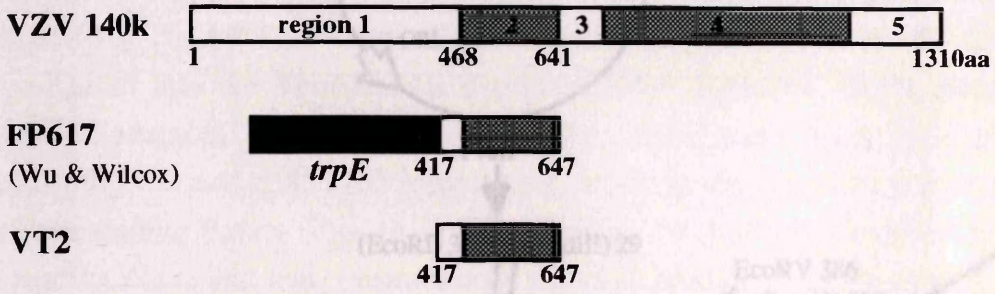
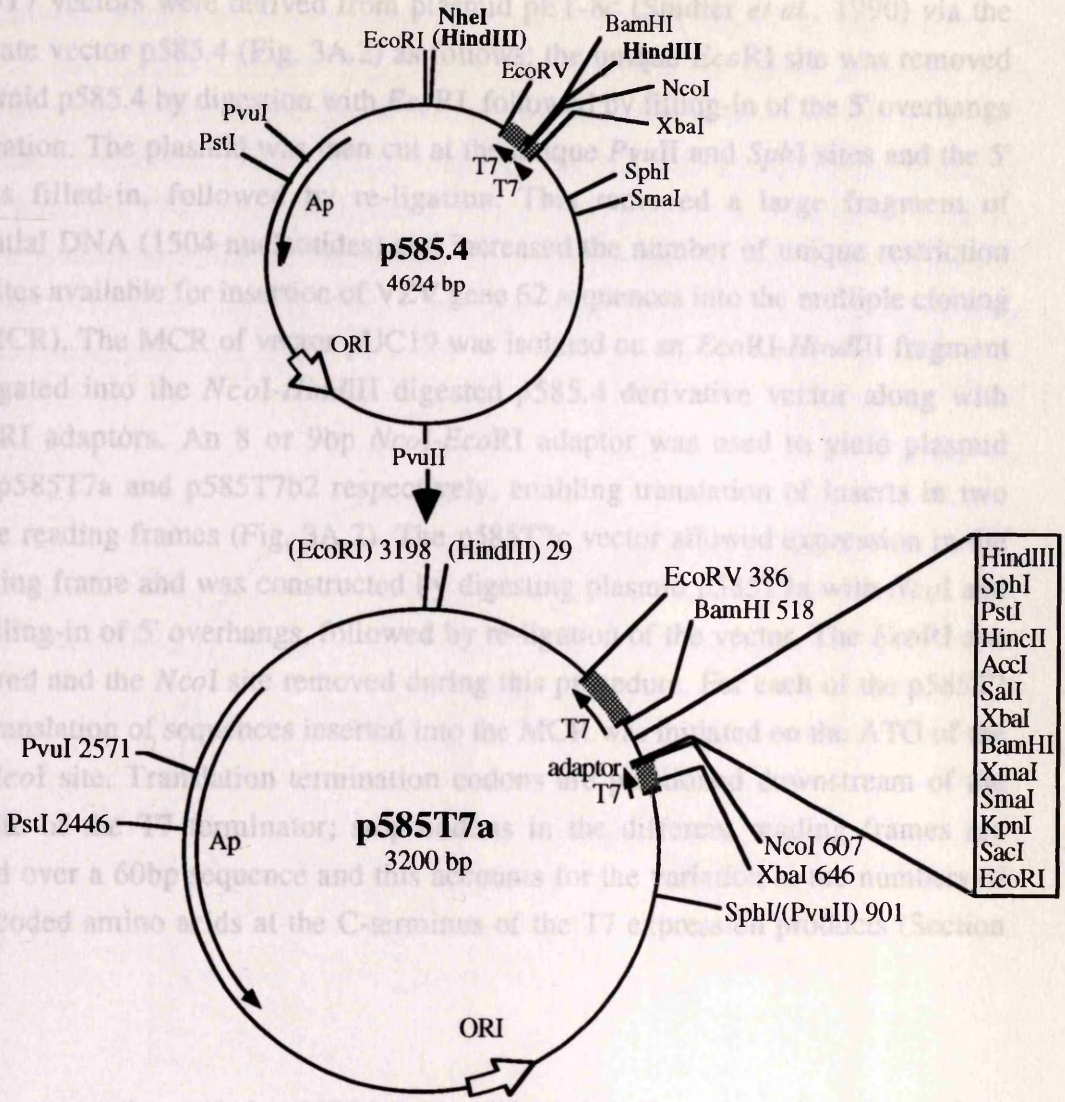


Figure 3A.1 Diagrammatic representations of VZV 140k and peptides encompassing its DNA binding domain. The 1310 amino acid VZV 140k protein (shown at the top) is divided into five regions based on the degree of amino acid similarity with HSV-1 Vmw175 (see Section 1C.1.8). The highly conserved regions 2 and 4 are shown shaded and the first and last amino acid coordinates of region 2 are shown. Peptides FP617 and VT2 encompass the 140k DNA binding domain and include region 2. The FP617 peptide contains 140k residues 417 to 647 fused to codons 1-323 of the *trp*E gene of the *E. coli* tryptophan operon, represented by the black box (Wu and Wilcox, 1991). The non-fusion VT2 peptide used in this study contains 140k residues 417 to 647.

During the course of the analyses described in this Section (3A), the DNA binding characteristics of the VZV 140k DNA binding domain were compared to those of the homologous domain of HSV-1 Vmw175. The aim was to identify possible variations in DNA binding activities that could account for the functional differences reported between the 140k and Vmw175 proteins in transient transfection assays (see Section 1C.2.3). The T7 bacterial expression system of Studier *et al.* (1990) was adopted for the expression of portions of the VZV gene 62 coding sequence, as this approach had proved highly effective for the expression of the Vmw175 DNA binding domain (Pizer *et al.*, 1991).

3A.2 Construction of the p585T7 series of T7 expression vectors

The p585T7 series of vectors were constructed to facilitate fusion of portions of the VZV gene 62 coding sequence to the T7 transcription-translation signals and to enable expression of inserts in the required reading frame.



	*Nco*I	*Eco*RI
p585T7a	5'-CC **ATG GCT** GAA TTC G-3'	
	3'-GG TAC CGA **CTT AA**G C-5'	
	*Nco*I	*Eco*RI
p585T7b2	5'-CC **ATG GCT** AGA ATT CG-3'	
	3'-GG TAC CGA **TCT TAA** GC-5'	
	(*Nco*I) *Eco*RI	
p585T7c	5'-CC **ATG** AAT TCG-3'	
	3'-GG TAC **TTA AG**C-5'	

Figure 3A.2 Construction of the p585T7 series of vectors. Only relevant restriction enzyme sites are shown. p585T7 vectors carry fragments of bacteriophage T7 DNA (indicated by shaded boxes) to direct expression in the direction indicated by the arrows; these include the $\phi 10$ promoter, the *s10* translation initiation region for the gene 10 protein and the Tϕ transcription termination signal which were originally inserted into the *Bam*HI site of pBR322 to make plasmid pET-8c (Studier *et al.*, 1990). Plasmid p585.4 was previously derived from pET-8c by digestion with *Hind*III, filling-in and re-ligation to remove the *Hind*III site, a novel *Nhe*I site was introduced in the process; also a new *Hind*III site was introduced between the T7 promoter and terminator regions (these differences to pET-8c are shown in bold). P585T7 vectors have the unique *Eco*RI site removed, a large *Sph*I-*Pvu*II fragment removed and the pUC19 MCR inserted between the *Nco*I and *Hind*III sites with the aid of *Nco*I-*Eco*RI adaptors (indicated by the black box) as described in the text. The reading frames of the three p585T7 vectors were determined by the version of adaptor inserted and are indicated at the bottom of the panel; adaptor derived sequences are shown in bold and the restriction enzyme target sequences are underlined.

The p585T7 vectors were derived from plasmid pET-8c (Studier *et al.*, 1990) via the intermediate vector p585.4 (Fig. 3A.2) as follows: the unique *Eco*RI site was removed from plasmid p585.4 by digestion with *Eco*RI, followed by filling-in of the 5' overhangs and re-ligation. The plasmid was then cut at the unique *Pvu*II and *Sph*I sites and the 5' overhangs filled-in, followed by re-ligation. This removed a large fragment of non-essential DNA (1504 nucleotides) and increased the number of unique restriction enzyme sites available for insertion of VZV gene 62 sequences into the multiple cloning region (MCR). The MCR of vector pUC19 was isolated on an *Eco*RI-*Hind*III fragment and co-ligated into the *Nco*I-*Hind*III digested p585.4 derivative vector along with *Nco*I-*Eco*RI adaptors. An 8 or 9bp *Nco*I-*Eco*RI adaptor was used to yield plasmid versions p585T7a and p585T7b2 respectively, enabling translation of inserts in two alternative reading frames (Fig. 3A.2). The p585T7c vector allowed expression in the third reading frame and was constructed by digesting plasmid p585T7a with *Nco*I and *Eco*RI, filling-in of 5' overhangs, followed by re-ligation of the vector. The *Eco*RI site was restored and the *Nco*I site removed during this procedure. For each of the p585T7 vectors, translation of sequences inserted into the MCR was initiated on the ATG of the adaptor *Nco*I site. Translation termination codons are positioned downstream of the *Bam*HI site of the T7 terminator; stop codons in the different reading frames are distributed over a 60bp sequence and this accounts for the variation in the numbers of vector-encoded amino acids at the C-terminus of the T7 expression products (Section 3B.3).

3A.3 Construction of the VZV 140k DNA binding domain expression vectors p585T7aVT2 and p585T7aVT2X

VZV gene 62 sequences spanning the 140k DNA binding domain were inserted into the p585T7a expression vector as follows: an *Eco*RI site introduced by linker insertion into codon 418 of plasmid pV17, was utilised for isolation of an *Eco*RI-*Bst*EII fragment carrying gene 62 codons 417 to 647. Plasmid pV17 was cut with *Bst*EII and the 5' overhang filled-in, before a second restriction digest with *Eco*RI. The resulting *Eco*RI-(*Bst*EII) fragment was inserted between the adaptor *Eco*RI site and the MCR *Sma*I site of vector p585T7a. The version of the *Nco*I-*Eco*RI adaptor in plasmid p585T7a maintained the correct 140k peptide reading frame. *Bam*HI digestion, followed by re-ligation of the plasmid construct removed DNA sequences between the *Bam*HI sites of the multiple cloning and T7 terminator regions. Plasmid p585T7aVT2 encoded a 256 residue peptide, designated VT2 (Fig. 3A.3). Plasmid p585T7aVT2X was derived from p585T7aVT2 by blunt end insertion of a duplex self-complementary 18mer *Xba*I linker oligonucleotide into the filled-in *Bam*HI site of plasmid p585T7aVT2. The linker

oligonucleotide contained stop codons (underlined) in all three reading frames as follows: 5'-AATTAATCTAGATTAATT-3' (the *Xba*I site is shown in bold). This process removed 16 vector-encoded residues from the C-terminal end of the expressed protein that were present in the VT2 polypeptide. The 240 residue expression product of plasmid p585T7aVT2X was designated VT2X. The p585T7aVT2X version of the plasmid was made in addition to p585T7aVT2, due to the finding that 'X' versions of the HSV-1 Vmw175 DNA binding domain peptides were generally more stable than the non-X versions (Pizer *et al.*, 1991).

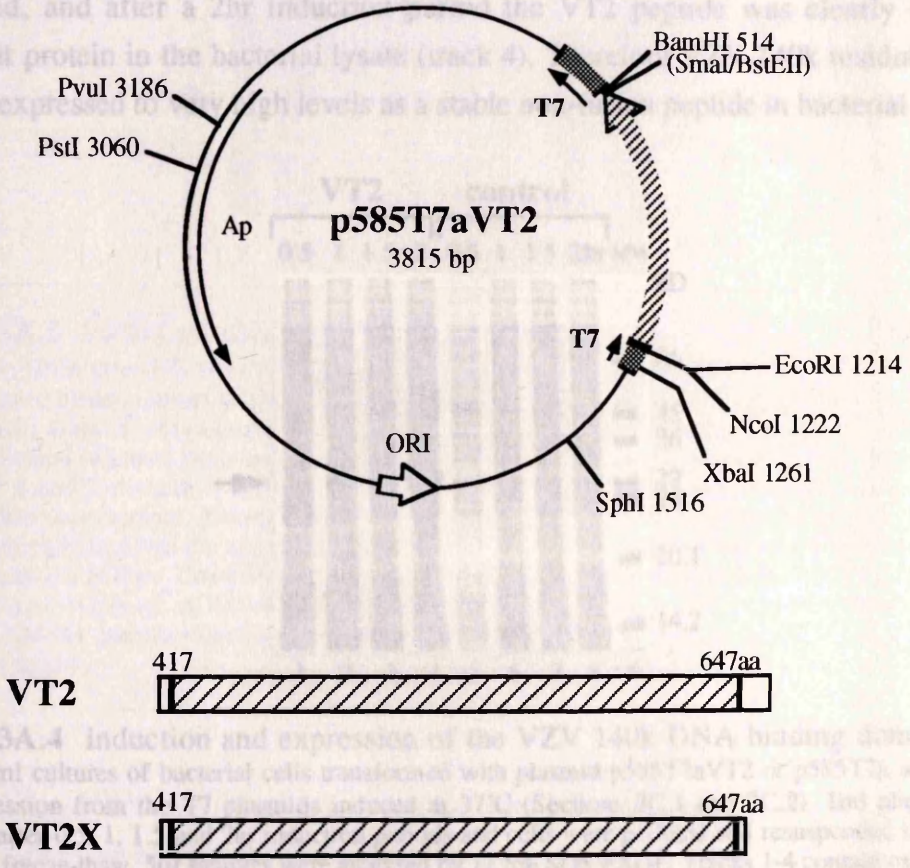


Figure 3A.3 Plasmid p585T7aVT2 and the VZV 140k DNA binding domain expression products VT2 and VT2X. Only selected restriction enzyme sites are marked on the p585T7aVT2 plasmid. The location of the T7 transcriptional initiation and termination signals are indicated by shaded boxes and arrows indicate the direction of expression. The black box represents the *Nco*I-*Eco*RI oligonucleotide adaptor. The striped box represents the VZV gene 62 coding region (codons 417-647) on an *Eco*RI-*Bst*EII fragment from plasmid pV17, inserted into the vector p585T7a as discussed in the text. Plasmid p585T7aVT2X differs from p585T7aVT2 by having an *Xba*I linker inserted into the *Bam*HI site 3' to the gene 62 coding sequences. Shown at the bottom are representations of the expression products of plasmid p585T7aVT2, peptide VT2; and the product of p585T7aVT2X, peptide VT2X. VT2 includes 4 N-terminal adaptor encoded residues (MAEF) followed by residues 417 to 647 of the 140k protein (striped box) and 21 vector-encoded residues (white boxes represent vector-encoded residues). VT2X is identical to peptide VT2 apart from having only 5 C-terminal vector-encoded residues (GDQLI). The 256 residue VT2 peptide has a predicted size of 28.4kD and the 240 residue peptide VT2X has a predicted size of 26.7kD.

3A.4 Bacterial expression and partial purification of the VZV 140k DNA binding domain

To determine whether VZV gene 62 codons 417 to 647 yield a stable domain when expressed as a non-fusion peptide, plasmid p585T7aVT2 was transformed into the BL21 (pLysS) DE3 strain of *E. coli*, a culture grown and expression from the T7 promoter induced as described in Sections 2C.1 and 2C.2. Samples of the culture were taken after various induction times at 37°C for analysis by SDS PAGE (Fig. 3A.4). A novel peptide of the approximate predicted size was apparent in the VT2 tracks at the position marked by the arrow. The VT2 band became more prevalent as the duration of induction was increased, and after a 2hr induction period the VT2 peptide was clearly the most abundant protein in the bacterial lysate (track 4). Therefore VZV 140k residues 417 to 647 are expressed to very high levels as a stable non-fusion peptide in bacterial cells.

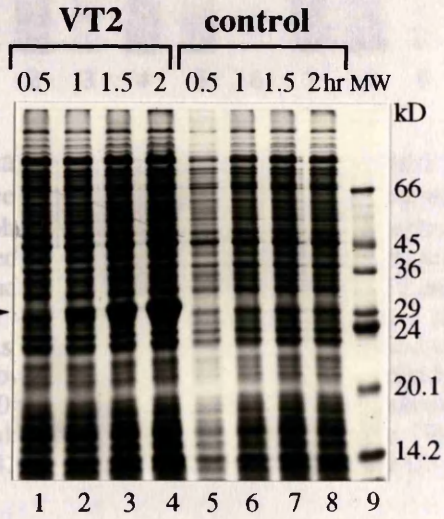


Figure 3A.4 Induction and expression of the VZV 140k DNA binding domain in *E. coli*. 10ml cultures of bacterial cells transformed with plasmid p585T7aVT2 or p585T7a were grown and expression from the T7 plasmids induced at 37°C (Sections 2C.1 and 2C.2). 1ml aliquots were removed after 0.5, 1, 1.5 and 2hr induction periods and cells were pelleted and resuspended in 20µl 1 X TE. After freeze-thaw, 5µl samples were analysed by 12.5% SDS PAGE. Tracks 1-4 contain extracts from cells transformed with plasmid p585T7aVT2 (VT2) and tracks 5-8 contain extracts from cells transformed with p585T7a (control). Track 9 contains low molecular weight markers: BSA, 66kD; albumin, 45kD; glyceraldehyde-3-phosphate dehydrogenase, 36kD; carbonic anhydrase, 29kD; trypsinogen, 24kD; trypsin inhibitor, 20.1kD; lactalbumin, 14.2kD. The position of peptide VT2 is marked by an arrow; the particular T7 expression construct used in this experiment expressed a VT2 peptide which included an extra 15 vector-derived residues at the C-terminus of the peptide having a predicted size of 30.1kD (271 amino acids) rather than the expected size of 28.4kD (256 amino acids). The form of the p585T7aVT2 vector described in Section 3A.3 was used in subsequent experiments.

A 2hr induction period did not appear to be deleterious to the level of VT2 expression. Trial experiments found that induction at 26°C tended to produce a more stable peptide, whilst not greatly reducing the peptide yield (not shown). Therefore a 2hr induction period at 26°C was used in all subsequent experiments. To determine whether the 140k

DNA binding domain was soluble, partially purified extracts of the VT2 and VT2X peptides were prepared as described in Section 2C.2. Figure 3A.5 shows an SDS PAGE analysis of samples taken at stages throughout the purification procedure; clearly both the VT2 and VT2X peptides were highly soluble. In the following DNA binding analyses, peptides VT2 and VT2X can be considered to be interchangeable as they were found to have identical DNA binding properties.

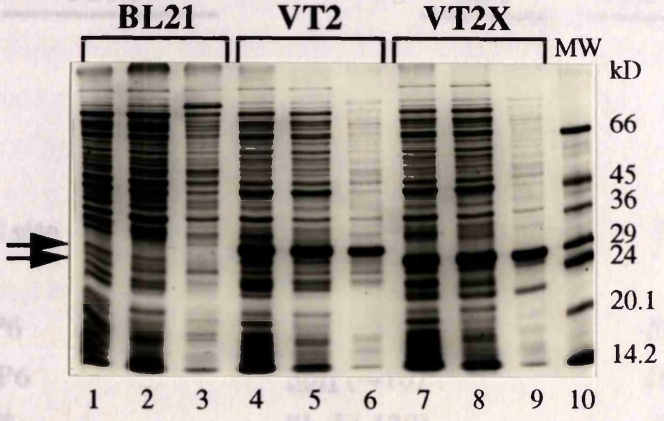
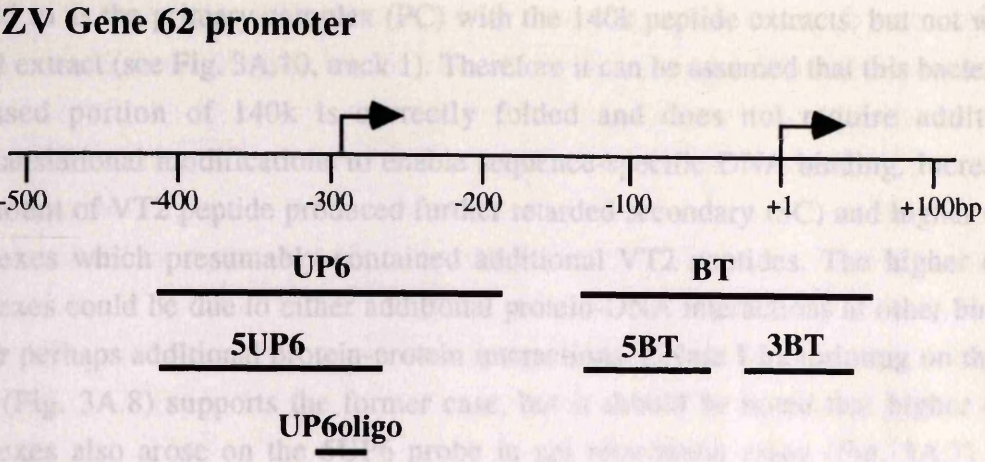


Figure 3A.5 Partial purification of the bacterially expressed VZV 140k DNA binding domain. Transformed *E. coli* were grown, expression from the T7 plasmids induced and crude extracts were prepared by ammonium sulphate precipitation (Section 2C.2); samples were analysed by 12.5% SDS PAGE. BL21 is an extract prepared from induced bacteria lacking an expression plasmid, VT2 and VT2X indicate extracts prepared from bacteria carrying plasmids p585T7aVT2 and p585T7aVT2X respectively. Tracks 1, 4 and 7 contain lysed bacteria after harvesting; tracks 2, 5 and 8 contain post polymin P precipitation supernatants. Tracks 3, 6 and 9 contain partially purified extracts after 35% ammonium sulphate precipitation and the amount of extract in these tracks corresponds to that produced from 5ml of induced bacterial culture. Track 10 contains low molecular weight standards: BSA, 66kD; albumin, 45kD; glyceraldehyde-3-phosphate dehydrogenase, 36kD; carbonic anhydrase, 29kD; trypsinogen, 24kD; trypsin inhibitor, 20.1kD; lactalbumin, 14.2kD. The positions of peptides VT2 (28.4kD) and VT2X (26.7kD) are marked by arrows.

3A.5 The VZV 140k DNA binding domain interacts with its own VZV gene 62 promoter in gel retardation assays

Initial investigations into the DNA binding activity of the 140k DNA binding domain peptides used probes derived from the VZV gene 62 promoter, as the autoregulation of the gene 62 promoter seen in transient transfection assays (Disney *et al.*, 1990) was likely to involve sequence-specific binding, by analogy to the situation with the HSV-1 Vmw175 protein (Section 1B.2.3). Moreover, Wu and Wilcox (1991) had recently detected interactions between the gene 62 promoter and their 140k DNA binding domain fusion peptide (FP617, Fig. 3A.1).



5' RE site	3' RE site	5' RE site
UP6	*Sal*I (-410)	*Rsa*I (-188)
5UP6	*Sal*I (-410)	*Dra*I (-270)
BT	*Xho*I (-130)	*Bgl*II (+57)
5BT	*Xho*I (-130)	*Taq*I (-51)
3BT	*Rsa*I (-23)	*Ava*I (+43)
UP6oligo	*Ava*I (-311)	*Bam*HI (-283)

Figure 3A.6 The VZV gene 62 promoter fragment probes. The numbers on the scale correspond to nucleotide positions in the gene 62 promoter. The locations of the experimentally determined mRNA start site at +1 (McKee *et al.*, 1990) and the previously proposed mRNA start site at -287 (Felser *et al.*, 1988) are indicated by the arrows (see Fig. 1B.4). Below the scale are shown the relative positions of the probe fragments derived from the gene 62 promoter; the table gives the identity and position of restriction enzyme (RE) sites used for their isolation (see Fig. 1B.4). Underlined restriction sites indicate the radiolabelled end of each fragment probe (Section 2D.1.1). Both ends of the UP6oligo were labelled. Fragment probes were isolated from plasmid p140, apart from the BT fragment which was isolated from plasmid p140BT which has a *Bgl*II site artificially introduced at nucleotide +57. In the case of the UP6oligo probe, the numbers refer to gene 62 promoter sequences (shown in bold below) included in the probe. The UP6oligo probe comprises a pair of 35mer oligonucleotides synthesised to yield *Ava*I and *Bam*HI sticky ends flanking the gene 62 sequences when annealed together; a sequence homology to the 5' motif of the HSV-1 Vmw175 consensus binding site is underlined:

5'-CCGAG**ATATTATATATCGTCTGTAGACACACGAT**G-3'
3'-C**TATAATATATAGCAGACATCTGTGTGCTA**CCTAG-5'

DNA binding by the partially purified 140k DNA binding domain peptides was tested by gel retardation assay using the UP6 and BT probes, which span large stretches of gene 62 promoter sequences (Fig. 3A.6). The shorter probes (5UP6, 3BT and 5BT) were subsequently used to further analyse locations of specific sequences recognised by the 140k peptide. All gel retardation assays included 1µg non-specific competitor (polydI:polydC) and typical results with the VT2 peptide are shown in Figure 3A.7. Each of the gene 62 fragment probes analysed gave a discrete protein-DNA complex

referred to as the primary complex (PC) with the 140k peptide extracts, but not with a control extract (see Fig. 3A.10, track 1). Therefore it can be assumed that this bacterially expressed portion of 140k is correctly folded and does not require additional post-translational modifications to enable sequence-specific DNA binding. Increasing the amount of VT2 peptide produced further retarded secondary (SC) and higher order complexes which presumably contained additional VT2 peptides. The higher order complexes could be due to either additional protein-DNA interactions at other binding sites or perhaps additional protein-protein interactions. DNase I footprinting on the BT probe (Fig. 3A.8) supports the former case, but it should be noted that higher order complexes also arose on the 5UP6 probe in gel retardation assay (Fig. 3A.7) even though only one binding site was apparent by DNase I footprinting in this region of the gene 62 promoter (Fig. 3A.8). It is possible that in addition to the sequence-specific binding activity of VT2, there exists transient or low-specificity VT2-DNA interactions that fail to be detected by DNase I footprinting analysis.

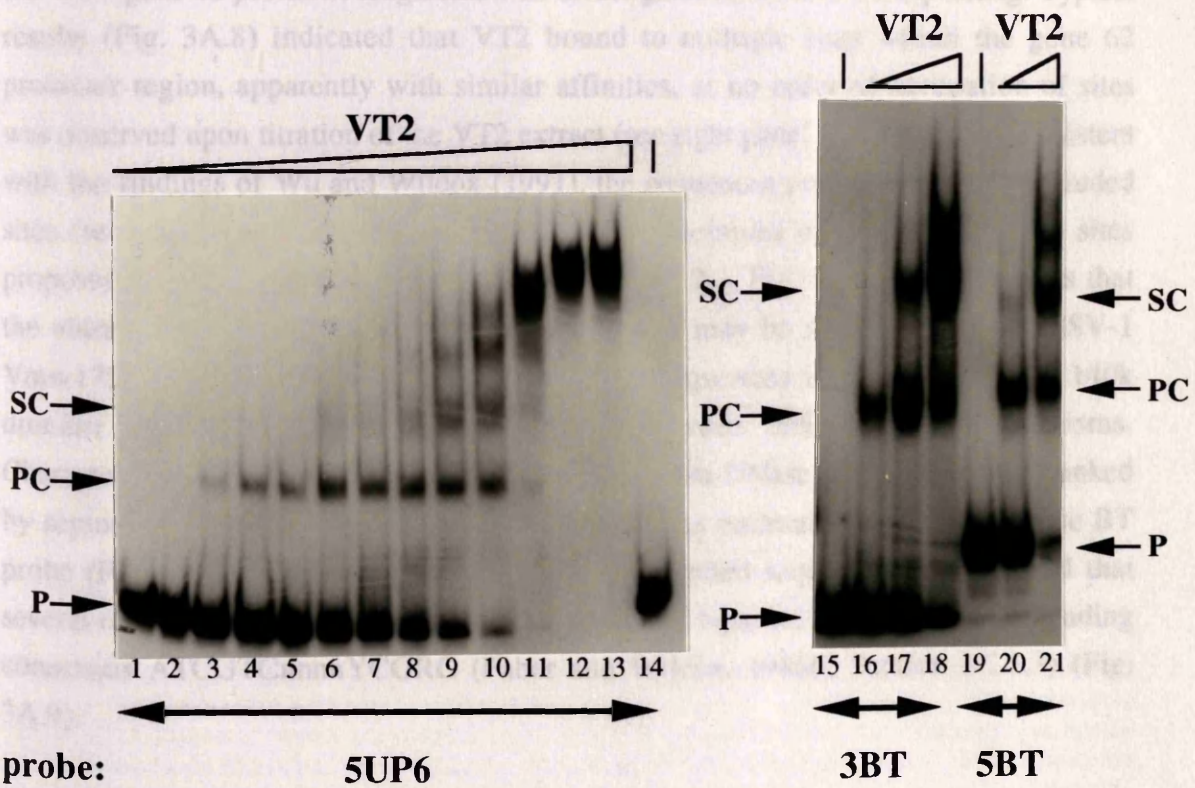


Figure 3A.7 Gel retardation analysis of the VZV gene 62 promoter region. Partially purified VT2 peptide was incubated with the indicated gene 62 fragment probes (described in Fig. 3A.6) and the products were resolved by native gel electrophoresis, as described in Section 2D.2. Tracks 14, 15 and 19 contained no protein extract. The remaining tracks included a VT2 extract; tracks 1-13 contained a range of amounts from 0.05 to 1.0µl; tracks 16 and 20 contained 0.5µl; 17 and 21 contained 1.0µl and 18 contained 2µl; the VT2 extract used had a VT2 concentration of approximately 0.2µg/µl as determined by comparison with BSA standards on SDS PAGE gels (not shown). The position of free probe is indicated by 'P', the primary complex by 'PC' and the secondary complex by 'SC', higher order complexes remain unlabelled.

The specificity of the VT2-DNA binding interaction was tested by competition experiments (data not shown) (see Section 2D.2); the primary VT2-DNA complex was 50% competed from the UP6oligo and 5UP6 DNA targets by a 200-fold molar excess of unlabelled UP6oligo, whereas the 5BT and 3BT probes, and also the IE3L probe (Fig. 3A.11) required slightly higher amounts of competitor to achieve a similar degree of complex reduction. Approximately 4μg of polydI:polydC or 750ng of pUC9 competed off 50% of the VT2-DNA complexes. These results suggest that the 140k DNA binding domain recognises these different promoter sequences with similar moderate affinities.

3A.6 The VZV 140k DNA binding domain recognises multiple VZV gene 62 promoter sequences in DNase I footprinting analysis

The specificity of the interaction between the 140k DNA binding domain peptide and the VZV gene 62 promoter fragments was investigated by DNase I footprinting. Typical results (Fig. 3A.8) indicated that VT2 bound to multiple sites within the gene 62 promoter region, apparently with similar affinities, as no order of occupation of sites was observed upon titration of the VT2 extract (see right panel of Fig. 3A.8). Consistent with the findings of Wu and Wilcox (1991), the sequences protected by VT2 included sites (referred to as U and P4) in the immediate vicinities of both mRNA start sites proposed for the gene 62 promoter (see Section 1C.2.3, Fig. 1B.4). This suggests that the autoregulatory mechanism used by VZV 140k may be similar to that of HSV-1 Vmw175, although the recognition of multiple sequences is particular to the 140k domain and indicates possible differences between their precise mechanisms. Characteristically, the sites of protection by VT2 from DNase I digestion were flanked by regions of DNase I hypersensitivity, and this was particularly apparent on the BT probe (Fig. 3A.8). Upon close analysis of the protected sequences it was found that several of the sites contained sequences similar to the bipartite HSV-1 Vmw175 binding consensus ATCGTCnnnnYCGRC (Faber and Wilcox, 1986b; Section 1C.1.3) (Fig. 3A.9).

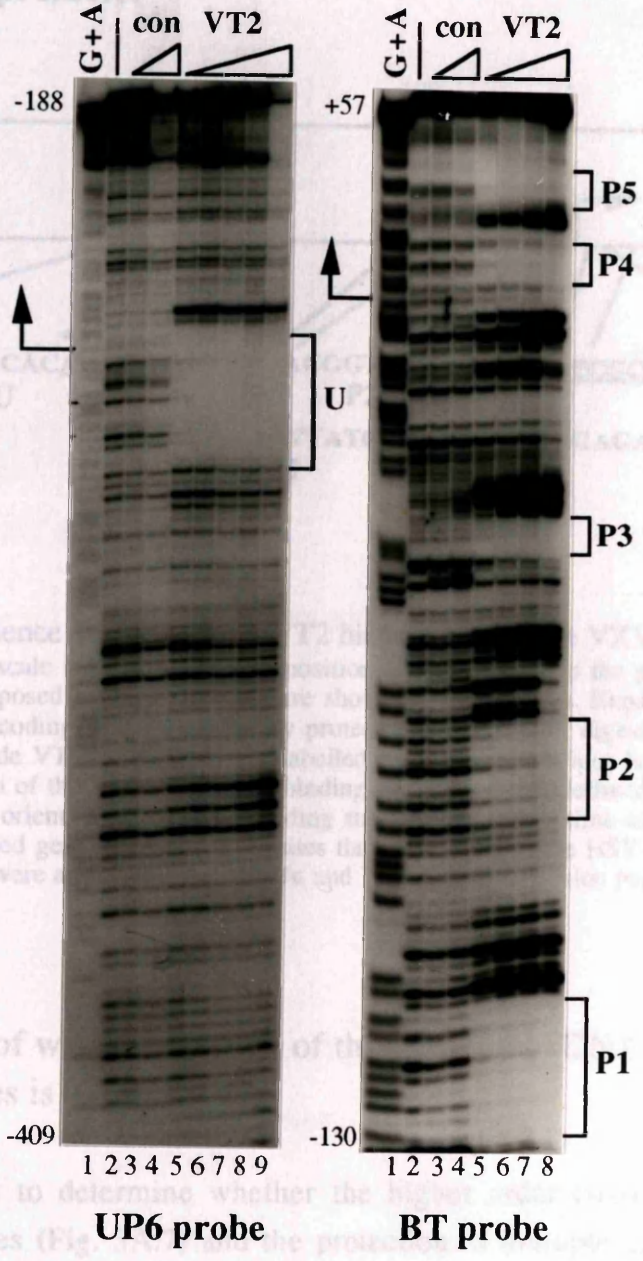


Figure 3A.8 DNase I footprints on the VZV gene 62 promoter region. DNase I footprinting assays were performed with end-labelled UP6 and BT DNA fragments as described in Section 2D.4. Details of probes are given in Fig. 3A.6 and the gene 62 nucleotides included in each probe are indicated on the footprints, in each case the non-coding strand was radiolabelled. Arrows indicate the proposed transcriptional initiation sites at nucleotides -287 and +1 (see Fig. 1B.4). Brackets representing regions of protection from DNase I digestion are labelled U (-302 to -280), P1(-107 to -97), P2 (-81 to -74), P3 (-52 to -47), P4 (+6 to +18) and P5 (+27 to +46). In both panels: track 1 contained the relevant Maxam and Gilbert G+A sequencing reaction, track 2 contained no extract, track 3 and 4 contained partially purified bacterial extract carrying the p585T7a vector. Tracks labelled VT2 contained partially purified bacterial VT2 extract. Triangles represent progressively increasing amounts of extract. On the UP6 probe: 0.4, 1, 2, 5 and 10µl VT2 extract was included in tracks 5-9. On the BT probe: 0.05, 0.2, 0.5 and 1µl of extract was included in tracks 5-8. Amounts of control extract corresponded to the same range of amounts of the VT2 extract used in each experiment.

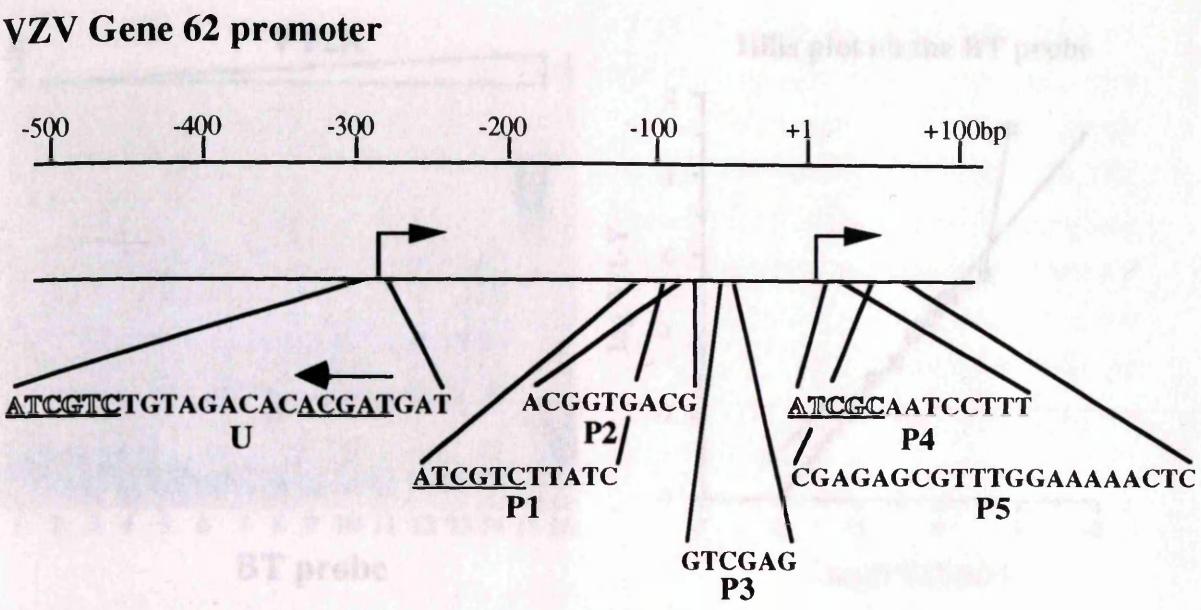
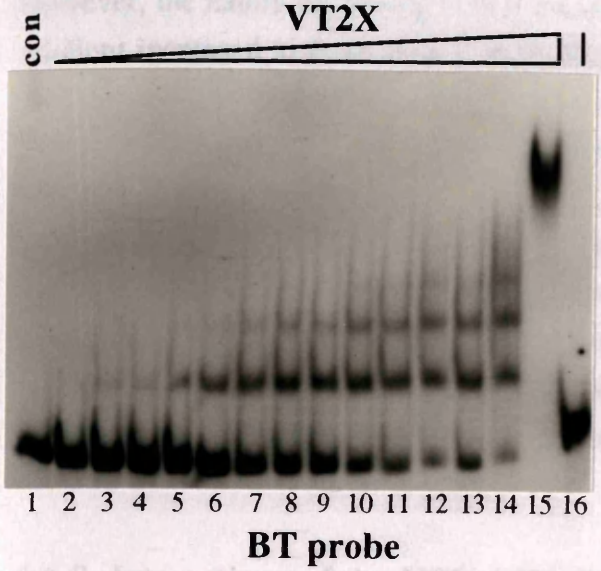
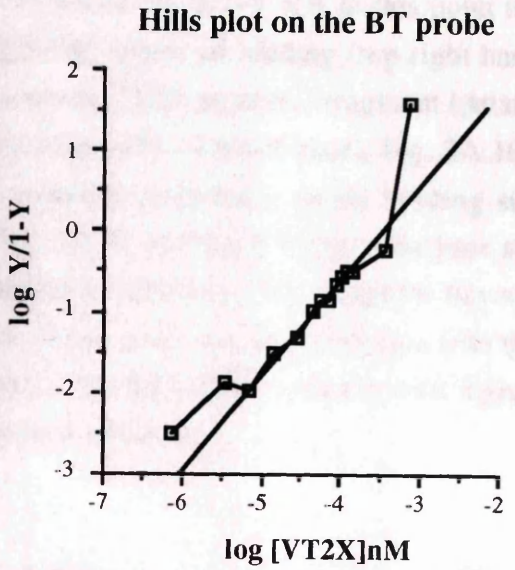


Figure 3A.9 Sequence and location of VT2 binding sites in the VZV gene 62 promoter. Numbers on the top scale indicate nucleotide positions corresponding to the gene 62 promoter. The locations of both proposed mRNA start sites are shown by bent arrows. Expansions of the gene 62 promoter sequences (coding strand) specifically protected from DNase I digestion by the 140k DNA binding domain peptide VT2 in Fig. 3A.8 are labelled U, P1 to P5. Obvious homologies to the motif found in the 5' portion of the HSV-1 Vmw175 binding consensus are underlined, the arrow indicates a motif in the opposite orientation on the non-coding strand. Shown in outline are the sequences in the vicinity of the proposed gene 62 mRNA start sites that are similar to the HSV-1 Vmw175 consensus sequence; these sites were also recognised by Wu and Wilcox's FP617 fusion peptide (Wu and Wilcox, 1991; Section 1C.2.2).

3A.7 Analysis of whether binding of the VZV 140k DNA binding domain to additional sites is cooperative

It was of interest to determine whether the higher order complexes seen in gel retardation analyses (Fig. 3A.7) and the protection of multiple binding sites in the DNase I footprinting (Fig. 3A.8) arose due to facilitated or cooperative binding of 140k DNA binding domain peptides. A wide range of VT2X peptide concentrations were incubated with the BT probe which has multiple binding sites for the 140k DNA binding domain (as discerned in Fig 3A.8), and the complexes were resolved by gel retardation assay. A Hills plot (top right panel, Fig. 3A.10) was derived from the information in the autoradiograph from one such experiment (top left panel, Fig. 3A.10). The part of the plot that reflects occupation of second, third and fourth binding sites (indicated by the best fit line in the graph at the top right panel of Fig. 3A.10) has a gradient of approximately 1 and indicates that the initial binding of additional VT2X peptides to the BT probe occurred in an independent or non-cooperative manner (see legend to Fig. 3A.10).





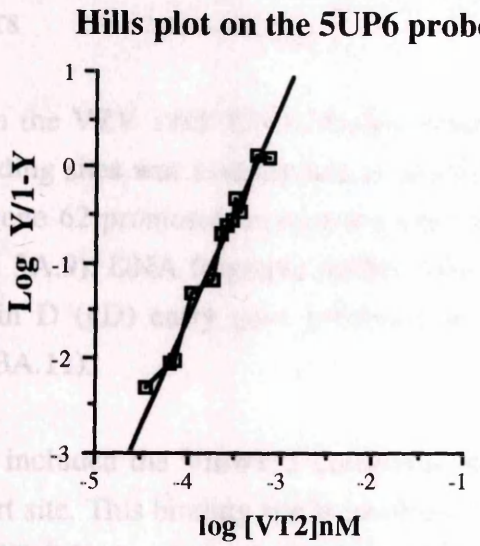


Figure 3A.10 Analysis into the nature of binding of additional VZV 140k DNA binding domain peptides. The top left hand panel shows a gel retardation analysis upon the BT probe (details given in Fig. 3A.6), tracks 2-15 included increasing amounts (0.001-1µl) of a partially purified VT2X peptide extract which had a VT2X concentration of 0.38µg/µl (determined by comparison with BSA standards on SDS PAGE gels). Track 1 contains a control extract from plasmid p585T7a transformed bacteria and track 16 contains no extract. The top right hand panel shows the corresponding Hills plot where Y is the fraction of binding sites occupied; the total number of binding sites in the BT probe was assumed to be 5 (from the 5 complexes apparent on longer exposure of the autoradiograph shown here and also from the DNase I footprinting assay in Figure 3A.8). This analysis was based on the assumption that the PC indicates one occupied binding site, SC indicates two occupied binding sites and so on. The fraction of binding sites occupied at each concentration of VT2X was calculated as follows: the scintillation count for an individual protein-DNA complex (see Section 2D.2) was multiplied by the number of sites occupied in that particular complex, the combined value for all complexes in a single track was then divided by 5 times the total value of scintillation counts of free and complexed probe. The best fit lines in both graphs were fitted by eye. A Hills plot gradient of more than one indicates cooperative binding and if the gradient is equal to one binding is non-cooperative. The lower panel shows a Hills plot from a gel retardation analysis of VT2 binding to the 5UP6 probe (details in Fig. 3A.6), very similar to that shown on the left hand panel of Figure 3A.7. Data was calculated on the assumption that the number of binding sites was 5 (from the 5 successive protein-DNA complexes formed in gel retardation analyses) on the 5UP6 probe. Re-calculation of the data for 4 or 6 binding sites produced very little alteration in the gradient of the plot shown here.

However, the nature of binding is best measured where log Y/1-Y = 0, at this point the gradient increased to more than 1, suggesting cooperativity of binding (top right hand panel, Fig. 3A.10). A corresponding analysis upon the 5UP6 promoter fragment (details in Fig. 3A.6) produced a Hills gradient of approximately 2 (lower panel, Fig. 3A.10). This particular region of the VZV gene 62 promoter included a single binding site apparent by DNase I footprinting analysis (Fig. 3A.8). Although further analyses are required, this initial data suggests that the loading of additional VT2 peptides beyond the initial occupation of specific binding sites is cooperative, and may correlate with the widening of regions of protection from DNase I digestion that occurred with higher peptide levels in the DNase I footprinting analyses (see below).

3A.8 Interactions of the VZV 140k DNA binding domain with the HSV-1 IE3 and gD promoters

The interaction between the VZV 140k DNA binding domain and functional HSV-1 Vmw175 consensus binding sites was investigated, as several of the VT2 binding sites identified in the VZV gene 62 promoter showed sequence homology to the Vmw175 binding consensus (Fig. 3A.9). DNA fragment probes from the HSV-1 IE3 promoter and HSV-1 glycoprotein D (gD) early gene promoter were used in DNA binding analyses (details in Fig. 3A.11).

The HSV-1 IE3 probes included the Vmw175 consensus binding site (Muller, 1987) spanning the mRNA start site. This binding site is involved in the autoregulation of the IE3 promoter (Section 1B.2.3). In view of the observation that VZV 140k does not repress the IE3 promoter (Disney *et al.*, 1990), the relationship between VT2 and the IE3 binding site was analysed, initially by gel retardation assay. Incubation of the VT2 or VT2X peptide with the IE3 probes gave specific protein-DNA complexes and additional multiple complexes arose on titration of peptide, as noted earlier with the VZV gene 62 promoter probes (Section 3A.5). Figure 3A.15, track 2 shows two such complexes on the IE3L probe, while track 3 shows a primary complex produced with an equivalent amount of the Vmw175 DNA binding domain peptide, I9X. Therefore the 140k DNA binding domain has an affinity for this region of the HSV-1 IE3 promoter similar to that of the Vmw175 domain. The 140k DNA binding domain *trpE* fusion protein (FP617) has been reported to have a binding activity of roughly one-tenth that of the equivalent Vmw175 fusion protein on HSV-1 IE3 and gD promoter fragment probes (Wu and Wilcox, 1991; Betz and Wydoski, 1993).

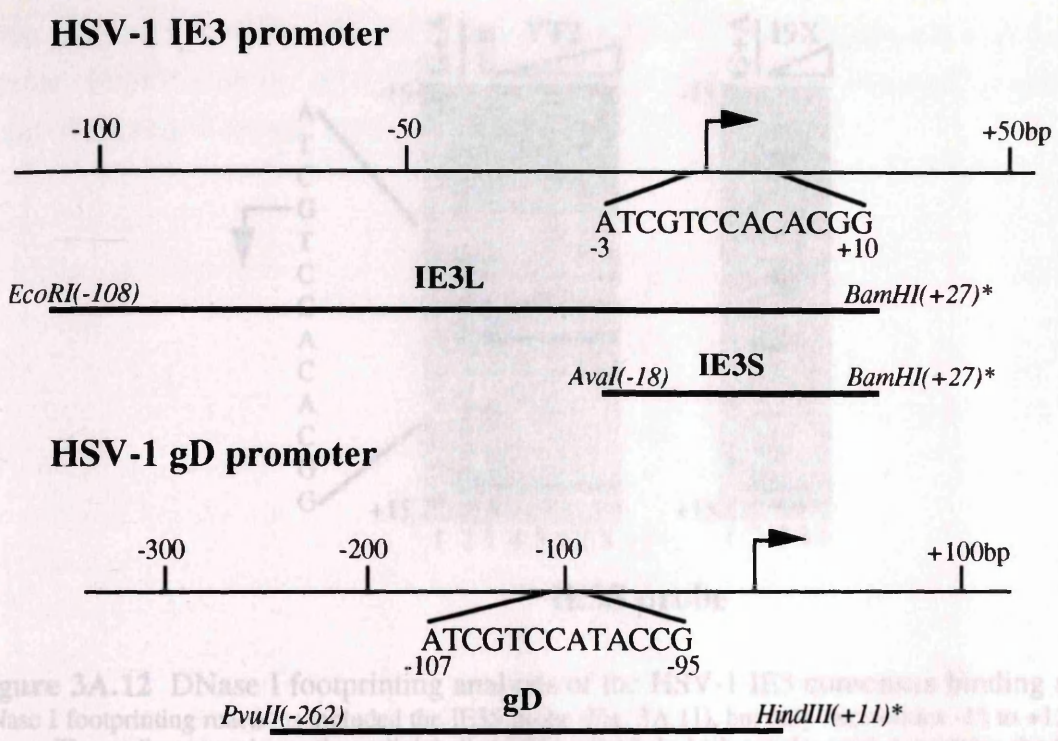


Figure 3A.11 HSV-1 IE3 and gD promoter fragment probes. Schematics of the HSV-1 DNA fragment probes used in this study. Numbers on the scales correspond to nucleotide positions in each promoter and the mRNA start sites are shown by bent arrows. The position and sequence of the IE3 (Muller, 1987) and the gD (Faber and Wilcox, 1986b) consensus Vmw175 binding sites are given. The thick lines represent sequences included in the IE3L, IE3S and gD probe fragments, the restriction sites used for their isolation are indicated; asterisks indicate the radiolabelled end of each fragment probe. IE3L and IE3S fragment probes were isolated from plasmids pIE3CAT or pUCIE3, the gD fragment probe was isolated from plasmid pgDCAT.

A high-resolution DNase I footprinting analysis on the HSV-1 IE3S fragment probe provided verification that the VZV 140k DNA binding domain specifically recognised the HSV-1 IE3 consensus site. The VT2 peptide protected, albeit faintly, both halves of this bipartite Vmw175 consensus site from DNase I digestion (Fig. 3A.12). As seen previously on the VZV gene 62 promoter (Fig. 3A.8), DNase I hypersensitivity was induced at the edges of the protected sequences as the amount of VT2 extract present increased. The footprint produced at this site by the VZV 140k DNA binding domain, VT2, differed in several respects to that produced by the corresponding domain of HSV-1 Vmw175 (peptide I9X, for details see Section 1C.1.10 and Fig. 3B.7). VT2 induced greater DNase I hypersensitivity to the 5' side of the binding site but provided less complete protection within the binding site itself (Fig. 3A.12). In particular, VT2 failed to protect the region between the 5' (ATCGTC) and 3' (CGG) motifs. Therefore, even though both peptides recognised the same IE3 consensus site, the details of their structural interactions must differ.

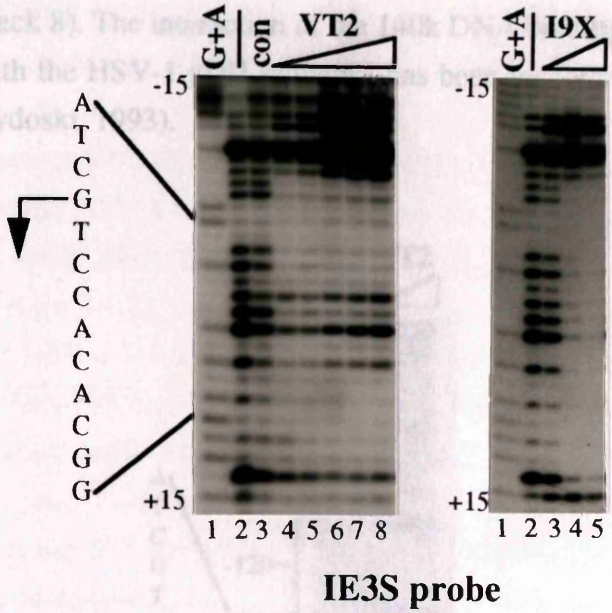


Figure 3A.12 DNase I footprinting analysis of the HSV-1 IE3 consensus binding site. DNase I footprinting reactions included the IE3S probe (Fig. 3A.11), but only nucleotides -15 to +15 are shown. The coding strand was the radiolabelled DNA strand. In both panels: track 1 contains the G+A Maxam and Gilbert sequencing reaction and track 2 contained no extract; the 'con' track included 5µl of an extract from bacteria containing plasmid p585T7a. On the left hand side: tracks 4-8 contain 0.4, 1, 2, 5 and 10µl of a partially purified bacterial VT2 extract respectively. On the right hand side: tracks 3-5 contain increasing amounts of partially purified I9X extract covering an equivalent range of peptide concentrations to the VT2X range used on the left. An expansion of the sequence of the IE3 Vmw175 consensus binding site is shown and the direction and position of transcriptional initiation is marked by the arrow. The I9X footprint was kindly provided by Dr R. D. Everett.

The gD promoter probe (Fig. 3A.11) included the Vmw175 consensus binding site (gDII site) that is known to contribute to the stimulation of transcription by HSV-1 Vmw175 *in vitro* (Tedder *et al.*, 1989). VZV 140k is a potent transactivator of the HSV-1 gD promoter in transient transfection assays (Disney, 1990) and the 140k DNA binding domain recognised the gD fragment probe in gel retardation assays (see Fig. 3E.6). Analysis of the specificity of the interaction between the VT2 peptide and the gD promoter identified multiple sequences weakly protected from DNase I digestion, that were flanked by DNase I hypersensitivity sites (Fig. 3A.13). This was reminiscent of the pattern observed on the VZV gene 62 promoter (Fig. 3A.8) and suggests that binding to multiple sites is unlikely to be a feature specific to the negative autoregulatory function of VZV 140k. The sites protected by VT2 included the gDII Vmw175 consensus sequence and at least four other sequences not recognised by Vmw175; these additional binding sites were centred around gD nucleotides -125, -85, -55 and -45 (Fig. 3A.13). DNase I footprinting with the highest amount of VT2 (Fig. 3A.13, track 5) resulted in a generalised widening of the protected sequences, perhaps due to lower specificity interactions. This effect was previously seen, although to a lesser extent, on the VZV gene 62 promoter probe BT (right hand panel of Fig. 3A.8, track 8) and also on the UP6

probe (Fig. 3B.6, track 8). The interaction of the 140k DNA binding domain *trpE* fusion peptide (FP617) with the HSV-1 gDII sequence has been recently reported by another group (Betz and Wydoski, 1993).

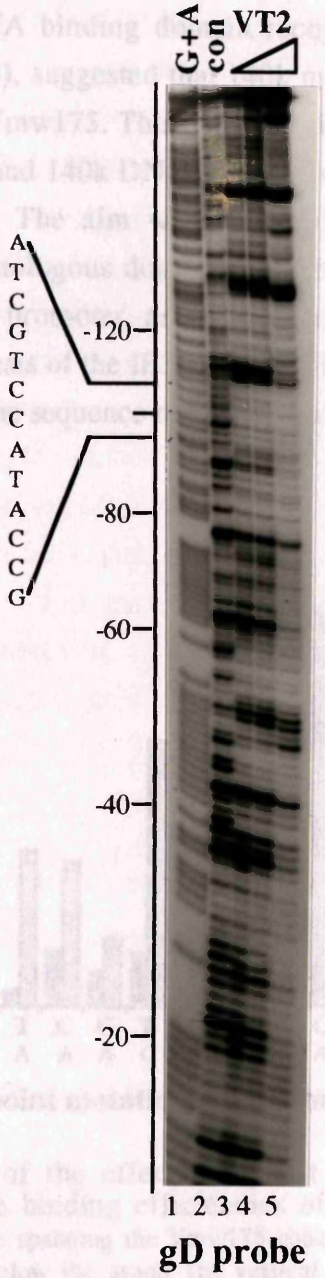


Figure 3A.13 DNase I footprinting analysis of the HSV-1 gD promoter region. The DNase I footprinting assay was performed with the end-labelled gD promoter target (details given in Fig. 3A.11). The coding strand was the radiolabelled strand. The numbers on the scale correspond to nucleotide positions in the gD promoter relative to the mRNA start site at +1; the position and sequence of the gDII consensus binding site are given (Faber and Wilcox, 1986b). Track 1 contained a G+A Maxam and Gilbert sequencing reaction. Track 2 included 1µl partially purified bacterial extract carrying the p585T7a vector. Tracks 3-5 included 0.2, 0.5 and 1µl partially purified bacterial VT2 extract.

3A.9 The VZV 140k DNA binding domain is less sequence-specific than the corresponding domain of HSV-1 Vmw175

The finding that several VT2 binding sites within the VZV gene 62 promoter region showed similarity to the HSV-1 Vmw175 consensus binding sequence (Fig. 3A.9), and also that the VZV 140k DNA binding domain recognises the HSV-1 gD and IE3 consensus sites (Section 3A.8), suggested that 140k may have a similar DNA binding specificity to that of HSV-1 Vmw175. This was investigated by comparing the binding specificities of the Vmw175 and 140k DNA binding domains by use of a mutagenised IE3 consensus binding site. The aim was to identify differences in the binding specificities of these two homologous domains that might account for their opposing functional effects on the IE3 promoter, as seen in transfection assays (Disney *et al.*, 1990). An extensive mutagenesis of the IE3 consensus binding site had been performed previously for an analysis of the sequence requirements for Vmw175 binding (Everett *et al.*, 1991a).

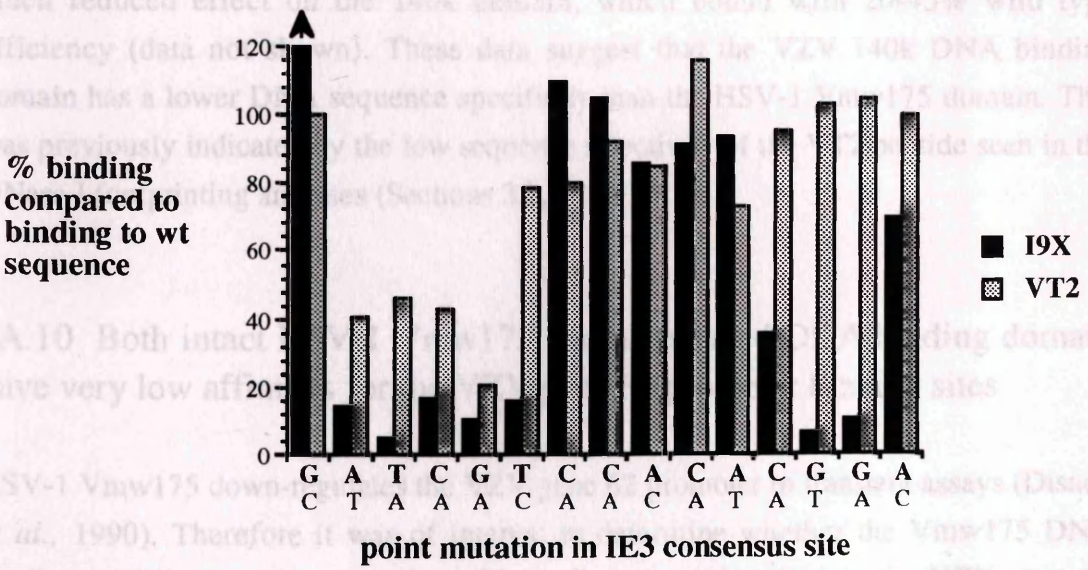


Figure 3A.14 Comparison of the effects of point mutations in the HSV-1 IE3 consensus binding site on the binding efficiencies of the 140k and Vmw175 DNA binding domains. The sequence spanning the Vmw175 consensus binding site in the HSV-1 IE3 promoter is shown in bold type below the graph The vertical bars above each base represent the percentage efficiency of peptide binding, when that base was mutated to the one shown in smaller type below (compared to the efficiency of binding to the wild type probe). The particular base substitution at each position was the one that was previously found to be most deleterious to the DNA binding efficiency of Vmw175 (Everett *et al.*, 1991a). Values were derived by quantitation of free and complexed probe bands resolved by gel retardation analysis (see Section 2D.2); the arrow indicates a value above 120%. Data used to produce the I9X binding profile was abstracted from Everett *et al.*, (1991a). Although the experiments using the VT2 and I9X peptides were performed independently, the reaction conditions and other parameters were identical. Analysis of VT2 binding efficiency was carried out with sets of mutagenised probe fragments isolated from the M13IE3MB series of plasmids, and probes prepared on two separate occasions gave highly reproducible results. Values used to produce the VT2 profile represent averages of the results provided by the two analyses.

The mutagenised progeny M13 phage resulting from this earlier study were utilised to isolate 'prime cut' *Ava*I-*Bam*HI IE3S fragment probes, containing single point mutations at positions throughout the IE3 consensus binding site. Single stranded M13 templates were re-sequenced to ensure the presence of the expected mutation, then the probes were isolated by the 'prime cut' method to yield DNA targets with equivalent specific activities suitable for comparative analyses (Section 2D.1.2). The mutant probes were used in gel retardation experiments with the partially purified VT2 bacterial extract, in order to determine the VT2 binding specificity for the HSV-1 IE3 binding site. The results are summarised in histogram form in Figure 3A.14, combined with the previously determined binding profile for the equivalent HSV-1 Vmw175 domain, peptide I9X (Everett *et al.*, 1991a). It is apparent that the 140k DNA binding domain has an overall reduced binding specificity for this IE3 consensus binding site compared to the Vmw175 domain, and it appears to be virtually insensitive to point mutations in the 3' portion of the bipartite sequence. Additionally, a point deletion of the first T (in brackets) of the A(T)CGTC motif reduced the I9X binding efficiency to 5% the level of binding to the wild type IE3 sequence (Everett *et al.*, 1991a). This same deletion has a much reduced effect on the 140k domain, which bound with 20-45% wild type efficiency (data not shown). These data suggest that the VZV 140k DNA binding domain has a lower DNA sequence specificity than the HSV-1 Vmw175 domain. This was previously indicated by the low sequence selectivity of the VT2 peptide seen in the DNase I footprinting analyses (Sections 3A.6 and 3A.8).

3A.10 Both intact HSV-1 Vmw175 and its isolated DNA binding domain have very low affinities for the VZV gene 62 promoter binding sites

HSV-1 Vmw175 down-regulates the VZV gene 62 promoter in transient assays (Disney *et al.*, 1990). Therefore it was of interest to determine whether the Vmw175 DNA binding domain interacts with the VT2 binding sites identified in the VZV gene 62 promoter, to further analyse whether the mechanistic basis of gene 62 repression was similar to autoregulation of the HSV-1 IE3 promoter by Vmw175. Approximately equivalent amounts of VT2X and the corresponding Vmw175 DNA binding domain peptide, I9X, were incubated with HSV-1 IE3 and VZV gene 62 promoter probes and the resultant complexes analysed by gel retardation assays. The 140k DNA binding domain, peptide VT2X, gave discrete complexes with the gene 62 and IE3 promoter probes (Fig. 3A.15, tracks 2, 7 and 12), and as expected, I9X specifically recognised the IE3 promoter probe (track 3). However, no obvious complexes specific to I9X were detected with either the UP6oligo or the BT probes that were derived from the gene 62 promoter (Fig. 3A.15, tracks 8 and 13 respectively).

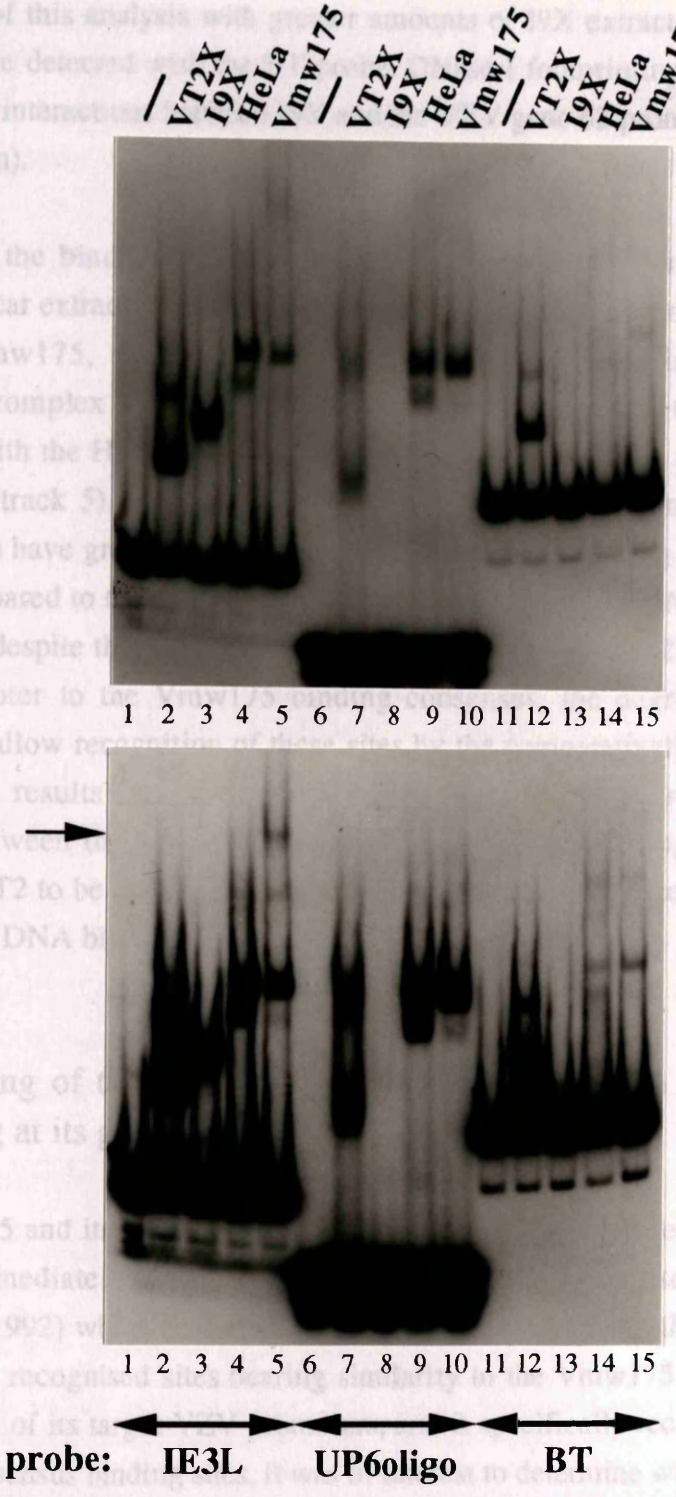


Figure 3A.15 Gel retardation analysis of HSV-1 IE3 and VZV gene 62 promoter regions. The IE3L DNA probe, derived from the HSV-1 IE3 promoter is described in Fig. 3A.11. The UP6oligo (including the U site) and BT (including sites P1-P5) probes were derived from the VZV gene 62 promoter and are described in Fig. 3A.6. Target DNA was incubated with approximately equivalent amounts of partially purified VZV VT2X or HSV-1 I9X bacterial extracts, or equivalent amounts of HeLa nuclear extracts with or without intact Vmw175 as indicated above the tracks. The Vmw175 protein was expressed from plasmid pI8 that has an insertion mutation in a region not involved in DNA binding (Paterson and Everett, 1988a; b). HeLa extracts were produced by T. Paterson. Two different autoradiographic exposures of the same gel are shown here, since prolonged exposure was necessary to visualise the complex between the HSV-1 IE3L fragment and intact Vmw175 (track 5), marked by the arrow.

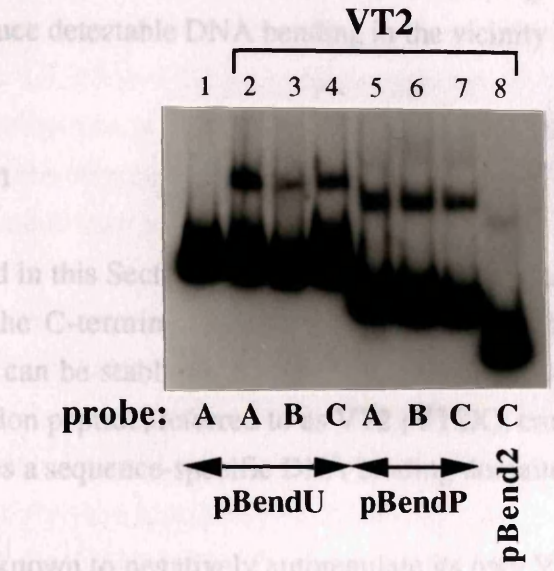
On repetition of this analysis with greater amounts of I9X extract, faint protein-DNA complexes were detected with the BT probe. DNase I footprinting analyses have also failed to detect interactions between I9X and the VZV gene 62 promoter probes BT and UP6 (not shown).

To verify that the binding results obtained for I9X reflect the situation with intact Vmw175, nuclear extracts from HeLa cells, transfected either with or without a plasmid expressing Vmw175, were incubated with the same probes as above. The only protein-DNA complex specific to intact Vmw175 produced in this gel retardation analysis was with the HSV-1 IE3L probe (marked by the arrow in the longer exposure of Fig. 3A.15, track 5). Therefore, both intact HSV-1 Vmw175 and its isolated DNA binding domain have greatly reduced affinities for the VT2 binding sites in the gene 62 promoter, compared to the binding observed with the VZV 140k DNA binding domain. It appears that despite the sequence similarity of several of the VT2 binding sites in the gene 62 promoter to the Vmw175 binding consensus, the degree of homology is insufficient to allow recognition of these sites by the comparatively specific Vmw175 peptide. These results are in agreement with those obtained from the specificity comparison between the VZV VT2 and HSV-1 I9X DNA binding domain peptides, which found VT2 to be have a reduced stringency for the IE3 consensus site compared to the Vmw175 DNA binding domain (Section 3A.9).

3A.11 Binding of the VZV 140k DNA binding domain does not induce DNA bending at its gene 62 promoter binding sites

HSV-1 Vmw175 and its isolated DNA binding domain both induce a significant DNA bend in the immediate vicinity of the HSV-1 IE3 and gDII consensus binding sites (Everett *et al.*, 1992) which may be of functional relevance. Since the VZV 140k DNA binding domain recognised sites bearing similarity to the Vmw175 consensus binding sequence in one of its target VZV promoters, and it specifically recognised the HSV-1 IE3 and gD consensus binding sites, it was of interest to determine whether VT2 binding would induce similar DNA conformational alterations. VT2 induced DNA bending was undetectable at the HSV-1 gD and IE3 consensus sites using the probes described by Everett *et al.* (1992) (data not shown). It could be argued that these heterologous binding sites are not directly functionally relevant to VZV 140k. Therefore the binding sites in the vicinities of both proposed VZV gene 62 mRNA start sites were tested for VT2 induced DNA bending. The method used to detect DNA bending depends on the principle that a bend near the end of a restriction fragment causes a less pronounced reduction in electrophoretic mobility as compared to a bend in the middle, as assayed by

gel retardation analysis. Regions of the gene 62 promoter, spanning the U site and also the P4/P5 140k binding sites (Fig. 3A.9) were cloned into the pBend2 plasmid (Zweib *et al.*, 1989) to yield plasmids pBendU and pBendP respectively. From these plasmids, probe DNA fragments were prepared by cutting with either *Bgl*II, *Xho*I or *Bam*HI.



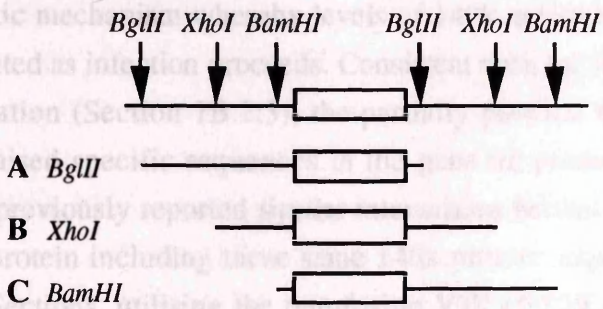


Figure 3A.16 Circular permutation gel retardation assay with VZV gene 62 promoter binding sites. Track 1 contained free probe while track 8 contained probe derived from plasmid pBend2 by *Bam*HI digestion. Probes in tracks 1, 2, 3 and 4 (which include site U; see Fig. 3A.9) were derived from the plasmid pBendU by cutting with the enzyme indicated by the letter below each track (A = *Bgl*II, B = *Xho*I, C = *Bam*HI) to yield 184bp DNA fragments. Similarly, probes in tracks 5, 6 and 7 (which include sites P4 and P5) were derived from plasmid pBendP to give 218bp DNA fragments. Incubation mixes for tracks 2-8 contained sufficient VT2 extract to produce only primary complexes and the amount of non-specific inhibitor (polydI:polydC) was increased to 4μg per reaction to reduce the extent of non-specific VT2 binding to the non-gene 62 promoter-derived probe regions. Non-specific DNA binding still occurred as indicated by the faint complex seen in track 8 containing the pBend2 probe with no introduced VT2 binding site, but this non-specific complex is a minor species compared to the specific primary complexes apparent in tracks 2-7. Beneath the gel is shown a schematic of a binding site-containing fragment (depicted by the box) inserted between the pBend2 derived tandemly repeated sequence. *Bgl*II, *Xho*I and *Bam*HI restriction sites are at equivalent positions in each of the repeats as shown. Single enzyme digestions gave fragments of identical lengths but with the position of the binding site circularly permuted with respect to the ends of the probes. Plasmid pBendU contained the blunt ended UP6oligo (Fig. 3A.6) inserted into the *Sal*I site located between the tandem repeats of plasmid pBend2 and included VZV gene 62 sequences -311 to -283. Plasmid pBendP contained an *Rsa*I-*Rsa*I fragment inserted into the pBend2 *Sal*I site and included VZV gene 62 promoter sequences -23 to +50.

The DNA fragments contained VT2 binding sites at circularly permuted positions within each set of three probes for use in gel retardation assays. An example of a typical DNA bending analysis is shown in Figure 3A.16. There was no significant difference in the mobility of the primary protein-DNA complex within each set of three probes (A, B, C). Therefore, in contrast to the HSV-1 Vmw175 DNA binding domain, the VZV 140k peptide does not induce detectable DNA bending in the vicinity of its binding sites.

3A.12 Discussion

The results presented in this Section (3A) demonstrate that a fragment of the VZV 140k protein, including the C-terminal portion of region 1 and the whole of the highly conserved region 2, can be stably expressed in isolation from other viral sequences in *E. coli*. This non-fusion peptide, referred to as VT2 (VT2X), contains 140k residues 417 to 647 and constitutes a sequence-specific DNA binding domain.

The 140k protein is known to negatively autoregulate its own VZV gene 62 promoter in transient transfection assays in selected cell types (Disney *et al.*, 1990). This presumably reflects a homeostatic mechanism whereby levels of 140k are stabilised and high level expression is prevented as infection proceeds. Consistent with the mechanism of HSV-1 Vmw175 autoregulation (Section 1B.2.3), the partially purified bacterially expressed VT2 peptide recognised specific sequences in the gene 62 promoter region. Wu and Wilcox (1991) had previously reported similar interactions between gene 62 sequences and a large fusion protein including these same 140k protein sequences. The analyses described in these Sections, utilising the non-fusion VT2 (VT2X) 140k DNA binding domain polypeptide, confirm and greatly extend the initial observations of Wu and Wilcox.

The 140k DNA binding domain peptide, VT2, specifically interacted with multiple sequences within the VZV gene 62 promoter region; at least six individual VT2 recognition sites were identified in the stretches of the promoter analysed. Preferential binding to any single site was not apparent, and preliminary data suggested that DNA binding of additional 140k DNA binding domain peptides proceeded in a cooperative manner (given sufficient peptide amounts). No obvious 140k binding consensus sequence could be derived from the VT2 binding sites identified in this analysis. However, several of the sequences recognised by VT2 in the gene 62 promoter contained homologies to the 5' portion of a consensus sequence that was originally derived by Faber and Wilcox (1986b) for some of the HSV-1 Vmw175 binding sites (Section 1C.1.3). Interestingly, two such Vmw175 consensus-like sites are located in the

immediate vicinities of the VZV gene 62 transcriptional initiation site at +1 (McKee *et al.*, 1990) and the previously proposed gene 62 mRNA start site further upstream at -287 (Felser *et al.*, 1988) (see Fig. 1B.4).

There is indirect evidence to suggest that the regulatory functions of HSV-1 Vmw175 and its PRV IE180 homologue are mediated by simultaneous interactions with DNA and TFIID (see Sections 1B.3.3 and 1C.3.1). Autoregulation of the HSV-1 IE3 promoter requires the strong interaction of Vmw175 with a specific Vmw175 consensus sequence spanning the IE3 mRNA start site, and appears to be a consequence of inhibition of Sp1 induced activation, rather than inhibition of basal transcription (see Section 1B.2.3). As such, Vmw175 occupation of the IE3 consensus binding site is unlikely to prohibit assembly of the transcriptional preinitiation complex, but may preclude further assembly of the active transcription complex or interfere with the activity of the TBP-associated factors (TAFs) for upstream activating factors. By analogy, the interactions of the VZV 140k DNA binding domain with the sequences in close proximity to the possible VZV gene 62 mRNA start sites are implicated in the autoregulation function of 140k. It has been proposed that the VZV 140k alphaherpesvirus homologues may all share a common autoregulatory mechanism as sequences similar to the HSV-1 Vmw175 consensus are also found overlapping the transcription start sites of the PRV IE and EHV IE genes (Wu and Wilcox, 1991; Smith *et al.*, 1992a). Furthermore, a fusion peptide containing the PRV IE180 residues corresponding to those in peptide VT2, specifically recognised the site spanning its own mRNA start site (Wu and Wilcox, 1991). In the case of the VZV gene 62 promoter, a thorough mutational analysis would be necessary to determine whether the sequences near the mRNA start sites, or any of the numerous other specific VT2 binding sites identified in this study, are required for down-regulation of the gene 62 promoter by the 140k protein.

The similarity of several VT2 binding sites to the Vmw175 consensus sequence suggests that VZV 140k may have a related sequence selectivity to that of HSV-1 Vmw175. This is further supported by the observation that the 140k DNA binding domain specifically recognised Vmw175 consensus binding sites in the HSV-1 promoters analysed. Recent experiments have implied that VZV 140k binding to the gDII consensus site contributes to the 140k-mediated transactivation of the HSV-1 gD promoter (Betz and Wydoski, 1993). VT2 interacts specifically with the consensus binding site spanning the mRNA start site of the HSV-1 IE3 gene, having an apparent binding affinity comparable to that of the corresponding domain of Vmw175. Yet evidence from transfection assays (Disney *et al.*, 1990) and also from analyses of the recombinant HSV-140 virus (Disney and Everett, 1990), suggested that intact 140k fails to repress the HSV-1 IE3 promoter, although the later observation appears to depend

upon the cell type used (J. Smiley, personal communication). Similarly, the EHV IE1 protein down-regulates its own promoter, presumably through interaction with the Vmw175 consensus-like site over its mRNA start site, yet fails to repress the HSV-1 IE3 promoter in transient assays (Smith *et al.*, 1992a). Slight sequence differences between the binding sites found at the mRNA start sites of the VZV gene 62, HSV-1 IE3 and EHV-1 IE genes may account for these findings, as each protein is likely to have its own particular sequence requirements for optimal recognition. The apparent differences between the details of the interactions of the 140k and Vmw175 DNA binding domains with the HSV-1 IE3 consensus binding site support this idea. Therefore the interaction with the heterologous binding site may not be appropriate to enable promoter repression by the models suggested above. Additionally, there is evidence to suggest that binding of Vmw175 alone to the IE3 consensus site is not sufficient for autoregulation, as a mutant Vmw175 protein containing a temperature sensitive lesion (due to the substitution A966T in the highly conserved region 4) was indistinguishable from the wild type in its ability to bind to the IE3 consensus binding site, yet failed to autoregulate at the NPT (Paterson *et al.*, 1990). According to the autoregulation mechanism suggested above, the lesion in region 4 may affect the interactions between Vmw175 and other factors (perhaps TFIID) such as to allow transcriptional initiation even in the presence of tightly bound Vmw175. Perhaps the precise positioning of the HSV-1 IE3 site with respect to the TATA box may not be correct for the optimal interaction between the VZV 140k protein and TFIID that is required for transcriptional repression. A close correlation has been observed between Vmw175-mediated negative autoregulation and the stereo-specific or distance-dependent placement of the IE3 binding site and the TATA box (Didonato and Muller, 1989).

HSV-1 Vmw175 is equally as effective as VZV 140k in the down-regulation of the VZV gene 62 promoter in transfection assays (Disney *et al.*, 1990), yet intact Vmw175 and also its isolated DNA binding domain displayed greatly reduced affinities for the binding sites therein, compared to that of the 140k DNA binding domain. It is possible that the results of the *in vitro* DNA binding analyses do not fully reflect the situation in infected or transfected cells. Down-regulation of the gene 62 promoter, by Vmw175 at least, may be more complicated than DNA binding alone and may involve interactions with other viral proteins or host transcription factors. It has recently been shown that the DNA binding activity of Vmw175 is facilitated by TFIID and TFIIB *in vitro* (Smith *et al.*, 1993). Perhaps the very low levels of Vmw175 binding to the VZV gene 62 promoter is sufficient to mediate transcriptional repression in the presence of these basal transcription factors *in vivo*.

It is of interest to note that although negative autoregulation of the VZV gene 62 promoter in Vero and BHK cells is well documented (Felser *et al.*, 1988; Disney *et al.*, 1990; this Thesis, Section 3F.4), the opposite effect has been demonstrated in T lymphocytes and neural cells (PC-12). In these particular cell lines VZV 140k, and also HSV-1 Vmw175, activated the VZV gene 62 promoter (Perera *et al.*, 1992b). T lymphocytes and neural cells are important sites for productive VZV growth during infection and therefore these cell lines were considered to be physiologically relevant. Tissue specific differences in the abundance or the complement of transcription factors may determine whether VZV 140k activates or represses its cognate gene 62 promoter.

An examination of the specificity of the VZV 140k DNA binding domain peptide for a Vmw175 consensus binding site found that sequences in the 3' portion of the consensus site were unimportant for VT2 binding. It is possible that the VT2 binding sites may have their own distinct 3' motif, but none is apparent for the sites identified in this study. The VT2 sites may transpire to fit a broad consensus that may be a more degenerate form of the consensus statistically determined for a wide range of Vmw175 sites (DiDonato *et al.*, 1991). However, it should be noted that the sites of DNase I protection by VT2 were of highly variable sizes. The results of a specificity analysis using the mutagenised HSV-1 IE3 consensus site found that the VZV 140k DNA binding domain has a lower DNA binding specificity than the equivalent domain of HSV-1 Vmw175. This reduced DNA binding specificity was reflected in the high number and also the variety of sequences recognised by VT2 in the VZV gene 62 and HSV-1 gD promoters. The low binding specificity of the VT2 peptide could explain the promiscuous transactivating phenotype of 140k in transfection assays (Section 1C.2.3), as the 140k protein would be predicted to have a low promoter selectivity. During infection, 140k may function by activating suboptimally utilised RNA polymerase II promoters in a similar manner to that previously suggested for the homologous PRV IE180 protein (Thali *et al.*, 1990). In this scenario, the sequence of the binding site is not the primary determinant of a target promoter. Instead, the low specificity of the 140k protein would make all physically accessible, TATA-containing promoters, that are not already fully activated (by the presence of an enhancer or multiple *cis*-acting elements) targets for activation by 140k; the most abundant class of target would therefore be the viral promoters.

VT2 binding to multiple sites as seen on the VZV gene 62 promoter is unlikely to be a feature specific to the autoregulatory function of 140k, since DNase I footprinting analysis of the HSV-1 glycoprotein gD promoter, which is strongly activated by 140k in transfection assays (Disney, 1990), identified multiple sites of VT2 protection from DNase I digestion and extensive DNase I hypersensitivity. In addition, the homologous

PRV IE180 protein binds multiple sequences with limited homology in the AdML and hsp70 promoters, and DNA binding is necessary for the activation of these promoters by PRV IE180 (Cromlish *et al.,* 1989). Even though multiple VT2-DNA interactions were apparent in the DNA binding analyses, it is possible that multiple binding of 140k proteins is not functionally relevant and may just be a consequence of the low specificity of the protein combined with the high amounts of peptide used in the analyses. During infection, the low specificity of the 140k protein may provide a mechanism to increase the probability of target promoter recognition. By analogy to a current functional model for HSV-1 Vmw175 (Section 1C.1.7), the position and orientation of the particular DNA binding site occupied by 140k would not greatly affect the efficiency of transactivation of the promoter.

The VZV 140k protein is a more potent transactivator of gene expression than HSV-1 Vmw175 in transfection assays (Cabirac *et al.,* 1990; Disney, 1990). This strong transactivation phenotype of 140k is not merely a function of the relative strengths of the HSV-1 IE3 and VZV gene 62 promoters (Disney, 1990), but could be attributable to the relative strengths of their activating domains. This remains to be quantified, although an isolated 140k potential activation domain activates transcription to a level comparable to that of the potent HSV VP16 activation domain (Cohen *et al.,* 1993). An alternative explanation suggested by this study is that multiple interactions of 140k proteins within the same promoter may have an additive or perhaps even a synergistic effect, enhancing the efficiency of activation. This could be facilitated by interactions with other viral or cellular proteins or via conformational effects on the promoter (see below).

Of possible functional significance are the alterations of DNA conformation induced by binding of the 140k DNA binding domain peptide, made apparent by the sites of DNase I hypersensitivity flanking the VT2 binding sites on all three promoters analysed. The intact Vmw175 protein and its isolated DNA binding domain significantly bend the DNA at its binding sites (Everett *et al.,* 1992), whereas detectable DNA bending was not induced by the 140k DNA binding domain peptide in the corresponding analysis. This lack of DNA bending at individual VT2 binding sites, and also the extensive nature of the DNase I hypersensitivity pattern, suggest that the conformational effects may extend along the length of the 140k target promoters, induced by multiple VT2 peptides binding to numerous sites. If this transpires to be the case for binding of the intact 140k protein during VZV infection, numerous possible functional implications of such a broad-ranging conformational effect can be envisaged, including modification of the interactions of transcription factors with DNA or with other proteins, or clearing the promoter of non-specific or inhibitory DNA binding proteins. Thus the low specificity

of binding of the 140k DNA binding domain, which could alternatively be termed 'sequence selectivity', may facilitate an unusual mechanism of transcriptional regulation.

The VZV 140k DNA binding interaction appears to be essential for the regulatory functions of 140k, as an insertion mutation into region 2 prevented 140k from transactivating or repressing gene expression (Disney *et al.*, 1990). Therefore future investigations into the mechanistic details of the 140k DNA binding interaction will be crucial to gain an understanding of the role of 140k in infection. The molecular basis of whether the outcome of 140k binding to its target promoters is transactivation or repression remains to be determined.

3B Determination of the VZV 140k protein sequence requirements for DNA binding

3B.1 Strategy

The DNA binding domain of HSV-1 Vmw175 has been closely defined by insertion, deletion and truncation analyses (Sections 1C.1.9 and 1C.1.10). However, no comparable studies on the VZV 140k protein have been undertaken. The position of sequences important for the 140k DNA binding activity was indirectly implied from a single 4 amino acid insertion into 140k residue 472 that disrupted the regulatory functions of the intact 140k protein (Disney *et al.*, 1990). Therefore, the protein sequences included in the 140k DNA binding domain peptide VT2 (amino acid residues 417 to 647) (Fig. 3A.1) and also in the FP617 fusion protein (Wu and Wilcox, 1991), had been selected mainly on the basis of homology to the HSV-1 Vmw175 DNA binding domain.

There are clear functional and sequence similarities between the VZV 140k and HSV-1 Vmw175 proteins, but there are also distinct differences in the details of their regulatory and DNA binding activities (Sections 1C.2.1, 1C.2.3 and 3A). The studies described in the following Sections aimed to discern whether the variations between these homologous IE proteins extend to their protein sequence requirements for DNA binding. It was of particular interest to investigate the requirement for the C-terminal portion of region 1 for 140k DNA binding activity, because this region is very poorly conserved between HSV-1 Vmw175 and VZV 140k (Figs. 1C.1 and 1C.3), but is essential for the DNA binding activity of Vmw175 (Section 1C.1.9 and 1C.1.10; Fig 1C.2).

In order to refine the minimum region of the VZV 140k protein that is sufficient for sequence-specific DNA binding, a series of plasmids expressing progressively smaller versions of the VT2 peptide were constructed, having truncations at either one or both ends. The gene 62 sequences selected for bacterial expression were based on the availability of convenient restriction enzyme sites. Care was taken when designing deletion constructs to have a 'favourable' codon following the ATG initiation codon, because the identity of the second codon can have a dramatic effect on the level of peptide expression in *E. coli* (Looman *et al.*, 1987). The experiments described in the following Sections analysed the DNA binding abilities of the truncated 140k peptides.

3B.2 Construction of plasmids expressing truncated versions of the VZV 140k DNA binding domain

Portions of the VZV gene 62 coding sequences were inserted into the MCR of the appropriate T7 expression vector of the p585T7 series (described in Section 3A.2), to maintain the correct reading frame. A summary of the gene 62 sequences inserted into the T7 expression vectors are given in Table 3B and the expression products are illustrated in Figure 3B.1. Plasmids pV17 and pV18 have an *Eco*RI site inserted on a 12bp linker into gene 62 codons 418 and 472 respectively (Disney, 1990) and were utilised for isolation of gene 62 sequences (nucleotides are given in brackets relative to the gene 62 mRNA start site at +1):

p585T7aVT1 - includes an *Eco*RI-*Sma*I fragment of plasmid pV17 (1322-2312bp) inserted into an *Eco*RI-*Sma*I digested p585T7a vector.

p585T7aVT2 and **p585T7aVT2X** - have been fully described in Section 3A.3.

p585T7aVT3 - an *Eco*RI-*Ava*I fragment of plasmid pV17 (1322-2071bp) with the *Ava*I 5' overhang filled-in, was inserted into the *Eco*RI- blunt *Bam*HI sites of vector p585T7a.

p585T7aVT4 - includes a *Bsp*HI-*Bam*HI fragment of plasmid p585T7aVT2 (1556-2114bp) inserted into the *Nco*I-*Bam*HI digested p585T7a vector (*Nco*I and *Bsp*HI digests produce complementary sticky ends).

p585T7aVT4X - derived from plasmid p585T7aVT4 by insertion of an *Xba*I stop linker into the *Bam*HI site, as previously described for the construction of plasmid p585T7aVT2X (Section 3A.3).

pT7VT4*ts*K - includes a *Bsp*HI-*Ava*I fragment of plasmid pVZV*ts*K (1556-2071bp) inserted into the *Nco*I-*Ava*I digested plasmid p585T7aVT2 (retaining 2072-2144bp).

p585T7aVT5 - an *Eco*RI-*Ava*I fragment of plasmid pV18 (1585-2071bp) was inserted into the *Eco*RI-*Ava*I digested plasmid p585T7aVT2 (retaining 2072-2114bp).

p585T7aVT6 - an *Eco*RI-*Bgl*I fragment of plasmid p585T7aVT2 (1322-1953bp) with the 3' *Bgl*I overhang removed, was inserted into the *Eco*RI-*Sma*I digested p585T7a vector.

p585T7aVT7 - an *Eco*RI-*Bgl*I fragment of plasmid pV18 (1585-1953bp) with the 3' *Bgl*I overhang removed, was inserted into the *Eco*RI-*Sma*I digested p585T7a vector.

p585T7aVT8 - an *Eco*RI-*Nco*I fragment of plasmid pV18 (1585-1937bp) with the 5' *Nco*I overhang filled-in, was inserted into the *Eco*RI-*Sma*I digested p585T7a vector.

p585T7aVT9 - an *Eco*RI-*Kpn*I fragment of plasmid pV18 (1585-1881bp) with the 3' *Kpn*I overhang removed, was inserted into the *Eco*RI-*Sma*I digested p585T7a vector.

p585T7aVT10 - includes an *Eco*RI-*Rsa*I fragment of plasmid p585T7VT5 (1585-1749bp) inserted into the *Eco*RI-*Sma*I digested p585T7a vector.

p585T7aVT11 - a *Ban*II-*Bam*HI fragment of plasmid p585T7aVT2 (1683-2114bp) with the 3' *Ban*II overhang removed, was inserted into the blunt *Nco*I-sticky *Bam*HI sites of vector p585T7a.

p585T7cVT12 - includes the *Rsa*I-*Kpn*I fragment of plasmid p585T7aVT5 (1750-1881bp) and the *Kpn*I-*Bam*HI fragment of plasmid p585T7aVT2 (1882-2114bp) co-ligated into the blunt *Eco*RI-sticky *Bam*HI sites of vector p585T7c.

a.Plasmid	b.Peptide	c.Nucleotides	d.Codons	e.N-term	f.C-term	g.Size(aa)	h.Size(kD)
p585T7aVT1	VT1	1322-2312	417-708	MAEF	21	317	35.2
p585T7aVT2	VT2	1322-2114	417-647	MAEF	21	256	28.4
p585T7aVT2X	VT2X	1322-2114	417-647	MAEF	GDQLI	240	26.7
p585T7aVT3	VT3	1322-2071	417-633	MAEF	21	242	26.9
p585T7aVT4	VT4	1556-2114	462-647	-	21	207	23.0
p585T7aVT4X	VT4X	1556-2114	462-647	-	GDQLI	191	21.2
pT7VT4*ts*K	VT4*ts*K	1556-2114	462-647	-	21	207	23.0
p585T7aVT5	VT5	1585-2114	472-647	MAEF	21	201	22.3
p585T7aVT6	VT6	1322-1953	417-594	MAEF	21	202	22.4
p585T7aVT7	VT7	1585-1953	472-594	MAEF	36	163	18.1
p585T7aVT8	VT8	1585-1937	472-588	MAEF	18	139	15.4
p585T7aVT9	VT9	1585-1881	472-570	MAEF	34	137	15.2
p585T7aVT10	VT10	1585-1749	472-525	MAEF	36	94	10.4
p585T7aVT11	VT11	1683-2114	504-647	M	21	166	18.4
p585T7cVT12	VT12	1750-2114	526-647	MN	21	145	16.1

Table 3B Specifications of the VZV 140k DNA binding domain truncation plasmids and their peptide products. Columns a. and b. give the nomenclature for the T7 plasmid constructs expressing the gene 62 sequences and their expression products respectively. Column c. gives the gene 62 nucleotides (numbered from the mRNA start site at +1) inserted into the expression vectors (inclusively), column d. indicates the 140k amino acid positions corresponding to residues included in the peptides. Columns e. and f. give the identity or number of vector derived amino acid residues at the N- and C-terminus of the peptide respectively. The variation in the number of vector-encoded residues to the C-terminal end of the peptides is in part due to the wide spacing between translational stop codons in the different reading frames in the T7 termination region (see Section 3A.2). Columns g. and h. give the total number of amino acid residues in the peptide and the predicted peptide molecular weight respectively. 'X' versions of peptides contain fewer C-terminal vector-encoded amino acids than the non-X counterparts, but are otherwise identical. The plasmid pT7VT4*ts*K contains a base change that results in the substitution A628V, equivalent to HSV-1 Vmw175 *ts*K substitution A475V.

3B.3 Bacterial expression of the VZV 140k DNA binding domain deletion peptides

The truncated VZV 140k DNA binding domain peptides were expressed in *E. coli* from the above T7 expression plasmids and partially purified extracts were prepared as described previously for the VT2 peptide (Section 3A.4). The majority of the truncated DNA binding domain peptides were expressed very efficiently, being the predominant species in the bacterial lysates, and were highly soluble in the partially purified extracts (see Fig. 3B.2, summarised in Fig. 3B.1). However, peptides with deletions encroaching into the C-terminal end of region 2 were expressed very poorly. These peptides, VT7-VT10, were not apparent in the Coomassie-stained gels of bacterial lysates and were barely visible in the Coomassie-stained partially purified extracts; their presence was verified by western blot analysis (not shown). Peptides VT6 and VT12 were both expressed to high levels, but VT6 was totally insoluble (data not shown).

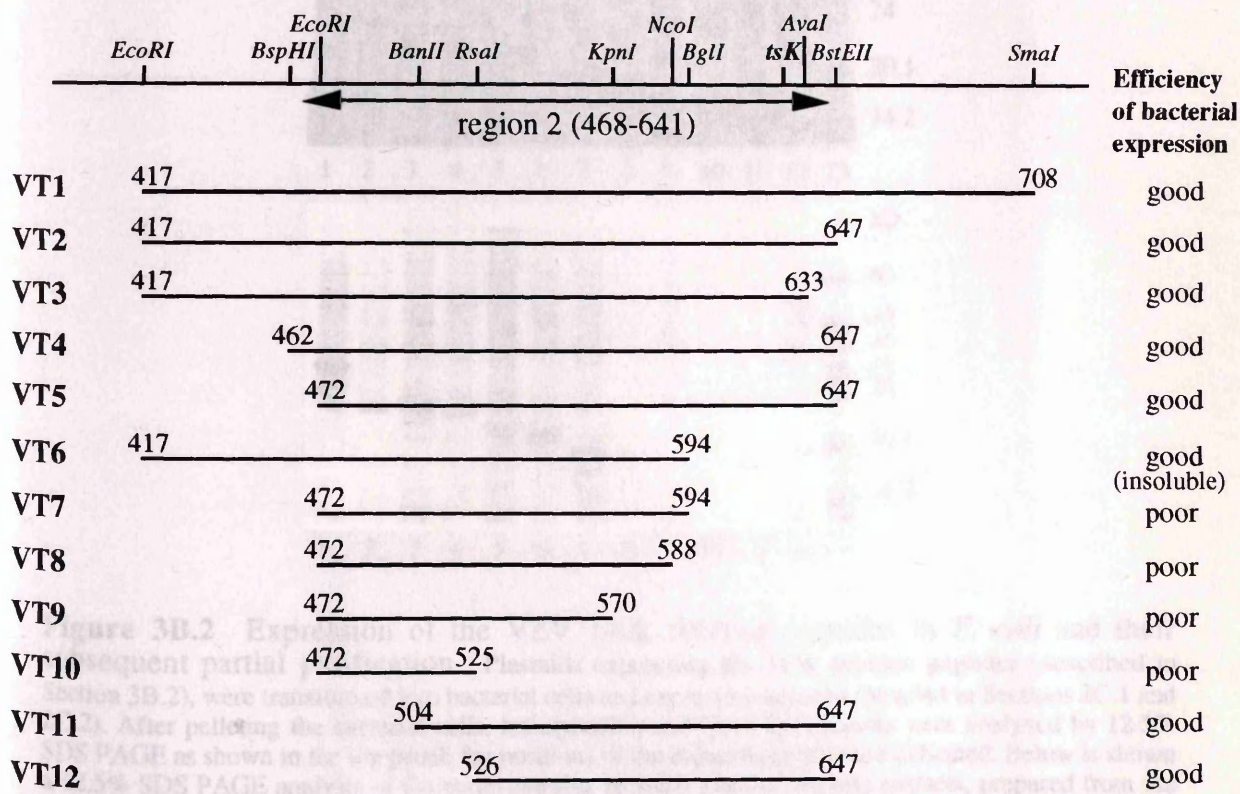


Figure 3B.1 Schematic representations of the VZV 140k DNA binding domain deletion peptides. The expressed regions of the 140k protein (VT1-VT12) are indicated by horizontal lines. The numbers of the 140k amino acid residues at the N- and C-terminal ends of the peptides are given and the corresponding restriction sites used for the isolation and cloning of gene 62 sequences into the T7 expression vectors are shown above. The *Eco*RI sites were introduced into the gene 62 sequences by insertion mutagenesis (Disney, 1990). The position of the highly conserved protein region 2, and the *ts*K mutation at VZV 140k amino acid 628 are also indicated. All constructs are drawn to scale. For each peptide, the efficiency of bacterial expression is indicated on the right; all peptides were highly soluble with the exception of peptide VT6.

The VT1 peptide included additional region 3 sequences, similar to the I9X and I10X DNA binding peptides of HSV-1 Vmw175 (Fig. 3B.7). This longer VT1 peptide was less stable than VT2 and tended to degrade, as demonstrated by the breakdown product approximately 7kD below the VT1 band in the bottom panel of Figure 3B.2, track 1. The VT2 peptide, and to a lesser extent the VT4 peptide, were also prone to slight proteolytic degradation and this was particularly apparent by western blot analyses of extracts (not shown). Construction of the 'X' versions of these peptides largely overcame this degradation problem (see Section 3A.3).

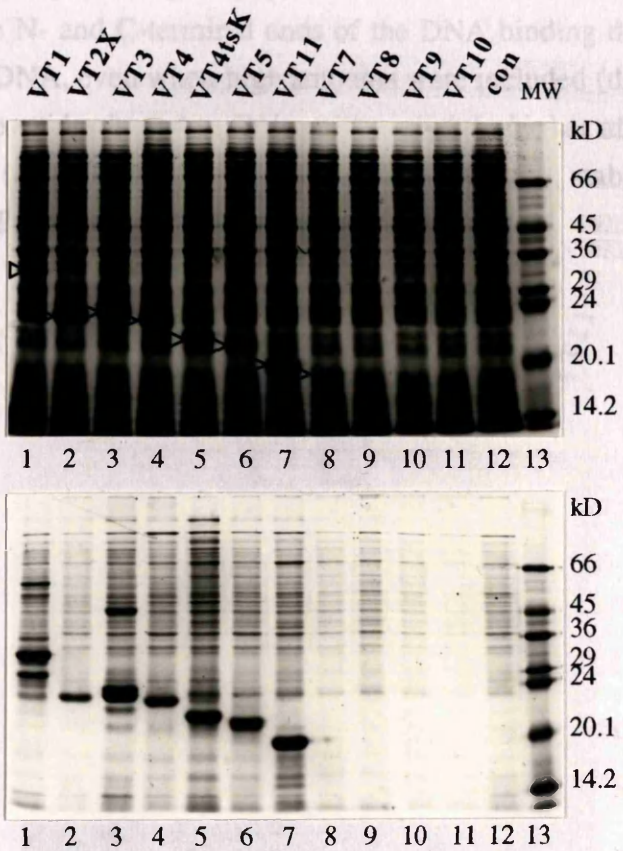


Figure 3B.2 Expression of the VZV 140k deletion peptides in *E. coli* and their subsequent partial purification. Plasmids expressing the 140k deletion peptides (described in Section 3B.2), were transformed into bacterial cells and expression induced (detailed in Sections 2C.1 and 2C.2). After pelleting the bacterial cells, resuspending and lysis, 2µl aliquots were analysed by 12.5% SDS PAGE as shown in the top panel; the positions of the induced peptides are indicated. Below is shown a 12.5% SDS PAGE analysis of the corresponding partially purified peptide extracts, prepared from the lysed bacteria above; the amount of extract loaded in each track corresponds to that from 4ml of bacterial culture. the extract in track 12 was produced from bacterial cells transformed with plasmid p585T7a. Details of the molecular weight standards in track 13 are as given previously (Fig. 3A.4). The sizes of the expressed polypeptides correspond well with the predicted sizes given in Table 3B. All extracts of peptide VT4*ts*K showed a prominent degradation product approximately 2kD smaller than the expected peptide, the particular VT4tsK extract shown here almost completely comprises the breakdown product. The version of the VT4*ts*K extract used for DNA binding analyses was approximately 50% degraded. Novel peptide bands in the appropriate positions for peptides VT7-10 were detectable upon close inspection of the stained gel shown in the bottom panel, but are lost upon photography.

3B.4 The DNA binding activity of VZV 140k resides within residues 472 to 633, as determined by gel retardation analysis

The DNA binding ability of each partially purified 140k deletion peptide was determined by gel retardation assay, an example of which is shown in Figure 3B.3. Peptides VT1 to VT5 gave well defined protein-DNA complexes, while equivalent amounts of VT7 and VT11 failed to bind to the DNA probe (Fig. 3B.3, tracks 7 and 8 respectively), even at greatly increased protein concentrations (data not shown). As expected, the VT12 peptide and peptides VT8-VT10, which contained further truncations into the N- and C-terminal ends of the DNA binding domain respectively, also failed to bind DNA, even when high amounts were included (data not shown). The VT2X and VT4X peptides bound to DNA with a slightly higher affinity than the VT2 and VT4 peptides (not shown), presumably due to the more stable nature of the 'X' peptides (Section 3B.3).

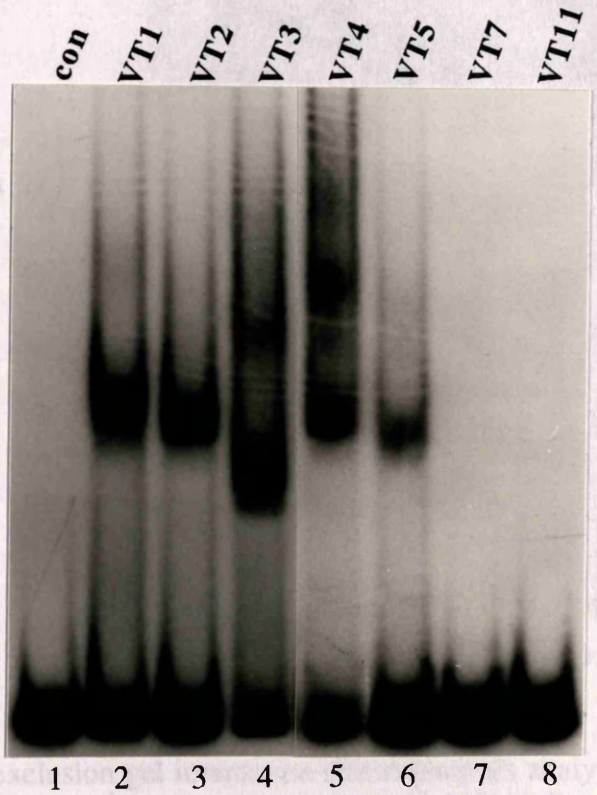


Figure 3B.3 Gel retardation assay with the VZV 140k DNA binding domain deletion peptides. Partially purified bacterial extracts of each deletion peptide were incubated with the HSV-1 IE3 promoter IES fragment probe (described previously in Fig. 3A.11) and resultant complexes resolved by native gel electrophoresis. Similar results were obtained with VZV gene 62 promoter fragment probes. The particular deletion polypeptide extract present in each incubation is indicated for tracks 2-8; track 1 contains an extract from bacteria carrying plasmid p585T7a. Primary protein-DNA complexes are apparent with peptides VT1-5, the VT3 and VT4 tracks also contain secondary protein DNA complexes of lower mobility.

The DNA binding ability of each deletion construct is summarised in Figure 3B.7. From this analysis it can be concluded that the minimal DNA binding domain of the 140k protein lies within residues 472 to 633. These sequences are contained entirely within the highly conserved region 2 (residues 468 to 641) of the 140k protein. Therefore the equivalent sequences from region 1 that are required in the minimal Vmw175 DNA binding domain, are not necessary for the DNA binding activity of 140k, as determined by gel retardation analysis.

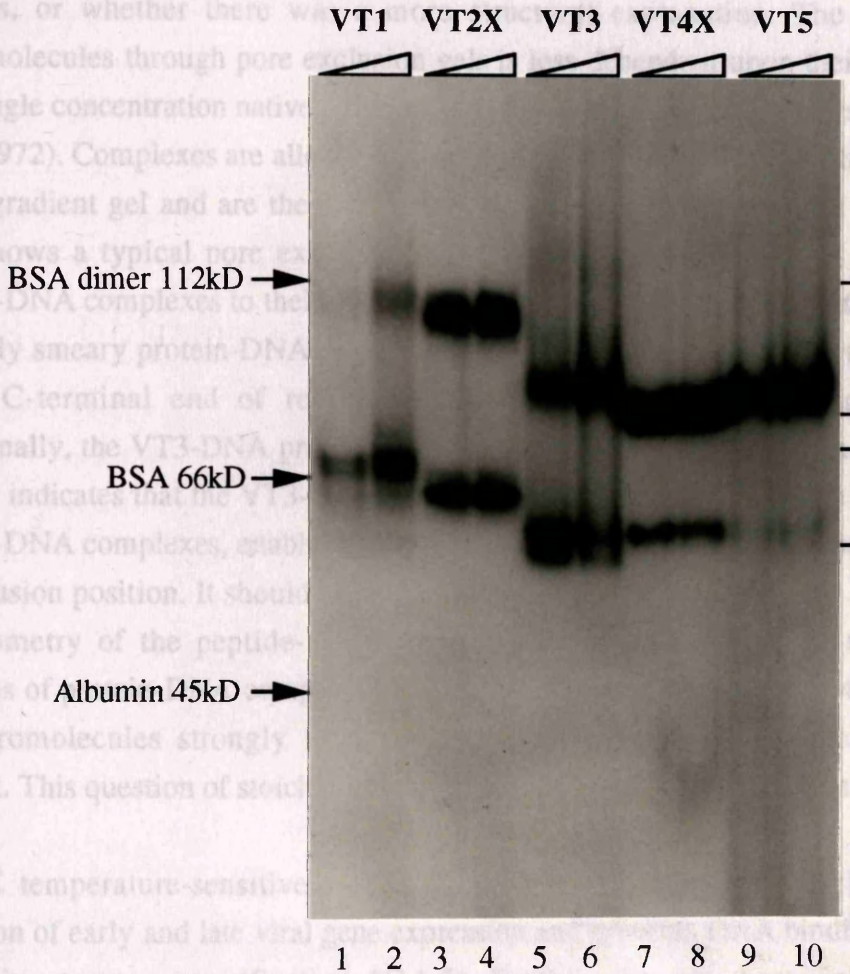


Figure 3B.4 Pore exclusion gel retardation electrophoresis analysis of the 140k VZV deletion peptide-DNA complexes. Partially purified 140k DNA binding domain deletion peptides were incubated with the VZV gene 62 promoter UP6oligo DNA probe (Fig. 3A.6) and resolved on a 4 to 24% gradient gel as described in Section 2D.3. Equivalent amounts of the indicated 140k deletion peptides were included in the analysis to obtain formation of the primary (PC) and secondary (SC) protein-DNA complexes, the free probe is not seen as it is run far off the gel in this type of analysis. The second track in each pair contained double the amount of extract included in the first. The gel was run for 40hr to allow complexes to reach their equilibrium positions. The relative migration positions of Coomassie-stained protein standards run on the same gel are indicated to the left.

It was consistently noticed during these studies that the 140k peptides lacking region 1 sequences (VT4 and VT5), readily produced secondary protein-DNA complexes; furthermore, it was unusual to obtain the primary complex in absence of the secondary complex, suggesting that the secondary complexes formed by these particular peptides have greater stability than their primary complexes. In addition, the mobilities of the VT3 peptide-DNA complexes were consistently observed to be greater than expected from the size of VT3 (Fig. 3B.3, track 4), as compared to the other deletion peptides. Pore exclusion electrophoresis (Section 2D.3) was employed to investigate whether the unusual mobility of the VT3-DNA complex was an artefact of the native gel retardation analysis, or whether there was a more structural explanation. The migration of macromolecules through pore exclusion gels is less dependent upon their charge, than with single concentration native gels that are used in gel retardation analysis (Andersson *et al.*, 1972). Complexes are allowed to migrate until they are trapped in the matrix sieve of the gradient gel and are therefore separated on the basis of size and shape. Figure 3B.4 shows a typical pore exclusion gel after electrophoresis of the 140k deletion peptide-DNA complexes to their equilibrium positions. The VT3 peptide tended to yield relatively smeary protein-DNA complexes, implying that the truncation to residue 633 at the C-terminal end of region 2, destabilises the DNA binding interaction. Additionally, the VT3-DNA primary complex had a higher mobility than predicted and perhaps indicates that the VT3-DNA complex has a more ellipsoid shape than the other peptide-DNA complexes, enabling it to move further through the matrix before reaching its exclusion position. It should be noted that it is not possible to predict accurately the stoichiometry of the peptide-DNA interaction from comparisons of the migration positions of protein-DNA complexes to the protein standards in Fig. 3B.4, as the shape of macromolecules strongly influences their equilibrium position reached on the gradient. This question of stoichiometry of binding is addressed in Section 3D.4.

The *ts*K temperature-sensitive (*ts*) lesion in region 2 of HSV-1 Vmw175 precludes activation of early and late viral gene expression and prevents DNA binding at the non-permissive temperature (Section 1C.1.5). Furthermore, the bacterially expressed Vmw175 DNA binding domain containing the *ts*K single amino acid substitution also has a temperature-sensitive DNA binding phenotype (Allen and Everett, 1994). The equivalent mutation had previously been introduced into the VZV gene 62 sequence on plasmid pVZV*ts*K. A version of the 140k DNA binding domain containing the A628V substitution, peptide VT4*ts*K, was tested to see if it displayed a temperature-sensitive phenotype (Fig. 3B.5). Compared to DNA binding at 4°C, all the peptides had greatly reduced binding efficiencies at 38.5°C, presumably due to peptide or complex instability at the higher temperature.

Consistent with this was the finding that the stability of HSV-1 Vmw175-DNA complexes was increased over 100-fold by shifting the temperature from 23 to 4°C (Kattar-Cooley and Wilcox, 1989). As mentioned above, the 'X' peptides had slightly higher binding efficiencies, presumably due to their reduced tendency to degrade. Comparison of the relative intensities of the different complexes at each temperature indicates that the *ts*K mutation does not produce a 140k DNA binding domain with a temperature-sensitive DNA binding phenotype (Fig. 3B.5). The very low levels of VT5 and VT3 binding at 38.5°C is probably not a temperature-sensitive phenotype, but is more likely due to enhanced instability of the peptide-DNA complexes, because even at 4°C these peptides yielded highly smeary complexes.

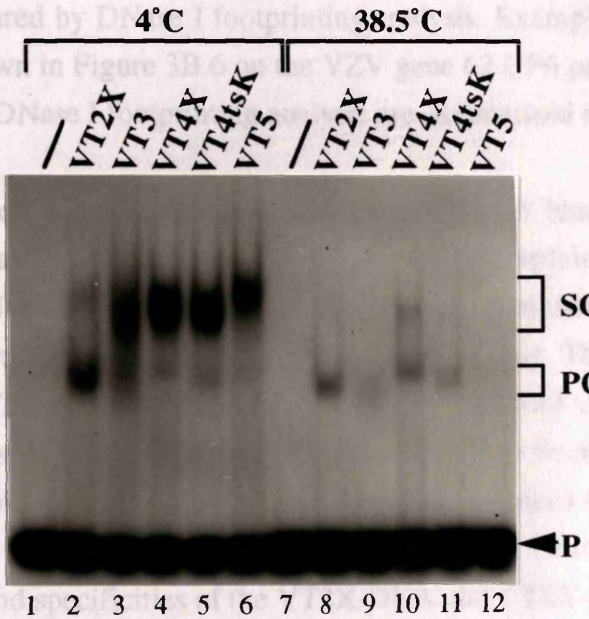


Figure 3B.5 Gel retardation analysis of the temperature-sensitivity of the VZV 140k deletion peptide-DNA interactions. The UP6oligo DNA target (Fig. 3A.6) used in this analysis contains a VT2 binding site identified by DNase I footprinting analysis (Section 3A.6). Equivalent amounts of partially purified bacterially expressed peptides were included in each track as indicated. Reactions in tracks 1-6 were incubated on ice, whereas those in tracks 7-12 were incubated in a 38.5°C water-bath. Samples were loaded directly onto a running gel and electrophoresed at RT. The positions of the free probe (P), the primary (PC) and secondary (SC) protein-DNA complexes are indicated. As seen here, it has been consistently observed that the relative mobilities of complexes including the UP6oligo probe are not as expected from the size of the peptides. The relative lack of sensitivity to temperature seen here for the DNA binding activity of the VT4*ts*K peptide was highly reproducible with several different preparations of the peptide and on all target DNA fragments tested.

3B.5 Region 1 sequences are required for DNase I footprinting by the VZV 140k DNA binding domain

Having determined by gel retardation analysis that truncated versions of peptide VT2 were able to bind DNA, DNase I footprinting was used to determine whether these shorter peptides had the same specificity of binding. The DNA binding interactions of peptides VT1, VT2, VT3, VT4 and VT5 (Fig. 3B.1) were compared on various VZV gene 62 and HSV-1 IE3 promoter fragments, as determined by DNase I footprinting. Peptides VT1 and VT3 produced identical footprinting patterns to VT2 on each DNA probe tested, while peptides VT4 and VT5 consistently failed to recognise any specific sequences. Therefore, even though region 1 sequences are dispensable for the DNA interaction as assayed by gel retardation, they are essential for recognition of specific sequences as measured by DNase I footprinting analysis. Examples of the two binding phenotypes are shown in Figure 3B.6 on the VZV gene 62 UP6 promoter fragment, and the findings of this DNase I footprinting analysis are summarised in Figure 3B.7.

Various attempts were made to identify features in the DNA binding characteristics of the 140k peptides lacking region 1 sequences, that might explain why these particular peptides bind to DNA in gel retardation analysis with apparent similar affinities to VT2, yet fail to give interactions detectable by DNase I footprinting. The binding efficiencies of the VT4 and VT5 peptides for the mutagenised HSV-1 IE3 consensus binding site were assayed, as described previously for the VT2 peptide in Section 3A.9. The binding profiles obtained for these peptides lacking region 1 sequences were not significantly different to that of the VT2 peptide shown in Figure 3A.14 (data not shown). The relative stabilities and specificities of the VT2X-DNA and VT4X-DNA complexes were compared as described in Section 2D.2. No obvious differences in the relative dissociation rates of the VT2X and VT4X peptides were noticed, although the VT4X peptide had a slightly lower DNA binding specificity than VT2X (data not shown). Whether this reduced specificity can explain the lack of protection from DNase I digestion provided by the 140k peptides lacking region 1 sequences is far from certain. However, a similar situation was discovered with the I10X version of the HSV-1 Vmw175 DNA binding domain peptides, which has a reduced amount of region 1 sequences and fails to give a DNase I footprint (see below).

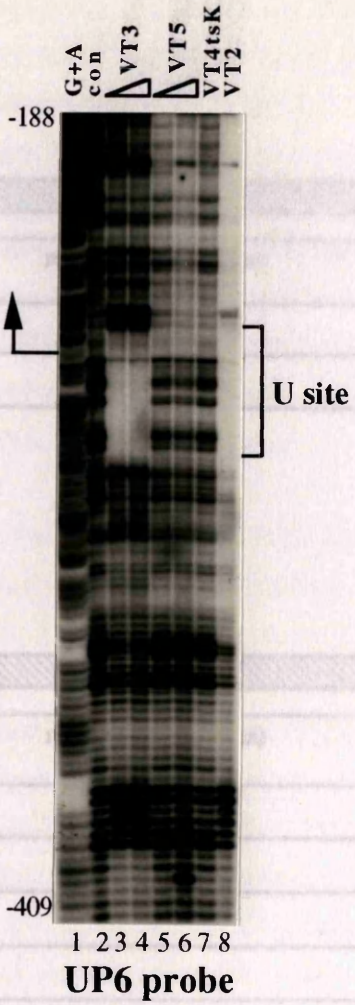


Figure 3B.6 DNase I footprinting analysis with the VZV 140k deletion peptides. The UP6 fragment probe was derived from the VZV gene 62 promoter, the details of which are given in Figure 3A.6. The position of the mRNA start site at +1 is indicated by the bent arrow and the gene 62 nucleotide coordinates at either end of the probe fragment are given relative to the mRNA start site (+1). Tracks 2, 4, 6, 7 and 8 contain equivalent amounts of extracts containing the indicated partially purified bacterially expressed peptides; tracks 3 and 5 contain 2/5 the amount of extract included in tracks 4 and 6 respectively. The 'con' extract was prepared from bacteria transformed with plasmid p585T7a. The position of the U binding site is indicated on the right (Fig. 3A.9). The amount of VT2 included in track 8 was such that produced a generalised broadening of the region of protection.

3B.6 Discussion

The results in this Section (3B) demonstrate that residues contained within the highly conserved region 2 of VZV 140k are sufficient for DNA binding, although flanking sequences may be additionally required for the full interaction with DNA, as discussed below.

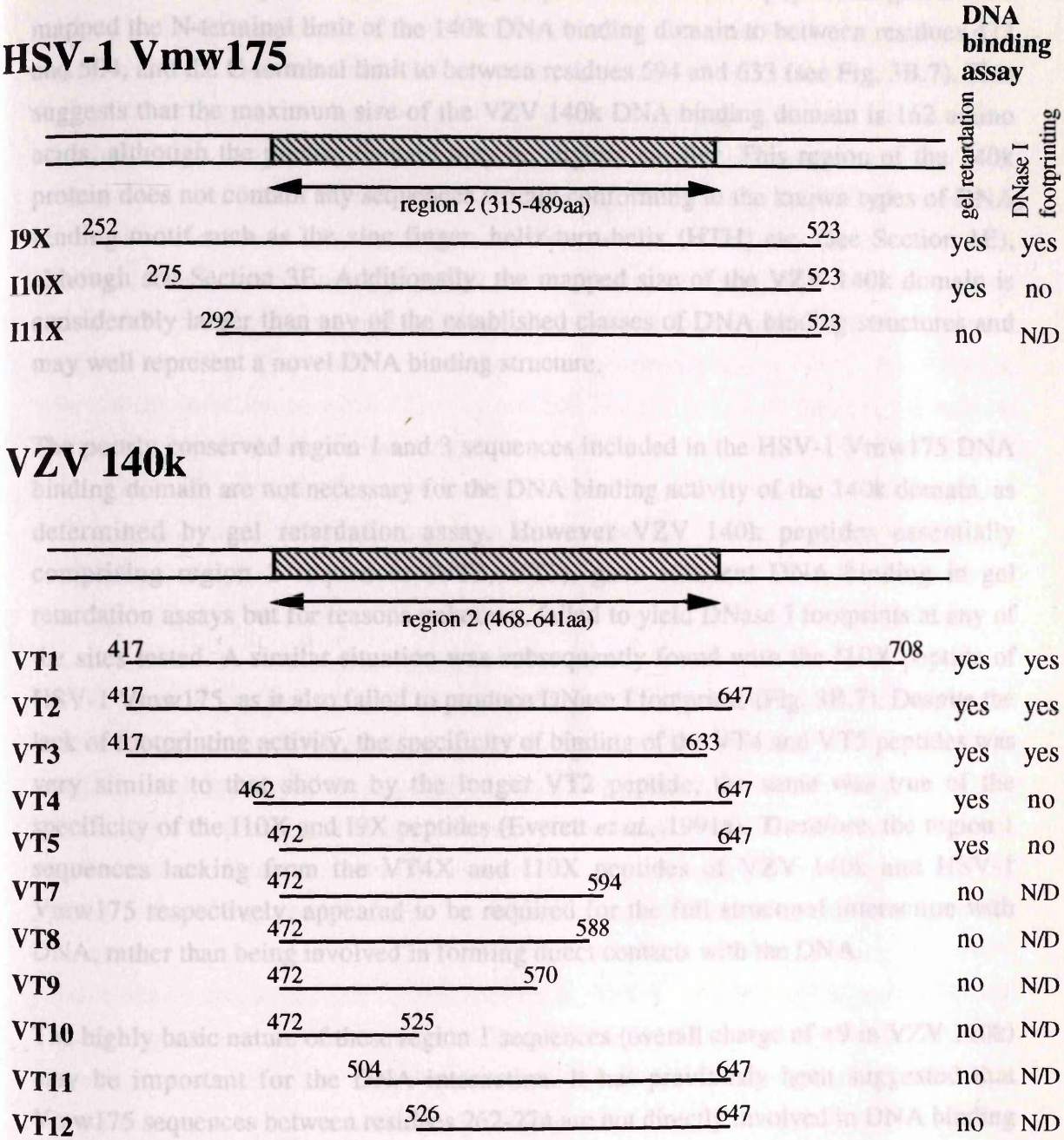


Figure 3B.7 The HSV-1 Vmw175 and VZV 140k DNA binding domain peptides and their DNA binding activities. The portion of each protein encompassing the DNA binding domain is shown diagrammatically. For each protein, the numbers of the first and last residues are given for the highly conserved region 2 (represented by a shaded box); protein regions 1 and 3 lie to the left and right respectively. Below these are shown horizontal lines that correspond to the protein sequences included in each of the deletion peptides. The residue at the N- and C-terminus of each peptide are indicated. All constructs are shown to scale and region 2 of the HSV-1 Vmw175 and VZV 140k constructs are accurately aligned beneath one another. The DNA binding abilities of the peptides are indicated, as determined by gel retardation (Section 3B.4) and DNase I footprinting assays (Section 3B.5); 'yes' indicates an ability to bind to DNA, 'no' indicates that DNA binding was not detected and 'N/D' indicates experiments not carried out. The HSV-1 Vmw175 peptides I9X, I10X and I11X are described by Pizer *et al.* (1991). The experiments to determine the I9X and I10X DNA binding activities were performed in parallel to the analyses described in this Section and are in agreement with those obtained previously (Pizer *et al.*, 1991; R. D. Everett, personal communication), the I11X result was taken from Pizer *et al.* (1991). [A loose copy of a figure showing these constructs is found at the back of this Thesis].

Gel retardation analysis of the bacterially expressed VZV 140k peptide fragments has mapped the N-terminal limit of the 140k DNA binding domain to between residues 472 and 504, and the C-terminal limit to between residues 594 and 633 (see Fig. 3B.7). This suggests that the maximum size of the VZV 140k DNA binding domain is 162 amino acids, although the precise domain may be slightly smaller. This region of the 140k protein does not contain any sequences strictly conforming to the known types of DNA binding motif such as the zinc finger, helix-turn-helix (HTH) etc. (see Section 1E), although see Section 3E. Additionally, the mapped size of the VZV 140k domain is considerably larger than any of the established classes of DNA binding structures and may well represent a novel DNA binding structure.

The poorly conserved region 1 and 3 sequences included in the HSV-1 Vmw175 DNA binding domain are not necessary for the DNA binding activity of the 140k domain, as determined by gel retardation assay. However VZV 140k peptides essentially comprising region 2 sequences (VT4, VT5), gave efficient DNA binding in gel retardation assays but for reasons unknown, failed to yield DNase I footprints at any of the sites tested. A similar situation was subsequently found with the I10X peptide of HSV-1 Vmw175, as it also failed to produce DNase I footprints (Fig. 3B.7). Despite the lack of footprinting activity, the specificity of binding of the VT4 and VT5 peptides was very similar to that shown by the longer VT2 peptide; the same was true of the specificity of the I10X and I9X peptides (Everett *et al.*, 1991a). Therefore, the region 1 sequences lacking from the VT4X and I10X peptides of VZV 140k and HSV-1 Vmw175 respectively, appeared to be required for the full structural interaction with DNA, rather than being involved in forming direct contacts with the DNA.

The highly basic nature of these region 1 sequences (overall charge of +9 in VZV 140k) may be important for the DNA interaction. It has previously been suggested that Vmw175 sequences between residues 262-274 are not directly involved in DNA binding and that the structure adopted by this region influences the DNA binding properties of Vmw175 (Wu and Wilcox, 1990). It is noteworthy that the ability to bind to specific DNA targets in gel retardation assay but not in DNase I footprinting has been previously observed with the intact HSV-1 Vmw175 protein (Papavassiliou and Silverstein, 1990). A recent one-step growth curve analysis of HSV-1 viruses with insertions at the I9, I10, or I11 positions of the HSV-1 IE3 gene, found HSV-I9 to grow as wild type whereas HSV-I11 was severely impaired (Allen and Everett, 1994). Interestingly, the HSV-I10 virus had intermediate growth characteristics which correlates with the findings of the DNA binding analyses, and indicates that the integrity of sequences around the I10 position of HSV-1 Vmw175 is required for normal function.

It was interesting to find that the VT5 peptide (140k residues 472-647) possessed DNA binding activity because a four amino acid insertion into residue 472 destroyed the regulatory functions of the intact 140k protein in transient transfection assays (Disney *et al.*, 1990). It has been observed previously that small insertions can have a more profound effect on the functions of the HSV-1 Vmw175 protein compared to deletions of extensive stretches of amino acids in the same region (Section 1C.1.9) and this may also be the case here. Alternatively the isolated DNA binding domain may have an enhanced DNA binding activity due to removal of masking protein sequences in the intact 140k protein. This could in part explain why it has been so difficult to detect sequence-specific DNA binding by the intact 140k protein (Disney, 1990). However, the effect of the insertion at reside 472 may indicate that the full DNA binding interaction, as measured by DNase I footprinting, is required for the regulatory functions of VZV 140k. The 140k peptides that were further truncated at the N-terminal end of region 2 (VT11, VT12) were efficiently expressed and soluble, but tended to aggregate (Section 3D.2), suggesting that the N-terminal portion of region 2 plays a role in the formation of the stable physical domain.

The 140k deletion peptide VT3 was capable of the full interaction with DNA, as demonstrated by the DNA binding analyses. However, the mobility of the VT3-DNA complex was greater than expected from its predicted molecular weight in gel retardation and pore exclusion gel electrophoresis analyses. It is possible that the deletion into the C-terminal end of region 2 is encroaching upon an important part of the protein and altering the conformation of the complex. Further truncations into the C-terminal end of region 2 (as with peptides VT6-VT10) resulted in very inefficient peptide expression in bacteria (or insolubility in the case of VT6) and suggests that the deleted sequences constitute a structurally essential part of the stable domain. The role played by sequences in the central portion of 140k region 2 has not been addressed in this study.

By combining all the available truncation analysis data for the HSV-1 Vmw175 DNA binding domain (see Section 1C.1.10), it would appear that Vmw175 sequences between residues 275 to 490 are sufficient for its DNA binding activity. Therefore, the 162 residue region of VZV 140k that mediates its DNA binding activity is significantly shorter than this equivalent 216 residue minimal DNA binding region of Vmw175. This indicates that the 140k protein has a lower protein sequence requirement for DNA binding compared to its HSV-1 counterpart.

The HSV-1 Vmw175 *ts*K mutation is found in a long stretch of residues that is 80% conserved between the 140k and Vmw175 DNA binding domains (Fig. 1C.3). In

contrast to the situation with Vmw175 *ts*K, the equivalent base substitution in the VZV gene 62 sequence does not produce a 140k DNA binding domain with a temperature-sensitive phenotype. Therefore, either the 140k protein is more tolerant to this mutation due to lower structural requirements for DNA binding, or the exact sequences that are important for DNA binding are not identical in these two homologous IE proteins.

3C Purification of the VZV 140k DNA binding domain as a dimer

3C.1 Strategy

It was desirable to have available purified preparations of the VZV 140k DNA binding domain peptides for physical and possible structural analyses. Once purified, the 140k DNA binding domain could also be used as an antigen for raising polyclonal antisera in rabbits. Therefore, a purification procedure was developed using the FPLC apparatus, based roughly on the previous purification of the HSV-1 Vmw175 DNA binding domain peptides (Everett *et al.*, 1991a). Like many protein domains that interact with DNA, the 140k DNA binding domain is highly basic, for example the VT2X peptide has a net charge of +16. This made it possible to separate the bacterially expressed 140k peptides from the majority of other components in the partially purified extracts by ion exchange chromatography. Further purification was achieved by a gel filtration chromatography step.

The gel filtration technique would also allow derivation of the oligomeric state of the 140k DNA binding domain peptides. As is the case for many DNA binding proteins, the native form of the HSV-1 Vmw175 protein is homodimeric (Metzler and Wilcox, 1985). Additionally, the isolated Vmw175 DNA binding domain exists as a stable dimer in solution (Everett *et al.*, 1991a), suggesting that sequences included within the DNA binding domain are involved in the dimerisation of the intact Vmw175 protein. It was of interest to see whether the counterpart VZV 140k domain was also dimeric by using gel filtration chromatography. Glycerol gradient and protein cross-linking analyses were also employed to investigate the native state of the 140k DNA binding domain.

3C.2 FPLC purification of the VZV 140k DNA binding domain to apparent homogeneity

The VT2 peptide (codons 417-647) was selected for purification, as this version of the VZV 140k DNA binding domain appeared to be capable of the full interaction with DNA (Section 3A). The optimal conditions for the purification were determined empirically and are detailed in Section 2C.5.

3C.2.1 *Anion exchange chromatography of VT2 (Mono Q column)*

Initially, the partially purified extract of the bacterially expressed VT2 peptide was loaded onto an anion exchange column (Mono Q) and a linear salt gradient applied (Section 2C.5.1). The basic VT2 peptide failed to interact with the positively charged matrix and flowed straight through this column. Thus the VT2 peptide was purified away from a significant proportion of the bacterial proteins which were retained on the column, as seen in Figure 3C.1.

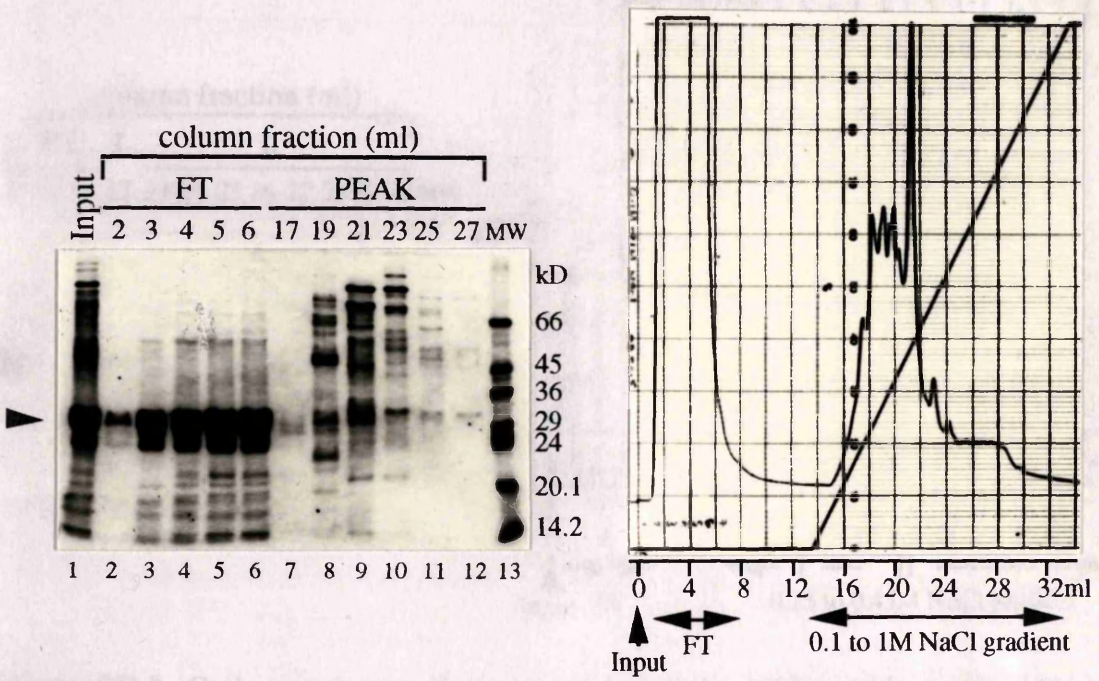


Figure 3C.1 Anion exchange chromatography of the VT2 extract on the Mono Q column. 1ml of a partially purified VT2 bacterial extract was fractionated on a Mono Q column as described in Section 2C.5.1. The left hand panel shows a Coomassie-stained 12.5% SDS PAGE analysis of selected column fractions (1ml) from the separation; track 1 contained 5µl of the input extract, tracks 2-12 contained 10µl of either flow-through (FT) fractions or fractions from the peaks eluting on the gradient, as indicated above each track. Track 13 contained molecular weight markers as detailed previously in the legend of Fig. 3A.4. The arrow marks the position of peptide VT2 in the flow-through fractions. The right hand panel shows the corresponding chart recorder uv trace for the Mono Q run (OD setting 0.5). The scale at the bottom indicates the elution volume, where each grid division on the horizontal axis represents 2ml (chart speed = 0.5cm/ml). The NaCl concentration of the elution buffer is indicated by the straight line where divisions on the vertical axis are as follows: 0 = 0.1M, 100 = 1.0M. A 0.1 to 1.0M NaCl gradient was run between 13ml to 33ml, as indicated by the diagonal line. After a slight delay due to the dead space of the column, the VT2 peptide eluted in the flow-through, which has an OD off the scale. The remaining peptides eluted on the gradient as seen from the uv trace.

3C.2.2 *Cation exchange chromatography of VT2 (Mono S column)*

The pooled Mono Q flow-through fractions were loaded directly onto a cation exchange column (Mono S, Section 2C.5.2). The VT2 peptide interacted with this column and was eluted on a 0.25 to 0.45M NaCl gradient (Fig. 3C.2); this narrow gradient was used to maximise the separation of VT2 from its proteolytic degradation fragments which eluted at similar salt concentrations to VT2. This Mono S column step succeeded in separating

the VT2 peptide away from the majority of bacterial peptides, although a major contaminant (apparent in Fig. 3C.2, track 4; approximate Mr of 24kD) co-eluted with VT2 in the lower salt concentration part of the VT2 peak and was probably a degradation product of VT2. The major peak eluting around 0.32M NaCl contained substantially purified VT2 peptide, although slight degradation products co-eluted with VT2, as seen in Figure 3C.2, tracks 5-8.

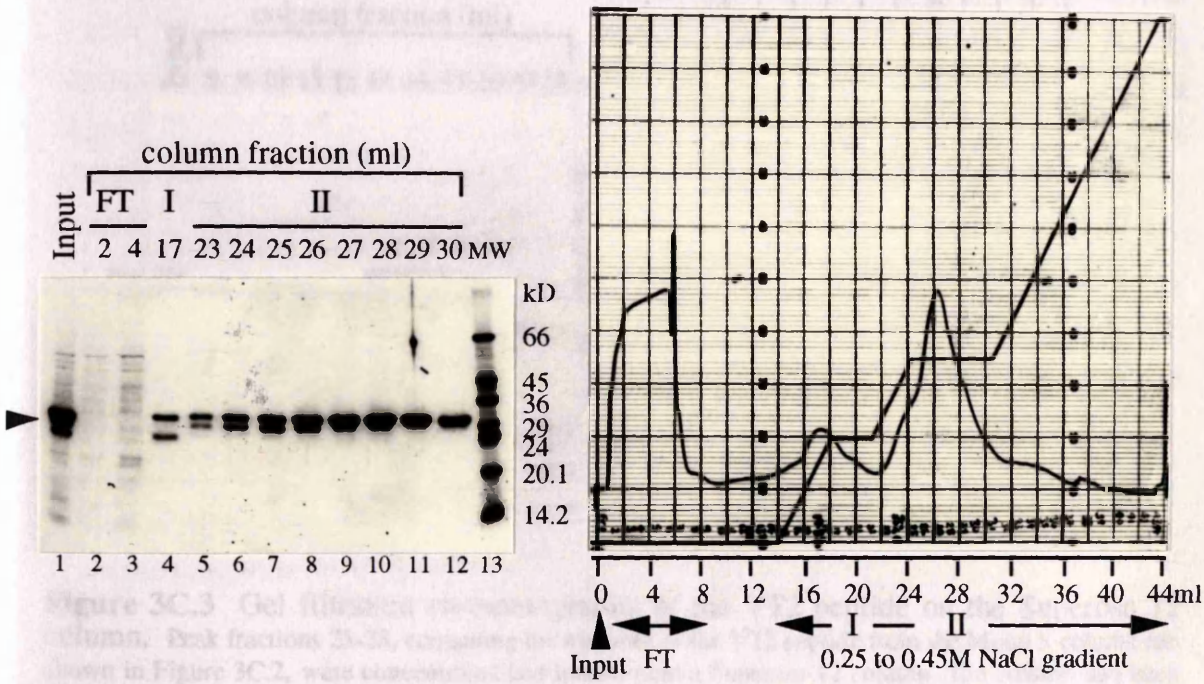


Figure 3C.2 Cation exchange chromatography of the VT2 peptide on the Mono S column. Flow through fractions 2-7 from the Mono Q column run shown in Figure 3C.1, were loaded directly onto a Mono S column and eluted as described in Section 2C.5.2. The left hand panel shows a Coomassie-stained 12.5% SDS PAGE analysis of selected column fractions (1ml) from the Mono S run; track 1 contained 5µl of the column input, while tracks 2-12 contained 10µl of the flow-through (FT) or peak fractions (peak I or II) indicated above the tracks. Track 13 contains molecular weight standards as detailed previously in the legend of Fig. 3A.4. The arrow marks the position of the VT2 peptide eluted in the peak fractions. The right hand panel gives the corresponding chart recorder uv trace for the Mono S run (OD setting 0.5). The scale at the bottom indicates the elution volume, each division on the horizontal axis represents 2ml (chart speed = 0.5cm/ml). The NaCl concentration of the elution buffer is indicated by the straight line, where divisions on the vertical axis are as follows; 0 = 0.25M, 100 = 0.45M. A 0.25 to 0.45M NaCl gradient was ran between 13.5ml to 43ml as indicated by the diagonal line. The gradient was stalled between approximately 18 and 21ml (at 0.29M NaCl), to ensure elution of all the 26kD contaminant peptide. The gradient was again stalled between approximately 24 and 31ml (at 0.32M NaCl) to obtain maximum yield of the purified VT2 peptide. The protein peaks are identified below the trace as flow-through (FT), peak I (co-elution of VT2 and 24kD degradation product) and peak II (VT2).

3C.2.3 *Gel filtration chromatography of VT2 (Superose 12 column)*

Lower molecular weight contaminants (around 16kD) were apparent after concentration of the Mono S peak fractions, as seen in track 1 of Figure 3C.3. These contaminants were separated from the VT2 peptide by gel filtration chromatography using a size exclusion column (Superose 12, Section 2C.5.3). VT2 eluted from the column as a sharp single peak, as seen on the chart recorder trace and appeared to be homogeneous by SDS

PAGE analysis, although slight degradation is apparent beneath the main VT2 band (Fig. 3C.3, track 8). DNA binding activity was maintained throughout the purification procedure (data not shown). In all the following analyses, the peptide in the Superose 12 peak fractions is referred to as the 'purified peptide'.

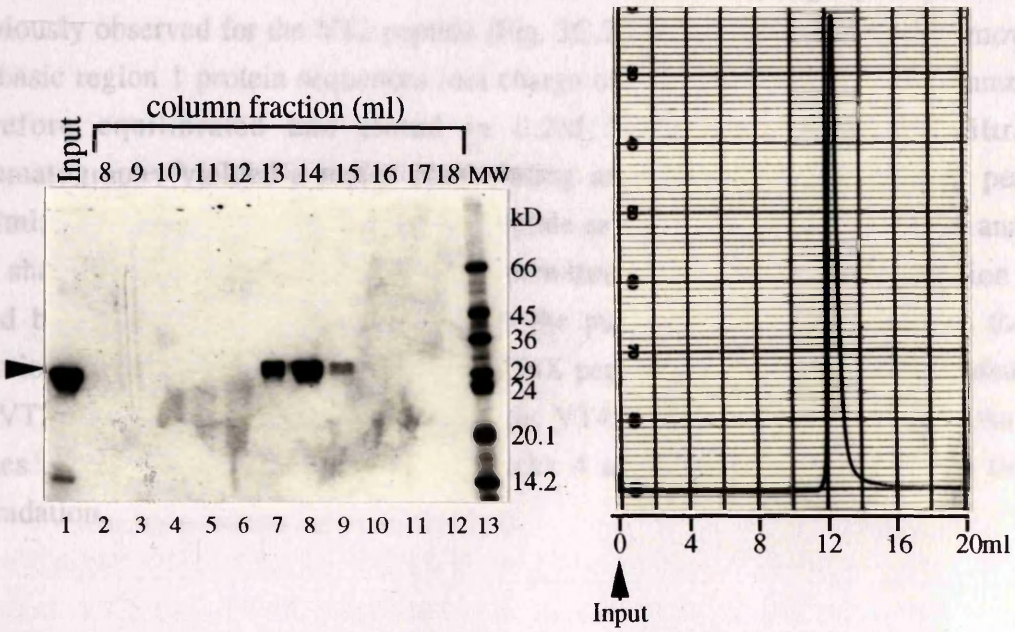


Figure 3C.3 Gel filtration chromatography of the VT2 peptide on the Superose 12 column. Peak fractions 25-28, containing the majority of the VT2 peptide from the Mono S column run shown in Figure 3C.2, were concentrated and loaded onto a Superose 12 column. The column had been equilibrated in 0.32M NaCl FPLC buffer and peptides were eluted in this same buffer, as described in Section 2C.5.3. The left hand panel shows a Coomassie-stained 12.5% SDS PAGE analysis of selected column fractions (1ml) from the Superose 12 run. Track 1 contained 2µl of the column input, while tracks 2-12 contained 10µl of the elution fractions indicated above the tracks. Track 13 contains molecular weight standards as detailed previously in the legend of Fig. 3A.4. The arrow marks the position of the VT2 peptide eluted in the peak fractions. The marks below the level of the VT2 band, particularly apparent in tracks 4-6, are a consequence of uneven destaining exaggerated by drying down of the gel. The right hand panel gives the corresponding chart recorder uv trace for the Superose 12 run (OD setting 0.5) and the scale below indicates the elution volume. The VT2 peptide eluted in fractions 13 and 14 with an elution peak at 12.62ml, the VT2 peptide gave the only detectable peak. Slight variations in the exact elution volume (+/- < 0.1ml) were obtained with different VT2 extracts.

3C.2.4 *Purification of the VT4X version of the 140k DNA binding domain*

Although the purification procedure described above yielded large amounts of highly purified 140k DNA binding domain, degradation products consistently co-purified with the VT2 peptide. This did not present a major problem for most analyses of the purified 140k domain, but these degradation products would prohibit any attempts to obtain crystals for structural studies in the future. The VT4 version of the 140k DNA binding domain had previously been found to be less prone to degradation during bacterial expression and partial purification. In addition, the 'X' version of the VT4 peptide appeared to be particularly highly stable (Section 3B.3); compare the relative lack of degradation of the VT4X peptide in Figure 3C.4, track 2 to an equivalent VT2 extract in

Figure 3C.1, track 1. For this reason, the VT4X peptide (140k codons 462 to 647) was purified with the following modifications of the above procedure: the Mono Q step was performed exactly as above, but the flow-through fractions were diluted to 0.05M NaCl and eluted from the Mono S column by a 0.05M to 0.45M NaCl gradient. Typically, VT4X eluted from the Mono S column at around 0.2M NaCl, slightly lower than that previously observed for the VT2 peptide (Fig. 3C.2), presumably due to the removal of the basic region 1 protein sequences (net charge of +9). The Superose 12 column was therefore equilibrated and eluted in 0.2M NaCl buffer. The gel filtration chromatography yielded a major peak eluting around 13.9ml and a lesser peak at 12.9ml; both peaks contained the VT4X peptide as determined by SDS PAGE analysis (not shown). Figure 3C.4 shows samples taken throughout the VT4X purification from lysed bacterial cells (track 1), through to the purified VT4X peptide from the gel filtration peak (track 5). Even though the VT4X peptide tended to be less degraded than the VT2 peptide, a slight shadow beneath the VT4X band did develop during the later stages of the purification (Fig. 3C.4, tracks 4 and 5), presumably due to limited degradation.

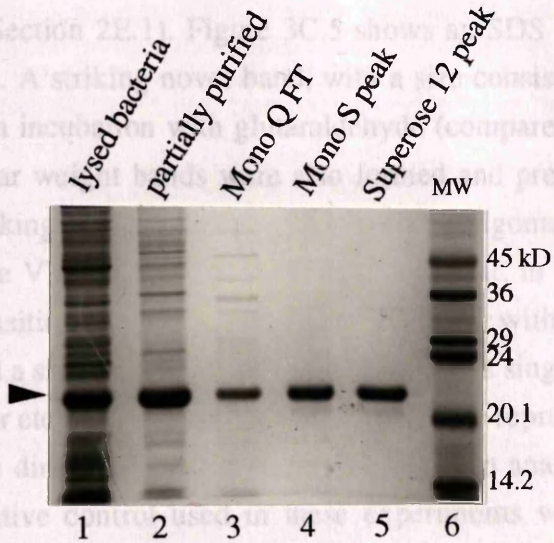


Figure 3C.4 Purification of the VZV 140k DNA binding domain peptide VT4X. Samples taken throughout the purification of the bacterially expressed VT4X peptide were analysed by 12.5% SDS PAGE and visualised by Coomassie-staining as follows: track 1 contained a sample of the soluble proteins after bacterial lysis and track 2 contained a sample of the partially purified VT4X peptide after a 35% ammonium sulphate cut, as described in Section 2C.2. Track 3 contained a sample of the combined flow-through (FT) fractions eluted from the Mono Q column, and this was the input onto the Mono S column. Track 4 contained a sample of the combined VT4X peak fractions eluted from the Mono S column, this was the input onto the Superose 12 column. Track 5 contained a sample of the final purified VT4X peptide from the combined Superose 12 peak fractions. Track 6 contained molecular weight markers as detailed previously (Fig. 3A.4). The arrow marks the position of peptide VT4X in tracks 1-5. The amount of each extract loaded was adjusted to contain roughly equivalent amounts of the VT4X peptide to highlight the relative purification at each stage of the procedure.

The purification of the HSV-1 Vmw175 DNA binding domain peptides also suffered from proteolytic degradation problems (R. D. Everett, personal communication). DNA binding activity was retained throughout the VT4X purification (not shown).

Protein standards of known molecular weight were also eluted through the Superose 12 gel filtration column: cytochrome C (13kD) eluted at 15.25ml and carbonic anhydrase (29kD) eluted at 14.11ml. From the elution volumes stated above for VT2 (28.4kD) at 12.62ml and for VT4X (21.1kD) at 13.9ml (and also at 12.9ml), it is apparent that these 140k DNA binding domain peptides are not monomeric. Comparison of the elution volumes for the 140k peptides with those of the protein standards suggests that the VZV 140k DNA binding domain may be dimeric; this was further analysed below.

3C.3 Glutaraldehyde cross-linking of VZV 140k DNA binding domain peptides

To investigate whether the 140k DNA binding domain has the ability to oligomerise, the purified VT2 and VT4X peptides were incubated with the cross-linking agent glutaraldehyde (see Section 2E.1). Figure 3C.5 shows an SDS PAGE analysis of the cross-linked products. A striking novel band, with a size consistent with being a VT2 dimer, is produced on incubation with glutaraldehyde (compare track 4 with track 1). Faint higher molecular weight bands were also formed and presumably reflect either non-specific cross-linking or a low level of higher order oligomers in the VT2 sample. The situation with the VT4X peptide was slightly different, in that a ladder of bands with decreasing intensities was produced upon incubation with glutaraldehyde. Each band in the ladder had a size consistent with the addition of a single VT4X monomer i.e. dimer, trimer, tetramer etc. (Fig. 3C.5, track 5); they may be representative of the VT4X oligomers (other than dimers) detected by the gel filtration analyses (Sections 3C.2.4 and 3C.4). The negative control used in these experiments was a protein standard (carbonic anhydrase) which is known to be monomeric in solution. Clearly, the 140k DNA binding domain has the intrinsic ability to oligomerise and at least a proportion of the 140k peptides in solution are likely to exist as dimers or higher order oligomers at any given time.

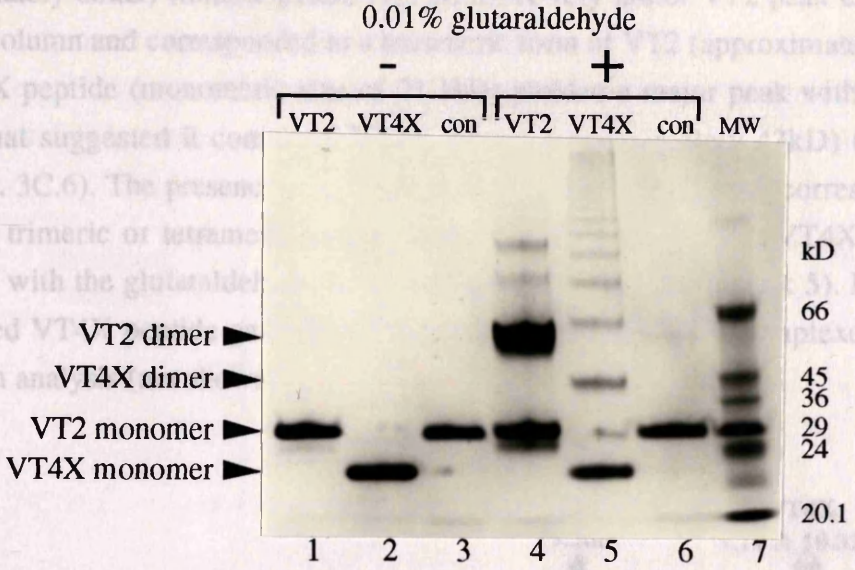


Figure 3C.5 Glutaraldehyde cross-linking analysis of the purified 140k DNA binding domain peptides. Approximately 4μg of peptides were incubated in the absence or presence of 0.01% glutaraldehyde, as described in Section 2E.1. The products were analysed by 12.5% SDS PAGE and visualised by Coomassie-staining. The identity of the peptide included in the incubations is indicated above each track. Incubations loaded in tracks 3 and 6 included a monomeric protein standard (carbonic anhydrase; Mr 29kD). Track 7 contains molecular weight standards, as described previously in the legend of Fig. 3A.4. The positions of VT2 and VT4X monomer and dimer bands are indicated on the left hand side, higher molecular weight bands in tracks 4 and 5 are apparent but remain unlabelled. Slight degradation of the VT2 peptide occurred on incubation with glutaraldehyde, apparent beneath the VT2 bands in track 4. Cross-linking by 0.01% glutaraldehyde is rarely more efficient than that seen in track 4; higher concentrations of glutaraldehyde lead to increased non-specific cross-linking.

3C.4 The native state of the VZV 140k DNA binding domain is dimeric

The multiplicity of 140k DNA binding domain peptides in the oligomers could not be accurately determined from the glutaraldehyde cross-linking analysis described above. The question also remains as to whether the oligomers are stable in solution, or whether an equilibrium exists between the monomeric and oligomeric units. The purified VT2 and VT4X peptides were analysed by analytical gel filtration and glycerol gradient centrifugation to address these questions.

A high resolution Superdex 75 gel filtration column was used to estimate the native sizes of the purified VT2 and VT4X peptides, as described in Section 2E.2. A range of protein standards of known molecular weights were analysed on the Superdex 75 column and their elution volumes were used to produce a calibration curve (left hand panel, Fig. 3C.6). The apparent size of the 140k DNA binding domain peptides were estimated from this curve; the purified VT2 peptide (monomeric size of 28.4kD) yielded a sharp peak having an elution volume consistent with it being a dimer in solution

(approximately 65kD) (middle panel, Fig. 3C.6). A very minor VT2 peak eluted faster from the column and corresponded to a tetrameric form of VT2 (approximately 110kD). The VT4X peptide (monomeric size of 21.1kD) yielded a major peak with an elution volume that suggested it contained VT4X dimers (approximately 43kD) (right hand panel, Fig. 3C.6). The presence of a faster eluting second VT4X peak corresponding to either the trimeric or tetrameric (approximately 72kD) form of the VT4X peptide is consistent with the glutaraldehyde cross-linking result (Fig. 3C.5, track 5). In addition, the purified VT4X peptide gave apparently trimeric protein-DNA complexes upon gel retardation analysis (not shown).

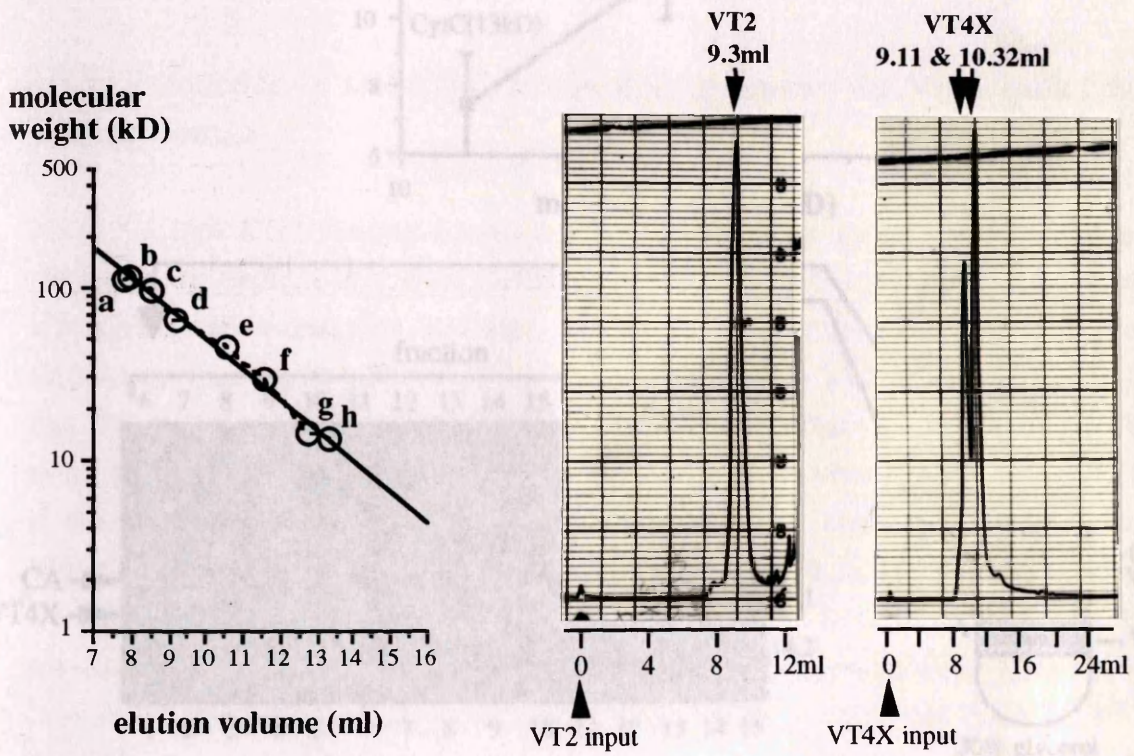


Figure 3C.6 Analytical gel filtration analysis of the purified VZV 140k DNA binding domain peptides. The left hand panel shows a calibration curve for the Superdex 75 column produced by plotting log molecular weight against elution volume for various protein standards. The accurate elution volumes were derived from integration of the peak data by the FPLC software and are as follows: a = BSA dimer (112kD) at 7.83ml; b = βgalactosidase (116.2kD) at 7.95ml; c = phosphorylase b (97.4kD) at 8.55ml; d = BSA monomer (66kD) at 9.2ml; e = albumin (45kD) at 10.56ml; f = carbonic anhydrase (29kD) at 11.59ml; g = αlactalbumin (14.2kD) at 12.77ml; h = cytochrome C (13kD) at 13.28ml. The dotted line joins all data points, the solid line is an exponential curve fitted by the graphics program; apparent molecular weights of the 140k peptides were estimated from the solid line. The middle panel shows the chart recorder uv trace obtained when approximately 50μg of purified VT2 was analysed on the Superdex 75 column (as described in Section 2E.2), at chart speed 0.5cm/ml and OD setting 0.05. The right hand panel shows the chart recorder uv trace obtained when approximately 50μg of purified VT4X peptide was analysed on the Superdex 75 column, at chart speed 0.25cm/ml and OD setting 0.2. The scale indicates column elution volumes corresponding to the uv trace. The identity of the peptides in the peaks were verified by SDS PAGE analysis of 0.25ml fractions collected throughout the column runs (not shown). The arrows at the top identify the peaks and their elution volumes.

From the gel filtration analysis, it was apparent that the 140k DNA binding domain predominantly exists in the dimeric state in solution. The failure to observe any detectable peaks eluting at the VT2 or VT4X monomer positions indicated that dimers of the 140k DNA binding domain were stable and were not in equilibrium with monomers.

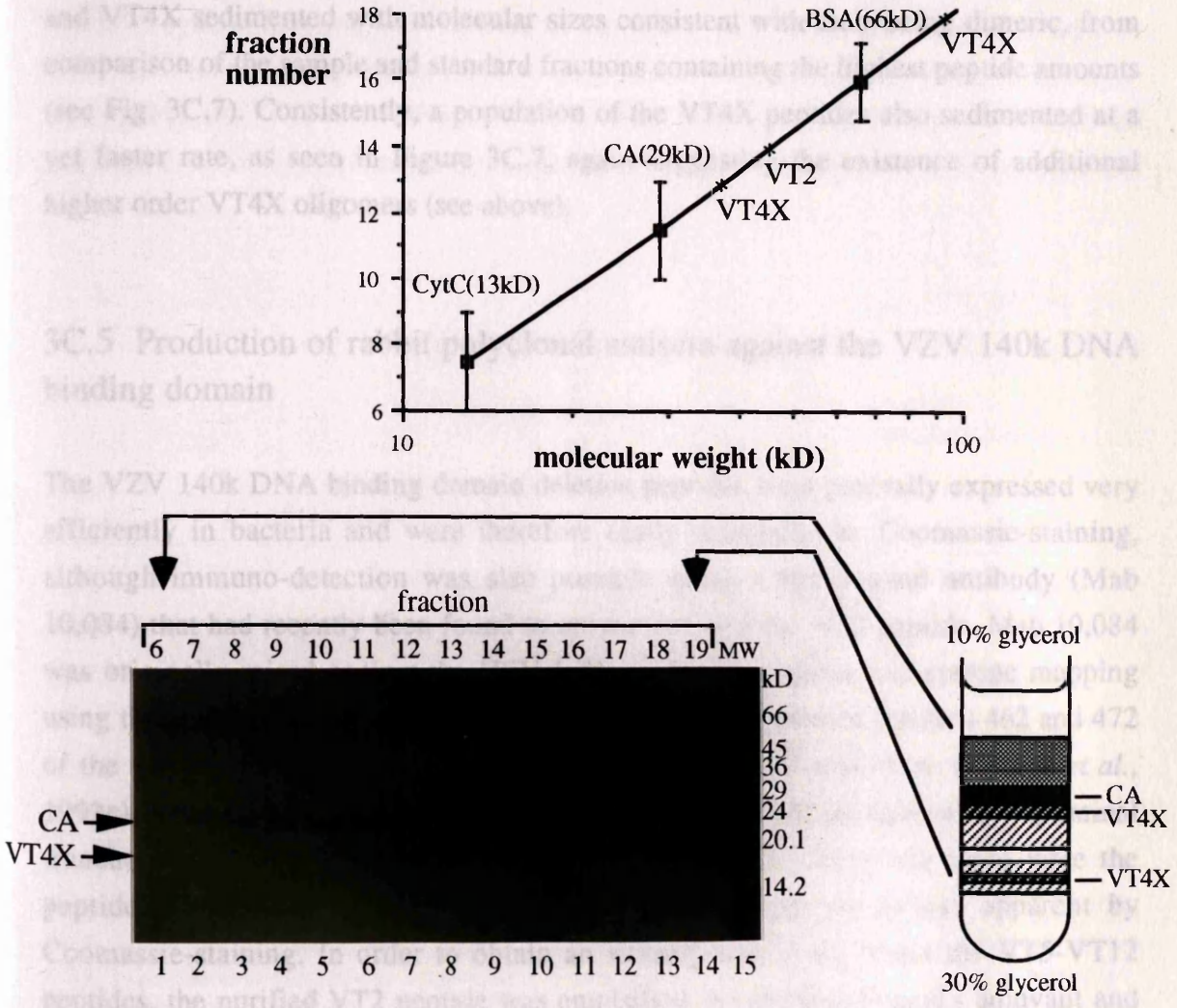


Figure 3C.7 Glycerol gradient centrifugation analysis of the VZV 140k DNA binding domain. Purified peptide samples and standards were spun on 10-30% glycerol gradients as described in Section 2E.3, and fractions analysed by SDS PAGE. At the top is shown a calibration curve for estimation of the native size of sample peptides. Protein standards: cytochrome C (CytC-13kD), carbonic anhydrase (CA-29kD) and bovine serum albumin (BSA-66kD), were analysed on a single gradient and the fractions containing the majority of each peptide plotted against their log molecular weights to yield the calibration curve shown here. Molecular weights of samples were estimated from this curve by comparing the fraction numbers containing the sample after centrifugation, asterisks denote the centre of the position of VT2 and VT4X peptides on the glycerol gradients. Samples were centrifuged with a single standard (usually CA) to allow comparison of gradients with the calibration curve. Below is shown a Coomassie-stained 13.75% SDS PAGE gel loaded with 20µl of selected fractions from a glycerol gradient analysis of purified VT4X peptide and CA. The identity of the fraction samples loaded in tracks 1-14 are given above the tracks, track 15 contained molecular weight standards (as detailed in the legend of Fig. 3A.4). A total of 25 fractions were obtained from the gradient (from the top downwards), the relative positions of fraction 6 and 19 are indicated on the diagram of the gradient on the right. The shading gives an approximate representation of the peptide sedimentation positions determined from the gel on the left, the lighter shading in each case represents lower concentrations of the relevant peptide (striped shading = VT4X, solid shading = CA).

The dimeric nature of the VZV 140k DNA binding domain was verified using glycerol gradient centrifugation analysis, as described in Section 2E.3. The sedimentation rates of the purified VT2 and VT4X peptides were compared with those of protein standards of known molecular weights analysed on parallel gradients. Even though the peptides were spread over a fairly wide range of gradient fractions, it was apparent that both VT2 and VT4X sedimented with molecular sizes consistent with their being dimeric, from comparison of the sample and standard fractions containing the highest peptide amounts (see Fig. 3C.7). Consistently, a population of the VT4X peptides also sedimented at a yet faster rate, as seen in Figure 3C.7, again suggesting the existence of additional higher order VT4X oligomers (see above).

3C.5 Production of rabbit polyclonal antisera against the VZV 140k DNA binding domain

The VZV 140k DNA binding domain deletion peptides were generally expressed very efficiently in bacteria and were therefore easily detectable by Coomassie-staining, although immuno-detection was also possible using a Monoclonal antibody (Mab 10,084) that had recently been found to cross-react with the VT2 peptide. Mab 10,084 was originally raised against the HSV-1 Vmw175 I9X peptide and epitope mapping using the 140k deletion peptides found the epitope to lie between residues 462 and 472 of the VZV 140k protein; these experiments are described elsewhere (Everett *et al.*, 1993c). Therefore, MAb 10,084 failed to recognise the 140k peptides with N-terminal truncations up to or beyond residue 472 (see Fig. 3B.1), ironically these were the peptides which also tended to be poorly expressed and not always apparent by Coomassie-staining. In order to obtain an antiserum that could detect the VT5-VT12 peptides, the purified VT2 peptide was emulsified in complete Freund's adjuvant and injected into two rabbits; subsequent injections used incomplete Freund's. The author acknowledges Miss Gillian McVey for raising the polyclonal antisera. Antiserum 108 was raised by 2 injections of VT2 and diluted out to 1:920 by western blot analysis; antiserum 109 was produced by a course of 4 injections and diluted out to 1:512,000 by western blot analysis. Antiserum 109 strongly recognised the complete series of truncated versions of the VZV 140k DNA binding domain by western blot analysis (not shown).

3C.6 Discussion

This Section (3C) describes an efficient 3 column step FPLC procedure for purification of the bacterially expressed VZV 140k DNA binding domain peptides, as outlined in Figure 3C.8. Due to the high levels of expression of the VT2 peptide obtained in the bacterial system (Section 3A.4), yields of purified peptide of around 10mg/litre culture could be obtained using this procedure. Unfortunately it has not been possible to purify completely the VT2 version of the 140k DNA binding domain away from all its degradation products, although this degradation problem was greatly reduced with the VT4X peptide.

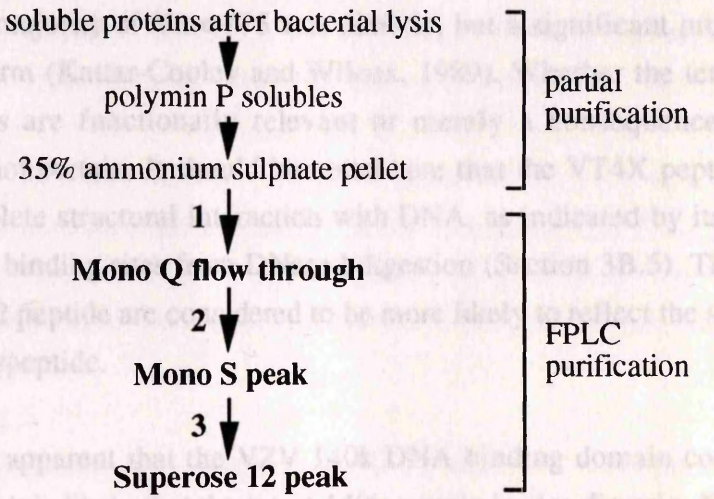


Figure 3C.8 Schematic outlining the complete purification procedure for the VZV 140k DNA binding domain peptides. The partial purification is described in Section 3A.4 and the FPLC purification is described in Section 3C.2. The numbers 1 and 2 denote ion exchange column steps, and number 3 indicates the gel filtration step.

In the past, various experimental approaches have found that the HSV-1 Vmw175 and PRV IE180 homologue proteins bind directly to their DNA target sequences (Michael *et al.*, 1988; Kattar-Cooley and Wilcox, 1989; Cromlish, 1989). Similarly, the purified 140k DNA binding domain peptides retained their DNA binding activity, indicating that the interaction of VZV 140k with DNA is also likely to be direct.

The physical analyses described in these Sections, using the purified VT2 and VT4X peptides, demonstrate that the VZV 140k DNA binding domain is predominantly dimeric in solution. The dimers appeared to be highly stable units, as they did not

detectably dissociate into monomers under the buffer conditions and the preparation procedures used for the purification, gel filtration and glycerol gradient analyses. The apparent high affinity of the monomer units for each other suggests that dimerisation of the 140k DNA binding domain is likely to occur very rapidly after translation, perhaps even concurrently with the polypeptide folding process, as previously proposed for the intact HSV-1 Vmw175 protein (Shepard *et al.*, 1991a).

Additionally, higher order oligomers (trimers or tetramers) of the 140k DNA binding domain were detected by these analyses, and this was particularly pronounced with the VT4X peptide. This again suggests that the region 1 sequences lacking from VT4X affect the structure of the 140k DNA binding domain (as discussed in Section 3B.6). Higher order oligomeric forms of the affinity purified HSV-1 Vmw175 protein have also been reported; the majority of Vmw175 was dimeric, but a significant proportion was in the tetrameric form (Kattar-Cooley and Wilcox, 1989). Whether the tetrameric forms of these proteins are functionally relevant or merely a consequence of the purification process is not certain. It should be noted here that the VT4X peptide was not capable of the complete structural interaction with DNA, as indicated by its failure to protect specific DNA binding sites from DNase I digestion (Section 3B.5). Therefore the properties of the VT2 peptide are considered to be more likely to reflect the situation with the intact 140k polypeptide.

From these data, it was apparent that the VZV 140k DNA binding domain contains a dimerisation surface and it is likely that the intact 140k protein is also dimeric. The sizes of the mapped DNA binding domains of VZV 140k and HSV-1 Vmw175 are very large compared to the established classes of DNA binding domains, and it is likely that they contain distinct sequences mediating their dimerisation and DNA recognition activities. In several of the defined classes of DNA binding domain, the amino acid sequences involved in dimerisation are adjacent to the DNA recognition motif. Both the helix-loop-helix class of proteins (for example *MyoD*; Murre *et al.*, 1989), and the leucine zipper family of DNA binding proteins (for example GCN4; Landschulz *et al.*, 1988), have a basic DNA recognition motif followed by an α-helical region that provides a dimerisation interface (see Section 1E.3). Sequences strictly conforming to these or other established types of DNA recognition and multimerisation motifs are not apparent within the DNA binding domains of the VZV 140k family of alphaherpesvirus proteins. Whether sequences within the DNA binding domain alone are sufficient for dimerisation of the intact 140k protein, or whether other protein sequences are additionally involved, remains to be determined.

3D Characterisation of the dimerisation activity of the VZV 140k DNA binding domain

3D.1 Strategy

The finding that the VZV 140k DNA binding domain existed as a stable dimer (Section 3C) presented various questions: which protein sequences are involved in dimerisation? Is it the dimeric form of the 140k DNA binding domain that interacts with DNA? If so, do both monomers specifically contact the DNA binding site? Various approaches were used to narrow down the location of sequences mediating dimerisation. Attempts were made to determine whether each of the truncated versions of the VZV 140k DNA binding domain (Fig. 3B.1) was dimeric using the physical analyses described in Section 3C (protein cross-linking, gel filtration chromatography, glycerol gradient centrifugation). Also, pairs of differently sized 140k DNA binding domain deletion peptides were co-translated, and heterodimer formation assayed by gel retardation analysis.

Due to the sequence similarity and dimeric nature of the VZV 140k and HSV-1 Vmw175 DNA binding domains, it was of interest to see whether these two homologous domains would heterodimerise. Having found that this was the case, a whole variety of experiments became possible: a series of insertion mutation versions of the HSV-1 Vmw175 DNA binding domain were co-translated with the VZV 140k DNA binding domain and heterodimer formation assayed as above; the aim was to identify insertions that might disrupt the subunit interaction. Additionally, heterodimer formation between an epitope-tagged HSV-1 Vmw175 DNA binding domain peptide and the whole spectrum of Vmw175 insertion and VZV 140k deletion peptides was assayed by co-immunoprecipitation analyses. This last approach yielded additional information, as it enabled detection of all heterodimers, not just those that bind to DNA.

3D.2 Physical analyses of the native state of the VZV 140k DNA binding domain deletion peptides

Various methods were used to try to establish whether each of the truncated versions of the VZV 140k DNA binding domain (Fig. 3B.1) existed as a dimer in solution. These analyses are briefly outlined here (data is not shown for negative results). The aim behind these experiments, and most of the following studies in these Sections, was to separate the sequences mediating the dimerisation and DNA binding activities of the VZV 140k and also the HSV-1 Vmw175 DNA binding domain.

3D.2.1 *Glutaraldehyde cross-linking of the VZV 140k DNA binding domain deletion peptides*

The intrinsic dimerisation activity of partially purified bacterial extracts of the truncated VZV 140k peptides was analysed by glutaraldehyde cross-linking (as in Section 3C.3). Peptides VT1-VT5 all produced cross-linked dimer bands, indicating that the sequences mediating dimerisation lie within region 2 of 140k. Interestingly, the peptides that yielded cross-linked dimers were also the versions of the DNA binding domain that were able to bind DNA (summarised in Fig. 3D.14). This suggested that dimerisation is a prerequisite for DNA binding. It was more difficult to determine whether the peptides with further N- or C-terminal deletions (VT7-VT12) had intrinsic dimerisation activities due to their general instability, and the low level expression of peptides VT7-VT10. When the sensitive ECL detection technique (see Section 2C.4.2) was used, additional cross-linked dimer bands were produced only by the VT11 peptide. The VT7-VT10 peptides were almost completely degraded upon cross-linking, while VT11 and VT12 were also significantly degraded, implicating sequences at both the N- and C-termini of region 2 as being necessary for the overall stability of the domain. Further analyses were required to support the cross-linking results as the instability of the VT7-VT12 peptides complicated their interpretation.

3D.2.2 *Analysis of the native state of VZV 140k peptides with N-terminal deletions of region 2 sequences*

The low level bacterial expression of the VT7-VT10 peptides made their analysis by gel filtration chromatography and glycerol gradient centrifugation impossible. However, peptides VT11 and VT12 were both expressed to high levels (see Section 3B.3). The majority of the VT11 peptide and all the VT12 peptide sedimented to the bottom of the glycerol gradients as aggregated complexes (in gradient fraction 25; see Fig. 3C.7 for an explanation of the technique). However, a population of the VT11 peptides reproducibly sedimented to a position on the gradient that was between the positions predicted for

VT11 monomers and dimers, as determined by comparison with protein standards (not shown). Although it is not possible for a protein to sediment at a faster rate than determined by its molecular weight, a protein with an open conformation will be more retarded on the gradient than might otherwise be expected. Therefore, it is likely that these VT11 peptides were partially unfolded, yet still dimeric. Additionally, the VT11 and VT12 peptides were not amenable to further purification by ion exchange chromatography (as in Section 3C.2), even though the deleted sequences were not particularly highly charged. Gel filtration chromatography (Superose 12 and Superdex 75 columns) found peptides VT11 and VT12 to elute in every column fraction, presumably due to peptide aggregation or precipitation on the column, despite elution in 0.5M NaCl buffers. Purification and analytical gel filtration chromatography in the presence of 10% glycerol at all stages was more successful; although the majority of peptide VT11 and VT12 eluted from the Superdex 75 column as very high molecular weight complexes, an additional VT11 peak was detected in the dimer position. However it remained possible that this was also a consequence of aggregation. These gel filtration results are consistent with the glycerol gradient centrifugation analysis (above). Although it appeared that a minority of the VT11 peptides were dimeric, it was difficult to determine unequivocally the native state of peptides VT11 and VT12 from these analyses (but see Section 3D.9). Clearly the deletion of sequences from the N-terminal part of region 2 had drastic conformational effects on the 140k DNA binding domain.

3D.2.3 *Analysis of the native state of VZV 140k peptides with C-terminal deletions of region 2 sequences, by use of the GST fusion system*

The poorly expressed C-terminal deletion versions of the 140k DNA binding domain were expressed in bacteria as glutathione-S-transferase (GST) fusion proteins (Ausubel *et al.*, 1993b). The VT7, VT9 and VT10 regions of 140k were stably expressed to high levels as GST fusion peptides; addition of N-terminal sequences presumably stabilised these 140k peptides, which were otherwise very poorly expressed. GST itself is a dimeric protein. Therefore the 140k portions of the fusion polypeptides had to be separated from GST to enable analysis of the native state of the 140k peptides. Specific sequences between the GST coding region and the MCR of the vector pGEX2TN3, resulted in fusion proteins with a protease (thrombin) recognition site immediately following the GST sequences (see Fig. 3D.1). Thrombin treatment of the purified peptides GSTVT7, GSTVT9 and GSTVT10 each yielded a cleavage product of the expected size (Section 2C.6; for VT10, see Fig. 3D.2, track 3). However, only the VT10 cleavage peptide gave a discrete peak upon gel filtration analysis and clearly eluted as a monomer (Fig. 3D.2). Additionally, the VT10 cleavage product failed to yield dimer bands by glutaraldehyde cross-linking analysis (not shown).

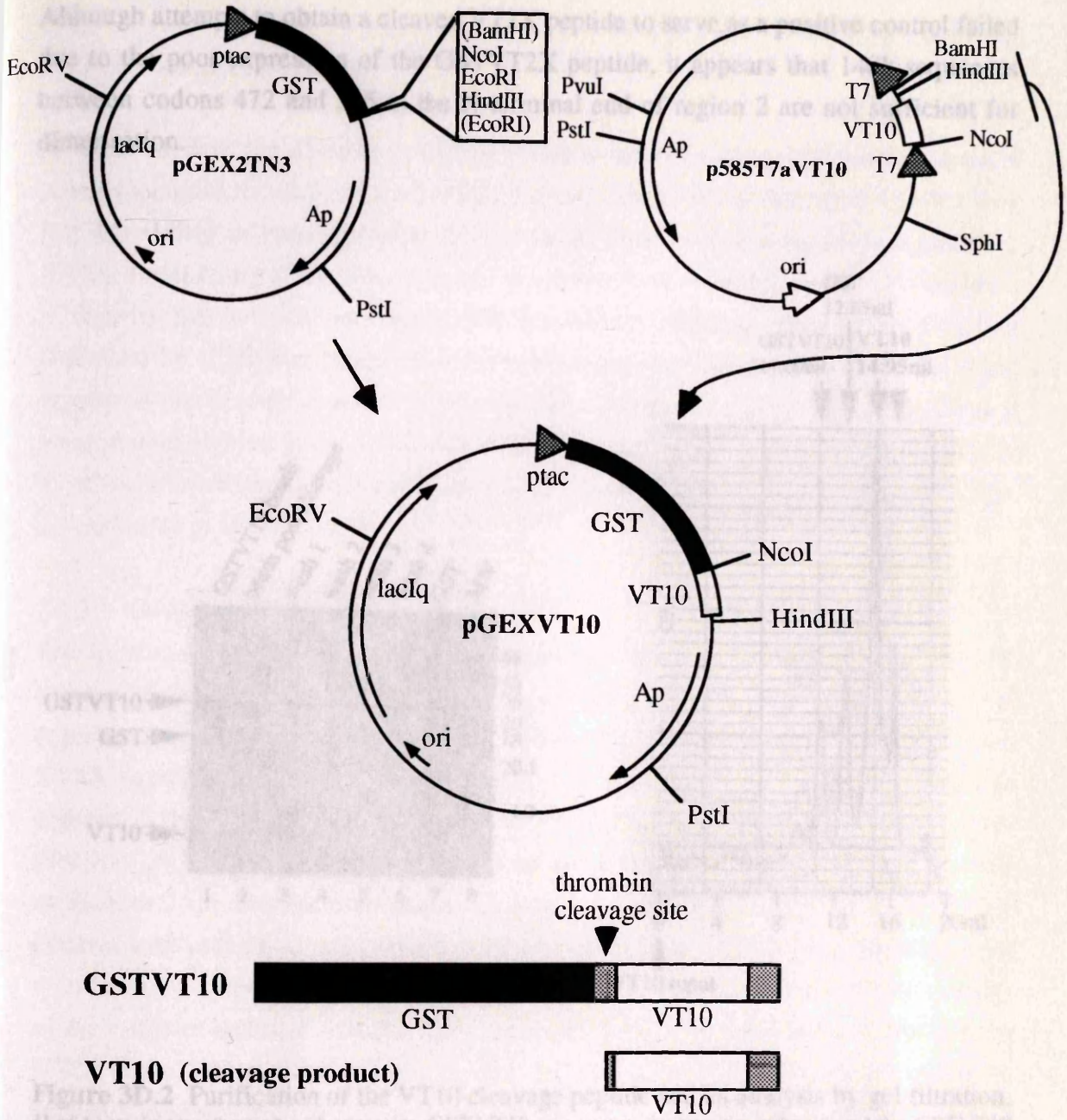


Figure 3D.1 Construction of the pGEXVT10 plasmid and its expression product, GSTVT10. At the top are shown the plasmids pGEX2TN3 and p585T7aVT10 that were used to construct plasmid pGEXVT10, only selected restriction enzyme sites are shown. Plasmid pGEX2TN3 was derived from the commercial pGEX2TN vector (Pharmacia), by altering the restriction sites downstream of the glutathione-S-transferase (GST) sequences (black box), as shown (R. D. Everett, personal communication). Shaded triangles represent expression signals. The *Nco*I-*Hind*III fragment of plasmid p585T7aVT10 (described in Section 3B.2) was inserted into the *Nco*I-*Hind*III digested pGEX2TN3 vector to yield plasmid pGEXVT10. Below are shown representations of peptides GSTVT10 and VT10 (cleavage product). The GSTVT10 expression product of plasmid pGEXVT10 comprises 234 GST residues (black box) fused to the VT10 region of 140k (residues 472-525) (white box). Between the GST and VT10 sequences are 11 vector-encoded amino acids which include a thrombin cleavage site [LVPR/GS]. To the C-terminus of the 140k sequences are 16 vector-encoded residues (shaded boxes represent vector-encoded residues). Thrombin cleavage yields the VT10 peptide with 3 N-terminal vector-encoded residues fused to 140k residues 472-525, followed by 16 C-terminal vector-encoded residues. The 315 residue GSTVT10 peptide has a predicted size of 35kD, the 73 residue VT10 cleavage product has a predicted size of 8.1kD.

Although attempts to obtain a cleaved VT2X peptide to serve as a positive control failed due to the poor expression of the GSTVT2X peptide, it appears that 140k sequences between codons 472 and 525 at the N-terminal end of region 2 are not sufficient for dimerisation.

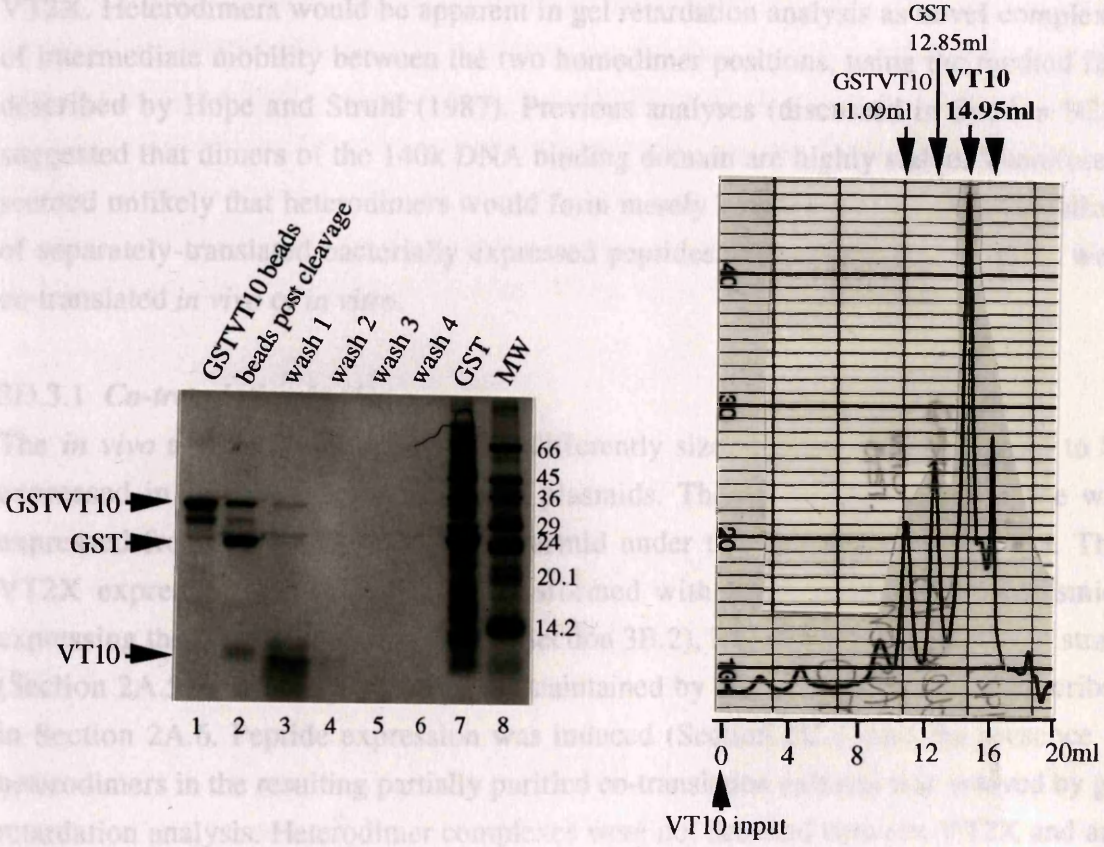


Figure 3D.2 Purification of the VT10 cleavage peptide and its analysis by gel filtration. 1l of bacteria transformed with plasmid pGEXVT10 was grown, expression induced and the GSTVT10 fusion peptide purified as described in Section 2C.6. The VT10 peptide was released from GSTVT10 by thrombin cleavage and samples taken before and after cleavage were analysed by 20% SDS PAGE (tricine buffer system) as shown on the left. Track 1 contained 1µl of purified GSTVT10-agarose beads, track 2 contained 1µl of the beads post thrombin cleavage (after the washes), tracks 3-6 contained 20µl of sequential washes of the beads after thrombin cleavage. Track 7 contained 5µl of a bacterial GST extract and molecular weight markers were loaded in track 8 (as described in legend to Figs. 3A.4). Wash 1 was concentrated and loaded in 200µl onto a Superdex 75 gel filtration column and eluted in 0.3M NaCl buffer, as described in Section 2E.2. Shown on the right is the chart recorder uv trace from one such column run (OD setting 0.05). The peptides eluting in each peak were identified by SDS PAGE analysis of 0.5ml fractions collected throughout the run, and are indicated above the trace with their respective elution volumes. A small amount of un-cleaved GSTVT10 and GST cleavage product remained in the preparation analysed here as seen in track 3 of the left hand panel, the minor unlabelled peak was thought to contain proteolytic fragments of VT10 peptide. The peak eluting at 14.95ml contained the VT10 cleavage product (8.1kD) as determined by SDS PAGE, this peak contained peptides of approximately 7kD by comparison with the calibration curve shown in Fig. 3C.6. An independent repeat of the full procedure described here gave a VT10 elution peak at 14.85ml.

3D.3 Co-translation of pairs of differently sized VZV 140k DNA binding domain deletion peptides

Another method for discerning whether each truncated 140k DNA binding domain peptide included the sequences required for dimerisation, was to determine whether they had the ability to heterodimerise with a larger dimerisation competent peptide i.e. VT2X. Heterodimers would be apparent in gel retardation analysis as novel complexes of intermediate mobility between the two homodimer positions, using the method first described by Hope and Struhl (1987). Previous analyses (discussed in Section 3C.6) suggested that dimers of the 140k DNA binding domain are highly stable. Therefore it seemed unlikely that heterodimers would form merely by mixing together preparations of separately-translated bacterially expressed peptides. Instead pairs of peptides were co-translated *in vivo* or *in vitro*.

3D.3.1 ***Co-translation in vivo***

The *in vivo* approach required the two differently sized 140k deletion peptides to be expressed in bacteria from compatible plasmids. Therefore the VT2X peptide was expressed from a pACYC184 based plasmid under the T7 expression signals. This VT2X expressing plasmid was co-transformed with each of the p585T7 plasmids expressing the truncated 140k peptides (Section 3B.2), into the appropriate *E. coli* strain (Section 2A.5) and both plasmids were maintained by antibiotic selection as described in Section 2A.6. Peptide expression was induced (Section 2C.2) and the presence of heterodimers in the resulting partially purified co-translation extracts was assayed by gel retardation analysis. Heterodimer complexes were not detected between VT2X and any of the range of deletion peptides (data not shown). Even co-translation of the dimeric VT2X and VT4X peptides failed to produce any additional bands, although it should be said that the resolution of this analysis was low. A logical explanation for the lack of heterodimer formation is that the pBR322 and pACYC184 based plasmids localise to different replication compartments within the bacterial cell, so that their expression products are liable to be translated at different spatial cellular locations. The results of this analysis support the idea that dimerisation occurs very rapidly after translation. In view of the success of the co-immunoprecipitation procedure (Section 3D.9), it should be possible to co-express an epitope-tagged DNA binding domain with each 140k deletion peptide in the manner described above, followed by co-immunoprecipitation, to assay more stringently for *in vivo* heterodimer production.

3D.3.2 *Co-translation in vitro*

An *in vitro* approach would facilitate formation of heterodimers by overcoming the possible problems of spatial separation of translation of peptides encoded by different plasmids (above). Transcription and translation of pairs of 140k DNA binding domain peptides were performed within a single reaction using the TNT T7 Coupled Reticulocyte Lysate System (Promega), as described in Section 2C.7. Each 140k DNA binding domain deletion peptide was co-translated with the larger DNA binding-competent, dimeric VT2X peptide, in the presence of a radiolabelled amino acid. All peptides were expressed from the p585T7 series of expression vectors (Section 3B.2), previously used for the bacterial expression (Section 3B.3). The efficiency of translation of the radiolabelled peptides was determined by SDS PAGE, before using the *in vitro* translated peptides in further analyses. Figure 3D.3 shows typical *in vitro* translation extracts of several of the 140k DNA binding domain deletion peptides.

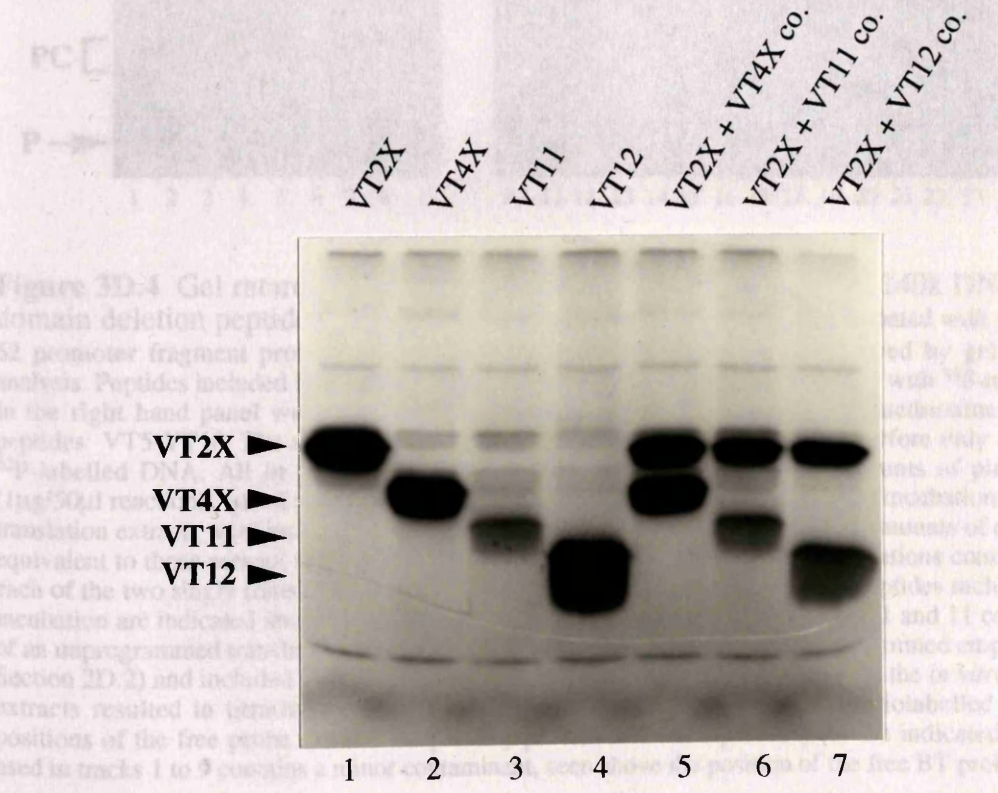


Figure 3D.3 *In vitro* translation extracts of VZV 140k DNA binding domain deletion peptides. Peptides VT2X, VT4X, VT11 and VT12 were translated individually (tracks 1-4) co-translated with peptide VT2X (tracks 5-7) using the TNT T7 Coupled Rabbit Reticulocyte Lysate System, as described in Section 2C.7. The gel above shows an autoradiograph of a 13.5% SDS PAGE analysis of 5µl aliquots from *in vitro* translation reactions. The identity of peptides expressed in each reaction are given above the tracks. Tracks labelled 'co.' indicate pairs of peptides co-translated within a single reaction. The position of each peptide is indicated on the left hand side. Comparative western blot analysis found 1µl of an *in vitro* translated extract to contain the same amount of VT2X peptide as 0.02µl of partially purified bacterial extract (approximately 5ng VT2X per µl *in vitro* translation extract) (not shown).

3D.4 The VZV 140k DNA binding domain binds to DNA as a dimer

For each 140k DNA binding domain deletion peptide, single translation and VT2X co-translation extracts (of the type shown in Fig. 3D.3), were incubated with the radiolabelled BT gene 62 promoter fragment (Fig. 3A.6) under the specific conditions detailed in Section 2D.2. Protein-DNA complexes were resolved by gel retardation analysis and Figure 3D.4 shows one such analysis.

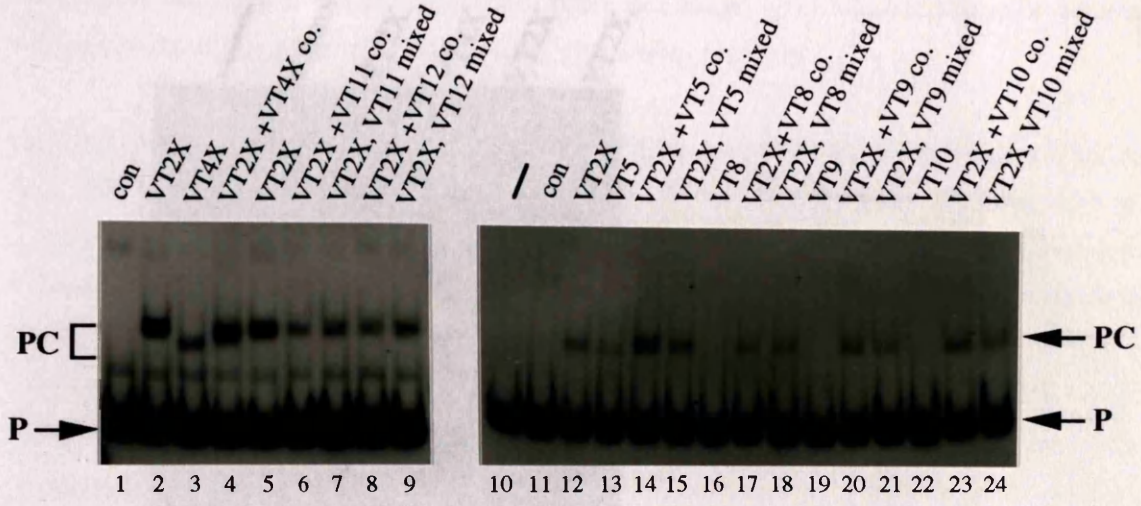


Figure 3D.4 Gel retardation analysis of the *in vitro* translated VZV 140k DNA binding domain deletion peptides. Aliquots of *in vitro* translation extracts were incubated with the BT gene 62 promoter fragment probe (see Fig. 3A.6 for details) and complexes resolved by gel retardation analysis. Peptides included in the analysis in the left hand panel were radiolabelled with ^{35}S-met, peptides in the right hand panel were labelled with ^{3}H-leu due to the low number of methionine residues in peptides VT5-VT10. The autoradiographs shown here are second films and therefore only detected the ^{32}P-labelled DNA. All *in vitro* translation reactions contained equivalent amounts of plasmid DNA (1µg/50µl reaction); 1µl of single translation extracts were included in the relevant incubations, 2µl of co-translation extracts were included in the incubations in the 'co.' tracks (to provide amounts of each peptide equivalent to those present in the singly translated incubations). The 'mixed' incubations contained 1µl of each of the two singly translated extracts. The identity of the *in vitro* translated peptides included in each incubation are indicated above the tracks. Track 10 contains the free probe, tracks 1 and 11 contained 1µl of an unprogrammed translation reaction. Optimal incubation conditions were determined empirically (see Section 2D.2) and included no specific competitor, as the plasmid DNA present in the *in vitro* translation extracts resulted in titration of the DNA binding domains away from the radiolabelled target. The positions of the free probe (P) and the primary protein-DNA complexes (PC) are indicated. The probe used in tracks 1 to 9 contains a minor contaminant, seen above the position of the free BT probe.

The singly *in vitro* translated peptides VT8 (track 16), VT9 (track 19), VT10 (track 22), VT11 and VT12 (not shown) did not bind to DNA, while peptides VT2X (tracks 2, 5 and 12), VT4X (track 3) and VT5 (track 13) interacted with the DNA probe. Therefore the behaviour of the *in vitro* translated peptides is in agreement with the previous DNA binding analysis of the bacterially expressed 140k DNA binding domain deletion peptides (Section 3B.4). Due to the similar sizes of the VT2X, VT4X and VT5 peptides, it was not possible to be certain whether heterodimer complexes were formed in the

co-translation extracts in tracks 4 and 14 of Figure 3D.4. Therefore, gel retardation analysis on 35cm gels was used to increase the resolution between the VT2X and VT4X protein-DNA complexes (Fig. 3D.5). When the two separately-translated proteins were incubated together, only two bands corresponding to the separate VT2X and VT4X protein-DNA complexes were observed (Fig. 3D.5, track 6).

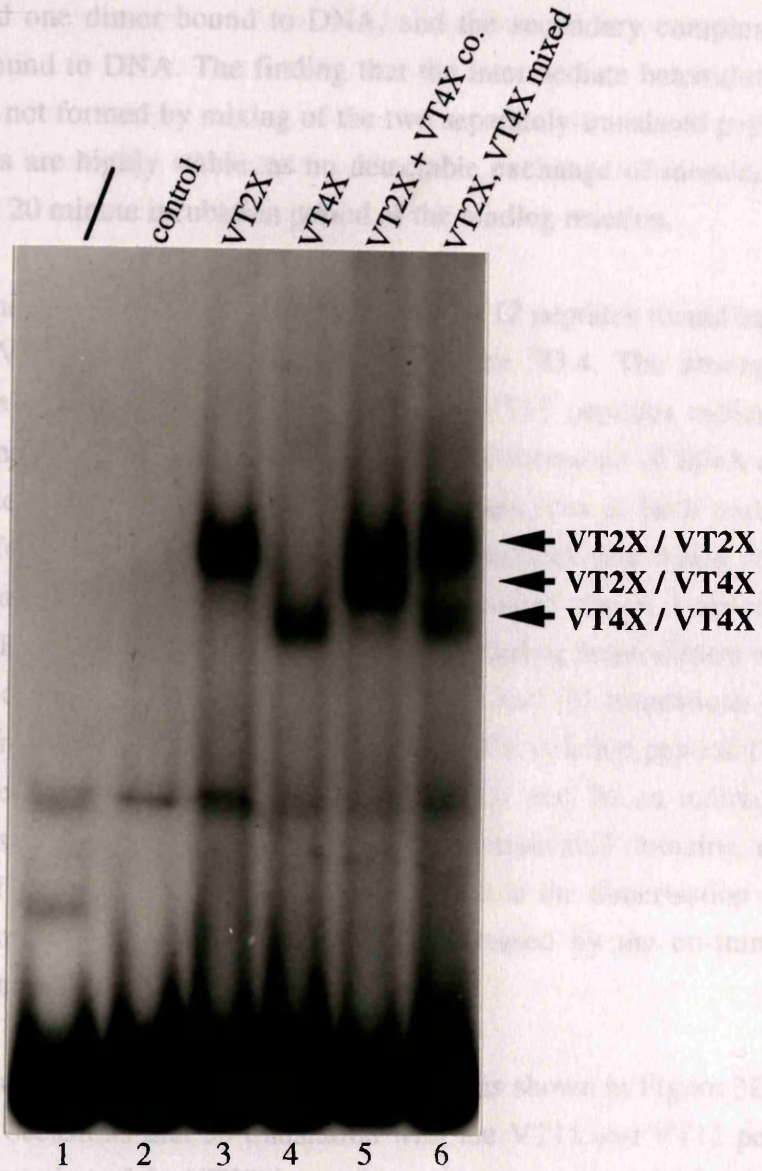


Figure 3D.5 Gel retardation analysis showing VT2X/VT4X heterodimers formed by *in vitro* co-translation. *In vitro* translation extracts were incubated with the radiolabelled gene 62 promoter BT fragment (details in Fig. 3A.6) and complexes resolved by gel retardation analysis on 35cm gels. The gel shown here is a second film autoradiograph of one such analysis. *In vitro* translations included 0.02µg total plasmid DNA per 50µl reaction. Incubations loaded in tracks 3 and 4 contained 2µl of single peptide translation extracts, track 5 contained 4µl of the co-translation extract and track 6 contained 2µl of each single translation extract. Track 1 contains the free probe and the incubation in track 2 contained 2µl of the unprogrammed translation extract. The positions and identity of peptide dimer-DNA complexes are indicated on the right. In the particular co-translation extract used in track 5, the VT2X peptide was more efficiently translated than VT4X, explaining why the VT4X homodimer complex is relatively faint here. The faster mobility faint bands are due to minor contaminants of the radiolabelled DNA.

However, when the two peptides were co-translated, a third intermediate band was observed (Fig. 3D.5, track 5), which must be due to heterodimer formation between the VT2X peptide and the smaller VT4X peptide. This result proves unequivocally that it is the dimer of the 140k DNA binding domain that recognises the target binding sites. Therefore all primary protein-DNA complexes mentioned so far in this Thesis comprised one dimer bound to DNA, and the secondary complexes comprised two dimers bound to DNA. The finding that the intermediate heterodimeric VT2X/VT4X band was not formed by mixing of the two separately-translated peptides confirms that the dimers are highly stable, as no detectable exchange of monomer subunits occurs within the 20 minute incubation period of the binding reaction.

Heterodimers with the relatively small VT8-VT12 peptides would have migrated faster than the VT2X homodimer complex in Figure 3D.4. The absence of heterodimer complexes between the VT2X and the VT8-VT12 peptides indicates that deletions encroaching on either end of region 2 prevented formation of DNA binding-competent heterodimers. Although it is possible that sequences at both ends of region 2 are required for dimerisation, these results are inconclusive since (a) truncated 140k peptides lacking either the N- or C-terminal end of region 2 failed to bind to DNA (Section 3B.4), so it is possible that non-DNA binding heterodimers were produced that remain undetected by this method of analysis, and (b) truncations into either end of region 2 previously destabilised the resultant 140k deletion peptide (Sections 3B.3 and 3D.2). Therefore failure to dimerise with VT2X may be an indirect consequence of general disruptions of the conformation of the truncated domains, rather than due to deletion of specific sequences directly involved in the dimerisation of the 140k DNA binding domain. These possibilities are addressed by the co-immunoprecipitation experiments described later (Section 3D.9).

Although not particularly apparent in the analysis shown in Figure 3D.4, it was noticed on several occasions that co-translation with the VT11 and VT12 peptides resulted in reduced formation of the VT2X homodimer complex, as compared to that seen when the separately-translated peptides were mixed. It was possible that non-DNA binding heterodimers form between these peptides and VT2X, resulting in a reduction of the VT2X homodimer band. This was further analysed by setting up a series of co-translation reactions, so as to yield increasing amounts of peptide VT11 and a constant amount of peptide VT2X (left hand panel of Fig. 3D.6). Volumes of each extract containing equal amounts of the VT2X peptide were analysed by gel retardation analysis. As suspected, co-translation in the presence of VT11 resulted in reduction of the VT2X homodimer complex; see the right hand panel of Figure 3D.6, tracks 4-9. However, this was not considered to be solely a consequence of formation of non-DNA

binding VT11/VT2X heterodimers because mixing the separately-translated VT11 and VT2X peptides also resulted in a significant reduction of the VT2X homodimer complex (compare track 10 to track 3 in right hand panel, Fig. 3D.6), albeit to a lesser degree than that occurring when VT2X and VT11 were co-translated in the same proportions (track 9). Previously, the bacterially expressed VT11 and VT12 peptides tended to aggregate either with themselves and / or with other proteins in general (Section 3D.2.2). The reduction of the VT2X homodimer complex in Figure 3D.6 is likely to reflect, at least in part, non-specific interactions between peptides VT2X and VT11; co-translation presumably increases the probability of these interactions. Identical findings were provided by the VT12 peptide (not shown).

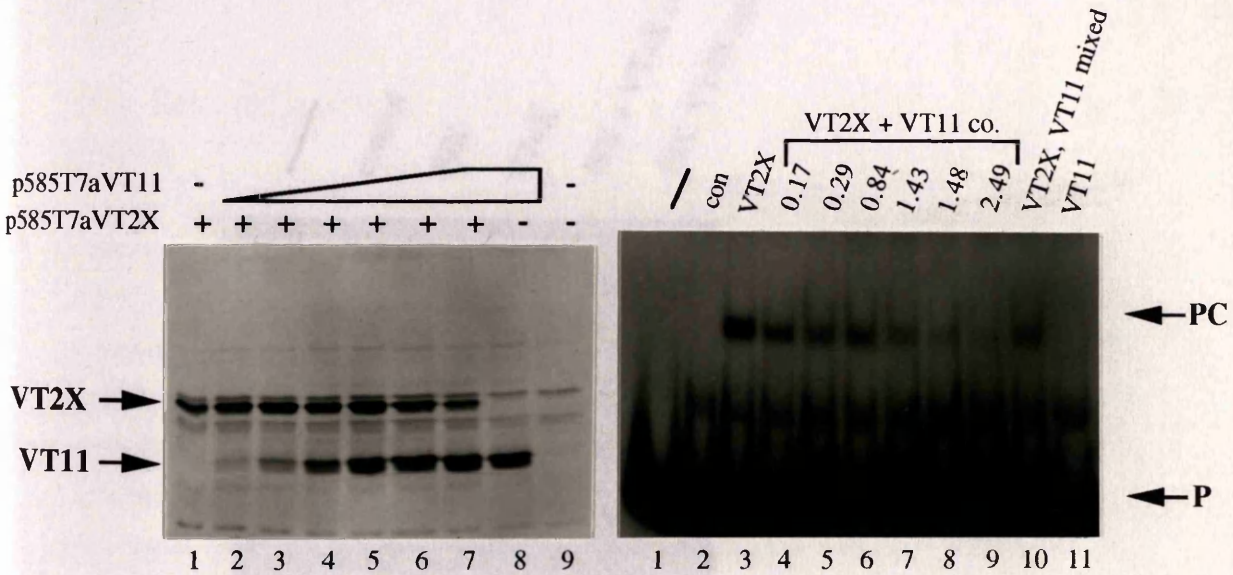


Figure 3D.6 Gel retardation analysis showing non-specific reduction of the VT2X homodimer complex by the VT11 peptide. The left hand panel shows an autoradiograph of a 12.5% SDS PAGE analysis of 2µl aliquots of *in vitro* translated extracts of the VT2X and VT11 peptides. A constant amount of plasmid p585T7aVT2X and increasing amounts of plasmid p585T7aVT11 were included in the reactions analysed in tracks 2-7. Track 1 and 8 included plasmids p585T7aVT2X and p585T7aVT11 alone respectively, track 9 is an unprogrammed lysate. The total amount of plasmid DNA in each translation reaction was equalised to 0.25µg/50µl reaction with plasmid p585T7a. The positions of the VT11 and VT2X peptides are indicated on the left, the relative amounts of VT2X and VT11 peptides in each track were quantitated by PhosphorImaging. The right hand panel shows the second film autoradiograph of the corresponding gel retardation analysis, produced when the *in vitro* translated extracts shown on the left were incubated with the radiolabelled gene 62 promoter BT fragment (details in Fig. 3A.6). Volumes of each translation extract containing equal amounts of the VT2X peptide were used in each incubation. The peptide concentration ratio of VT11/VT2X in the co-translated extracts is indicated above tracks 4-9. The incubations in tracks 10 and 11 contained an amount of VT11 peptide equivalent to the amount of VT11 included in track 9. The extra band above the position of the BT probe is a non-specific lysate complex (P = free probe, PC = primary protein-DNA complex).

3D.5 The VZV 140k DNA binding domain heterodimerises with the corresponding domain of HSV-1 Vmw175

Given that the DNA binding domains of the VZV 140k and the HSV-1 Vmw175 proteins are both dimeric and that they share a fairly high level of amino acid homology (Fig. 1C.3), it seemed possible that these homologous domains may dimerise in a similar manner. To test this, the HSV-1 Vmw175 and VZV 140k DNA binding domains were co-translated *in vitro,* followed by gel retardation analysis of the co-translation extracts.

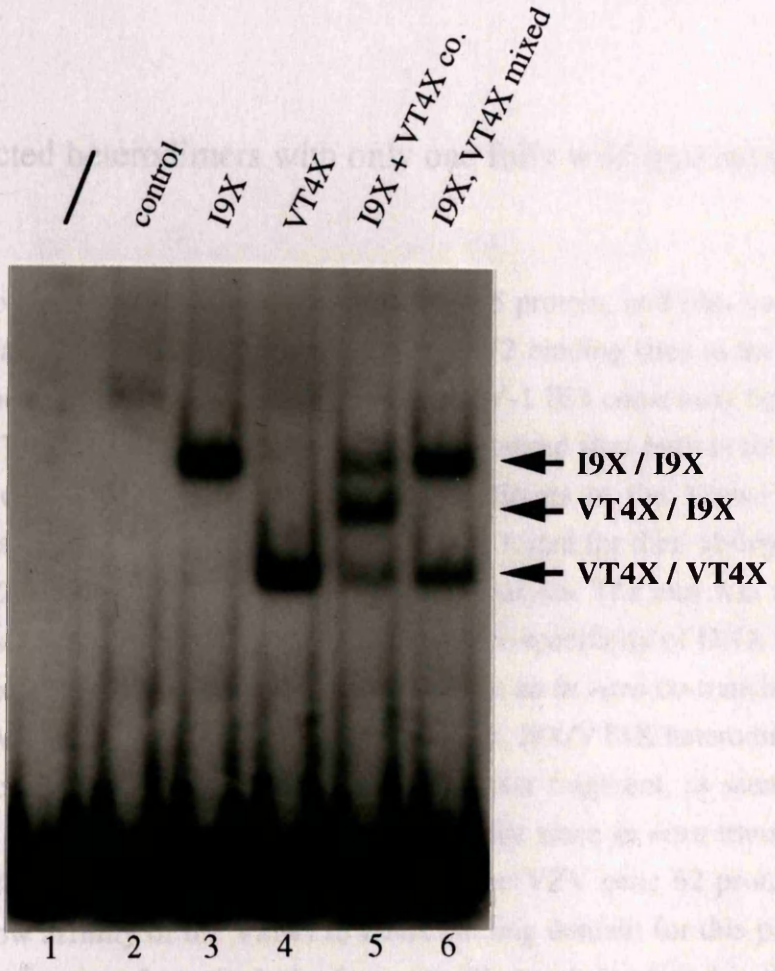


Figure 3D.7 Gel retardation analysis of heterodimer formation between the DNA binding domains of HSV-1 Vmw175 and VZV 140k. *In vitro* translation extracts were incubated with the radiolabelled HSV-1 IE3 promoter fragment IE3S (details in Fig. 3A.11) and complexes resolved by gel retardation analysis. Shown here is a second film autoradiograph of one such analysis. The VT4X peptide basically comprises region 2 of VZV 140k. Peptide I9X (expressed from plasmid pT7I9X; Pizer *et al.*, 1991) spans region 2 of HSV-1 Vmw175, but also includes some region 1 and region 3 sequences. Incubations loaded in tracks 3 and 4 contained 2µl of single peptide translation extracts, track 5 contained 4µl of the VT2X-I9X co-translation extract, track 6 contains 2µl of each single translation extract. Track 1 contains the free probe and the incubation in track 2 contained 2µl of an unprogrammed translation extract. The positions and identity of peptide dimer-DNA complexes are indicated on the right.

The Vmw175 peptide I9X and the 140k peptide VT4X were utilised due to their large size difference (see Fig. 3B.7), and a DNA probe containing the HSV-1 IE3 consensus binding site was used for the gel retardation analysis. Figure 3D.7, track 5 clearly shows the formation of VT4X/I9X heterodimers of intermediate mobility between the two homodimer bands; the relative amounts of the three complexes were in the expected 1:2:1 proportions. The VT4X/I9X heterodimer was absent when the separately-translated peptides were mixed after translation (Fig. 3D.7, track 6). The ease of formation of the heterodimers between the 140k and Vmw175 DNA binding domain peptides suggests that the VZV 140k and HSV-1 Vmw175 proteins have very similar structural requirements, and also perhaps similar sequence requirements for dimerisation.

3D.6 Selected heterodimers with only one fully wild type subunit can bind to DNA

It was previously noticed that the HSV-1 Vmw175 protein, and also its isolated DNA binding domain, had very low affinities for the VT2 binding sites in the VZV gene 62 promoter, compared to their affinities for the HSV-1 IE3 consensus binding site (see Fig. 3A.15). The VZV 140k domain however, recognised sites both in the VZV gene 62 promoter and the HSV-1 IE3 promoter. Heterodimers of the Vmw175/140k DNA binding domains produced *in vitro* (as above) were tested for their ability to bind to the VZV gene 62 promoter probe in gel retardation analysis. The aim was to see whether both subunits of the dimer contribute to the sequence-specificity of DNA binding. Track 5 in the top panel of Figure 3D.8, clearly shows that an *in vitro* co-translation extract of the I9X and VT4X peptides yielded I9X homodimer, I9X/VT4X heterodimer and VT4X homodimer complexes with the HSV-1 IE3 promoter fragment, as seen previously in Figure 3D.7. In a parallel experiment, aliquots of the same *in vitro* translation extracts were tested for their DNA binding activities on the VZV gene 62 promoter fragment probe. The low affinity of the Vmw175 DNA binding domain for this particular probe fragment was apparent from the lack of any specific bands in track 4 in the lower panel of Figure 3D.8. However the I9X/VT4X heterodimer clearly still recognised the gene 62 promoter binding sites, as a band in a position consistent with the production of a heterodimer-DNA complex was formed upon co-translation of the two homologous DNA binding domains (lower panel of Fig. 3D.8, track 5). This result suggests that the specific interaction of the single VT4X monomer with its DNA recognition site was sufficient to mediate the DNA binding activity of the I9X/VT4X heterodimer. Other possible explanations are discussed in Section 3D.10.

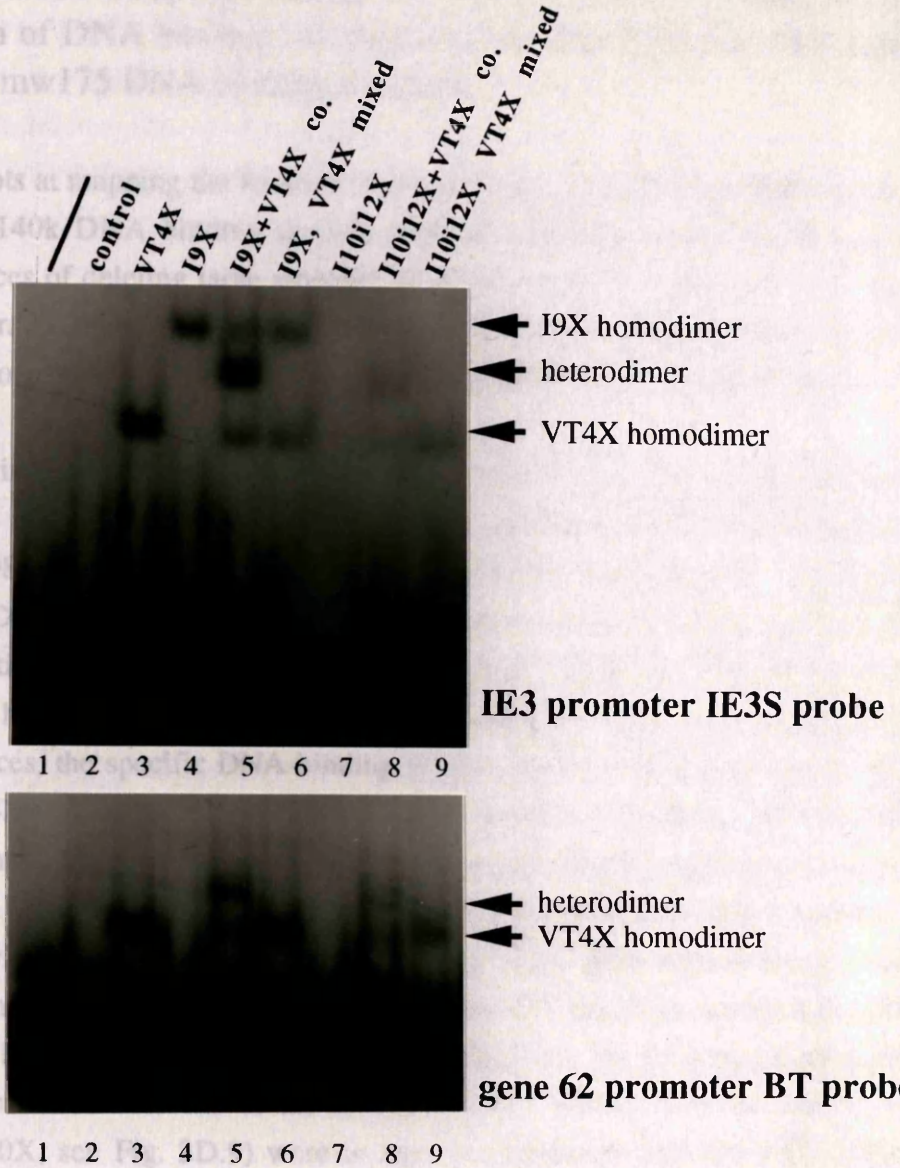


Figure 3D.8 Gel retardation analysis of Vmw175/140k DNA binding domain heterodimers on the HSV-1 IE3 and VZV gene 62 promoter fragments. *In vitro* translation extracts, as indicated above the tracks, were incubated with the radiolabelled DNA fragments and complexes resolved by gel retardation analysis. Shown here are second film autoradiographs. The top panel shows an analysis of the HSV-1 IE3 promoter fragment IE3S probe (details in Fig. 3A.11), the lower panel shows an analysis of aliquots of the identical *in vitro* translation extracts on the VZV gene 62 promoter BT fragment (details in Fig. 3A.6). The VT4X peptide spans the VZV 140k DNA binding domain, and peptide I9X spans the HSV-1 Vmw175 DNA binding domain. The I10i12X peptide spans the Vmw175 DNA binding domain and includes an insertion at amino acid 310 (see Section 3D.7). The amounts of extract included in each incubation were comparable to those used in the analysis in Figure 3D.4. In both panels: track 1 contained the free probe and track 2 contained 2µl of an unprogrammed translation extract. The positions and identity of peptide dimer-DNA complexes are indicated on the right hand side.

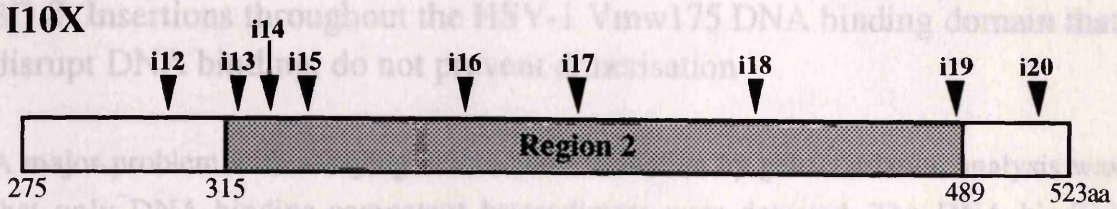
3D.7 Insertion mutations throughout region 2 of HSV-1 Vmw175 disrupt formation of DNA binding-competent heterodimers of the VZV 140k and HSV-1 Vmw175 DNA binding domains

The attempts at mapping the location of dimerisation sequences described so far, using the VZV 140k DNA binding domain deletion peptides, were complicated by the consequences of deleting large stretches of sequences which appeared to be necessary for the overall stability of the domain. Therefore the more subtle approach of looking at the effects of small insertion mutations on heterodimer formation was adopted.

Numerous insertion mutations had previously been introduced at positions throughout the HSV-1 Vmw175 protein, for the mapping of functional domains (Paterson and Everett, 1988a; b). Many of the insertions within the mapped DNA binding domain prevented DNA binding of the intact Vmw175 protein, and it was possible that some of these mutations affected DNA binding indirectly by interfering with dimerisation. The VT4X/I9X heterodimer experiment (described in Section 3D.6) suggested that in certain circumstances, the specific DNA binding property of the heterodimer can be provided by the VT4X monomer alone. Therefore it was possible that similar heterodimers with altered forms of the Vmw175 DNA binding domain subunit might also bind to DNA. Such DNA binding-competent mutant heterodimers would identify positions within region 2 which are involved in DNA binding rather than dimerisation. Sequences encoding the DNA binding domains of the Vmw175 insertion mutation proteins had previously been transferred to T7 expression plasmids (R. D. Everett, unpublished). Insertion mutation versions of the isolated HSV-1 Vmw175 DNA binding domain (peptide I10X, see Fig. 3D.9) were *in vitro* co-translated with the VZV 140k DNA binding domain peptide (VT4X) and heterodimer formation assayed by gel retardation analysis, as in Section 3D.4. The DNA binding activities of the individually translated insertion mutation versions of the I10X domain closely corresponded to those previously observed with the intact Vmw175 protein (Paterson and Everett, 1988b), as summarised in Figure 3D.9. This verifies the validity of the *in vitro* approach using the isolated DNA binding domain peptides.

The i12 insertion at the C-terminal end of region 1 of HSV-1 Vmw175 greatly reduced the DNA binding activity of the intact Vmw175 protein (Paterson and Everett, 1988b; see Fig. 3D.9). In accordance, the *in vitro* translated Vmw175 DNA binding domain including the i12 insertion (peptide I10i12X), produced no apparent complex with either of the DNA probe fragments used in Figure 3D.8, track 7. Although the results in Figure 3D.8 are quite faint, it can be seen that the I10i12X/VT4X heterodimers recognised both

the HSV-1 IE3 and the VZV gene 62 promoter probes with apparent affinities comparable to the VT4X homodimers (Fig. 3D.8, track 8). This result is consistent with the idea that in certain circumstances only one fully wild type monomer of the dimer is required for the interaction with DNA in a specific manner. Furthermore, it appears that sequences within the C-terminal end of region 1 of Vmw175 are involved in DNA binding rather than dimerisation of the DNA binding domain.



DNA binding activity

Insertion mutation	Intact Vmw175	Isolated Vmw175 DNA binding domain	
		homodimer	heterodimer
i12 (310)	+/-	-	+++
i13 (324)	+++	+++	++
i14 (329)	-	-	-
i15 (337)	-	-	-
i16 (373)	-	-	-
i17 (398)	+	-	-
i18 (438)	+	-	-
i19 (494)	+++	+++	+++
i20 (518)	++	++	+

Figure 3D.9 DNA binding activities of the *in vitro* translated insertion mutation versions of the HSV-1 Vmw175 DNA binding domain. At the top is shown a representation of the I10X DNA binding domain peptide of HSV-1 Vmw175 (Pizer *et al.*, 1991). Region 2 is highly conserved with the VZV 140k protein and is shown shaded. The first and last amino acid residues of the I10X peptide and Vmw175 region 2 are given. The position of the 4 amino acid insertion mutations, i12 to i20, are given in the first column of the table and also indicated on the diagram of the I10X peptide. The second column of the table gives the previously published DNA binding activities of HeLa nuclear extracts containing insertion mutation versions of the intact Vmw175 protein (Paterson and Everett, 1988b). The DNA binding activities of the insertion mutation versions of the HSV-1 Vmw175 DNA binding domain, expressed *in vitro* from the pT7I10i12X-pT7I10i20X series of plasmids, are also given: the third column gives the DNA binding activity of singly translated I10X derived peptides, as determined by gel retardation analysis of the IE3S probe. The final column gives the ability to form heterodimer-DNA complexes following *in vitro* co-translation with the VZV 140k DNA binding domain peptide VT4X. The symbols - to +++ provide an arbitrary scale to indicate the relative intensities of protein-DNA complexes obtained in gel retardation analyses.

However, I10X peptides containing single insertion mutations at positions throughout region 2 (i14 to i18), failed to bind to DNA as either I10X homodimers or as I10X/VT4X heterodimers (Fig. 3D.9). Therefore either sequences throughout region 2 are required for dimerisation, or the insertion mutations induce conformational alterations which inhibit dimerisation indirectly. Alternatively heterodimers may form, but fail to bind DNA (see below and Section 3D.10).

3D.8 Insertions throughout the HSV-1 Vmw175 DNA binding domain that disrupt DNA binding, do not prevent dimerisation

A major problem with assaying heterodimer formation by gel retardation analysis was that only DNA binding-competent heterodimers were detected. The DNA binding activity of the HSV-1 Vmw175 protein appears to be highly sensitive to conformational changes (Section 1C.1.9). Therefore it is likely that some heterodimers which include a truncated or mutated monomer fail to bind to DNA. Therefore the approach of co-immunoprecipitation of heterodimers from solution was adopted. A panel of monoclonal antibodies raised against the DNA binding domain of Vmw175 (Everett *et al.*, 1993c) was tested in immunoprecipitation experiments. Unfortunately, all those tested were very inefficient (not shown). Therefore an epitope-tagged Vmw175 DNA binding domain was constructed. MAb 11,060 (Everett *et al.*, 1993b) had recently been found to recognise strongly the HSV-1 Vmw110 protein in immunoprecipitations (Meredith *et al.*, 1994). Sequences spanning the MAb 11,060 epitope were fused to the N-terminus of the Vmw175 DNA binding domain, as described in Figure 3D.10. The author is very grateful to Maggie Elliott for assistance with cloning of the epitope.

Having established that the HSV-1 Vmw175 DNA binding domain (I10X peptide) could heterodimerise with the epitope-tagged NXI10X domain (not shown), NXI10X was co-translated *in vitro* with each of the insertion mutation versions of I10X (see Fig. 3D.9), followed by co-immunoprecipitation analysis, as described in Section 2E.4. The insertions had no apparent effect on peptide solubility, as all remained in the soluble fraction after pre-clearing of the *in vitro* translation extracts (upper panel, Fig. 3D.11).

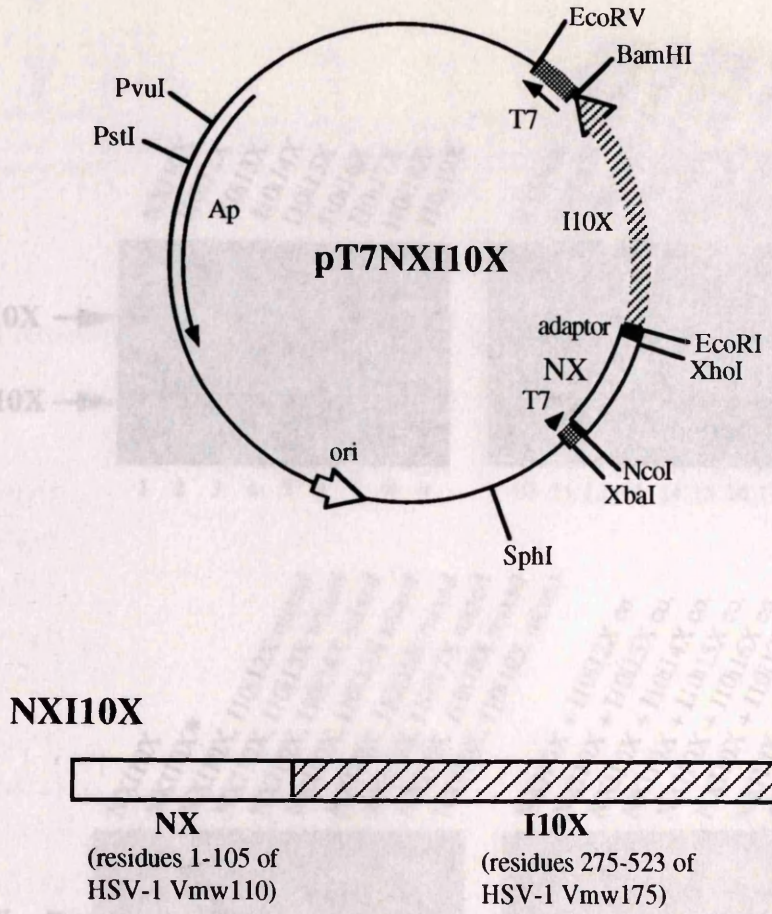


Figure 3D.10 The epitope-tagged HSV-1 Vmw175 DNA binding domain. At the top is shown T7 expression plasmid pNXI10X, only selected restriction enzyme sites are indicated. An *Eco*RI-*Pst*I fragment of plasmid pJI10X (including Vmw175 codons 275-523; striped box) was fused to a *Pst*I-*Xho*I fragment of plasmid pT7110 (including Vmw110 codons 1-105; white box) with the aid of an *Xho*I-*Eco*RI adaptor (black box) which also restores the reading frame. Below is shown a representation of the NXI10X expression product of plasmid pT7NXI10X. Shading of portions of the HSV-1 Vmw175 and Vmw110 coding sequences are as above. Translation of peptide NXI10X begins on the first ATG of the HSV-1 Vmw110 ORF and a single vector-encoded residue is at the C-terminal end of the peptide, which has a total predicted size of 40kD. The authenticity of the expressed peptide was verified by western blot analysis of bacterially expressed NXI10X peptide, using MAbs recognising the Vmw110 epitope and a Vmw175 epitope respectively.

It was apparent that none of the insertions in the Vmw175 DNA binding domain prevented dimerisation with the epitope-tagged domain, as all were specifically co-immunoprecipitated by MAb 11,060 (tracks 11-18 in the lower panel of Figure 3D.11). Additionally, the interaction was a consequence of specific heterodimerisation as none of the I10X insertion mutation peptides were co-immunoprecipitated when the separately-translated peptides were mixed together (lower panel of Fig. 3D.11, tracks 3-10).

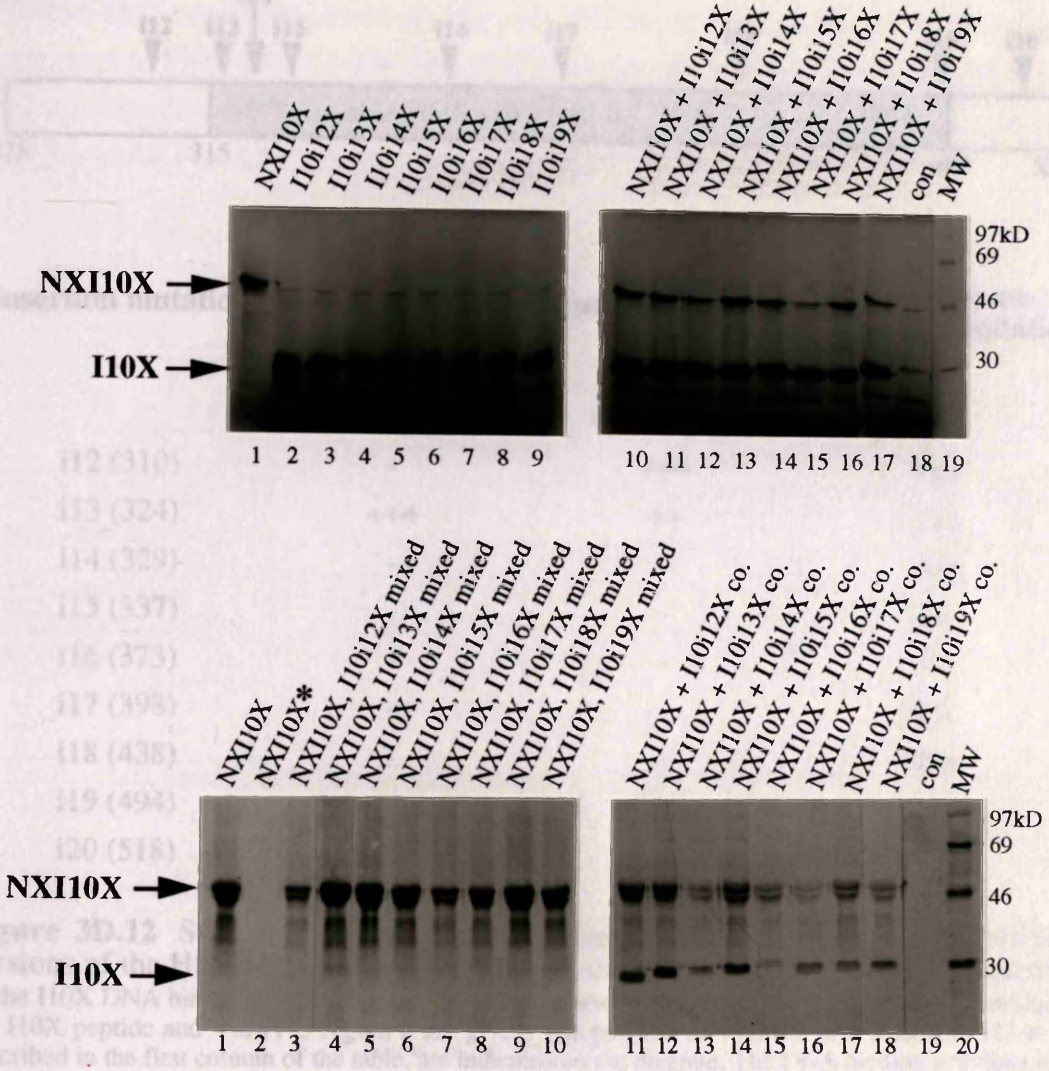
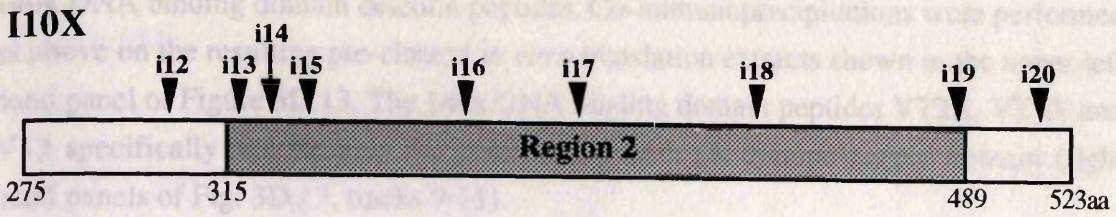


Figure 3D.11 Co-immunoprecipitation analysis of the insertion mutation versions of the HSV-1 Vmw175 DNA binding domain by an epitope-tagged domain. Insertion mutation versions of the HSV-1 Vmw175 DNA binding domain were *in vitro* translated individually, or co-translated in pairs with the epitope-tagged NXI10X peptide (Fig.3D.10). The reaction products were 'pre-cleared' to reduce background, as described in Section 2E.4. The top panels show autoradiographs of 12.5% SDS PAGE analyses of 20µl aliquots of pre-cleared extracts of the *in vitro* translated peptides, as indicated above each track. The lower panels show autoradiographs of 12.5% SDS PAGE analyses of a co-immunoprecipitation assay of the pre-cleared *in vitro* translated extracts shown above. MAb 11,060 recognises the N-terminal Vmw110 epitope of peptide NXI10X and was used for all immunoprecipitations, * with the exception of track 2 which included MAb 10,503 which is specific for a C-terminal epitope of Vmw110. The peptides included in the immunoprecipitation reactions were either co-translated or mixed after translation, as indicated above the tracks in the lower panel. 'con' indicates an unprogrammed translation reaction. 'MW' indicates ^{14}C radiolabelled molecular weight markers: phosphorylase b, 97kD; bovine serum albumin, 69kD; ovalbumin, 46kD; carbonic anhydrase, 30kD. The positions of peptides NXI10X (40kD) and I10X (28.2kD) are indicated on the left, the Vmw110 portion of the NXI10X fusion peptide was very highly acidic and caused the peptide to have a lower mobility than expected on SDS PAGE analysis. MAb 11,060 was highly efficient at immunoprecipitating the NXI10X peptide; the amount of peptides included in the immunoprecipitation reactions was equivalent to the amount of pre-cleared extracts analysed in the top panel (both sets of gels were exposed to film for three days). The split nature of the NXI10X peptide in the lower panels was probably due to immunoglobulin polypeptides migrating at the same position as NXI10X.



Insertion mutation	DNA binding activity		Dimeric by co-immunoprecipitation
	homodimer	heterodimer	
i12 (310)	-	+++	yes
i13 (324)	+++	++	yes
i14 (329)	-	-	yes
i15 (337)	-	-	yes
i16 (373)	-	-	yes
i17 (398)	-	-	yes
i18 (438)	-	-	yes
i19 (494)	+++	+++	yes
i20 (518)	++	+	N/D

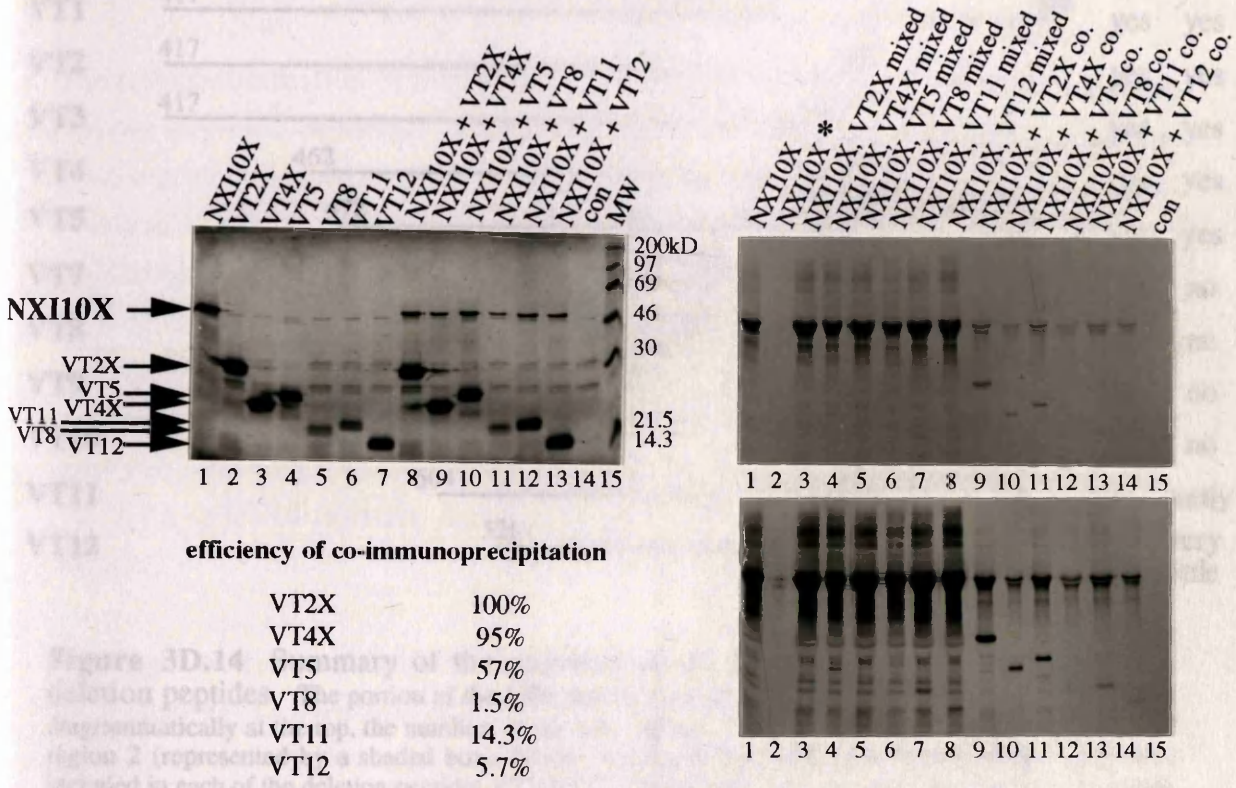
Figure 3D.12 Summary of the activities of the *in vitro* translated insertion mutation versions of the HSV-1 Vmw175 DNA binding domain. At the top is shown a representation of the I10X DNA binding domain peptide of HSV-1 Vmw175. The first and last amino acid residues of the I10X peptide and Vmw175 region 2 are given. The position of 4 amino acid insertions i12 to i20, described in the first column of the table, are indicated on the diagram. The DNA binding activities of the *in vitro* translated insertion mutation versions of the HSV-1 Vmw175 DNA binding domain as homodimers and as VT4X heterodimers are given in columns two and three, as given previously in Figure 3D.9. The third column gives the ability of each peptide to dimerise with the epitope-tagged NXI10X peptide, as determined in Fig. 3D.11. N/D indicates an experiment not performed.

From these studies it is apparent that the DNA binding interaction was more mutation sensitive than the monomer interaction, as many of the insertions within region 2 of HSV-1 Vmw175 disrupted DNA binding yet none affected dimerisation (summarised in Fig. 3D.12).

3D.9 Deletion of either the N- or C-terminus of VZV 140k region 2 disrupts dimerisation

Taking advantage of the previous finding that the VZV 140k and HSV-1 Vmw175 DNA binding domains can heterodimerise (Section 3D.5), the epitope-tagged HSV-1 Vmw175 DNA binding domain peptide was *in vitro* co-translated with each of the VZV

140k DNA binding domain deletion peptides. Co-immunoprecipitations were performed as above on the resulting pre-cleared *in vitro* translation extracts shown in the upper left hand panel of Figure 3D.13. The 140k DNA binding domain peptides VT2X, VT4X and VT5 specifically and strongly heterodimerised with the epitope-tagged domain (right hand panels of Fig. 3D.13, tracks 9-11).



efficiency of co-immunoprecipitation

VT2X	100%
VT4X	95%
VT5	57%
VT8	1.5%
VT11	14.3%
VT12	5.7%

Figure 3D.13 Co-immunoprecipitation of the VZV 140k DNA binding domain deletion peptides following *in vitro* co-translation with the epitope-tagged HSV-1 Vmw175 DNA binding domain. VZV 140k DNA binding domain deletion peptides were *in vitro* translated individually, or co-translated in pairs with the epitope-tagged NXI10X peptide (Fig. 3D.10). The top left hand panel shows an autoradiograph of a 12.5% SDS PAGE analysis of 20µl aliquots of the pre-cleared extracts (see Section 2E.4) of the *in vitro* translated ^{35}S labelled peptides, as indicated above each track. The right hand panel shows two different autoradiographic exposures of a 12.5% SDS PAGE analysis of co-immunoprecipitation assays of the pre-cleared extracts shown on the left. MAb 11,060 was used for all immunoprecipitations, * with the exception of track 2 which included MAb 10,503 as a negative control. The peptides included in the immunoprecipitation reactions were either co-translated or mixed after separate translation, as indicated above the tracks in the right hand panels. 'con' indicates an unprogrammed translation reaction. The relative positions of *in vitro* translated peptides in all panels are indicated on the far left. The amount of NXI10X peptide in tracks 1 and 3-8 in the right hand panels, is approximately 10 times the amount present in tracks 9-14. This reflects the higher amount of NXI10X peptide in the pre-cleared extract used in these particular immunoprecipitations (track 1, left hand panel) which is approximately 10 times greater than that in the co-translation extracts (tracks 8-13 on the left). Instability of the NXI10X fusion peptide is apparent following the longer exposure. 'MW' indicates ^{14}C radiolabelled molecular weight markers: myosin, 200kD; phosphorylase b, 97kD; bovine serum albumin, 69kD; ovalbumin, 46kD; carbonic anhydrase, 30kD; trypsin inhibitor, 21.5kD; lysozyme, 14.3kD. The bottom left hand panel gives the corresponding quantitated co-immunoprecipitation efficiencies for the 140k DNA binding domain deletion peptides, as determined by PhosphorImaging. The values are calculated as ratio of VT peptide/NXI10X for tracks 9-14 (right hand panel), minus the corresponding background value (ratio of VT peptide/NXI10X for tracks 3-8, right hand panel) expressed as a percentage of maximum co-immunoprecipitation obtained with the VT2X peptide (100%).

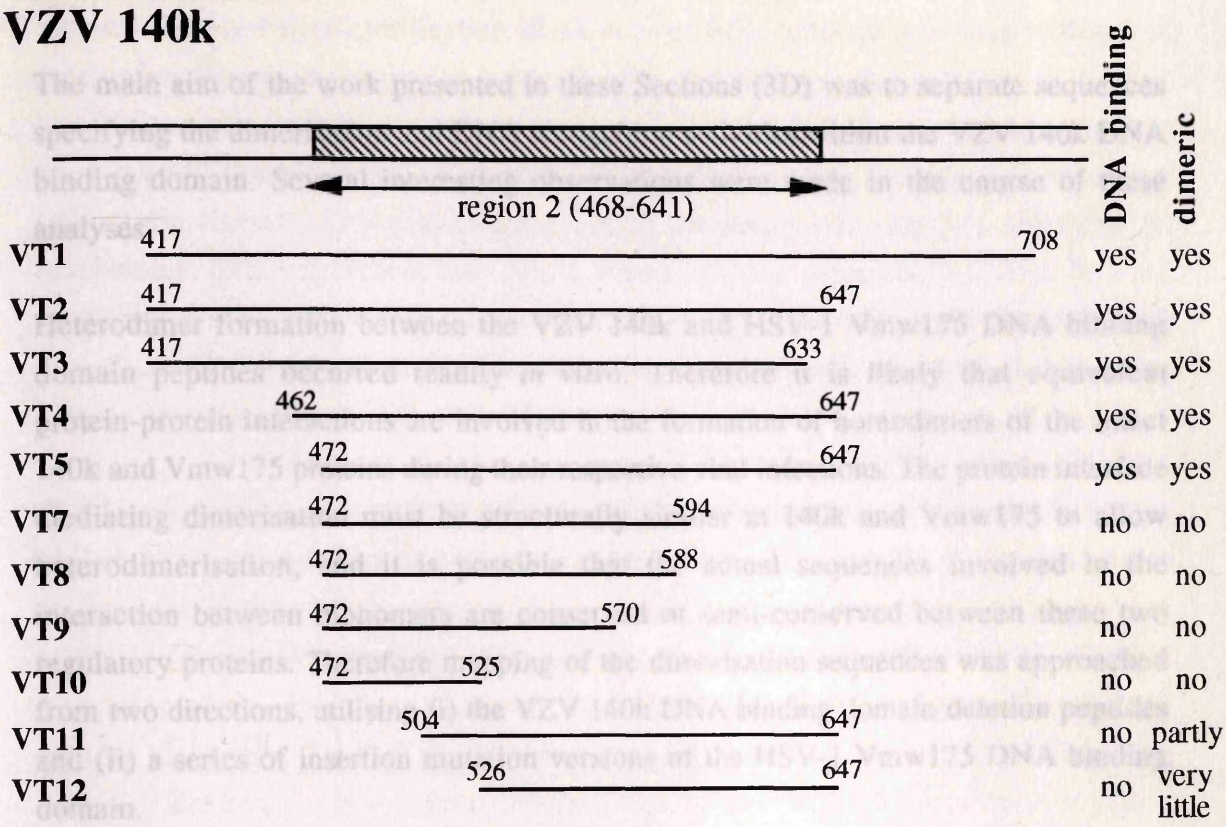


Figure 3D.14 Summary of the activities of the VZV 140k DNA binding domain deletion peptides. The portion of the 140k protein encompassing the DNA binding domain is shown diagrammatically at the top, the numbers of the first and last residues are given for the highly conserved region 2 (represented by a shaded box). Below, horizontal lines correspond to the protein sequences included in each of the deletion peptides VT1-VT12. The residues at the N- and C-termini of each peptide are indicated. The DNA binding abilities of the peptides are indicated as determined by gel retardation analysis of the bacterially expressed peptides (Section 3B.4) and the *in vitro* translated peptides (Section 3D.4). The ability of each peptide to dimerise is also indicated, as determined by a combination of the native state analyses on the bacterially expressed peptides (Section 3D.2), and heterodimerisation analyses of the *in vitro* translated peptides (Sections 3D.4 and 3D.9).

Co-immunoprecipitation of VT2X also occurred in the presence of 1M NaCl (not shown), indicating that the interaction is not mediated by DNA and is highly stable. The N-terminal deletion peptides VT11 and VT12 also heterodimerised, albeit to a lesser extent (right hand panels of Fig 3D.13, tracks 13 and 14). The VT11 result supports the findings of the previous analyses of the native state of the bacterially expressed VT11 peptide, which found VT11 to be partly dimeric in solution (Section 3D.2). Since co-immunoprecipitation of the C-terminal deletion peptide VT8 was undetected after prolonged exposure of the gel shown in the right hand panels of Figure 3D.13 (track 12), the quantitated value is considered to be negligible. The summary of this and all the other dimerisation analyses of the VZV 140k DNA binding domain deletion peptides is found in Figure 3D.14.

3D.10 Discussion

The main aim of the work presented in these Sections (3D) was to separate sequences specifying the dimerisation and DNA recognition activities within the VZV 140k DNA binding domain. Several interesting observations were made in the course of these analyses.

Heterodimer formation between the VZV 140k and HSV-1 Vmw175 DNA binding domain peptides occurred readily *in vitro*. Therefore it is likely that equivalent protein-protein interactions are involved in the formation of homodimers of the intact 140k and Vmw175 proteins during their respective viral infections. The protein interface mediating dimerisation must be structurally similar in 140k and Vmw175 to allow heterodimerisation, and it is possible that the actual sequences involved in the interaction between monomers are conserved or semi-conserved between these two regulatory proteins. Therefore mapping of the dimerisation sequences was approached from two directions, utilising (i) the VZV 140k DNA binding domain deletion peptides and (ii) a series of insertion mutation versions of the HSV-1 Vmw175 DNA binding domain.

Both approaches found that dimerisation had lower protein sequence requirements than DNA binding. In particular, the HSV-1 Vmw175 insertion mutation peptides i14-i18 failed to bind to DNA as either homodimers or heterodimers with a wild type subunit, yet all readily dimerised. Similarly, in a previous study insertions at Vmw175 residue 320 or 338 prevented DNA binding but did not disrupt dimerisation of the intact protein produced during viral infection (Shepard and DeLuca, 1991a). The DNA binding domain may be able to accommodate or compensate for the effects of insertions upon the individual facets of the dimerisation interface. Alternatively a short discrete sequence that has not been affected by the insertions may mediate dimerisation. It is improbable that all the insertions into Vmw175 region 2 are directly disrupting DNA binding, and it is more likely that the specific structure of the domain required for DNA recognition is highly sensitive to minor conformational alterations.

From the studies with the VZV 140k DNA binding domain deletion peptides, it was apparent that sequences towards the C-terminus of 140k region 2 (between residues 588 and 633) are essential either directly or indirectly for dimerisation. A long tract of residues within these region 2 C-terminal sequences is highly conserved (approximately 80% identity; see Section 1C.1.8) between VZV 140k and HSV-1 Vmw175 and secondary structure prediction programmes suggest a highly α–helical nature to this

tract. The well studied helix-loop-helix and leucine zipper dimerisation motifs are almost totally α–helical (see Section 1E.3), making the C-terminal portion of region 2 of these alphaherpesvirus proteins a particularly favourable candidate for mediating dimerisation. A single amino acid mutational analysis of these conserved sequences in the Vmw175 domain failed to identify any substitutions that disrupted dimerisation (Allen and Everett, 1994); although some of the mutations were not amenable to biochemical characterisation, they might re-pay analysis using the powerful *in vitro* translation and co-immunoprecipitation methods described here. It is also possible that, as suggested for the 4 amino acid insertions (see above), the dimerisation activity of the domain may tolerate minor changes of this type.

The N-terminal portion of VZV 140k region 2 (between residues 472 and 526) also contributes to normal dimerisation of the DNA binding domain. 140k sequences within amino acids 504 to 647 (peptide VT11) are sufficient for dimerisation, although at a lower efficiency than the full DNA binding domain (i.e. peptide VT2X). Further truncation of the domain to residues 526-647 (peptide VT12) reduced dimerisation activity yet further. The fact that dimers containing the VT11 and VT12 peptides were detected, albeit at a very low level, suggests that deletion of the N-terminus of region 2 has indirect effects on the subunit interaction. Also, the unusual behaviour of these particular peptides (Section 3D.2) supports the argument that deletion of N-terminal sequences causes general disruption of the conformation of the domain, rather than directly destroying the subunit interaction. In agreement, sequences between 140k residues 472 and 525 were not sufficient for dimerisation, as seen from the gel filtration analysis of the VT10 GST cleavage peptide (Section 3D.2.3).

Heterodimers comprising the wild type VZV 140k domain (VT4X) and a severely DNA binding compromised mutant HSV-1 Vmw175 peptide (I10i12X) recognised DNA with an apparent affinity equivalent to that of the wild type VT4X homodimers. Also a heterodimer comprising 140k/Vmw175 DNA binding domain peptides, interacted with VZV gene 62 promoter sites for which the Vmw175 peptide homodimer had an extremely low affinity. These data suggest that in certain circumstances, only one fully wild type subunit of the dimer is sufficient for the specific interaction with the recognition site. This raises the possibility that both subunits of the dimer are able to bind DNA independently, however there is no evidence for this occurring; gel retardation experiments using probes of different sizes in separate and mixed incubations gave no evidence of complexes containing both probe molecules (data not shown). The observation that in certain circumstances only one wild type subunit is sufficient for DNA binding could be explained by several models:

(A) *Only one subunit normally contacts the DNA binding site* (Fig. 3D.15, panel A). The second monomer does not directly contribute to the specificity or affinity of the DNA binding interaction but stabilises the overall protein-DNA complex. Evidence from substitution and deletion mutational analyses of the HSV-1 Vmw175 consensus binding sequence (Pizer *et al.*, 1991; Everett *et al.*, 1991a) and also footprinting data (Didonato and Muller, 1989) suggested that both half-sites of the consensus are specifically recognised by the wild type homodimers. The lack of a direct repeat or any obvious dyad symmetry in the bipartite Vmw175 consensus presents some difficulties for the simple model in which each monomer makes equivalent sequence-specific interactions with each half of the binding site. In B form DNA the two half-sites (underlined) of the ATCGTCnnnnYCGRC consensus are predicted to lie on the same face of the helix and could in theory be specifically contacted by a single monomer. In a similar scenario, the trimeric *Drosophila* heat shock factor (HSF) can form stable protein-DNA complexes in which only two of the three identical monomers contact the DNA (Perisic *et al.*, 1989).

(B) *Binding of the first monomer facilitates binding of the second monomer* (Fig. 3D.15, panel B). The high affinity binding of the wild type monomer allows cooperative binding of the second monomer, by reducing its energy requirement for DNA binding. In this case, the low affinity of the mutant monomer would not significantly reduce the DNA binding affinity of the heterodimer while the mutant homodimer would fail to bind DNA. The recognition site for the yeast GCN4 transactivator protein has non-equivalent half-sites; the GCN4 interaction with the left half-site is strong compared to the relatively weak right half-site (Oliphant *et al.*, 1989). The nature of this binding site facilitates cooperative binding of the second monomer of the GCN4 dimer to the weak half-site and a similar situation may hold for the Vmw175 and 140k binding sites.

(C) *Dimerisation with a wild type monomer induces a conformational change in the mutant monomer* (Fig. 3D.15, panel C). Binding to the wild type monomer induces the mutant monomer to adopt a DNA binding-competent conformation either before or after binding to DNA. In the presence of DNA, many DNA binding proteins adopt stabilised structures, for example the λ Cro protein undergoes a large quaternary change when the dimer binds to its operator sequence (Brennan *et al.*, 1990). Structural changes in the wild type Vmw175 monomer upon DNA binding may induce analogous conformational changes in the mutant Vmw175 subunit, restoring its ability to recognise target sequences specifically.

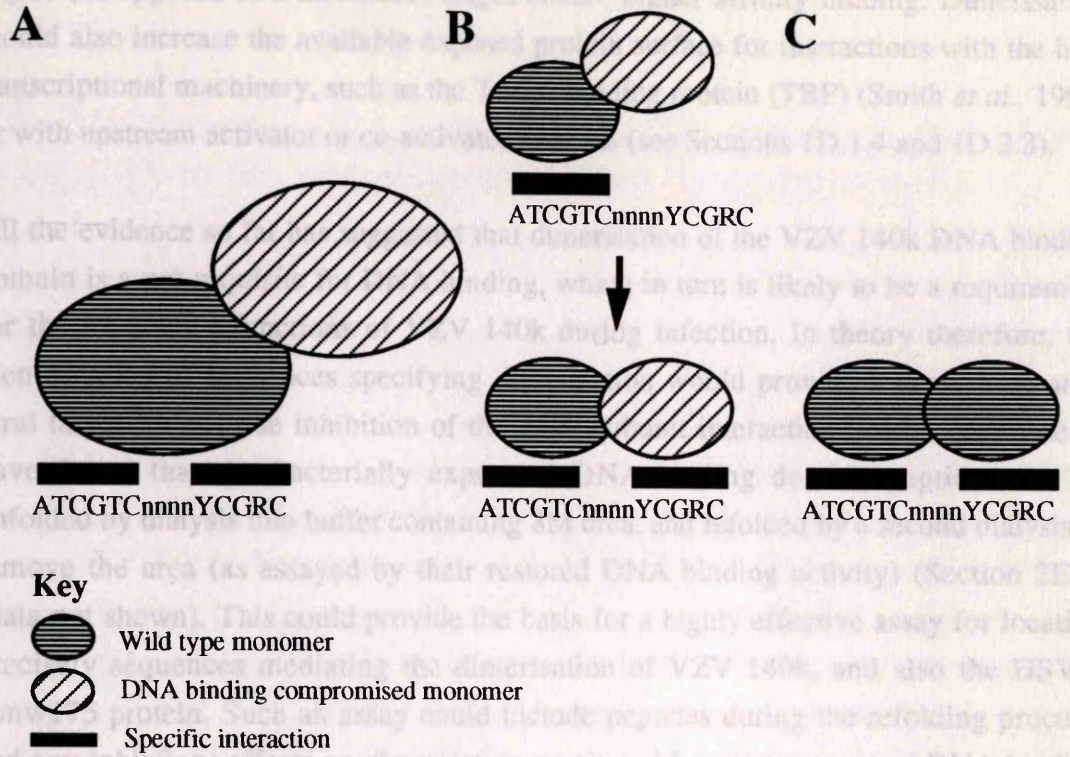


Figure 3D.15 Models for the DNA binding interaction of the dimeric DNA binding domain. Each oval represents a DNA binding domain monomer. The binding sequence shown in each section represents the HSV-1 Vmw175 consensus binding site of Faber and Wilcox (1986b) (see Section 1C.1.3). Panels A, B and C illustrate possible modes of binding site recognition, as discussed in the text. The dimer represented comprises a wild type monomer and a DNA binding compromised monomer (i.e. a monomer that has a low binding affinity for the site shown, or containing a mutation (i.e. i12) that is deleterious for DNA binding). **A.** Only one subunit of the dimer specifically recognises both half-sites of the Vmw175 consensus binding site. **B.** Binding of one monomer facilitates binding of the second monomer by a two step process, one monomer interacting with each half-site. **C.** Heterodimerisation induces a conformational change in the DNA binding compromised monomer, such that it adopts a wild type conformation that enables recognition of the binding site.

These three models for explaining the DNA binding interaction are not mutually exclusive. However, heterodimers including a version of the Vmw175 DNA binding domain containing insertion mutation i14, i15, i16, i17 or i18, failed to bind to DNA. If these particular insertion mutations induce profound conformational changes in the domain, then it can be envisaged how the altered domain structure could prevent or destabilise heterodimer-DNA interactions occurring by any of the three models presented above.

Besides facilitating cooperative DNA binding of the second subunit of a dimer molecule, dimerisation could confer several other possible advantages upon the VZV 140k and HSV-1 Vmw175 proteins. The dimeric nature of these proteins may allow a higher level of sequence selectivity, and the large surface area buried upon binding of a

dimer (as opposed to a monomer) might ensure higher affinity binding. Dimerisation would also increase the available exposed protein surface for interactions with the host transcriptional machinery, such as the TATA binding protein (TBP) (Smith *et al.*, 1993) or with upstream activator or co-activator proteins (see Sections 1D.1.4 and 1D.2.3).

All the evidence so far has suggested that dimerisation of the VZV 140k DNA binding domain is a pre-requisite for DNA binding, which in turn is likely to be a requirement for the regulatory functions of VZV 140k during infection. In theory therefore, the identification of sequences specifying dimerisation would provide a favourable anti-viral target for peptide inhibition of the 140k subunit interaction. Initial experiments have found that the bacterially expressed DNA binding domain peptides can be unfolded by dialysis into buffer containing 8M urea, and refolded by a second dialysis to remove the urea (as assayed by their restored DNA binding activity) (Section 2E.5) (data not shown). This could provide the basis for a highly effective assay for locating precisely sequences mediating the dimerisation of VZV 140k, and also the HSV-1 Vmw175 protein. Such an assay could include peptides during the refolding process, and any inhibitory effects on dimerisation monitored by measurement of DNA binding efficiency.

3E Mutagenesis of a short VZV 140k sequence that shows homology to the homeodomain DNA recognition helix

3E.1 Strategy

Structural studies and sequence comparisons have found that many DNA binding proteins can be grouped into classes that use related structural motifs for DNA recognition (see Section 1E). Like several other important transactivator proteins, for instance SV40 large T antigen (Arthur *et al.*, 1988) and NF-κB (Kieran *et al.*, 1990), the VZV 140k DNA binding domain does not contain sequences that fit neatly into any of the recognised classes of DNA binding motif. However, a short stretch of residues within the HSV-1 Vmw175 DNA binding domain, showing significant homology to the DNA recognition helix (helix 3) of the homeodomain, has been identified by sequence analysis (Allen, 1993). As described in Section 1E.2.2, the homeodomain is a eukaryotic DNA binding motif that is structurally related to the prokaryotic helix-turn-helix (HTH) motif found in proteins such as the λ repressor. The homeodomain typically comprises 60 residues arranged in three α–helical segments. Helices 2 and 3 comprise the HTH, and the third helix is the principal determinant for DNA recognition.

The region of HSV-1 Vmw175 that shows homology to the homeodomain DNA recognition helix, is highly conserved within VZV 140k and the other alphaherpesvirus transactivator homologues. The possible functional implications of the conservation of this region was investigated by mutagenesis of the 140k sequence. Three single amino acid substitutions were introduced at positions predicted to be important for forming DNA contacts, by analogy to the known homeodomain-DNA structures. The DNA binding activities of the mutant 140k DNA binding domain peptides were investigated in order to determine if the homology region is directly involved in DNA recognition. In addition, the effects of the mutations on the regulatory functions of the intact 140k protein were analysed by transient transfection assays.

3E.2 The WLQN region of VZV 140k that shows homology to the homeodomain DNA recognition helix

To date, dozens of homeodomains have been described, falling into at least 25 distinct families (Scott *et al.*, 1989; Laughon, 1991). Residues at several positions are significantly conserved throughout all the homeodomain families, but the remaining residues show considerable divergence. The homeodomain homology region was first identified by sequence alignment of a weight matrix compiled for a variety of homeodomains to the HSV-1 Vmw175 DNA binding domain (Allen, 1993). Subsequently, similar homeodomain consensus sequence alignments with VZV 140k and the other alphaherpesvirus homologue proteins identified the same region (data not shown). The variation within the homeodomain sequences creates significant divergence in any consensus derived for homeodomain sequences. However, individual homeodomain sequences also aligned convincingly to the same alphaherpesvirus sequences. Figure 3E.1 shows the DNA recognition helix (helix 3) of the prototype *Drosophila Antennapedia* (*Antp*) homeodomain aligned with the homology region identified within the VZV 140k alphaherpesvirus homologues. The level of sequence homology between the first two homeodomain helices and the alphaherpesvirus proteins was not considered significant and is therefore not shown in the alignment in Figure 3E.1.

		------recognition helix--------
		H BHHBB PBP P
		* ** * *
		45 50 55
Antp		ERQIKIWFQN RRMKWKK
VZV 140k	**537**	DRDPIAWL**QN** P**K**LTGVN
HSV-1 Vmw175	**384**	DAEAMGWLQN PRVVPGD
PRV IE180	**554**	AGEAMSWLQN PRMQAPD
EHV-1 IE1	**487**	SREAMSWLQN SKLSGQD

Figure 3E.1 Alignment of the DNA recognition helix of the *Drosophila Antp* homeodomain to the alphaherpesvirus homologues of VZV 140k. The top sequence shows the *Antp* DNA recognition helix (homeodomain helix 3). *Antp* amino acids are numbered according to convention to indicate their position in the 60-residue homeodomain (Laughon, 1991). Positions corresponding to the hydrophobic core (H), base contact (B) and phosphate contact (P) amino acids of the *Antp* (Otting *et al.*, 1990), MATα2 (Wolberger *et al.*, 1991) and *en* (Kissinger *et al.*, 1990) homeodomain-DNA structures are indicated above the *Antp* sequence. Positions that are invariant or nearly invariant within all known homeodomains are indicated by asterisks. Aligned beneath are the sequences identified within region 2 of the alphaherpesvirus proteins (VZV 140k, HSV-1 Vmw175, pseudorabies virus IE180 and equine herpesvirus 1 IE1) that show homology to the homeodomain DNA recognition helix. The position of the first amino acid in each sequence is given. The VZV 140k residues that were subject to mutation are shown in bold. The alignment between the *Antp* homeodomain and the full length alphaherpesvirus proteins was produced using the GCG PileUp and Pretty programmes (Section 2G).

It should be noted that helices 1 and 2 are also highly variable between the homeodomains; in essence, the homeodomain is a structural fold carrying the DNA recognition helix that lies across the major groove of DNA. The sequences shown in Figure 3E.1 are approximately in the centre of region 2 of each alphaherpesvirus protein (as indicated for VZV 140k in the lower panel of Fig. 3E.3). The asterisks in Figure 3E.1 indicate positions that are invariant or almost invariant within the DNA recognition helix of all homeodomains. These residues are also conserved or semi-conserved within the alphaherpesvirus proteins. Especially highly conserved is the quartet of residues WLQN, hence this short region of homology to the homeodomain is referred to as the WLQN region. The structures determined for three homeodomain-DNA complexes (Kissinger *et al.*, 1990; Otting *et al.*, 1990; Wolberger *et al.*, 1991), demonstrate that the invariant homeodomain residues either mediate specific DNA contacts or maintain the hydrophobic core that stabilises the homeodomain protein fold (as indicated in Fig. 3E.1).

To investigate further the relevance of the sequence homology between the homeodomain and the WLQN region of the alphaherpesvirus proteins, a weight matrix was created for the 60 residues of the VZV 140k, HSV-1 Vmw175, PRV IE180 and EHV-1 IE1 proteins that corresponded to their hypothetical homeodomains (encompassing the WLQN region). When this weight matrix was used to search the Swissprot protein sequence database, six of the ten highest scoring alignments were to homeodomains of recognised homeobox proteins (discounting the first four alignments to the viral IE proteins). The scores for these ten alignments were between 4.51-5.83; they represent the differences in standard deviation units between each normalised comparison score and the mean normalised comparison score for sequences unrelated to the weight matrix profile.

A site directed mutagenic approach was adopted to investigate whether this potentially interesting WLQN region of sequence conservation is involved in the DNA recognition function of the VZV 140k protein.

3E.3 Mutagenesis, expression and purification of the VZV 140k DNA binding domain

In vitro mutagenesis was used to alter charges on single specific VZV 140k residues, while maintaining the size of amino acid side chains to minimise alterations of the protein conformation. Two single conserved residues within the WLQN motif itself were mutated: 140k residue Gln 545 was replaced with Leu (Q545L), and residue Asn

546 was replaced with Leu (N546L). In the known homeodomain-DNA structures (Kissinger *et al.*, 1990; Otting *et al.*, 1990; Wolberger *et al.*, 1991), the corresponding residues of the homeodomain, residues 50 and invariant Asn 51 (Fig. 3E.1), both point into the major groove of DNA and hydrogen bond with DNA bases. In addition, genetic and biochemical analyses have shown that the residue at homeodomain position 50 determines the differential binding specificity of the homeodomain-DNA interaction (Hanes and Brent, 1989; Treisman *et al.*, 1989; Hanes and Brent, 1991).

Thirdly, VZV 140k residue Lys 548 was replaced with Glu (K548E), due to the conservation of a basic residue at this position in the alphaherpesvirus proteins and all homeodomains. The corresponding homeodomain residue Arg 53 is known to form critical hydrogen bonds with two phosphate groups of the DNA backbone, thus stabilising the protein-DNA interaction.

Synthetic oligonucleotides including the relevant base changes (Fig. 3E.2) were used to direct the mutagenesis, using the single stranded bacteriophage M13mp18EK template (see Fig 3E.3), as described in Section 2B.10. The resultant bacteriophage M13Q545L, M13N546L and M13K548E were utilised to derive plasmids pVT2XQ545L, pVT2XN546L and pVT2XK548E (collectively referred to as pVT2X[substitution]) (Fig. 3E.3).

```
                 542                                   552
                  A  W  L  Q  N  P  K  L  T  G  V
Wild type     5- GCATGGCTCCAGAACCCCAAGCTGACCGGTGTC -3
 140k         3- CGTACCGAGGTCTTGGGGTTCGACTGGCCACAG -5

                    W  L  L  N  P  K  L  T  G
Oligo 1       3-  TACCGAGGACTTGGGGTTCGAATGGCCACA  -5
                                  HindIII

                    W  L  Q  L  P  K  L
Oligo 2       3-  ACCGAGGTCGACGGGTTCGAC          -5
                      PvuII

                          N  P  E  L  T  G
Oligo 3       3-       TCTTGGGGCTCGAGTGGCCACA   -5
                                SacI
```

Figure 3E.2 Oligonucleotides used for mutagenesis of the WLQN region of VZV 140k. At the top is shown the wild type VZV 140k protein sequence (residues 542-552) and the corresponding sequences of both VZV gene 62 DNA strands which encompass the region of mutagenesis. Below are given the sequences of the three oligonucleotides (corresponding to the gene 62 non-coding strand) used for the oligonucleotide-directed mutagenesis, and the relevant coding potentials. The nucleotides that differ from the wild type sequence and the resultant amino acid alterations are shown in bold. Restriction enzyme sites introduced by taking advantage of the degenerate nature of the genetic code are also indicated and were used in screening for the mutated gene 62 sequences. Oligo 1 facilitated the Q545L substitution; oligo 2, the N546L substitution; oligo 3, the K548E substitution.

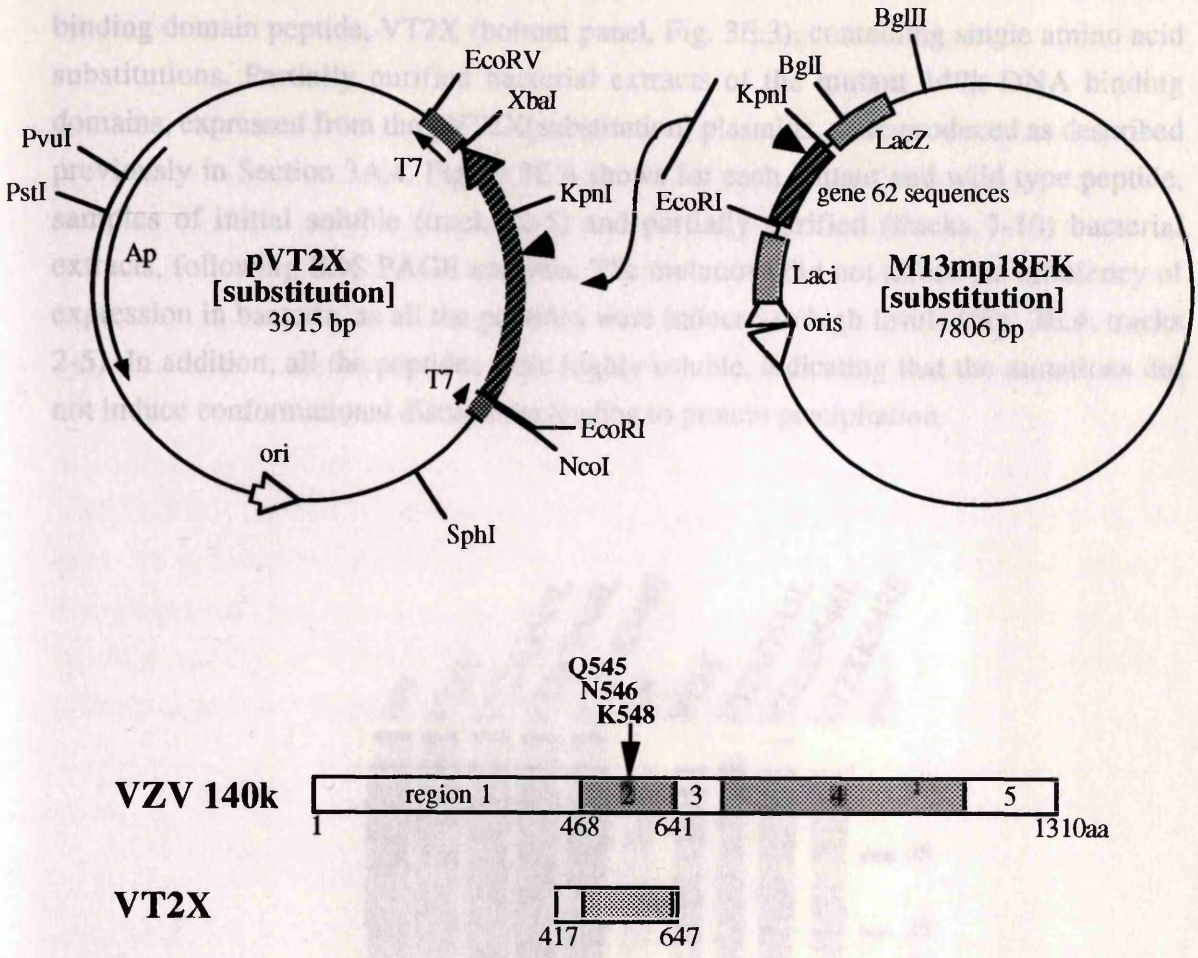


Figure 3E.3 Vectors used for mutagenesis of the VZV 140k DNA binding domain and its subsequent expression. Only selected restriction enzyme sites are shown. Mutagenesis template bacteriophage M13mp18EK was constructed by insertion of an *Eco*RI-*Kpn*I fragment spanning codons 417 to 570 of VZV gene 62 (isolated from plasmid p585T7aVT2X), between the *Eco*RI and *Kpn*I sites of the MCR of vector M13mp18. Oligonucleotide-directed mutagenesis of the gene 62 sequences in vector M13mp18EK was performed as described in Section 2B.10. Mutagenesis yielded bacteriophage M13Q545L, M13N546L and M13K548E (collectively referred to as M13mp18EK[substitution]). The restriction sites introduced by the mutagenesis facilitated identification of the appropriate mutagenised progeny bacteriophage, and the presence of the correct mutations was confirmed by direct plasmid sequencing. The arrow head on both vectors indicates the approximate site of the introduced base changes. *Eco*RI-*Kpn*I fragments were re-isolated from replicative forms of the M13mp18EK[substitution] bacteriophage. Each mutated *Eco*RI-*Kpn*I fragment was swapped for the wild type *Eco*RI-*Kpn*I fragment, by insertion between the *Eco*RI and *Kpn*I sites of plasmid p585T7aVT2X. This cloning procedure produced plasmids pVT2XQ545L, pVT2XN546L and pVT2XK548E which encode the mutated 140k DNA binding domain under the control of T7 expression signals (collectively referred to as pVT2X[substitution]). In the bottom panel are shown diagrammatic representations of VZV 140k and its DNA binding domain peptide, VT2X (Section 3A.3). The region 2 and 4 sequences that are highly conserved between VZV 140k and HSV-1 Vmw175 are shown shaded, amino acid coordinates of the boundaries of the conserved regions are given (inclusively). The VT2X peptide contains VZV 140k residues 417 to 647, encompassing the highly conserved region 2. The arrow indicates the approximate positions of the three residues, whose substitution is discussed in the text.

The pVT2X[substitution] T7 expression plasmids encoded versions of the 140k DNA binding domain peptide, VT2X (bottom panel, Fig. 3E.3), containing single amino acid substitutions. Partially purified bacterial extracts of the mutant 140k DNA binding domains, expressed from the pVT2X[substitution] plasmids, were produced as described previously in Section 3A.4. Figure 3E.4 shows for each mutant and wild type peptide, samples of initial soluble (tracks 2-5) and partially purified (tracks 7-10) bacterial extracts, following SDS PAGE analysis. The mutations did not affect the efficiency of expression in bacteria, as all the peptides were induced to high levels (Fig. 3E.4, tracks 2-5). In addition, all the peptides were highly soluble, indicating that the mutations did not induce conformational disruptions leading to protein precipitation.

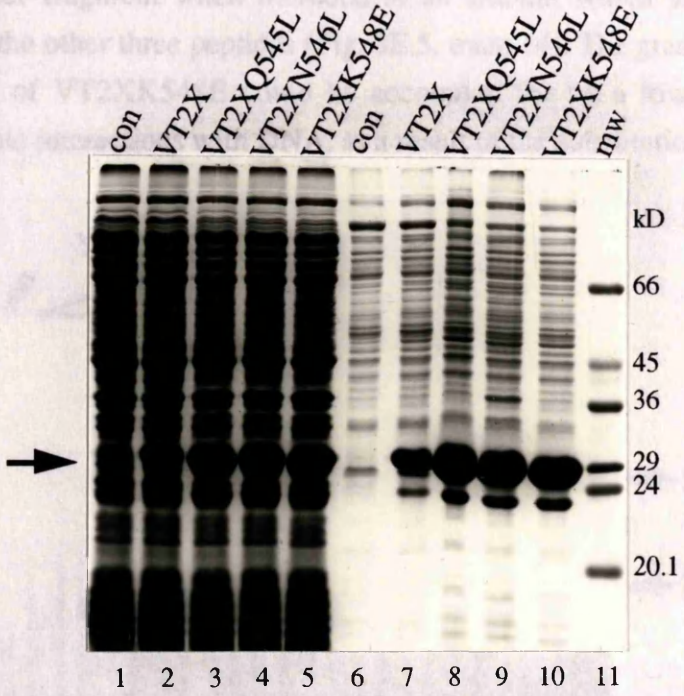


Figure 3E.4 Bacterial expression and partial purification of the mutated VZV 140k DNA binding domain. Shown here is a Coomassie-stained 12.5% SDS PAGE analysis of the partial purification of bacterially expressed wild type and mutant VT2X peptides. Expression of the 140k DNA binding domain peptides was induced in *E. coli* from the appropriate T7 expression plasmids (as described in Sections 2C.1 and 2C.2). Tracks 1-5 contained 2µl of the total soluble proteins from 100ml of each induced bacterial culture after resuspension in 1ml buffer, bacterial lysis and centrifugation (Section 2C.2). Tracks 6-10 contained 2µl of partially purified extracts after 35% ammonium sulphate precipitates were resuspended in 300µl buffer. 'con' indicates an extract prepared from induced bacteria carrying the p585T7a plasmid. The arrow indicates the position of the 26.7kD VT2X peptide; a smaller proteolytic degradation product is apparent beneath. Track 11 contains molecular weight standards: BSA, 66kD; albumin, 45kD; glyceraldehyde-3-phosphate dehydrogenase, 36kD; carbonic anhydrase, 29kD; trypsinogen, 24kD; trypsin inhibitor, 20.1kD.

3E.4 Mutation of a single lysine residue within the WLQN region severely impairs the DNA binding activity of the VZV 140k DNA binding domain

Initially, the mutant 140k DNA binding domain peptides were analyzed for their abilities to recognise their own VZV gene 62 promoter in gel retardation analyses. Incubation of an equivalent range of amounts of each peptide with the gene 62 promoter BT fragment yielded primary protein-DNA complexes for wild type VT2X, VT2XQ545L and VT2XN546L peptides (Fig. 3E.5, tracks 3-5, 6-8, 9-11 respectively) and secondary complexes were produced with higher amounts of these peptides. The VT2X, VT2XQ545L and VT2XN546L peptides had comparable DNA binding affinities, as determined by competition experiments (data not shown). The VT2XK548E peptide produced a very faint, smeary protein-DNA complex with the gene 62 promoter fragment when included at an amount which yielded secondary complexes with the other three peptides (Fig. 3E.5, track 14). The greatly reduced DNA binding activity of VT2XK548E could be accounted for by a lower DNA binding affinity or unstable interactions with DNA, as a result of the substitution.

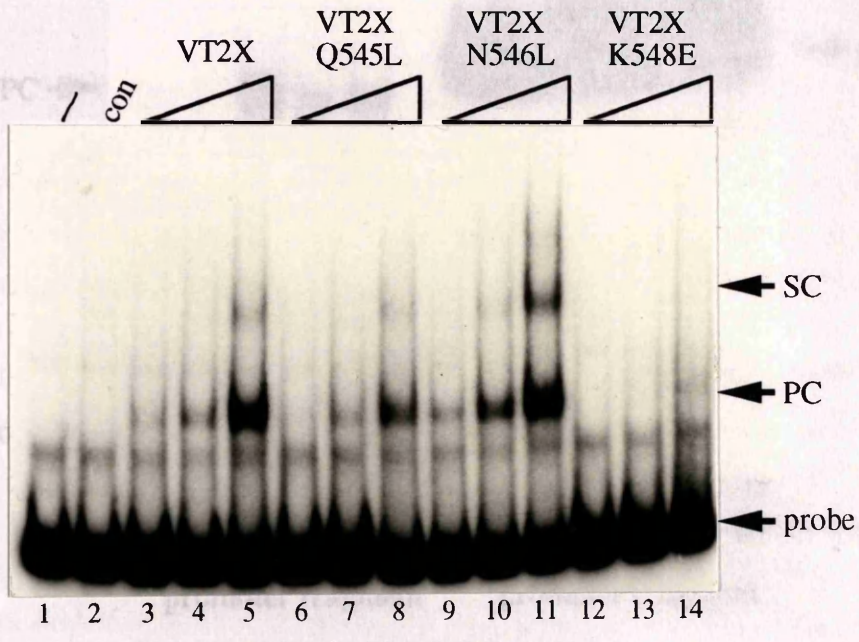


Figure 3E.5 Gel retardation analysis of the VZV gene 62 promoter with the mutant VZV 140k DNA binding domain peptides. The radiolabelled BT DNA fragment was derived from the VZV gene 62 promoter (details given in Fig. 3A.6). The probe was incubated with approximately 25ng, 50ng or 100ng of the partially purified bacterial extracts of wild type or mutant VT2X peptide indicated above tracks 3-14. Protein amounts were determined by densitometry of Coomassie-stained SDS PAGE gels. The incubation reaction for track 1 contained no extract and for track 2 contained a control extract prepared from bacteria carrying the p585T7a vector. Complexes were resolved by gel retardation analysis as described in Section 2D.2, followed by autoradiography. The positions of the free DNA probe, primary (PC) and secondary (SC) protein-DNA complexes are indicated on the right hand side. A contaminating probe band is seen in all tracks, migrating above the free BT probe.

Additionally, gel retardation assays were performed with DNA fragments derived from the HSV-1 IE3 and glycoprotein D (gD) promoters, to analyse whether the effects of mutations in the WLQN region on DNA binding were specific to sites within the VZV gene 62 promoter. Both of the HSV-1 fragment probes that were employed included a Vmw175 consensus binding site recognised by the wild type 140k DNA binding domain (Section 3A.8). As seen with the gene 62 promoter fragment, mutations Q545L and N546L had little effect on the apparent DNA binding affinity of VT2X for either the IE3S or the gD probe (Fig. 3E.6).

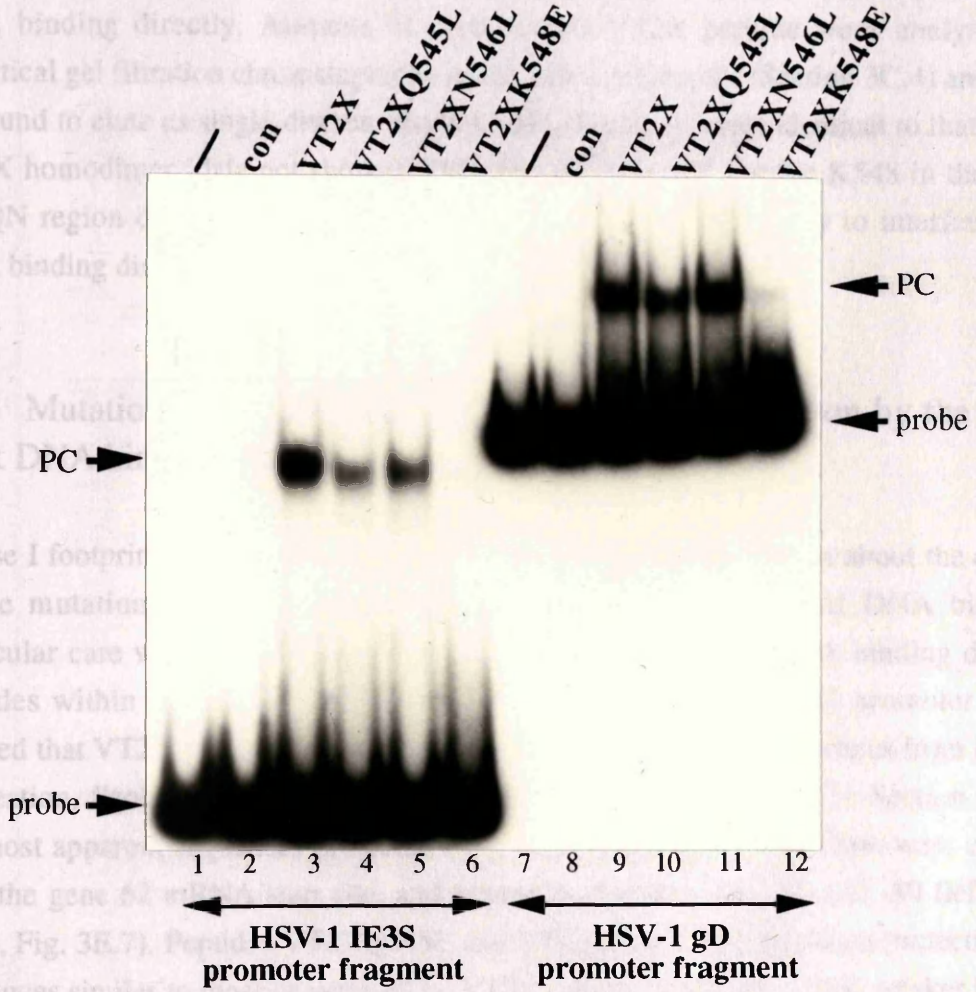


Figure 3E.6 Gel retardation analysis of HSV-1 promoter fragments with the mutant VZV 140k DNA binding domain peptides. Radiolabelled DNA fragments were derived from the HSV-1 IE3 gene promoter (probe IE3S) (tracks 1-6) and the HSV-1 gD early gene promoter (tracks 7-12) (details given in Fig. 3A.11). The IE3S and gD probes were incubated with approximately 40ng and 30ng (respectively) of the partially purified bacterial extracts of the VT2X derived peptides, as indicated above the tracks. The incubation reactions for tracks 1 and 7 contained no extract, and for tracks 2 and 8 contained a control extract prepared from bacteria carrying the p585T7a vector. The components of the reactions were resolved by electrophoresis, as in the previous analysis (Fig. 3E.5). This experiment used an equivalent amount of each peptide that gave only the primary protein-DNA complex, at the positions indicated at each side.

No detectable binding of peptide VT2XK548E to the HSV-1 IE3 promoter fragment was observed (Fig. 3E.6, track 6) and a barely detectable complex was seen on the gD promoter fragment (Fig. 3E.6, track 12). The greatly reduced DNA binding activity of VT2XK548E on all three probes analysed, indicated that the effects of this substitution reflect a generalised disruption of the VZV 140k DNA binding interaction.

The VZV 140k DNA binding domain is a stable dimer and dimerisation appears to be a prerequisite for DNA binding (Section 3C). Therefore, it was necessary to eliminate the possibility that mutation K548E might have disrupted dimerisation rather than affected DNA binding directly. Aliquots of each mutant VT2X peptide were analyzed by analytical gel filtration chromatography as described previously (Section 3C.4) and were all found to elute as single discrete species, with elution volumes identical to that of the VT2X homodimer (data not shown). Therefore, mutation of residue K548 in the 140k WLQN region did not affect the subunit interaction, and was likely to interfere with DNA binding directly.

3E.5 Mutations in the WLQN region alter DNA recognition by the VZV 140k DNA binding domain

DNase I footprinting analysis was used to provide further information about the effects of the mutations in the 140k WLQN region on the specificity of DNA binding. Particular care was taken to ensure equal input of the different DNA binding domain peptides within each experiment. Footprinting on the VZV gene 62 promoter probe showed that VT2X provided characteristic protection of multiple sequences from DNase I digestion, flanked by marked sites of DNase I hypersensitivity (as in Section 3A.6); the most apparent sequences protected by VT2X from DNase I digestion were centred over the gene 62 mRNA start site, and around nucleotides -30, -50 and -80 (left hand panel, Fig. 3E.7). Peptides VT2XQ545L and VT2XN546L both provided protection for sequences similar to those recognised by VT2X, albeit more faintly. This weaker DNase I protection was not merely a consequence of reduced DNA binding affinity (see above). These two mutant peptides also induced sites of DNase I hypersensitivity in similar positions to those produced with VT2X, although closer examination of the VT2X, VT2XQ545L and VT2XN546L tracks reveals slight differences in the details of these hypersensitivity sites. The VT2XK548E peptide failed to produce any footprinting pattern. However, on comparison with the control track, it is apparent that a degree of DNase I hypersensitivity was induced by the VT2XK548E peptide at similar positions to those produced by the other three peptides.

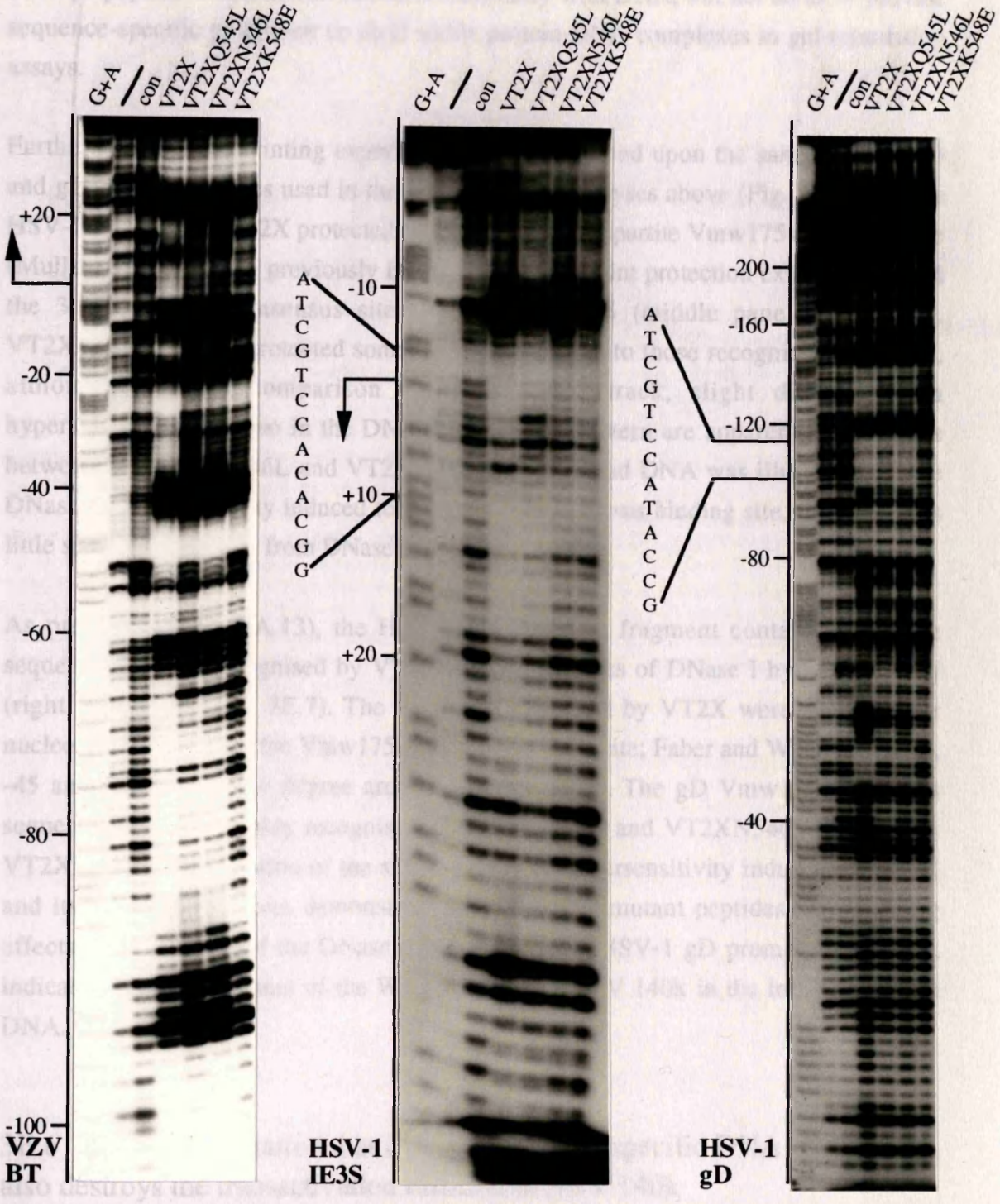


Figure 3E.7 DNase I footprinting analyses of the mutant VZV 140k DNA binding domain peptides. DNase I footprinting assays were performed with end-labelled promoter DNA fragments and partially purified bacterial extracts (as described in Section 2D.4). The radiolabelled promoter DNA fragment used in each experiment is indicated at the bottom left of each panel (details given in Fig. 3A.6 for the VZV gene 62 BT probe, in Fig. 3A.11 for the HSV-1 probes). Radiolabelling of the BT fragment was on the bottom DNA strand, the IE3S and gD fragments were labelled on the top strand. The wild type and mutant VT2X tracks contained approximately 1.6µg of the DNA binding domain peptide indicated. Tracks labelled 'G+A', '___' and 'con' contain Maxam and Gilbert sequencing reactions, no extract and induced an extract from bacteria carrying the p585T7a vector respectively. The numbered scale on the left of each panel identifies the positions of nucleotides included within each DNA fragment, numbered with respect to the mRNA start site (+1). The bent arrows indicate the position and orientation of the proposed mRNA start sites. The positions and sequences of the HSV-1 Vmw175 consensus binding sites in the IE3 and gD promoters are indicated.

Perhaps peptide VT2XK548E interacts transiently with DNA, but not so as to provide sequence-specific protection or yield stable protein-DNA complexes in gel retardation assays.

Further DNase I footprinting experiments were performed upon the same HSV-1 IE3 and gD promoter probes used in the gel retardation analyses above (Fig. 3E.6). On the HSV-1 IE3S probe, VT2X protected both halves of the bipartite Vmw175 consensus site (Muller, 1987) (as seen previously in Fig. 3A.12) and faint protection extended beyond the 3' end of the consensus site to nucleotide +25 (middle panel, Fig. 3E.7). VT2XQ545L weakly protected some similar sequences to those recognised by VT2X, although on close comparison with the VT2X track, slight differences in hypersensitivity and also in the DNase I protection pattern are apparent. Interactions between the VT2XN546L and VT2XK548E peptides and DNA was illustrated by the DNase I hypersensitivity induced to the 5' of the consensus binding site, but there was little specific protection from DNase I digestion.

As previously (Fig. 3A.13), the HSV-1 gD promoter fragment contained multiple sequences weakly recognised by VT2X, flanked by sites of DNase I hypersensitivity (right hand panel, Fig. 3E.7). The sequences protected by VT2X were centred over nucleotides -130, -100 (the Vmw175 consensus binding site; Faber and Wilcox, 1986b), -45 and also to a lesser degree around nucleotide -85. The gD Vmw175 consensus sequence was also weakly recognised by VT2XQ545L and VT2XN546L, but not by VT2XK548E. Examination of the sites of DNase I hypersensitivity induced by VT2X and its mutant derivatives demonstrates that the three mutant peptides had differing effects on the details of the DNase I footprint on the HSV-1 gD promoter fragment, indicating the involvement of the WLQN region of VZV 140k in the interaction with DNA.

3E.6 The lysine mutation that disrupts sequence-specific DNA recognition, also destroys the transactivation function of VZV 140k

Having established that amino acid substitutions within the WLQN region affected, to varying degrees, the interaction of the isolated VZV 140k DNA binding domain with DNA, it was of interest to determine the effects of these mutations, if any, on the transactivation function of the 140k protein. Each mutation was incorporated into the intact 140k protein as described in Figure 3E.8.

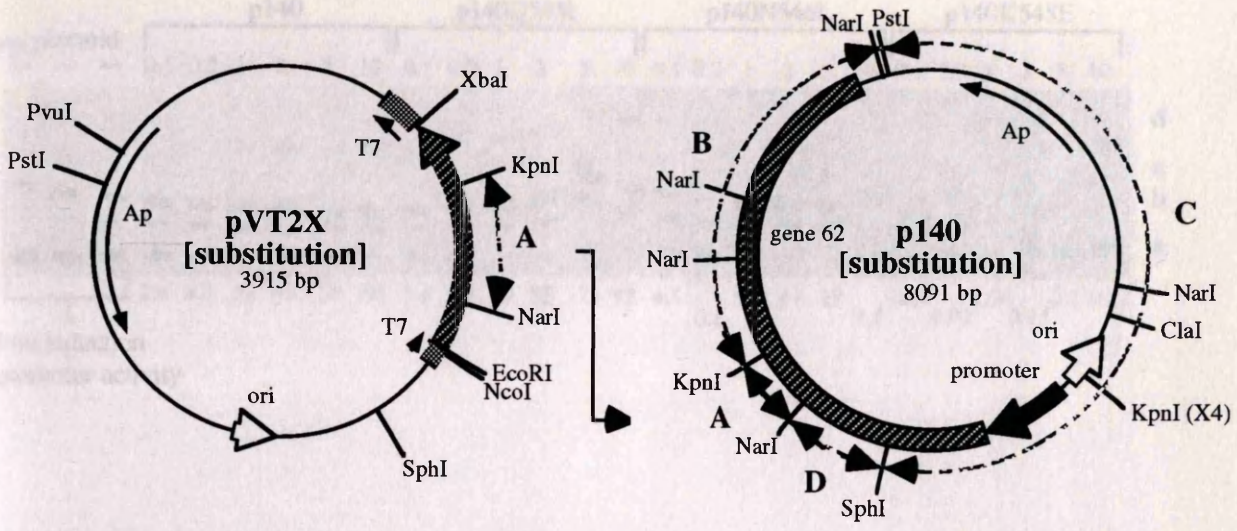
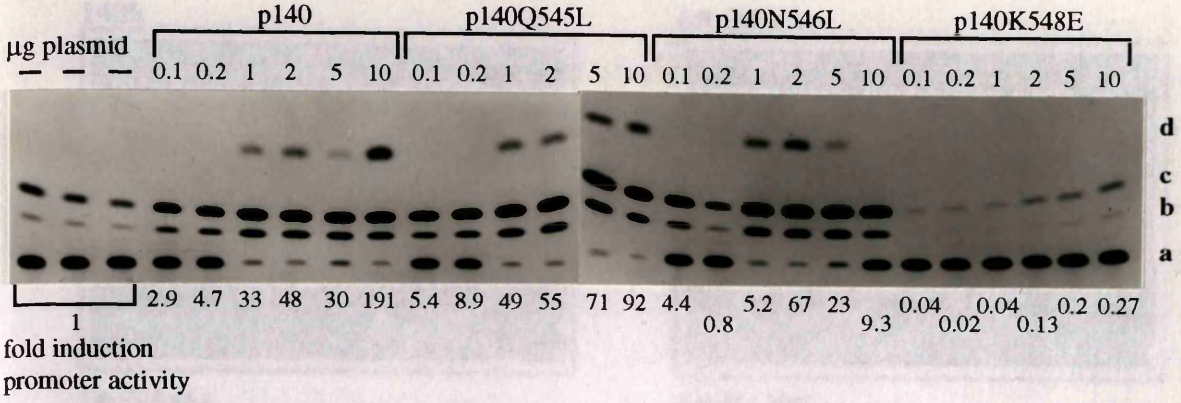


Figure 3E.8 Introduction of the WLQN region mutations into the intact VZV 140k protein. Nucleotide changes introduced by oligonucleotide-directed mutagenesis (Section 3E.3) were incorporated into the intact VZV gene 62 open reading frame in plasmid p140 as follows: *Nar*I-*Kpn*I fragments (A) isolated from plasmids pVT2XQ545L, pVT2XN546L and pVT2XK548E were each co-ligated with three fragments (*Kpn*I-*Pst*I fragment (B), 2414bp; *Pst*I-*Sph*I fragment (C), 4572bp; *Sph*I-*Nar*I fragment (D), 632bp) isolated from plasmid p140. The resultant plasmids p140Q545L, p140N546L and p140K548E (collectively referred to as p140[substitution]) were identified by extensive restriction enzyme digestion analyses using the restriction sites introduced by the mutagenesis procedure. The p140[substitution] plasmids encode the entire VZV 140k coding region including the specific base substitutions, under the control of its own VZV gene 62 promoter.

The abilities of each mutant 140k protein to transactivate the heterologous HSV-1 gD promoter was determined by transient transfection assay (CAT assay), to allow direct comparisons with the DNA binding results obtained on the gD promoter (above). The transactivation efficiency of each mutant 140k protein was compared to that obtained for wild type 140k, which is known to be a potent transactivator of the HSV-1 gD promoter (Disney, 1990). Figure 3E.9 shows the results of two such experiments where a range of concentrations of each effector plasmid and a fixed amount of the gD promoter-CAT reporter construct were transiently co-transfected into HeLa cells. Due to the variable nature of this type of assay, titration of the effector plasmids allowed a more accurate interpretation of the effects of each protein on promoter activity, as compared to using a single plasmid amount. As anticipated, wild type VZV 140k very strongly transactivated the HSV-1 gD promoter in a dose dependent manner (Fig. 3E.9). 140Q545L and 140N546L had reduced transactivation strengths, and this was more pronounced with protein 140N546L. The 140K548E protein consistently failed to activate gene expression from the gD promoter. The lack of function of the 140K548E protein was not specific to the HSV-1 gD promoter but represented a general phenomenon: 140K548E failed to transactivate the HSV-1 IE1 and *tk* promoters, and also the SV40 early promoter-enhancer regions (data not shown).



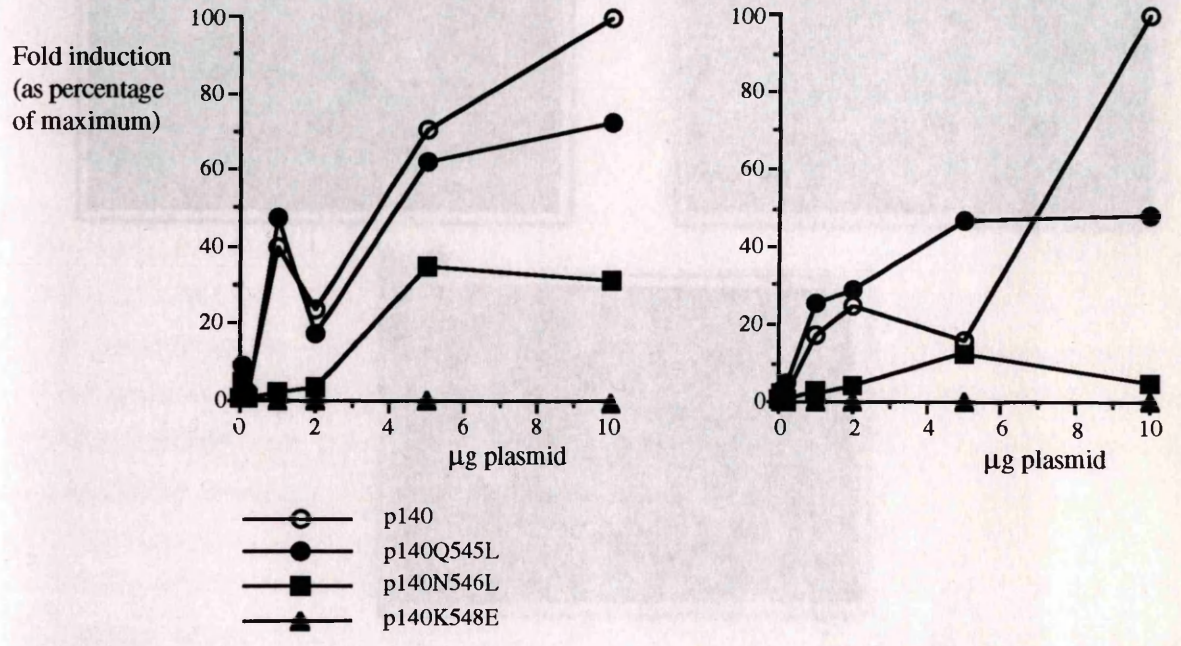


Figure 3E.9 Transactivation efficiencies of the mutant VZV 140k proteins. At the top are shown autoradiographs of CAT assay TLC plates from an experiment measuring transactivation of the HSV-1 gD promoter-leader sequences, by the VZV 140k protein and its mutant derivatives. Increasing amounts of each effector plasmid (as indicated) were co-transfected with 4µg of the reporter construct pgDCAT into HeLa cells by the calcium phosphate precipitation method (Section 2F.2). The CAT activity of the resultant HeLa cell extracts were determined by CAT assay as described in Section 2F.3. The identity of the ^{14}C-labelled spots are indicated on the left as follows: a - chloramphenicol, b and c - tautomers of monoacetyl chloramphenicol, d - diacetyl chloramphenicol; a is the substrate and b, c and d are products of the acetyl transferase reaction catalysed by the chloramphenicol acetyl transferase enzyme (CAT). Fold induction promoter activity induced by each effector plasmid over the level produced with gDCAT alone is indicated below each track. Fold induction was calculated by dividing the proportion conversion from substrate to product per microgram protein obtained at a particular plasmid amount (Section 2F.4), by the control value obtained when no effector plasmid was present (the control value typically represented an average of three individual transfections). Below are plotted the results of the CAT assay shown above (right hand graph), and also the results of another independent repeat of the experiment (left hand graph). The fold induction gD promoter activity for each transfection is expressed as a percentage of the maximum induction by 10µg of plasmid p140 (100%) and plotted against the amount of effector plasmid; the maximum fold induction was 117 for the graph on the left, and 191 for the graph on the right. The symbols corresponding to the effector plasmids expressing the VZV gene 62 sequences are as explained below the graphs.

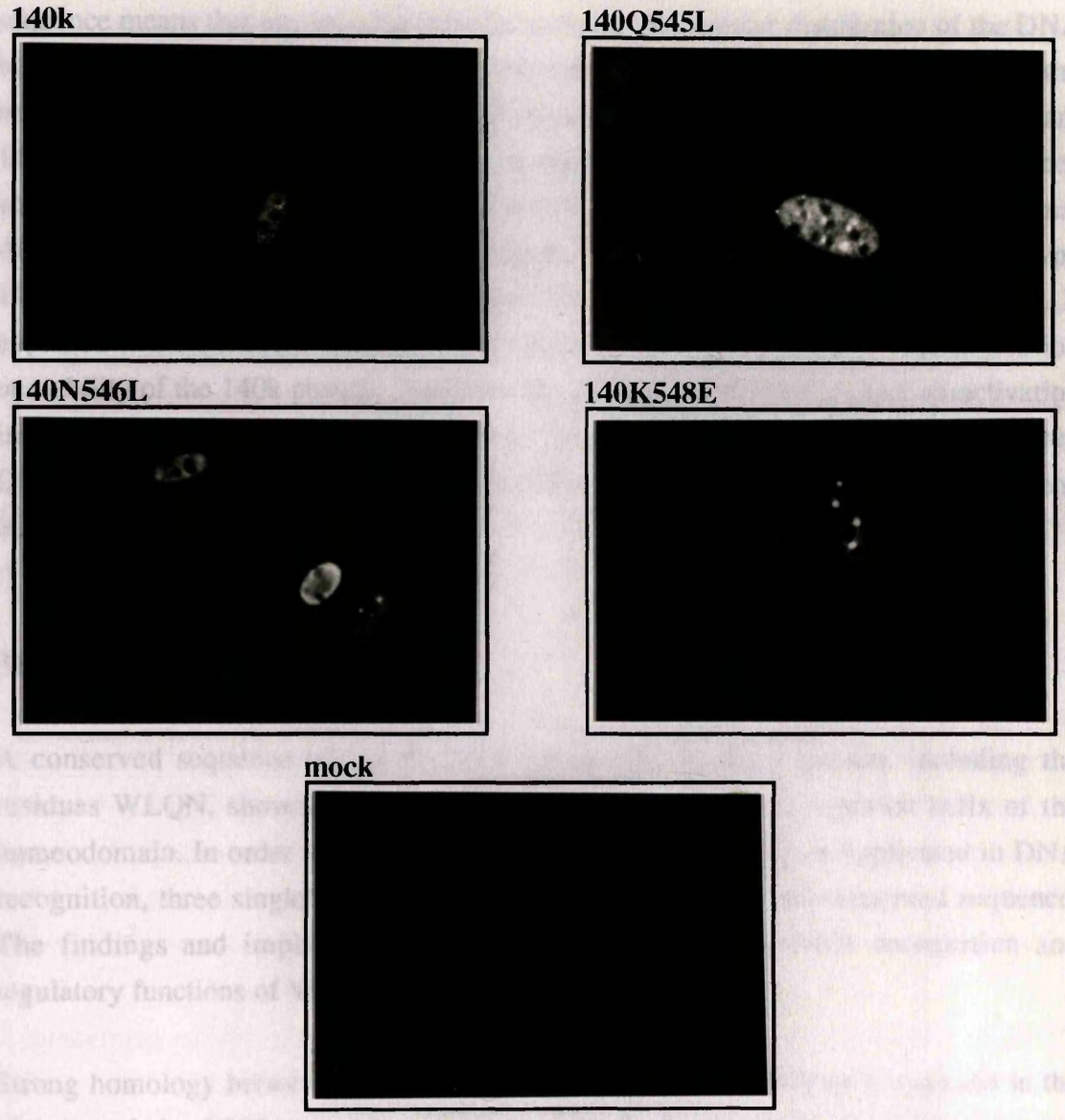


Figure 3E.10 Immunofluorescent staining of VZV 140k wild type and mutant proteins in BHK cells. Plasmids expressing protein 140k (p140), 140Q545L (p140Q545L), 140N546L (p140N546L), 140K548E (p140K548E) and no plasmid were transfected into BHK cells by lipofection (Section 2F.5). The 140k derived proteins were stained with polyclonal rabbit antiserum 109 to the VZV 140k DNA binding domain (Section 2F.6) and visualised by immunofluorescence microscopy. Granular and diffuse nuclear staining characteristic of the VZV 140k protein is provided by all the 140k derived peptides, as seen in the top four panels.

Even though the isolated VZV 140k DNA binding domain peptides containing mutations within the WLQN region were stable, soluble dimers (see Sections 3E.3 and 3E.4), it is conceivable that these mutations may have altered the expression or stability of the intact 140k protein. In this case, the reduced transactivation activities of the mutant 140k proteins seen in the transient transfection assays may have been due to protein instability. To clarify this issue, the immunofluorescence profiles of the mutant 140k proteins were compared to that of wild type 140k. The separation of the 140k DNA binding domain and the putative nuclear localisation signal in the primary

sequence means that any instability is likely to alter the cellular distribution of the DNA binding domain region of the 140k polypeptides. All four proteins localised to the nucleus and the characteristic distribution of nuclear staining was seen with each mutant 140k protein (Fig. 3E.10). For each set of transfections, similar intensity fluorescence was observed in comparable numbers of cells for the wild type and mutant proteins. Additionally, the mutant peptides gave bands of similar size and stability as wild type 140k in western blot analysis of extracts from transfected BHK cells (not shown). It appeared that the mutations in the WLQN region were not deleterious to the expression or stability of the 140k protein. Therefore, the differential efficiencies of transactivation induced by the mutant 140k proteins were likely to be a consequence of their altered DNA binding activities, as identified in the DNA binding analyses described in Sections 3E.4 and 3E.5.

3E.7 Discussion

A conserved sequence within the VZV 140k DNA binding domain, including the residues WLQN, showed significant similarity to the DNA recognition helix of the homeodomain. In order to determine whether the WLQN region is implicated in DNA recognition, three single 140k residues were mutated within this conserved sequence. The findings and implications of this mutagenesis for the DNA recognition and regulatory functions of VZV 140k are discussed here.

Strong homology between VZV 140k and the homeodomain was only apparent in the vicinity of the DNA recognition helix (homeodomain helix 3), but this does not automatically rule out the possibility of the 140k DNA binding domain including a homeodomain-like structure. The homeodomain protein fold is highly conserved between homeodomains that show very low sequence homology, as illustrated by a structural comparison of the yeast MATα2 (α2) and *Drosophila engrailed* (*en*) homeodomain-DNA complexes (Wolberger *et al.*, 1991). Recently, structural studies have identified numerous highly divergent homeodomains or eukaryotic HTH proteins (see Section 1E.2.3). The structures of non-classical homeodomains, such as the dimeric hepatocyte transcription factor HNF1/LFB1, have revealed their almost identical spatial arrangement of helices to those found in the corresponding HTH motifs of the *en* and α2 homeodomains, despite their very low level of sequence identity outside of the DNA recognition helix (reviewed by Brennan, 1993).

Unlike the homeodomain recognition helix, the WLQN region of VZV 140k is unlikely to be folded into a single uninterrupted linear α-helix, due to the presence of proline

residues at amino acids 540 and 547, but it could quite possibly adopt two helical turns between residues 540 and 547 and resume a helical conformation after residue 547 (see Fig. 3E.1). Typically, α–helices are kinked by 26° around a proline residue (Barlow and Thornton, 1988); examples of proline-induced α–helix kinking are seen in the *Eco*RV endonuclease structure (Winkler *et al.*, 1993), in the fourth helix of the divergent homeodomain (POU-specific domain) of Oct-1 (Assa-Munt *et al.*, 1993; Dekker *et al.*, 1993), and very recently the crystal structure of a heat shock factor (HSF) DNA binding domain identified a 29° proline-induced kink in helix 2 of a variant HTH motif (Harrison *et al.*, 1994). Applying the secondary structure prediction rules of Garnier *et al.* (1978) and Chou and Fasman (1974), only VZV 140k sequences corresponding to the hypothetical homeodomain helix 1 were predicted to fold into an α–helix (not shown). However, it should be noted that these types of prediction are only 62% accurate on average (Garnier *et al.*, 1978).

The bacterially expressed VZV 140k DNA binding domain provided a well characterised (Section 3A) and convenient tool for analysing the consequences of mutations in the WLQN region on DNA binding. By analogy to the findings of a comparative analysis of the HSV-1 Vmw175 protein and its isolated DNA binding domain (Everett *et al.*, 1991a), the effects of point mutations in the WLQN region on the DNA binding activity of the 140k DNA binding domain peptide are likely to reflect their probable consequences on the intact 140k protein.

Replacement of VZV 140k residue lysine 548 with glutamic acid (K548E) drastically reduced the sequence-specific DNA binding activity of the mutant 140k DNA binding domain, and resulted in a non-functional intact 140k protein, as determined by transient transfection assays. This residue occupies a position in the WLQN homology region equivalent to an invariant basic homeodomain residue (arginine 53) which forms two important contacts with the DNA phosphate backbone in the known structures. Given that no gross changes in dimerisation state, nuclear localisation, stability, solubility or efficiency of expression were detected as a result of the K548E mutation, these findings suggest that DNA binding is essential for the transactivation function of the VZV 140k protein. The relationship between the DNA binding and regulatory functions of the equivalent protein of HSV-1, Vmw175, is slightly controversial (see Section 1C.1.7). Detailed mutational and functional analyses have shown that mutations within the DNA binding domain that abrogated the DNA binding activity of Vmw175 also disrupted its regulatory functions (Paterson and Everett, 1988b; Shepard *et al.*, 1989), with one exception. This solitary case comprised a pseudorevertant of an inactive Vmw175 protein which had restored regulatory activities (in the context of viral infection) and impaired DNA binding properties (Shepard and DeLuca, 1991b). However, the isolated

DNA binding domain of the Vmw175 pseudorevertant protein was able to bind to DNA, at least at room temperature (K. E. Allen, N. DeLuca, R. Everett, unpublished data).

Although substitution K548E had pronounced effects on the DNA binding and regulatory functions of VZV 140k, the consequences of two further single mutations in the WLQN region were more subtle. Replacement of 140k residue glutamine 545 with leucine (Q545L) and asparagine 546 with leucine (N546L) had no apparent effect on DNA binding affinity as measured by gel retardation analysis. However, the specific details of their DNA binding interactions, as determined by DNase I footprinting analysis, significantly differed from those of either wild type VT2X or VT2XK548E. Mutations Q545L and N546L each reduced the transactivation efficiency of the 140k protein, the N546L mutation resulting in a more pronounced reduction in activation. The effects of these two mutations on the transactivation function are most easily explained as a consequence of their slightly altered interactions with DNA, although they may be caused by other factors not addressed in this study, such as subtle conformational changes within the domain resulting in defects in other macromolecular interactions. The corresponding homeodomain residues 50 and 51 are involved in determining the specificity of the homeodomain-DNA binding interaction. However, no change in specificity as a result of altering VZV 140k residue 545 or 546 was apparent, but would not necessarily be expected since the VT2X peptide has a fairly low sequence specificity (Section 3A.9) and may be tolerant to alteration of these hypothetical 'specificity-determining' residues. The effects of these three 140k mutations on DNA binding support the idea that this WLQN homology region is directly involved in the DNA binding interaction. Verification of the importance of the WLQN region will be provided by the analysis of the effects of these, and further amino acid substitutions, on the life cycle of the alphaherpesviruses.

The homeodomain contains a HTH motif. However, unlike the isolated HTH unit, the homeodomain forms a stable folded structure that is capable of binding to DNA by itself (Sauer *et al.*, 1988; Qian *et al.*, 1989; Affolter *et al.*, 1990). This 60 residue domain is much smaller than the minimum defined 140k DNA binding domain (162 residues) (Section 3B.4), although the structures of larger variant homeodomains have recently been determined. For example, the 'winged-helix' DNA binding motif of proteins such as HNF-γ3 (Clark *et al.*, 1993, reviewed by Brennan, 1993) is well over 100 amino acids in size.

The DNA binding activity of the HSV-1 Vmw175 protein is highly sensitive to insertion mutations throughout its entire DNA binding domain (see Section 1C.1.9 and this Thesis, Section 3D.7), and it is possible that other additional sequences within these

large domains are directly involved in DNA binding. Indeed, a HTH motif has been proposed on the basis of secondary structure predictions further C-terminal to the WLQN region of the HSV-1 Vmw175 protein (Shepard *et al.*, 1989) (Section 1C.1.9). Several substitutions within this helical region resulted in a temperature-sensitive DNA binding phenotype, while two other single mutations in the proposed 'turn' disrupted DNA binding (Allen and Everett, 1994). Although it is a possibility that this predicted HTH sequence may encode a structural scaffold which holds the WLQN region in such a way that it can be stably presented to DNA, it is not uncommon for proteins containing a homeodomain to couple additional DNA binding motifs or domains to increase their affinity or specificity of DNA binding. For example, prokaryotic HTH proteins such as λ repressor, use an extended region of peptide chain to wrap around the DNA (Clarke *et al.*, 1991). Similarly, the POU proteins have a conserved 65-75 residue segment, the POU-specific domain, on the N-terminal side of the homeodomain (Herr *et al.*, 1988) which contacts the DNA sequence ATGC, adjacent to the homeodomain target sequence (Ruvkun and Finney, 1991). Thus each of the two DNA binding motifs of the POU protein contacts one half of a non-palindromic bipartite target sequence; the bipartite nature of the non-palindromic HSV-1 Vmw175 consensus binding site is suggestive of an analogous DNA binding mechanism for the alphaherpesvirus proteins. Other DNA binding domains found in association with homeodomains include the paired domain (Treisman *et al.*, 1991), C_2H_2 zinc fingers and even additional homeodomains. The *Drosophila zfh-2* protein, containing three homeodomains and 16 zinc fingers, is an extreme example of this domain coupling (Fortini *et al.*, 1991). The involvement of several independent structural elements in the interaction with DNA would be consistent with the large size of the mapped DNA binding domains of the VZV 140k and HSV-1 Vmw175 proteins.

An interesting difference between the DNA recognition helix of the homeodomain and the corresponding sequences of the alphaherpesvirus proteins is that the latter lack basic C-termini. In this way the alphaherpesvirus WLQN region resembles the shorter DNA recognition helix of the prokaryotic HTH motif, which is equivalent to the N-terminal half of the homeodomain DNA recognition helix. The homeodomain-DNA interaction is stabilised by phosphate contacts by the basic residues in the C-terminal part of the recognition helix. Homeodomains are invariably monomeric, whereas the HTH proteins bind to DNA as dimers, perhaps to compensate for their lower DNA binding affinities resulting from the absence of the stabilising C-terminal basic residues. Comparisons can therefore be drawn between the prokaryotic HTH motif and the WLQN region of the VZV 140k protein, which has a single C-terminal basic residue (instead of the homeodomains' four) and binds DNA as a dimer.

Lastly, it may transpire that many DNA recognition motifs are variations on the same theme (Suzuki, 1993) and DNA binding proteins have evolved different structural scaffolds to present their DNA recognition motif to probe the DNA. The DNA recognition region of the VZV 140k family of alphaherpesvirus proteins may be an example of a divergent motif that has retained a minimum set of residues necessary for the interaction with DNA. The differences between the potential WLQN recognition sequence of the 140k protein and the known types of recognition motifs may facilitate its particular DNA binding characteristics, such as its low DNA binding specificity which allows it to bind multiple recognition sequences within a presumably very wide range of VZV promoters. The details of the involvement of the WLQN region in the DNA binding interaction awaits future structural analyses.

3F Activities of a hybrid HSV-1 Vmw175 protein containing the VZV 140k DNA binding domain in place of its own

3F.1 Strategy

There is evidence to suggest that VZV 140k and HSV-1 Vmw175 are functional counterparts in their respective alphaherpesviruses (Section 1C.2.1), although differences in detail between the two proteins have been identified in transient transfection assays. For example, 140k and Vmw175 were not entirely interchangeable in IE promoter repression assays; whereas HSV-1 Vmw175 down-regulated the VZV gene 62 promoter, VZV 140k actually slightly activated the HSV-1 IE3 promoter (in our assay system at least) (Disney *et al.*, 1990). In addition, 140k was a more potent and a more promiscuous transactivator of gene expression than Vmw175 in transient transfection experiments (see Section 1C.2.3).

Numerous differences between the DNA binding characteristics of the isolated VZV 140k and HSV-1 Vmw175 DNA binding domains have been identified (see Section 3A) which might account for the functional variations seen between the two intact proteins in the transient transfection experiments. To test this idea, a hybrid HSV-1 Vmw175 protein was constructed with the VZV 140k DNA binding domain in place of its own (protein 175R2DS). Transient transfection assays were used to compare the regulatory activities of the 175R2DS domain swap protein with those of the intact 140k and Vmw175 proteins. Additionally, the hybrid coding sequences were recombined into the HSV-1 genome and the effect of the domain swap on the efficiency of viral growth in tissue culture was determined.

3F.2 Construction of a hybrid HSV-1 Vmw175 protein containing the VZV 140k DNA binding domain (175R2DS)

Due to the lack of structural information on the HSV-1 Vmw175 and VZV 140k proteins, the details of the exchange of sequences encoding the DNA binding domains were designed using the information available from mutational analyses and sequence comparisons. The I9X and VT2X DNA binding domain peptides of Vmw175 and 140k respectively, appeared to be equivalent in that both were stable domains capable of the

full interaction with DNA, as determined by gel retardation and DNase I footprinting analyses (see Fig. 3B.7). Therefore the VT2X region of VZV 140k was inserted in place of the I9X region of HSV-1 Vmw175. Additionally, the relevant 140k region 3 sequences were included in the domain swap due to the presence of such sequences in the I9X peptide (for details, see Fig. 3F.1). VZV gene 62 codons 417-673 (spanning the 140k DNA binding domain) were amplified by PCR, due to the lack of convenient restriction enzyme sites for isolation of the required VZV gene 62 sequences.

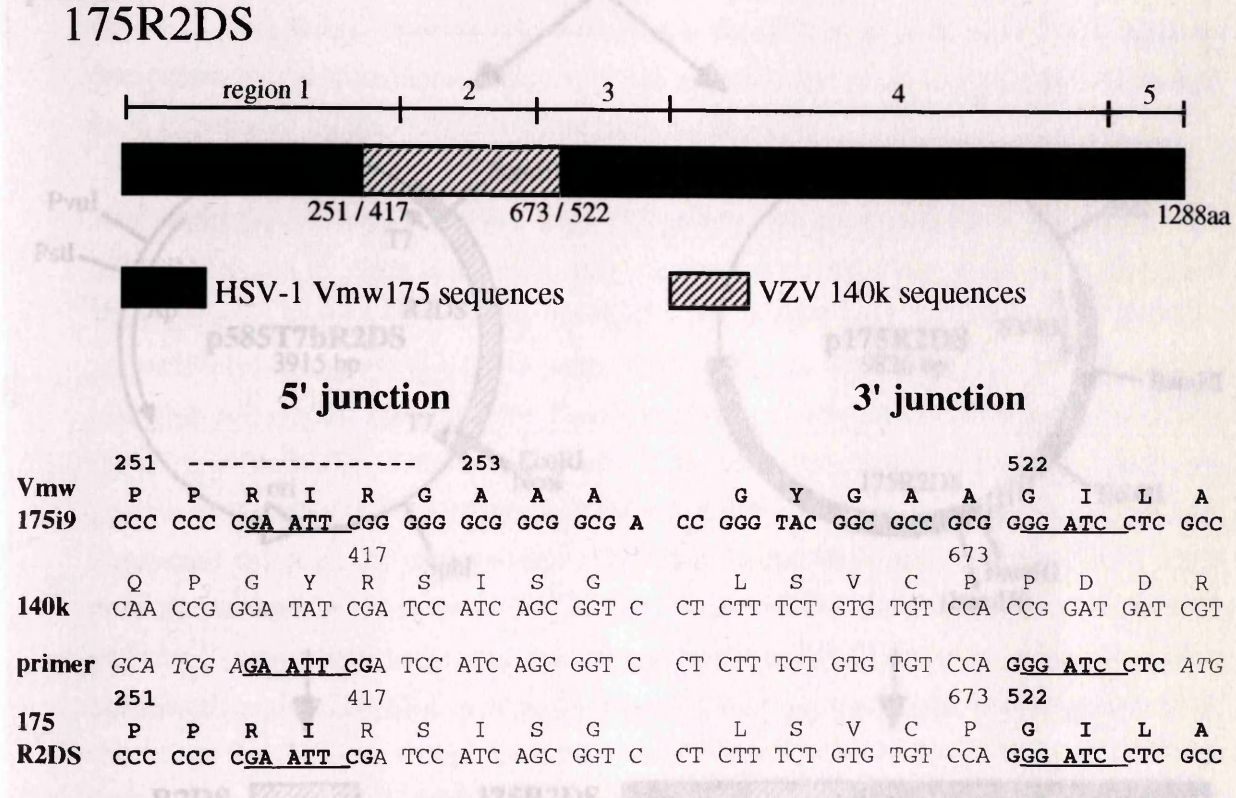


Figure 3F.1 Details of the 175R2DS domain swap protein. At the top is shown a schematic representation of the 175R2DS protein, which contains the DNA binding domain of VZV 140k (residues 417 to 673) replacing the corresponding region of HSV-1 Vmw175 (residues 252 to 521). Protein regions 1 to 5 are indicated on the scale above (Section 1C.1.8). Given beneath are details of the junctions between the Vmw175 and 140k coding sequences. The DNA sequence (top strand) and corresponding coding potential are shown for the Vmw175i9, 140k and 175R2DS proteins in the regions spanning the 5' and 3' fusion junctions. Vmw175i9 is a version of the Vmw175 protein including a 4 amino acid insertion as described below, and is encoded by gene IE3i9. Selected 140k and Vmw175 amino acids are numbered as appropriate. The sequences of the PCR primers used for amplification of the gene 62 sequences encoding the 140k DNA binding domain are also shown; the 3' junction PCR primer had the complementary sequence to that given here. All Vmw175i9 derived sequences are shown in bold, the 140k derived sequences are shown in normal text, and random sequences at the end of the primers are shown in italics. Underlining indicates restriction enzyme sites: the 5' primer was designed to include an *Eco*RI site which allows cloning of the PCR amplified fragment into the *Eco*RI site of the IE3i9 gene; the *Eco*RI site had been previously introduced on a 12bp linker (indicated by the dashed line) into codon 252 of the IE3 gene (Paterson and Everett, 1988a); this accounts for the discontinuity in the numbering of the Vmw175 amino acid residues. The 3' primer included a *Bam*HI site to allow cloning of the PCR product into the *Bam*HI site at codon 523 of the HSV-1 IE3 sequence. The 5' junction created an exact fusion between proteins Vmw175i9 and 140k, according to the alignment of the 140k and Vmw175 sequences of McGeoch *et al.* (1986). The 3' junction maintained the exact number of region 3 residues that are present in protein Vmw175i9.

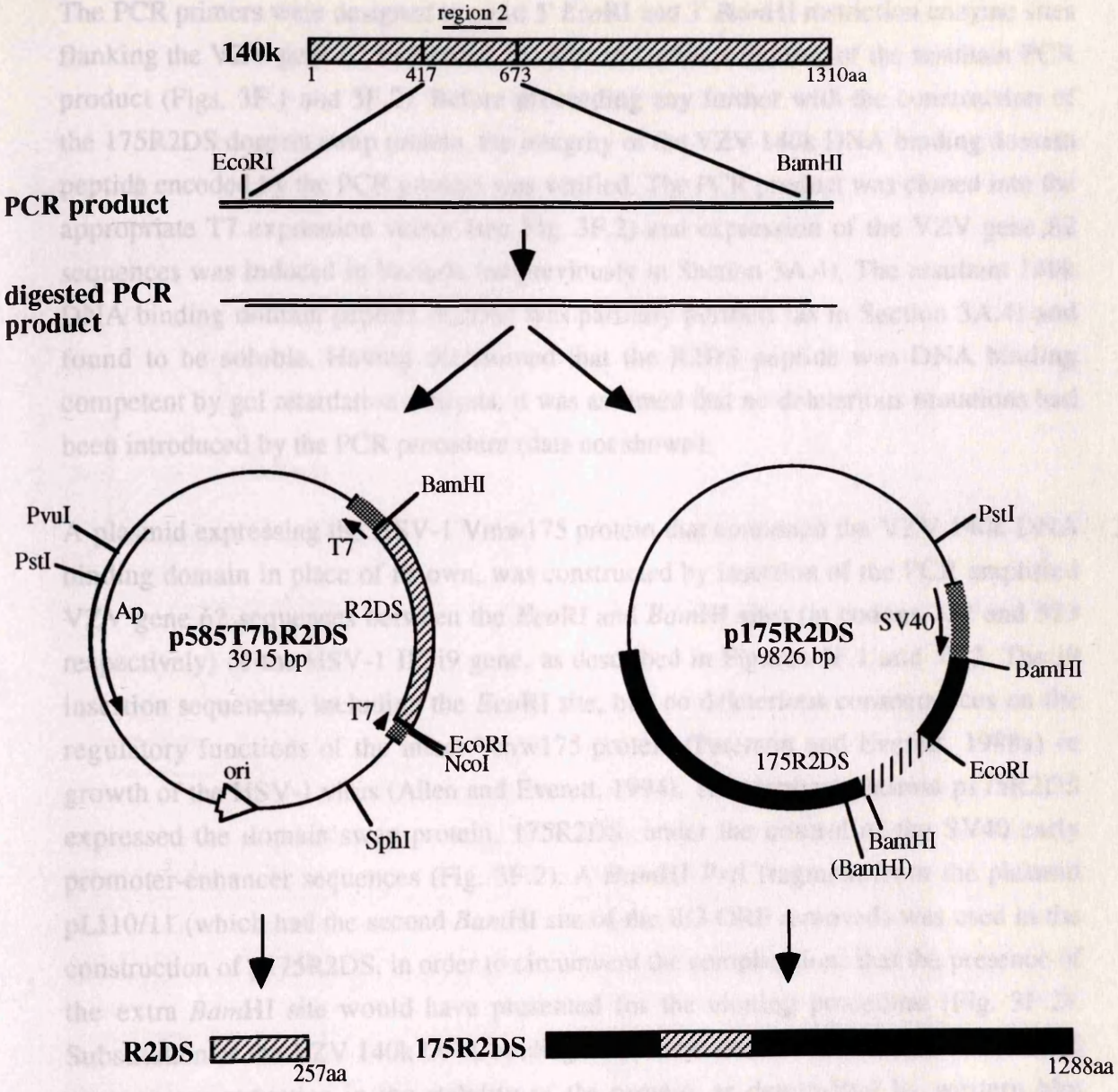


Figure 3F.2 Construction of plasmids expressing the R2DS VZV 140k DNA binding domain peptide and the 175R2DS hybrid protein. The top part shows schematic diagrams of the VZV 140k protein, and the PCR product encompassing VZV gene 62 codons 417 to 673; the PCR product was synthesised on the p140 plasmid template using the primers given in Figure 3F.1. After amplification and purification of the PCR product as described in Section 2B.8, it was cleaved with restriction enzymes *Eco*RI and *Bam*HI to yield sticky ends. The p585T7bR2DS plasmid was produced by insertion of the digested PCR product between the *Eco*RI and *Bam*HI sites of the MCR of T7 expression plasmid p585T7b2 (Section 3A.2). The p175R2DS plasmid was produced by co-ligation of the digested PCR product with the *Pst*I-*Eco*RI fragment (1925bp) of plasmid pI9 (including the remaining 5' sequences of the Vmw175 coding region) and the *Bam*HI-*Pst*I fragment (7084bp) of plasmid pLI10/11 (including the remaining 3' sequences of the Vmw175 coding region). Only selected restriction enzyme sites are shown. The corresponding expression products are shown schematically beneath each plasmid; the plasmid p585T7bR2DS expressed the 140k DNA binding domain peptide (R2DS), and plasmid p175R2DS expressed the hybrid Vmw175 protein containing the 140k DNA binding domain (protein 175R2DS). The total number of viral amino acid residues in each protein are indicated. The shading scheme for portions of protein sequences are consistent with that used in Fig. 3F.1.

The PCR primers were designed to yield 5' *Eco*RI and 3' *Bam*HI restriction enzyme sites flanking the VZV gene 62 sequences, in order to facilitate cloning of the resultant PCR product (Figs. 3F.1 and 3F.2). Before proceeding any further with the construction of the 175R2DS domain swap protein, the integrity of the VZV 140k DNA binding domain peptide encoded by the PCR product was verified. The PCR product was cloned into the appropriate T7 expression vector (see Fig. 3F.2) and expression of the VZV gene 62 sequences was induced in bacteria (as previously in Section 3A.4). The resultant 140k DNA binding domain peptide (R2DS) was partially purified (as in Section 3A.4) and found to be soluble. Having determined that the R2DS peptide was DNA binding competent by gel retardation analysis, it was assumed that no deleterious mutations had been introduced by the PCR procedure (data not shown).

A plasmid expressing the HSV-1 Vmw175 protein that contained the VZV 140k DNA binding domain in place of its own, was constructed by insertion of the PCR amplified VZV gene 62 sequences between the *Eco*RI and *Bam*HI sites (at codons 252 and 523 respectively) of the HSV-1 IE3i9 gene, as described in Figures 3F.1 and 3F.2. The i9 insertion sequences, including the *Eco*RI site, had no deleterious consequences on the regulatory functions of the intact Vmw175 protein (Paterson and Everett, 1988a) or growth of the HSV-1 virus (Allen and Everett, 1994). The resultant plasmid p175R2DS expressed the domain swap protein, 175R2DS, under the control of the SV40 early promoter-enhancer sequences (Fig. 3F.2). A *Bam*HI-*Pst*I fragment from the plasmid pLI10/11 (which had the second *Bam*HI site of the IE3 ORF removed) was used in the construction of p175R2DS, in order to circumvent the complications that the presence of the extra *Bam*HI site would have presented for the cloning procedure (Fig. 3F.2). Substitution of the VZV 140k DNA binding domain for that of HSV-1 Vmw175 caused no apparent reduction in the stability of the protein, as determined by western blot analysis of extracts from transiently transfected BHK cells (data not shown).

3F.3 Transactivation properties of the hybrid 175R2DS protein

The phenotype of the hybrid 175R2DS protein, was compared with those of intact VZV 140k and HSV-1 Vmw175 in transient transfection assays. It was of interest to determine the transactivation strength of the hybrid 175R2DS protein, as the 140k protein is known to be a much more potent transactivator than Vmw175 (Cabirac *et al.*, 1990; Disney, 1990). The strong transactivation phenotype of VZV 140k may be due to a specific feature of the binding activity of the 140k DNA binding domain, such as its lower sequence specificity (Section 3A.9) which presumably accounts for its ability to bind to multiple sequences within a given DNA fragment (Figs. 3A.8 and 3A.13).

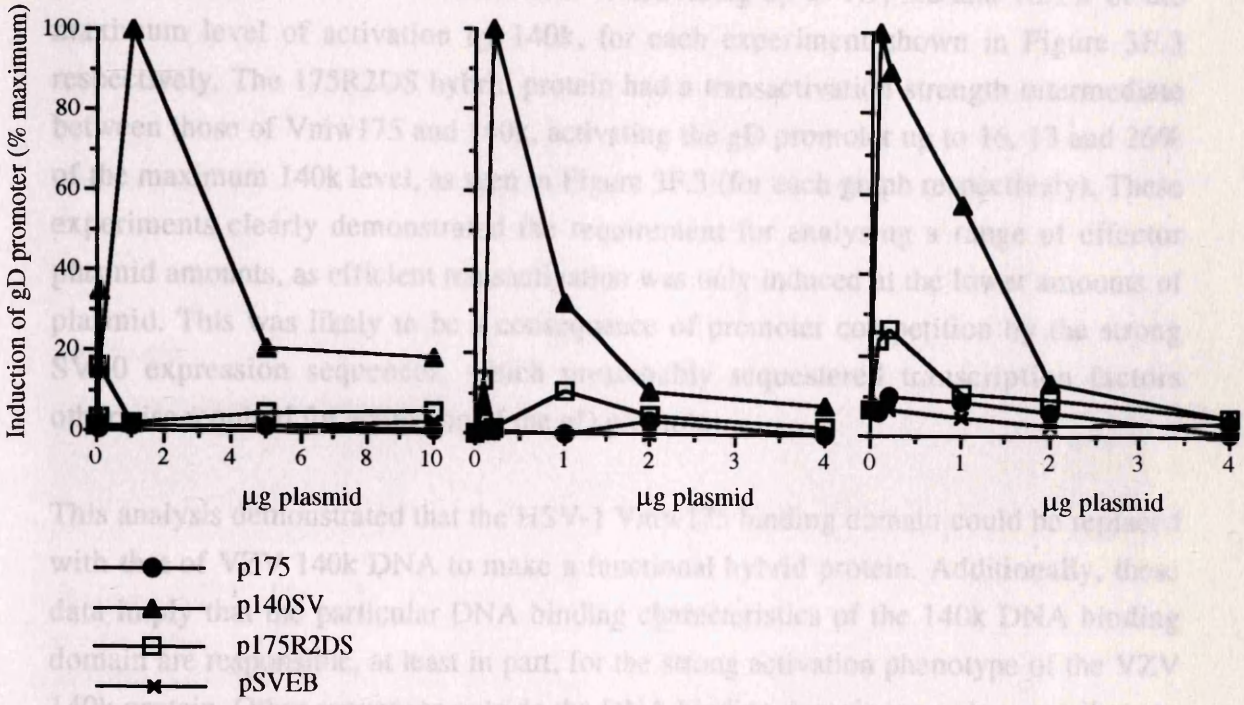


Figure 3F.3 Transactivation of the HSV-1 gD promoter in HeLa cells. The reporter construct pgDCAT (4µg) (carrying the CAT gene under control of the gD promoter of HSV-1) was transfected with increasing amounts of effector plasmids expressing Vmw175, 140k, 175R2DS or no coding region (plasmid p175, p140SV, p175R2DS or pSVEB respectively) and the resulting CAT activity was assayed as described in the legend to Figure 3E.9. The relative CAT activities (expressed as percentage of the maximum gD promoter induction by plasmid p140SV) from three independent repeats of such an experiment are shown above, plotted against amount of effector plasmid. The symbols representing effector plasmids are explained in the key beneath. The maximum fold gD promoter induction over the level resulting from transfection of plasmid pgDCAT alone was 69, 1093 and 15, for each experiment (from left to right).

The Vmw175, 140k and 175R2DS proteins were expressed from identical SV40 early promoter-enhancer sequences, to allow direct comparison of their relative transactivation strengths. Additionally, a plasmid which had the identical SV40 expression signals, but lacking a coding region (plasmid pSVEB), was included in these experiments to indicate the effects of promoter competition alone on the efficiency of gene expression from the reporter construct. As previously (Section 3E.6), transactivation of the HSV-1 glycoprotein D (gD) early promoter was determined from the reporter construct pgDCAT, following co-transfection into HeLa cells with increasing amounts of each effector plasmid. The results of three such experiments measuring transactivation of the HSV-1 gD promoter are shown in Figure 3F.3. Although variation in transfection efficiencies between the three experiments prevented statistical analysis of the data, it is clear that the relative transactivation strengths of the

Vmw175 and 140k and 175R2DS proteins were highly reproducible (Fig. 3F.3). Expression from the gD promoter was most strongly activated by VZV 140k, consistent with an earlier analysis of this promoter (Disney, 1990). Whereas the HSV-1 Vmw175 protein was a much weaker transactivator, activating up to 1.9, 3.2 and 10.1% of the maximum level of activation by 140k, for each experiment shown in Figure 3F.3 respectively. The 175R2DS hybrid protein had a transactivation strength intermediate between those of Vmw175 and 140k, activating the gD promoter up to 16, 13 and 26% of the maximum 140k level, as seen in Figure 3F.3 (for each graph respectively). These experiments clearly demonstrated the requirement for analysing a range of effector plasmid amounts, as efficient transactivation was only induced at the lower amounts of plasmid. This was likely to be a consequence of promoter competition by the strong SV40 expression sequences, which presumably sequestered transcription factors otherwise required for activation of the gD promoter.

This analysis demonstrated that the HSV-1 Vmw175 binding domain could be replaced with that of VZV 140k DNA to make a functional hybrid protein. Additionally, these data imply that the particular DNA binding characteristics of the 140k DNA binding domain are responsible, at least in part, for the strong activation phenotype of the VZV 140k protein. Other sequences outside the DNA binding domain may also contribute to the transactivation strength of 140k. Alternatively, unfavourable conformational effects caused by the domain swap could explain the lower transactivation efficiency of the 175R2DS hybrid, as compared to the 140k protein (see Section 3F.6).

Individually, HSV-1 immediate-early proteins Vmw175 and Vmw110 (see Section 1C.1.6) are very weak transactivators, only activating the gD promoter to 1-3 fold basal activity, following transfection into HeLa cells (Everett, 1986; this study, data not shown). However, co-transfection of Vmw175 and Vmw110 resulted in synergistic activation of gene expression, as a greater than additive increase in transcription resulted when the activators were present simultaneously (Everett, 1984b) (Section 1C.1.6). This synergistic activation of gene expression had not been observed between VZV 140k and HSV-1 Vmw110 in our assay system (Disney, 1990; this study, data not shown). However, the 175R2DS hybrid protein was able to activate the gD promoter in synergy with Vmw110. Synergistic activation resulted from co-transfection of the p175R2DS plasmid with increasing amounts of a plasmid expressing Vmw110 (plasmid p111) into HeLa cells, and also in the reciprocal experiment (Fig. 3F.4). Again, as with the transactivation experiments (Fig. 3F.3), large increases in activation were induced by addition of low amounts of (in this case) the second effector plasmid, resulting in further activation of the gD promoter over the combined effects of each separate effector plasmid. For example, co-transfection of 0.1μg of plasmid p111 with p175R2DS

resulted in an increase in transactivation of approximately 7 fold over the combined activation by the two separately transfected plasmids, (compare tracks 3 and 4, Fig. 3F.4), even though transactivation due to plasmid p111 alone was negligible (track 11). The finding that Vmw175 and 175R2DS, but apparently not 140k, acted in synergy with HSV-1 Vmw110, suggested that the DNA binding domains of 140k and Vmw175 are not sufficient for determining whether the protein can activate gene expression in synergy with Vmw110. However, synergistic activation may not occur when one of the transactivators works at very high efficiency, such as VZV 140k (see Section 3F.6).

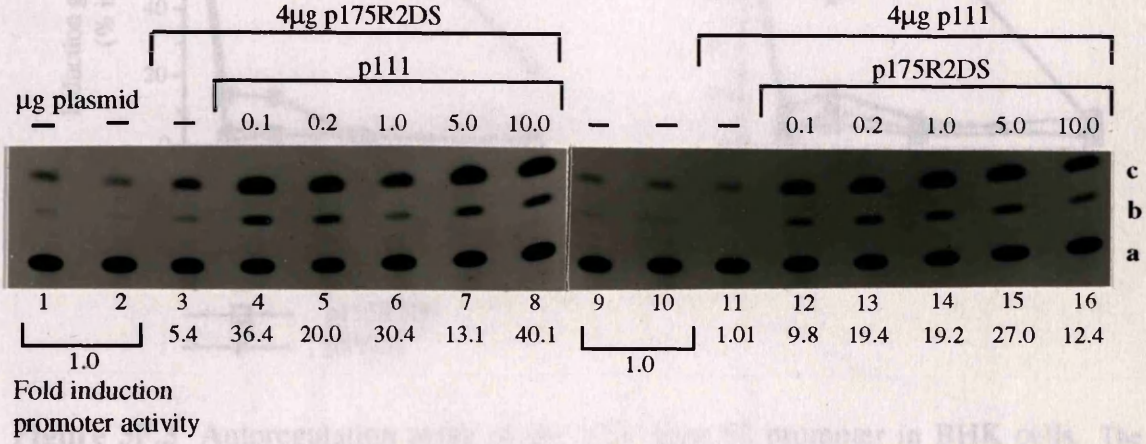


Figure 3F.4 Synergistic activation of the gD promoter by the 175R2DS hybrid protein with HSV-1 Vmw110 in HeLa cells. Increasing amounts of a plasmid expressing HSV-1 Vmw110 (p111) were co-transfected with 4µg of plasmid p175R2DS (left half of the Fig.). The reciprocal titration of increasing amounts of plasmid p175R2DS co-transfected with 4µg of plasmid p111 was also performed (right half of the Fig.), all transfections included 4µg of the target construct pgDCAT. Shown here are autoradiographs of TLC plates from CAT assays of the HeLa extracts resulting from such experiments. Fold induction of the gD promoter over the basal level of gD promoter activity for each extract (as indicated below each track) was calculated as described in the legend to Figure 3E.9. The letters on the right hand side indicate the positions of the radiolabelled substrate (a) and products (b and c), as described previously (legend to Fig. 3E.9).

3F.4 Analysis of the phenotype of the hybrid 175R2DS protein in autoregulation assays

As mentioned above (Section 3F.1), VZV 140k and HSV-1 Vmw175 affect differently the activities of their cognate promoters, as determined by transient transfection autoregulation assays (also see Section 1C.2.3). To investigate whether the DNA binding domain alone determines the specificity of transcriptional repression or whether other sequences modify this function, the phenotype of the 175R2DS hybrid protein in autoregulation assays was analysed. The cell lines and conditions used for these transient transfection assays were in keeping with those of the previous analysis of the 140k and Vmw175 proteins (Disney *et al.*, 1990).

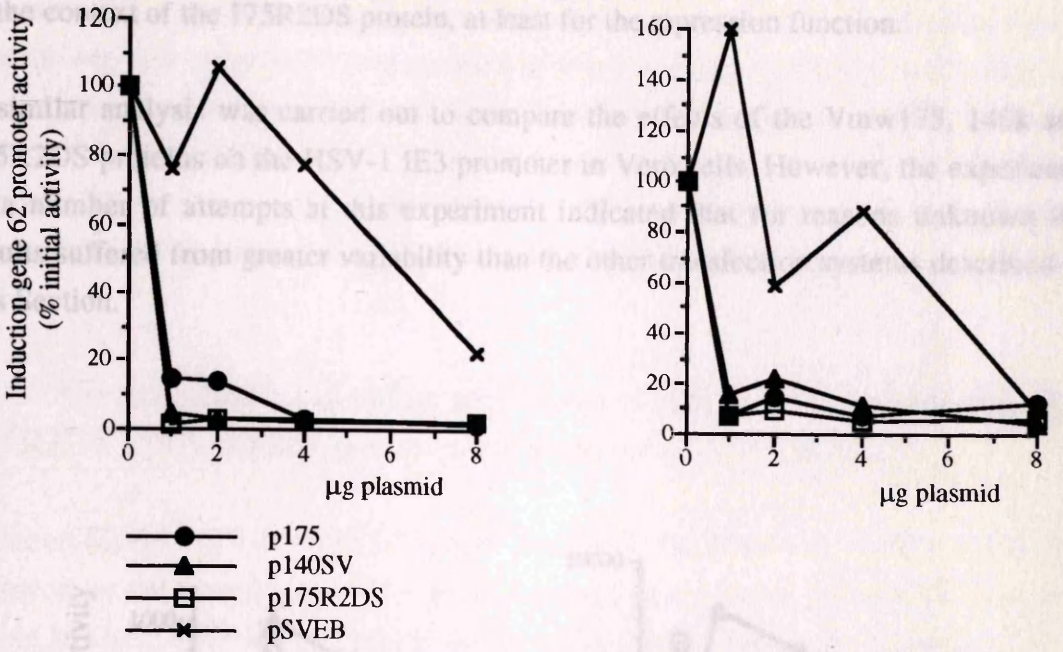


Figure 3F.5 Autoregulation assay of the VZV gene 62 promoter in BHK cells. The reporter construct p140CAT (4µg) (carrying the CAT gene under the control of the VZV gene 62 promoter), was co-transfected into BHK cells with 2µg of plasmid pMC1 (expressing HSV-1 Vmw65) and increasing amounts of an effector plasmid: p175, p140SV, p175R2DS or pSVEB (see legend to Fig. 3F.3.). CAT activities were quantitated and corrected for the protein concentration of BHK extracts as described previously (Fig. 3E.9). Gene 62 promoter induction is expressed as percentage of the gene 62 promoter activity when only plasmids pMC1 and p140CAT were transfected (100%). The data obtained from two repeats of such an experiment are shown graphically here, plotted against amount of transfected effector plasmid (as identified in the key below). The pSVEB plots indicate the level of repression due to promoter competition alone.

Consistent with the previous report (Disney *et al.*, 1990), both HSV-1 Vmw175 and VZV 140k repressed gene expression from the VZV gene 62 promoter in BHK cells (Fig. 3F.5). Inclusion of a plasmid (pMC1) which expresses the HSV-1 Vmw65 transactivator of immediate-early promoters (see Section 1B.2.2) was necessary to increase the initial level of gene 62 promoter activity in order to allow down-regulation to be detected. The control titration of plasmid pSVEB (carrying the SV40 expression signals but no coding region) was particularly important in these autoregulation assays, in order to differentiate between true repression due to the effector protein and repression caused by promoter competition effects which resulted from increasing the concentration of the SV40 promoter-enhancer sequences. The 175R2DS hybrid repressed the gene 62 promoter to the same degree as both Vmw175 and 140k (Fig. 3F.5). Although the analysis of the gene 62 promoter could not yield further information about the differential functional specificities of the 140k and Vmw175 DNA binding

domains (as they have the same phenotype on this promoter), it is apparent that the hybrid protein is fully functional for repression in this transient transfection system. This suggests that the integrity of the R2DS 140k DNA binding domain was fully maintained in the context of the 175R2DS protein, at least for the repression function.

A similar analysis was carried out to compare the effects of the Vmw175, 140k and 175R2DS proteins on the HSV-1 IE3 promoter in Vero cells. However, the experience of a number of attempts at this experiment indicated that for reasons unknown the results suffered from greater variability than the other transfection systems described in this Section.

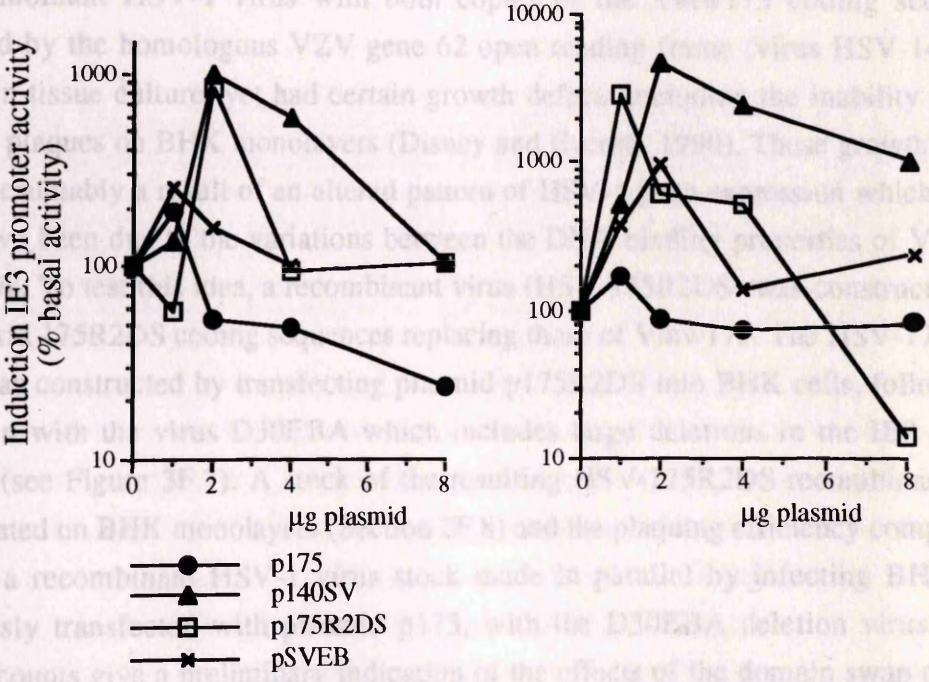


Figure 3F.6 Autoregulation assay of the HSV-1 IE3 promoter in Vero cells. The graphs represent the results of two repeats of an experiment measuring the effects of Vmw175, 140k and 175R2DS on the activity of the HSV-1 IE3 promoter. Experiments were performed exactly as for the VZV gene 62 promoter analysis (Fig. 3F.5) except that no pMC1 plasmid was included in transfections and the target construct was pIE3CAT (having the CAT gene under the control of the HSV-1 IE3 promoter sequences). For each transfection extract, the relative CAT activity was determined (as described in Figure 3E.9), expressed as a of the percentage of the basal level of gene expression (100%) and plotted against effector plasmid amount included in the transfection (using a log scale on the x-axis). The key below the graphs identifies the effector plasmid represented by each symbol, plasmid pSVEB included no coding region. In order to determine whether a particular amount of effector plasmid represses or activates gene expression, the IE3 promoter activity is compared to the value obtained using the equivalent amount of plasmid pSVEB.

Consistent with the previous findings using this system (Disney *et al.*, 1990), HSV-1 Vmw175 weakly repressed gene expression from its own IE3 promoter, while VZV 140k strongly activated this promoter (particularly apparent at around 2μg transfected plasmid) (Fig. 3F.6). Transactivation of the IE3 promoter by the 175R2DS hybrid protein was also observed at low amounts of effector plasmid, although activation was to a lesser extent than that seen with 140k. The Vmw175, 140k and 175R2DS proteins also had similar relative effects on the IE3 promoter activity in BHK cells (data not shown). These data suggested that the 140k DNA binding domain is important for determining the specificity of repression of the 175R2DS hybrid protein.

3F.5 Construction of a viable recombinant HSV-1 virus expressing the 175R2DS DNA binding domain swap protein (HSV-175R2DS)

A recombinant HSV-1 virus with both copies of the Vmw175 coding sequences replaced by the homologous VZV gene 62 open reading frame (virus HSV-140) was viable in tissue culture, yet had certain growth defects including the inability to form distinct plaques on BHK monolayers (Disney and Everett, 1990). These growth defects were presumably a result of an altered pattern of HSV-1 gene expression which in turn may have been due to the variations between the DNA binding properties of Vmw175 and 140k. To test this idea, a recombinant virus (HSV-175R2DS) was constructed with the hybrid 175R2DS coding sequences replacing those of Vmw175. The HSV-175R2DS virus was constructed by transfecting plasmid p175R2DS into BHK cells, followed by infection with the virus D30EBA which includes large deletions in the IE3 reading frames (see Figure 3F.7). A stock of the resulting HSV-175R2DS recombinant virus was titrated on BHK monolayers (Section 2F.8) and the plaquing efficiency compared to that of a recombinant HSV-1 virus stock made in parallel by infecting BHK cells previously transfected with plasmid p175, with the D30EBA deletion virus. These plaque counts give a preliminary indication of the effects of the domain swap on virus viability and hence the functional importance of the DNA binding characteristics specific to the 140k DNA binding domain (identified in Section 3A).

The results showed that the initial isolate of the HSV-175R2DS virus gave plaques on BHK monolayers with a titre of approximately 10^2pfu/ml, compared to 10^4pfu/ml of a virus stock obtained from a parallel recombination experiment using plasmid p175, which carries the wild type HSV-1 IE3 gene (see Fig. 3F.7).

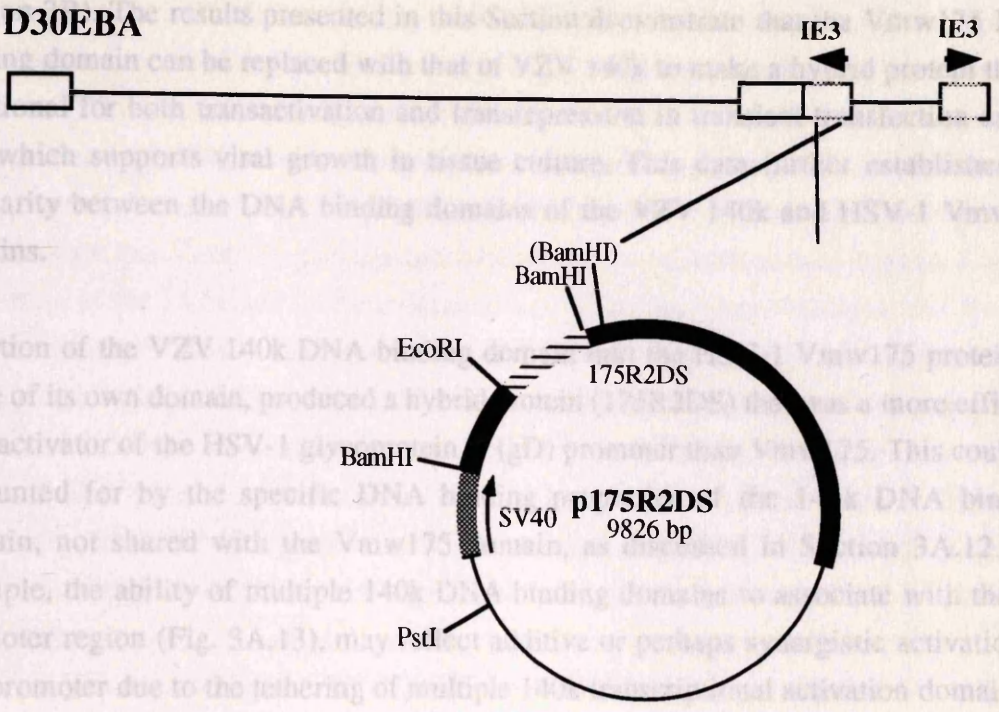


Figure 3F.7 Construction of the recombinant HSV-175R2DS virus. The HSV-175R2DS virus was produced by recombination between the D30EBA virus and the p175R2DS plasmid, as described in Section 2F.7. Virus D30EBA has a large deletion in both copies of the IE3 gene in the short repeats of the genome (Paterson and Everett, 1990). The p175R2DS plasmid encoding the domain swap protein is described in Section 3F.2. The recombinant HSV-1 virus was constructed in this same manner by inserting the wild type IE3 sequences, carried on plasmid p175, into virus D30EBA. Care was taken to include identical amounts of plasmid DNA in each transfection experiment and to infect with equal amounts of D30EBA.

A negative control stock produced in parallel from cells transfected with the pUC9 plasmid and infected with D30EBA contained no virus. Therefore a virus which expresses the hybrid Vmw175 protein with the VZV 140k DNA binding domain is viable in tissue culture, although it appeared to grow less efficiently than HSV-1. At this stage in the project the constraints of time prevented further characterisation of the HSV-175R2DS virus, but this has been undertaken by others during the writing of this Thesis. This further work showed that, of a sample of 12 plaques grown from the above HSV-175R2DS virus stock, all 12 had the recombinant gene at both IE3 loci of the genome. Further characterisation of the HSV-175R2DS virus will be presented elsewhere (Tyler *et al.*, 1994).

3F.6 Discussion

The DNA binding domains of the VZV 140k and HSV-1 Vmw175 proteins have been mapped to protein sequences that are highly conserved between the two proteins (see

Section 3B). The results presented in this Section demonstrate that the Vmw175 DNA binding domain can be replaced with that of VZV 140k to make a hybrid protein that is functional for both transactivation and transrepression in transient transfection assays and which supports viral growth in tissue culture. This data further establishes the similarity between the DNA binding domains of the VZV 140k and HSV-1 Vmw175 proteins.

Insertion of the VZV 140k DNA binding domain into the HSV-1 Vmw175 protein, in place of its own domain, produced a hybrid protein (175R2DS) that was a more efficient transactivator of the HSV-1 glycoprotein D (gD) promoter than Vmw175. This could be accounted for by the specific DNA binding properties of the 140k DNA binding domain, not shared with the Vmw175 domain, as discussed in Section 3A.12. For example, the ability of multiple 140k DNA binding domains to associate with the gD promoter region (Fig. 3A.13), may reflect additive or perhaps synergistic activation of this promoter due to the tethering of multiple 140k transcriptional activation domains to the promoter. This situation might be analogous to the synergistic activation mediated by other activator proteins such as Sp1, HSV-1 VP16 and EBV ZEBRA, that occurs as a consequence of multiple promoter occupancy or multimerisation of activation domains within the protein (Courey *et al.*, 1989; Emami and Carey, 1992; Chi and Carey, 1993).

However, the DNA binding properties of the VZV 140k DNA binding domain alone may not explain the strong transactivation phenotype of the 140k protein, as the hybrid 175R2DS protein was a less efficient transactivator than the intact 140k protein. There are several possible explanations for the transactivation strength of the hybrid protein being intermediate between that of Vmw175 and 140k. One of the more obvious explanations is that the VZV 140k protein may have a stronger transcriptional activation domain. This cannot be determined until the HSV-1 Vmw175 activation domain has been characterised. A potent activation domain has recently been identified within the first 90 amino acids of VZV 140k (Perera *et al.*, 1993; Cohen *et al.*, 1993) with a potency comparable to that of the activation domain of HSV-1 VP16 (Cohen *et al.*, 1993); the VP16 activation sequences are commonly held to be among the most potent. The activation domain of HSV-1 Vmw175 is not conserved with 140k, at least as far as location within the protein is concerned, as the first 90 amino acids of Vmw175 are dispensable for HSV-1 growth in tissue culture (DeLuca and Schaffer, 1987).

Alternatively, it is possible that the domain swap was not completely without consequence to the overall structure of the protein, leading to a reduction in the transactivation efficiency of the 175R2DS hybrid protein. Although the isolated R2DS domain was DNA binding competent, it is possible that in the background of Vmw175

sequences, the 140k DNA binding domain (R2DS) does not adopt the optimal conformation required for the efficient transactivation function. Region 2 of the 140k and Vmw175 proteins have very similar sizes and predicted primary sequences, whereas the requirement for the amount of region 1 and 3 sequences is very different between the two proteins (Section 1C.1.8). Mutational analyses have found that the regulatory functions of the Vmw175 protein are unaffected by insertions into region 1 and 3 sequences in the vicinities of the domain swap fusion junctions (Paterson and Everett, 1988a; Shepard *et al.*, 1989); so it is unlikely that the junctions themselves are deleterious to the functions of 175R2DS. Region 1 and 3 sequences are poorly conserved between 140k and Vmw175, and will probably fold differently in the two proteins. The tertiary conformations adopted by the 140k region 1 and 3 sequences included in the 175R2DS protein, may not be fully compatible with the conformations of the remaining Vmw175 protein regions, so reducing the transactivation efficiency of 175R2DS. It would be useful to construct the complementary domain swap (a VZV 140k protein including the HSV-1 Vmw175 DNA binding domain), in order to determine whether the intermediate transactivation strength of 175R2DS was due to conformational disruptions, or due to the absence of other specific 140k sequences required in addition to the DNA binding domain.

Like HSV-1 Vmw175, the 175R2DS hybrid protein activated gene expression in synergy with the HSV-1 immediate-early protein Vmw110. Little is known about the precise function of the Vmw110 protein (Section 1B.3.4); there is no evidence for Vmw110 binding directly to DNA although there is weak evidence for indirect binding (Hay and Hay, 1980). It is possible that Vmw110 and Vmw175 increase the efficiency of distinct rate-limiting steps in the transcription process, resulting in a greatly enhanced transactivation efficiency overall. There appeared to be a correlation between the extent of synergy with Vmw110 and the transactivation strength of the other effector protein, suggesting that addition of the second activator protein allows target promoter induction to reach maximum capacity. For example, the degree of synergy between the 175R2DS protein and Vmw110 (additional activation of 7 fold the combined individual effects) was not as marked as that seen previously with the weaker Vmw175 transactivator (around 11 fold the combined individual effects; Everett, 1986). Consistent with this idea, synergistic activation has not been detectable between the Vmw110 and VZV 140k proteins in our assay system. It is likely that the 140k protein is such a strong transactivator that the gD promoter was already fully activated by the 140k protein, in which case addition of further activators (i.e. Vmw110) would have little effect. This would be consistent with a findings of Thali *et al.* (1990); they demonstrated that the pseudorabies virus homologue of VZV 140k only stimulates promoter activity when the promoter is not already fully activated by other *trans*-inducing factors. Synergistic

activation of the VZV and HSV-1 *tk* promoters has been reported by VZV 140k with Vmw110 in HeLa and CV-1 cells (Cabirac *et al.*, 1990). This could be due to differences between the systems or alternatively the lower basal activity of the *tk* promoters might allow further activation beyond the level achieved by 140k alone.

As discussed previously (Section 1B.2.3), HSV-1 Vmw175 negatively autoregulates its own HSV-1 IE3 promoter. The fact that, like VZV 140k, the hybrid 175R2DS protein transactivated the HSV-1 IE3 promoter suggested that the DNA binding domain is important for determining the specificity of repression of these proteins. Presumably, the specific DNA binding characteristics of each DNA binding domain determined the outcome of the 140k and Vmw175 proteins binding to the HSV-1 IE3 promoter. The previously identified differences between the interactions of the Vmw175 and 140k DNA binding domains with the IE3 consensus site (Fig. 3A.12) may be relevant to the different effects that the two proteins have on the activity of the IE3 promoter.

The HSV-1 virus including the VZV 140k DNA binding domain (virus HSV-175R2DS) was viable, further demonstrating the functional similarity of the 140k and Vmw175 DNA binding domains. Any alterations in growth properties of the HSV-175R2DS virus are being explored and any differences found may be accounted for by the differences in the specific DNA binding activities of the 140k and Vmw175 DNA binding domains (Section 3A).

It is important to realise that the results obtained from transient transfection assays are notoriously variable (Everett, 1988b). For example, different consequences of the Vmw175 and 140k proteins on their cognate promoters have been reported in cell lines other than those used here (Perera *et al.*, 1992b). Therefore great caution is required when attempting to extrapolate the results of transient transfection assays to the situation during viral infection. However for the purposes of the analyses described here, transient transfection assays were utilised solely as a tool to compare the relative regulatory activities of the Vmw175, 140k and 175R2DS proteins. This allowed analysis of whether DNA binding characteristics unique to the 140k DNA binding domain were of functional relevance. Reliable information regarding the situation during infection would only be obtained from characterisation of the HSV-175R2DS virus.

Chapter 4: Discussion

The experiments described in this Thesis were designed to investigate the properties and functions of the DNA binding domain of the VZV 140k protein. Interpretations of the individual results can be found at the end of Sections 3A-3F. Section 4A integrates these findings and discusses their implications for the functions of VZV 140k. A model for the regulatory mechanism of VZV 140k is presented in Section 4B, taking into account the data presented in this Thesis and the previously published research on the other alphaherpesvirus homologue proteins (Section 1C). Finally, Section 4C discusses some possible avenues for future research leading on from these studies.

4A Overview and implications of the presented research

4A.1 The DNA binding characteristics of the VZV 140k DNA binding domain and their possible functional significance

The initial part of this project sought to characterise the DNA binding activity of the VZV 140k protein and to examine whether the functional differences between VZV 140k and its HSV-1 counterpart, Vmw175, that are apparent in transient transfection assays (Section 1C.2.3), are a consequence of possible variations in the DNA binding activities of the two proteins. As the full length VZV 140k protein had proved difficult to characterise (Disney, 1990), the approach chosen was to define the DNA binding properties of a bacterially expressed peptide fragment encompassing the 140k DNA binding domain. It is highly likely that the DNA binding activity of the isolated VZV 140k DNA binding domain reflects that of the intact 140k protein, by analogy to the very close correlation between the DNA binding characteristics of HSV-1 Vmw175 and its isolated DNA binding domain (Everett *et al.*, 1991a).

In common with the equivalent domain of HSV-1 Vmw175, the isolated VZV 140k DNA binding domain interacted with DNA in a direct and sequence-specific manner. Furthermore, it appears that DNA binding is an essential component of the regulatory

function of VZV 140k because a mutation (K548E) that disrupted DNA binding activity destroyed 140k transactivation (Section 3E). Importantly, mutation K548E had no apparent effect on other physical properties of the isolated domain or the full length protein (i.e. solubility, efficiency of expression, dimerisation, nuclear localisation). The involvement of DNA binding in the regulatory functions of VZV 140k has also been implied by the increased 140k-mediated transactivation that resulted from multimerisation of a 140k target binding site in the promoter of a reporter construct used in transient transfection experiments (Betz and Wydoski, 1993).

The VZV 140k DNA binding domain was found to interact specifically with a wide variety of sequences in all the target promoter fragments analysed and several of the 140k binding sites showed similarity to the HSV-1 Vmw175 consensus binding sequence (Section 3A). This indication that the VZV 140k and HSV-1 Vmw175 proteins may have similar DNA binding specificities was supported by the ability of the VZV 140k domain to interact specifically with Vmw175 consensus binding sites within HSV-1 promoters. This similar sequence specificity was not completely unexpected since the 140k and Vmw175 DNA binding domains share a high degree of predicted amino acid sequence identity (see Figs. 1C.1 and 1C.3).

The DNA binding characteristics of the 140k DNA binding domain, not shared with the Vmw175 domain, may account for the differences in the regulatory functions of VZV 140k and HSV-1 Vmw175 (Section 1C.2.3). The ability to bind multiple, fairly closely positioned binding sites within its target promoters was particular to the 140k DNA binding domain, and was a reflection of its lower DNA binding specificity, as compared to that of HSV-1 Vmw175 (Section 3A.9). This low specificity of DNA binding could account for both the wide range of 140k target promoters and also its powerful activation phenotype (see Sections 3A.12 and 4B). Furthermore, the distinct alterations to the DNA conformation induced by binding of the 140k and Vmw175 domains may have relevance to the different transactivation potencies of the two proteins; whereas the Vmw175 domain induced a significant DNA bend in the vicinity of its binding sites (Everett *et al.*, 1992), binding of the 140k domain induced marked DNase I hypersensitivity along the length of its target promoters.

The differences between the details of the interactions of the 140k and Vmw175 DNA binding domains with the HSV-1 IE3 consensus binding site (Section 3A.8) might explain the observation that VZV 140k activates the IE3 promoter, whereas HSV-1 Vmw175 down-regulates it (Section 1C.2.3). However, the fact that the Vmw175 DNA binding domain failed to recognise any sequences within the VZV gene 62 promoter, yet down-regulates this promoter in transient assays (Disney *et al.*, 1990) indicates that

there is no direct correlation between the ability to bind DNA *in vitro* and repression as measured by transient transfection experiments. It is likely that interactions with additional cellular factors, that do not occur in the *in vitro* binding analyses, are involved in repression by Vmw175; this would be consistent with the findings of Gu *et al.* (1993).

Characterisation of the activities of the hybrid 175R2DS protein, which contained the 140k DNA binding domain in the background of the HSV-1 Vmw175 protein, allowed the functional relevance of the differences between the DNA binding characteristics of the 140k and Vmw175 DNA binding domains to be assessed (Section 3F). The comparative analysis of the regulatory properties of the Vmw175 and 175R2DS proteins in transient transfection experiments indicated that the specific DNA binding properties of the VZV 140k DNA binding domain were responsible, at least in part, for the potent transactivation phenotype of VZV 140k and also for its specificity of repression. However, the intermediate nature of the regulatory activities of the hybrid protein between those of 140k and Vmw175 suggested that VZV 140k sequences, other than those within the DNA binding domain, may also play a part in determining the specific functional properties of the 140k protein. The possibility also remains that the domain swap between Vmw175 and 140k resulted in unfavourable conformational alterations; analysis of the complementary domain swap protein would differentiate between these two alternative explanations.

The viability of the HSV-175R2DS virus in tissue culture demonstrates clearly the homologous nature of the DNA binding domains of the VZV 140k and HSV-1 Vmw175 proteins. However, the reduced growth efficiency of HSV-175R2DS further supports the functional relevance of the differences between the DNA binding activities of the Vmw175 and 140k DNA binding domains detected *in vitro*. Thus, the specific DNA binding characteristics of the VZV 140k and HSV-1 Vmw175 transcriptional regulatory proteins are likely to contribute to the different characteristics of these two viral infections.

4A.2 The structure of the VZV 140k DNA binding domain

The VZV 140k and HSV-1 Vmw175 DNA binding domains occupy similar positions within each protein; both span protein region 2 (approximately 175 amino acids) that was originally defined on the basis of its high degree of sequence homology between 140k and Vmw175 (McGeoch *et al.*, 1986). Interestingly, the relatively poorly conserved sequences from the C-terminal end of region 1 were additionally required for

the full DNA binding interactions of the 140k and Vmw175 DNA binding domains, although this is to varying degrees in the two proteins (Section 3B). For instance, a VZV 140k DNA binding domain peptide lacking virtually all of region 1 (peptide VT4) bound to DNA with the same efficiency and specificity as the full domain (i.e. peptide VT2) in gel retardation analyses, whereas at least 30 amino acids of region 1 were additionally required for the HSV-1 Vmw175 DNA binding domain to exhibit the equivalent DNA binding activity. However, only those 140k and Vmw175 DNA binding domain peptides which included approximately 50-60 amino acids from region 1 were able to protect specific DNA sequences from DNase I digestion. These data demonstrate that the poorly conserved region 1 sequences are indirectly involved in the DNA binding interaction. This was previously implied by the findings of an insertion mutational analysis of the intact HSV-1 Vmw175 protein (Shepard *et al.*, 1989; Section 1C.1.9) and truncation analysis of the Vmw175 DNA binding domain (Wu and Wilcox, 1990; Section 1C.1.10).

The smaller size of the minimal 140k DNA binding domain indicates that the VZV 140k and HSV-1 Vmw175 proteins have different structural requirements for DNA binding, and this was further evidenced by the lack of a temperature-sensitive DNA binding phenotype of a 140k domain including the mutation equivalent to HSV-1 *ts*K. Interestingly, at 162 residues, the minimal VZV 140k DNA binding domain is larger than any of the documented DNA binding domains, with the exception of the C-terminal domain of the TATA binding protein (180 residues) (see Section 1E). No obvious extended homologies to the consensus sequences of the established types of DNA binding motifs are apparent within the VZV 140k or HSV-1 Vmw175 DNA binding domains, although this does not exclude the possibility that sequences within these alphaherpesvirus domains may adopt a similar fold to one of the previously recognised DNA binding motifs, as occurs with variant HTH proteins (Section 1E.2.3).

Of interest was a short conserved sequence (the WLQN region), located approximately in the centre of region 2 of the major alphaherpesvirus regulatory proteins, that shows homology to the DNA recognition helix of the homeodomain (Section 3E). The effect of mutations within the WLQN region of VZV 140k have demonstrated a possible direct role for these sequences in DNA recognition. The fact that the K548E mutation greatly reduced DNA binding activity of the 140k DNA binding domain indicates the importance of this region; however, more suggestive of a role for the WLQN region in DNA recognition is the fact that two further mutations (Q545L and N546L) altered the details of the specific DNA binding interaction. Verification of the role of the WLQN region in DNA recognition can only come from structural analysis of the DNA binding domain.

The intriguing question remains as to the requirement for the very highly conserved tract of sequences towards the C-terminus of region 2 of all the alphaherpesvirus homologue proteins (see Fig. 1C.3). The strong helical secondary structure predictions for this region make it a favourable candidate for either a dimerisation interface or a DNA recognition element; a substitution analysis confirmed the importance of this region but no dimerisation defective derivatives were isolated (Sections 1C.1.4 and 1C.1.9).

4A.3 The dimeric nature of the VZV 140k DNA binding domain

Like many sequence-specific DNA binding domains, the VZV 140k domain was found to be a stable dimer in solution, and to bind to DNA as a dimer (Sections 3C and 3D). It shares these properties with the HSV-1 Vmw175 protein and its isolated DNA binding domain, and functional studies have indicated that the dimerisation property of Vmw175 is significant *in vivo* (see Section 1C.1.4). Dimerisation itself is unlikely to play a regulatory role in the activities of 140k in the manner seen with the bHLH and bZIP proteins (Section 1E.3), as the dimers of the 140k DNA binding domain were extremely stable in solution. It is more likely that the dimeric nature of the 140k protein serves to increase the efficiency of the interaction with DNA, or its transactivation efficiency, by bringing additional 140k transcriptional activation domains to the promoter. The lack of any apparent palindromic nature of the VZV 140k and HSV-1 Vmw175 DNA target sequences is interesting because it indicates that, in contrast to the situation with other dimeric DNA binding proteins, the 140k monomers may not interact symmetrically with each half-site of the target DNA binding sequence. Furthermore, it appears that only one subunit of the dimer is sufficient for initial recognition of the binding site, even though both half-sites of the HSV-1 IE3 consensus binding site were protected by the 140k DNA binding domain from DNase I digestion. The alternative models for the dimer-DNA interaction accommodating these data are discussed in Section 3D.10. In the case of the 140k DNA binding domain, it is tempting to interpret the specificity data on the IE3 binding site (Section 3A.9) as indicating that the second monomer of the dimer binds non-specifically to the 3' half-site.

Although the entire minimal VZV 140k DNA binding domain (162 residues) was required for fully efficient dimerisation, the data indicate that the protein sequence and structural requirements for dimerisation are lower than those for DNA binding. Sequential deletions into the N-terminal end of 140k region 2 prevented DNA binding, yet only reduced the efficiency of dimerisation (see Fig. 3D.14). This suggests that sequences at the N-terminus of region 2 play an indirect role in the subunit interaction, perhaps stabilising the conformation of the domain. The fact that the VZV 140k and

HSV-1 Vmw175 DNA binding domains readily heterodimerised clearly demonstrates that the two proteins employ very similar modes of subunit interaction, and therefore the finding that insertions throughout the Vmw175 DNA binding domain had no apparent effect on the monomer interaction is applicable to the situation with the VZV 140k protein. It appears that dimerisation of these DNA binding domains is mediated by either a short stretch of amino acids that was missed by the insertion mutations or by a large interface that can accommodate the individual mutations.

Since the integrity of the entire domain was required either directly or indirectly for the DNA binding activity, it has not been possible using the current approaches to define completely individual sequences within the DNA binding domain of either 140k or Vmw175 that mediate the DNA binding and dimerisation interactions. Furthermore the results indicated that the majority of the DNA binding domain was also necessary for dimerisation, but this may reflect the requirement of sequences for the overall stability of the domain. It is probably that a large proportion of the sequences included in the DNA binding domain are required to form a scaffold for presentation of the DNA binding and dimerisation interfaces, rather than being directly involved in the macromolecular interactions. The localisation of the dimerisation interface will therefore require the identification of a subtle mutation that disrupts dimerisation of the domain without destroying its conformation. However it might be inferred from the fact that monomers of the DNA binding domains were never detected in solution, that such a monomeric mutant domain would be highly unstable and not amenable to analysis. As with the DNA recognition function of 140k, the details of the monomer interaction are only likely to be provided by structural analysis of the DNA binding domain.

4B A Model for transcriptional regulation by VZV 140k

Information concerning the mechanisms of transcriptional regulation by VZV 140k is scarce, but it is likely that parallels can be drawn with the other alphaherpesvirus homologue proteins. Possible mechanisms for VZV 140k function will be discussed on this basis.

The weight of evidence suggests that DNA binding is required for the regulatory functions of HSV-1 Vmw175 (see Sections 1B.2.3 and 1C.1.7), PRV IE180 (Cromlish *et al.,* 1989) and VZV 140k (this Thesis, Section 3E). However, unlike the majority of cellular transcription factors, these alphaherpesvirus proteins have degenerate DNA binding specificities (less pronounced with HSV-1 Vmw175), that could be more accurately termed as sequence-selectivity. Therefore it is unlikely that the DNA binding specificity of VZV 140k is itself responsible for selecting its target promoters. It is more probable that VZV 140k will bind to and transactivate any accessible class II promoter that is not already fully activated, according to the mechanism suggested for the PRV IE180 protein by Thali *et al.* (1990) (Section 1C.3). The viral promoters are the most abundant class of targets according to this model, and would account for the ability of VZV 140k to transactivate all the VZV IE, E and L promoters that have been analysed (Section 1C.2.3). The exception to this is the VZV gene 62 promoter, which is down-regulated by VZV 140k in certain cell types (see below) (Section 1C.2.3).

The HSV-1 Vmw175 protein appears to be able to activate gene expression from a variety of different positions within the vicinity of a promoter (Section 1C.1.7). The requirement for this property is not clear, but may solely be a consequence of its degenerate DNA binding specificity. As the 140k DNA binding domain is even less sequence-specific than Vmw175, and does not show any clear preference for particular sites within its target promoters, it is likely that 140k also functions in a position independent manner. If sufficient 140k is present, multiple 140k proteins may bind within the same promoter region in a possibly cooperative manner (Section 3A.7). Synergistic activation would result from multiple 140k occupancy in the promoter region and may even account for the potent transactivation phenotype of the VZV 140k protein seen in transient transfection assays. The 140k protein is highly abundant in these transiently transfected tissue culture cells, but it is not known whether this is also the case within VZV infected cells.

By analogy to the apparent situation with the HSV-1 Vmw175 and PRV IE180 proteins (Section 1B.3.3), VZV 140k-mediated transactivation probably involves contacts with the TFIID transcription factor complex and as such might enhance transcription by a mechanism similar to cellular transactivators (Section 1D.2.3). Deletion of the Sp1 binding sites of the HSV-1 *tk* promoter have indicated that Vmw175 can operationally substitute for Sp1 (Imbalzano *et al.*, 1991). This could occur if the transactivation mechanisms of Sp1 and Vmw175 are through a common pathway, perhaps involving contacts with a common protein. As Sp1 has recently been shown to contact a specific TAF co-activator component of TFIID (Section 1D.1.4), it is tempting to speculate that transactivation by Vmw175 (and perhaps the other alphaherpesvirus homologues) may also be mediated by contacting the Sp1 TAF. In agreement, the TAFs have also been implicated in the negative autoregulatory mechanism of HSV-1 Vmw175 (see Section 1B.2.3).

The autoregulatory mechanism of VZV 140k is liable to involve binding to the site(s) that has been identified adjacent to the gene 62 mRNA start site(s). When so positioned, 140k could interfere with the activity of cellular regulatory factors in a manner similar to that proposed for HSV-1 Vmw175 (Gu *et al.*, 1993; Section 1B.2.3). As the 140k DNA binding domain appeared to have a similar affinity for its different target sites, the inhibitory influence of 140k bound at the mRNA start site would have to be dominant over the stimulatory effect of 140k bound elsewhere within the gene 62 promoter region.

As with the apparent situation with the HSV-1 Vmw175 DNA binding activity (Section 1C.1.3), it is possible that the DNA binding activity (and hence the transactivation and repression functions of VZV 140k) might be regulated by post-translational modifications such as phosphorylation. Differential interactions with other cellular or viral factors may also play a part in determining its activities and promoter selectivity.

4C Further work

The work presented in this thesis has provided important insight into the activities of the VZV 140k protein, and in particular the properties of its DNA binding domain. However, numerous areas for potential investigation remain and several of these are outlined here.

It will be important to investigate further the relevance of the WLQN region for DNA recognition by analysing the consequences of other carefully selected mutations within this region of the VZV 140k DNA binding domain. Introduction of the WLQN mutations into the DNA binding domains of the other alphaherpesvirus homologue proteins could be used to indicate whether the conservation of these sequences reflects an evolutionarily conserved DNA recognition interface. Mutant forms of the HSV-1 Vmw175 DNA binding domain would be of particular value; the higher sequence-specificity of the Vmw175 DNA binding interaction would allow directional changes in binding specificity to be detected, using the random binding site selection method. By this approach it should be possible to identify the specific bases of the recognition site that are contacted by particular amino acids of the DNA recognition interface of Vmw175. Another question that could easily be answered, and would provide a better understanding of the mode of DNA recognition of these alphaherpesvirus proteins, is whether these DNA binding domains contact bases in the major or minor groove of DNA.

A fundamental objective should be to obtain X-ray crystallographic data for the VZV 140k DNA binding domain. This would provide answers to many questions, including the locations and structure of the DNA recognition and dimerisation interfaces, and would also elucidate the manner in which the dimer binds to an apparently asymmetrical target DNA sequence. The efficiency of the bacterial expression system and the purification procedure outlined in this Thesis will hopefully enable the structure of the DNA binding domain to be determined in the not too distant future.

The amino acid sequence of the poorly characterised region 4 of the alphaherpesvirus major regulatory proteins is even more highly conserved than their DNA binding domains, and as such is likely to be functionally important. During the course of this study an attempt to express the VZV 140k region 4 sequences as a GST fusion protein, and also as a non-fusion protein in *E. coli,* met with solubility problems. It would be

worthwhile to select a larger fragment of the protein for expression, perhaps including portions of region 3 and 5 sequences. Various approaches could be employed to investigate possible protein-protein interactions mediated by the 'tagged' sequences encompassing region 4, such as affinity chromatography or the co-immunoprecipitation technique that proved to be highly successful in Section 3D.

The approach of studying individual protein domains has many merits, but it is always important to relate their findings back to the situation with the intact protein, preferably in a physiological environment. Detailed characterisation of the properties of the HSV-175R2DS virus, in comparison with wild type HSV-1, might shed light on the functional relevance of the differences in the DNA binding characteristics of the Vmw175 and 140k DNA binding domains. In the near future, important mechanistic information regarding the transactivation activities of the alphaherpesvirus major regulatory proteins will no doubt be provided by reconstituted transcription systems, in the manner that the autoregulatory function of Vmw175 has been addressed (Gu *et al.*, 1993). It would be interesting to analyse further the role of DNA binding by studying the activities of the hybrid 175R2DS protein and also the WLQN mutant 140k proteins in this type of *in vitro* transcription system. The interplay between the alphaherpesvirus major regulatory proteins and the basal transcription machinery, the TAFs and also cellular *cis*-acting factors (such as Sp1) could all be addressed by such an *in vitro* approach. Such studies will be essential to gain a full understanding of the mechanism of action of these important viral proteins.

References

Abmayr,S.M., Feldman,L.D. and Roeder,R.G. (1985) In vitro stimulation of specific RNA polymerase II-mediated transcription by the pseudorabies virus immediate early protein. Cell, **43**, 821-829.

Abmayr,S.M., Workman,J.L. and Roeder,R.G. (1988) The pseudorabies immediate early protein stimulates in vitro transcription by facilitating TFIID : promoter interactions. Genes. Dev., **2**, 542-553.

Ace,C.I., McKee,T.A., Ryan,J.M., Cameron,J.M. and 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.

Ackermann,M., Braun,D.K., Pereira,L. and Roizman,B. (1984) Characterization of herpes simplex virus 1 α proteins 0, 4, and 27 with monoclonal antibodies. J. Virol., **52**, 108-118.

Affolter,M., Percival-Smith,A., Muller,M., Leupin,W. and Gehring,W.J. (1990) DNA binding properties of the purified *Antennapedia* homeodomain. Proc. Natl. Acad. Sci. USA, **87**, 4093-4097.

Alber,T. (1992) Structure of the leucine zipper. Curr. Opin. Genet. Dev., **2**, 205-210.

Allen,K.E. (1993) A mutational analysis of the DNA binding domain of the herpes simplex virus immediate early protein Vmw175. Ph.D Thesis, University of Glasgow.

Allen,K.E. and Everett,R.D. (1994) Mutations in a highly conserved tract in the HSV-1 Vmw175 DNA binding domain affect DNA binding, transcriptional activation and virus viability. Manuscript in preparation.

Allison,L.A. and Ingles,C.J. (1989) Mutations in RNA polymerase II enhance or suppress mutations in GAL4. Proc. Natl. Acad. Sci. USA, **86**, 2794-2798.

Allison,L.A., Moyle,M., Shales,M. and Ingles,C.J. (1985) Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. Cell, **42**, 599-610.

Anderson,A.S., Francesconi,A. and Morgan,R.W. (1992) Complete nucleotide sequence of the Marek's disease virus ICP4 gene. Virology, **189**, 657-667.

Andersson,L.-O., Borg,H. and Mikaelsson,M. (1972) Molecular weight estimations of proteins by electrophoresis in polyacrylamide gels of graded porosity. FEBS Lett., **20**, 199-202.

Anthony-Cahill,S.J., Benfield,P.A., Fairman,R., Wasserman,Z.R., Brenner,S.L., Stafford,W.F.III, Altenbach,C., Hubbell,W.L. and DeGrado,W.F. (1992) Molecular characterization of helix-loop-helix peptides. Science, **255**, 979-983.

apRhys,C.M.J., Ciufo,D.M., O'Neil,L.E.A., Kelly,T.J. and Hayward,G.S. (1989) Overlapping octamer and TAATGARAT motifs in the VF65-response elements in herpes simplex virus immediate early promoters represent independent binding sites for cellular nuclear factor III. J. Virol., **63**, 2798-2812.

Arkenasis,M., Campadelli-Fiume,G. and Roizman,B. (1988) Regulation of glycoprotein D synthesis: Does α 4, the major regulatory protein of herpes simplex virus 1, regulate late genes both positively and negatively? J. Virol., **62**, 148-158.

Arnosti,D.N., Preston,C.M., Hagmann,M., Schaffner,W., Hope,R.G., Laughlan,G. and Luisi,B. (1993) Specific transcriptional activation *in vitro* by the herpes simplex virus protein VP16. Nucleic Acids Res., **21**, 5570-5576.

Arthur,A.K., Hoss,A. and Fanning,A. (1988) Expression of simian virus 40 T-antigen in *Escherichia coli:* localization of T-antigen origin DNA-binding domain to within 129 amino acids. J. Virol., **62**, 1999-2006.

Arvin,A.M., Koropchak,C.M., Williams,B.R.G., Grumet,F.C. and Foung,S.K.H. (1986) Early immune response in healthy and immunocompromised subjects with primary varicella-zoster virus infection. J. Infect. Dis., **154**, 422-429.

Arvin,A.M., Sharp,M., Smith,S., Koropchak,C.M., Diaz,D.S., Kinchington,P.R., Ruyechan,W.T. and Hay,J. (1991) Equivalent recognition of a varicella zoster virus immediate-ealy protein (IE62) and glycoprotein I by cytotoxic T-lymphocytes of either the $CD4^+$ or $CD8^+$ phenotype. J. Immunol., **146**, 257-264.

Asano,Y., Itakura,N., Hiroishi,Y., Hirose,S., Ozaki,T., Yazaki,T., Yamanishi,K. and Takahashi,M. (1985) Viral replication and immunologic responses in children naturally infected with varicella-zoster virus and in varicella vaccine recipients. J. Infect. Dis., **152**, 863-868.

Assa-Munt,N., Mortishire-Smith,R.J., Aurora,R., Herr,W. and Wright,P.E. (1993) The solution structure of the Oct-1 POU-specific domain reveals a striking similarity to the bacteriophage λ repressor DNA-binding domain. Cell, **73**, 193-203.

Auble,D.T. and Hahn,S. (1993) An ATP-dependent inhibitor of TBP binding to DNA. Genes Dev., **7**, 844-853.

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1993a) Current Protocols in Molecular Biology, Vol **1**, Greene Publishing Assoc. and Wiley-Interscience, Brooklyn, New York.

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1993b) Current Protocols in Molecular Biology, Vol **2**, Greene Publishing Assoc. and Wiley-Interscience, Brooklyn, New York.

Bailey,A.C. and Thompson,R. (1992) A sequence-specific DNA-binding protein recognising a GA-rich element cooperates with Oct-1 at the herpes simplex virus type 1 IE3 promoter. Intervirol., **34**, 74-85.

Barlow,D.J. and Thornton,J.M. (1988) Helix geometry in proteins. J. Mol. Biol., **201**, 601-619.

Barlow,P.N., Luisi,B., Milner,A., Elliott,M. and Everett,R. (1994) Structure of the C_3HC_4 domain by ^{1}H-nuclear magnetic resonance spectroscopy; a new structural class of zinc-finger. J. Mol. Biol., in press.

Barnett,B.C., Dolan,A., Telford,E.A.R., Davison,A.J. and McGeoch,D.J. (1992) A novel herpes simplex virus gene (UL49A) encodes a putative membrane protein with counterparts in other herpesviruses. J. Gen. Virol., **73**, 2167-2171.

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

Beard,P., Faber,S., Wilcox,K.W. and Pizer,L.I. (1986) Herpes simplex virus immediate early infected-cell polypeptide 4 binds to DNA and promotes transcription. Proc. Natl. Acad. Sci. USA, **83**, 4016-4020.

Becker,Y. (1988) Computer-assisted primary and secondary structure analysis of DNA polymerases of herpes simplex, Epstein-Barr and varicella zoster virus reveal conserved domains with some homology to DNA binding domains in *E. coli* DNA pol I. Virus Genes, **1**, 351-367.

Ben-Porat,T., Hoffman,P., Brown,L., Feldman,L. and Blankenship,M.L. (1982) Partial characterization of temperature-sensitive mutants of pseudorabies virus. Virology, **122**, 252-267.

Ben-Zeev,A. and Becker,Y. (1977) Requirement of host cell RNA polymerase II in the replication of herpes simplex virus in α-amanitin sensitive and resistant cell lines. Virology, **76**, 246-253.

Berger,S.L., Cress,W.D., Cress.A., Triezenberg,S.J. and Guarente,L. (1990) Selective inhibition of activated but not basal transcription by the acidic activation domain of VP16: Evidence for transcriptional adaptors. Cell, **61**, 1199-1208.

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

Betz,J.L. and Wydoski,S.G. (1993) Functional interaction of varicella zoster virus gene 62 protein with the DNA sequence bound by herpes simplex virus ICP4 protein. Virology, **195**, 793-797.

Bexavanis,A.D. and Vinson,C.R. (1993) Interactions of coiled coils in transcription factors: Where is the specificity? Curr. Opin. Genet. Dev., **3**, 278-285.

Biggin,M., Farrell,P.J. and Barrell,B.G. (1984) Transcription and DNA sequence of the *BamHI* L fragment of B95-8 Epstein-Barr virus. EMBO J., **3**, 1083-1090.

Blaho,J.A. and Roizman,B. (1991) ICP4, the major regulatory protein of herpes simplex virus, shares features common to GTP-binding proteins and is adenylated and guanylated. J. Virol., **65**, 3759-3769.

Blaho,J.A., Michael,N., Kang,V., Aboul-Ela,N., Smulson,M.E., Jacobson,M.K. and Roizman,B. (1992) Differences in the poly(ADP-ribosyl)ation pattern of ICP4, the herpes simplex virus major regulatory protein, in infected cells and in isolated nuclei. J. Virol., **66**, 6398-6407.

Blair,E.D., Blair,C.C. and Wagner,E.K. (1987) Herpes simplex virus virion stimulatory protein mRNA leader contains sequence elements which increase both virus-induced transcription and mRNA stability. J. Virol., **61**, 2499-2508.

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

Breg,J.N., van Opheusden,J.H.J., Burgering,M.J., Boelens,R. and Kaptein,R. (1990) Structure of Arc repressor in solution: evidence for a family of β-sheet DNA-binding proteins. Nature, **346**, 586-589.

Brennan,R.G. (1993) The winged-helix DNA-binding motif: Another helix-turn-helix takeoff. Cell, **74**, 773-776.

Brennan,R.G. and Matthews,B.W. (1989) The helix-turn-helix DNA binding motif. J. Biol. Chem., **264**, 1903-1906.

Brennan,R.G., Roderick,S.L., Takeda,Y. and Matthews,B.W. (1990) Protein-DNA conformational changes in the crystal structure of a λ cro-operator complex. Proc. Natl. Acad. Sci. USA, **87**, 8165-8169.

Brou,C., Chaudhary,S., Davison,I., Lutz,Y., Wu,J., Egly,J.-M., Toro,L. and Chambon,P. (1993) Distinct TFIID complexes mediate the effect of different transcriptional activators. EMBO J., **12**, 489-499.

Brown,S.M., Ritchie,D.A. and Subak-Sharpe,J.H. (1973) Genetic studies with herpes simplex virus type 1. The isolation of temperature-sensitive mutants, their arrangement into complementation groups and recombination analysis leading to a linkage map. J. Gen. Virol., **18**, 329-346.

Buckmaster,A.E., Scott,S.D., Sanderson,M.J., Boursnell,M.E.G., Ross,N.L.J. and Binns,M.M. (1988) Gene sequence and mapping data from Marek's disease virus and herpesvirus of turkeys: implications for herpesvirus classification. J. Gen. Virol., **69**, 2033-2042.

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

Burley,S.K. (1994) DNA-binding motifs from eukaryotic transcription factors. Curr. Opin. Struct. Biol., **4**, 3-11.

Bzik,D.J. and 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.

Cabirac,G.F., Mahalingam,R., Wellish,M. and Gilden,D.H. (1990) Trans-activation of viral tk promoters by proteins encoded by varicella zoster virus open reading frames 61 and 62. Virus Res., **15**, 57-68.

Cabral,G.A., Courtney,R.J., Schaffer,P.A. and Marciano-Cabral,F. (1980) Ultrastructural characterization of an early, nonstructural polypeptide of herpes simplex virus type 1. J. Virol., **33**, 1192-1198.

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

Cai,C. and Schaffer,P.A. (1991) A cellular function can enhance gene expression and plating efficiency of a mutant defective in the gene for ICP0, a transactivating protein of herpes simplex virus type 1. J. Virol., **65**, 4078-4090.

Campbell,M.E.M. and Preston,C.M. (1987) DNA sequences which regulate the expression of the pseudorabies virus major immediate early gene. Virology, **157**, 307-316.

Campbell,M.E.M., Palfreyman,J.W. and 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.

Caughman,G.B., Staczek,J. and O'Callaghan,D.J. (1985) Equine herpesvirus type 1 infected cell polypeptides: Evidence for immediate early / early / late regulation of viral gene expression. Virology, **145**, 49-61.

Ceska,T.A., Lamers,M., Monaci,P., Nicosia,A., Cortese,R. and Suck,D. (1993) The X-ray structure of an atypical homeodomain present in the rat liver transcription factor LFB1/HNF1 and implications for DNA binding. EMBO J., **12**, 1805-1810.

Chasman,D.I., Flaherty,K.M., Sharp,P.A. and Kornberg,R.D. (1993) Crystal structure of yeast TATA-binding protein and model for interaction with DNA. Proc. Natl. Acad. Sci. USA, **90**, 8174-8178.

Chen,S., Mills,L., Perry,P., Riddle,S., Wobig,R., Lown,R. and Millette,R.L. (1992) Transactivation of the major capsid protein gene of herpes simplex virus type 1 requires a cellular transcription factor. J. Virol., **66**, 4304-4314.

Cheng,Y.-C., Tsou,T.Y., Hackstadt,T. and Mallavia,L.P. (1979) Induction of thymidine kinase and DNase in varicella-zoster virus-infected cells and kinetic properties of the virus-induced thymidine kinase. J. Virol., **31**, 172-177.

Chesnut,J.D., Stephens,J.H. and Dahmus,M.E. (1992) The interaction of RNA polymerase II with the adenovirus-2 major late promoter is precluded by phosphorylation of the C-terminal domain of subunit IIa. J. Biol. Chem., **267**, 10500-10506.

Cheung,A.K. (1989) DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies virus. Nucleic Acids Res., **17**, 4637-4646.

Chi,T. and Carey,M. (1993) The ZEBRA activation domain: modular organization and mechanism of action, Mol. Cell. Biol., **13**, 7045-7055.

Chou,P.Y. and Fasman,U.D. (1974) Predictions of protein conformation. Biochem., **13**, 222-245.

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

Clarke,N.D., Beamer,K.J., Goldberg,H.R., Berkower.,C. and Pabo,C.O. (1991) The DNA binding arm of λ repressor: Critical contacts from a flexible region. Science, **254**, 267-270.

Clarke,P., Brunschwig,A. and Gilden,D.H. (1993) DNA sequence of a simian varicella virus gene that encodes a homolgue of varicella zoster virus IE62 and herpes simplex virus ICP4. Virology, **197**, 45-52.

Clements,J.B., Watson,R.J. and Wilkie,N.M. (1977) Temporal regulation of herpes simplex virus type 1 transcription: location of transcripts on the viral genome. Cell, **12**, 275-285.

Coen,D.M., Weinheimer,S.P. and McKnight,S.L. (1986) A genetic approach to promoter recognition during trans induction of viral gene expression. Science, **234**, 53-59.

Cohen,J.I. and Seidel,K.E. (1993) Generation of varicella-zoster virus (VZV) and viral mutants from cosmid DNAs: VZV thymidylate synthetase is not essential for replication *in vitro*. Proc. Natl. Acad. Sci. USA, **90**, 7376-7380.

Cohen,J.I., Heffel,D. and Seidel,K. (1993) The transcriptional activation domain of varicella-zoster virus open reading frame 62 protein is not conserved with its herpes simplex virus homolog. J. Virol., **67**, 4246-4251.

Cohrs,R., Mahalingam,R., Dueland,A.N., Wolf,W., Wellish,M. and Gilden,D.H. (1992) Restricted transcription of varicella-zoster virus in latently infected human trigeminal and thoracic ganglia. J. Infect. Dis., **166** (Suppl. 1), 24-29.

Conaway,R.C. and Conaway,J.W. (1993) General initiation factors for RNA polymerase II. Annu. Rev. Biochem., **62**, 161-190.

Conley,A.J., Knipe,D.M., Jones,P.C. and Roizman,B. (1981) Molecular genetics of herpes simplex virus. VII. Characterization of a temperature-sensitive mutant produced by in vitro mutagenesis and defective in DNA synthesis and accumulation of γ polypeptides. J. Virol., **37**, 191-206.

Corden,J.L. (1990) Tails of RNA polymerase II. Trends Biochem. Sci., **15**, 383-387.

Corden,J., Cadena,D., Ahearn,J. and Dahmus,M. (1985) A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. Proc. Natl. Acad. Sci. USA, **82**, 7934-7938.

Cordingley,M.G., Campbell,M.E.M. and Preston,C.M. (1983) Functional analysis of a herpes simplex virus type 1 promoter: identification of far-upstream regulatory sequences. Nucleic Acids Res., **11**, 2347-2365.

Corsalo,C.M. and Pearson,M.L. (1981) Enhancing the efficiency of DNA-mediated gene transfer in mammalian cells. Somatic Cell Genet., **7**, 603-616.

Costanzo,F., Campadelli-Fiume,G., Foa-Tomasi,L. and Cassai,E. (1977) Evidence that herpes simplex virus DNA is transcribed by the cellular RNA polymerase B. J. Virol., **21**, 996-1001.

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

Courtney,R.J. and Benyesh-Melnick,M. (1974) Isolation and characterization of a large molecular-weight polypeptide of herpes simplex virus type 1. Virology, **62**, 539-551.

Courtney,R.J., Schaffer,P.A. and Powell,K.L. (1976) Synthesis of virus-specific polypeptides by temperature-sensitive mutants of herpes simplex virus type 1. Virology, **75**, 306-318.

Cousens,D.J., Greaves,R., Goding,C.R. and 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,W.D. and Triezenberg,S.J. (1991) Critical structural elements of the VP16 transcriptional activation domain. Science, **251**, 87-90.

Croen,K.D., Ostrove,J.M., Dragovic,L.J. and Straus,S.E. (1988) Patterns of gene expression and sites of latency in human nerve ganglia are different for varicella-zoster and herpes simplex viruses. Proc. Natl. Acad. Sci. USA, **85**, 9773-9777.

Cromlish,W.A., Abmayr,S.M., Workman,J.L., Horikoshi,M. and Roeder,R.G. (1989) Transcriptionally active immediate-early protein of pseudorabies virus binds to specific sites on class II gene promoters. J. Virol., **63**, 1869-1876.

Croston,G.E. and Kadonaga,J.T. (1993) Role of chromatin structure in the regulation of transcription by RNA polymerase II. Curr. Opin. Cell. Biol., **5**, 417-423.

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

Croston,G.E., Laybourn,P.J., Paranjape,S.M. and Kadonaga,J.T. (1992) Mechanism of transcriptional antirepression by GAL4-VP16. Genes Dev., **6**, 2270-2281.

Cunningham,C. and Davison,A.J. (1993) A cosmid-based system for constructing mutants of herpes simplex virus type 1. Virology, **197**, 116-124.

Dalrymple,M.A., McGeoch,D.J., Davison,A.J. and Preston,C.M. (1985) DNA sequence of the herpes simplex virus type 1 gene whose product is responsible for transcriptional activation of immediate early promoters. Nucleic Acids Res., **13**, 7865-7879.

Darst,S.A., Edwards,A.M., Kubalek,E.W. and Kornberg,R.D. (1991) Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution. Cell, **66**, 121-128.

Davidson,I. and Stow,N.D. (1985) Expression of an immediate early polypeptide and activation of a viral origin of DNA replication in cells containing a fragment of herpes simplex virus DNA. Virology, **141**, 77-88.

Davison,A.J. (1984) Structure of the genome termini of varicella-zoster virus. J. Gen. Virol., **65**, 1969-1977.

Davison,A.J. (1991) Varicella-zoster virus. The fourteenth Fleming Lecture. J. Gen. Virol., **72**, 475-486.

Davison,A.J. (1992) Channel catfish virus: a new type of herpesvirus. Virology, **186**, 9-14.

Davison,A.J. and McGeoch,D.J. (1986) Evolutionary comparisons of the S segments in the genomes of herpes simplex virus type 1 and varicella-zoster virus. J. Gen. Virol., **67**, 597-611.

Davison,A.J. and Scott,J.E. (1983) Molecular cloning of the varicella-zoster virus genome and derivation of six restriction endonuclease maps. J. Gen. Virol., **64**, 1811-1814.

Davison,A.J. and Scott,J.E. (1985) DNA sequence of the major inverted repeat in the varicella-zoster virus genome. J. Gen. Virol., **66**, 207-220.

Davison,A.J. and Scott,J.E. (1986) The complete DNA sequence of varicella-zoster virus. J. Gen. Virol., **67**, 1759-1816.

Davison,M.-J., Preston,V.G. and McGeoch,D.J. (1984) Determination of the sequence alteration in the DNA of the herpes simplex virus type 1 temperature-sensitive mutant *ts*K. J. Gen. Virol., **65**, 859-863.

Davison,A.J., Walter,D.J. and Edson,C.M. (1985) Identification of the products of a varicella-zoster virus glycoprotein gene. J. Gen. Virol., **66**, 2237-2242.

Davison,A.J., Edson,C.M., Ellis,R.W., Forghani,B., Gilden,D., Grose,C., Keller,P.M., Vafai,A., Wroblewska,Z. and Yamamishi,K. (1986) New common nomenclature for the glycoprotein genes of varicella zoster virus and their glycosylated products. J. Virol., **57**, 1195-1197.

Defechereax,P., Melen,L., Baudoux,L., Merville-Louis,M.-P., Rentier,B. and Piette,J. (1993) Characterization of the regulatory functions of varicella-zoster virus open reading frame 4 gene product. J. Virol., **67**, 4379-4385.

Dekker,N., Cox,M., Boelens,R., Verrijzer,C.P., van der Vliet,P.C. and Kaptein,R. (1993) Solution structure of the POU-specific DNA-binding domain of Oct-1. Nature, **362**, 852-855.

DeLuca,N.A. and Schaffer,P.A. (1985) Activation of immediate-early, early, and late promoters by temperature-sensitive and wild-type forms of herpes simplex virus type 1 protein ICP4. Mol. Cell. Biol., **5**, 1997-2008.

DeLuca,N.A. and Schaffer,P.A. (1987) Activities of herpes simplex virus type 1 (HSV-1) ICP4 genes specifying nonsense peptides. Nucleic Acids Res., **15**, 4491-4511.

DeLuca,N.A. and Schaffer,P.A. (1988) Physical and functional domains of the herpes simplex virus transcriptional regulatory protein ICP4. J. Virol., **62**, 732-743.

DeLuca,N.A., Courtney,M.A. and Schaffer,P.A. (1984) Temperature-sensitive mutants in herpes simplex virus type 1 ICP4 permissive for early gene expression. J. Virol., **52**, 767-776.

DeLuca,N.A., McCarthy,A.M. and Schaffer,P.A. (1985) Isolation and characterization of deletion mutants of herpes simplex virus type 1 in the gene encoding immediate-early regulatory protein ICP4. J. Virol., **56**, 558-570.

Dennis,D. and Smiley,J.R. (1984) Transactivation of a late herpes simplex virus promoter. Mol. Cell. Biol., **4**, 544-551.

Deshmane,S.L. and Fraser,N.W. (1989) During latency, herpes simplex virus type 1 DNA is associated with nucleosomes in a chromatin structure. J. Virol., **63**, 943-947.

DiDonato,J.A. and Muller,M.T. (1989) DNA binding and gene regulation by the herpes simplex virus type 1 protein ICP4 and involvement of the TATA element. J. Virol., **63**, 3737-3747.

DiDonato,J.A., Spitzner,J.R. and Muller,M.T. (1991) A predictive model for DNA recognition by the herpes simplex virus protein ICP4. J. Mol. Biol., **219**, 451-470.

Disney,G.H. (1990) A functional analysis of the product of varicella-zoster virus gene 62. Ph.D. Thesis, University of Glasgow.

Disney,G.H. and Everett,R.D. (1990) A herpes simplex virus type 1 recombinant with both copies of the Vmw175 coding sequences replaced by the homologous varicella-zoster virus open reading frame. J. Gen. Virol., **71**, 2681-2689.

Disney,G.H., McKee,T.A., Preston,C.M. and Everett,R.D. (1990) The product of varicella-voster virus gene 62 autoregulates its own promoter. J. Gen. Virol., **71**, 2999-3003.

Dixon,R.A.F. and Schaffer,P.A. (1980) Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding herpes simplex virus type 1 immediate early protein VP175. J. Virol., **36**, 189-203.

Dobersen,M.J., Jerkofsky,M. and Greer,S. (1976) Enzymatic basis for the selective inhibition of varicella-zoster virus by 5-halogenated analogues of deoxycytidine. J. Virol., **20**, 478-486.

Dolin,R., Reichman,R.C., Mazur,M.H. and Whitley,R.J. (1978) Herpes zoster-varicella infection in immunosuppressed patients. Ann. Intern. Med., **89**, 375-388.

Drapkin,R., Merino,A. and Reinberg,D. (1993) Regulation of RNA polymerase II transcription. Curr. Opin. Cell. Biol., **5**, 469-476.

Dueland,A.N., Martin,J.R., Devlin,M.E., Wellish,M., Mahalingam,R., Cohrs,R., Soike,K.F. and Gilden,D.H. (1992) Acute simian varicella infection. Clinical, laboratory, pathologic and virologic features. Lab Invest., **66**, 762-773.

Dumas,A.M., Geelen,J.L.M.C., Maris,W. and van der Noordaa,J. (1980) Infectivity and molecular weight of varicella-zoster virus DNA. J. Gen. Virol., **47**, 233-235.

Dumas,A.M., Geelen,J.L.M.C., Weststrate,M.W., Wertheim,P. and van der Noordaa,J. (1981) *Xba*I, *Pst*I and *Bgl*II restriction enzyme maps of the two orientations of the varicella-zoster virus genome. J. Virol., **39**, 390-400.

Dynlacht,B.D., Hoey,T. and Tjian,R. (1991) Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. Cell, **66**, 563-576.

Ecker,J.R. and Hyman,R.W. (1982) Varicella-zoster virus DNA exists as two isomers. Proc. Natl. Acad. Sci. USA, **79**, 156-160.

Eisenberg,S.P., Coen,D.M. and McKnight,S.L. (1985) Promoter domains required for expression of plasmid-borne copies of the herpes simplex virus thymidine kinase gene in virus-infected mouse fibroblasts and microinjected frog oocytes. Mol. Cell. Biol., **5**, 1940-1947.

ElKareh,A., Murphy,A.J.M., Fichter,T., Efstratiadis,A. and Silverstein,S. (1985) "Transactivation" control signals in the promoter of the herpesvirus thymidine kinase gene. Proc. Natl. Acad. Sci. USA, **82**, 1002-1006.

Ellenberger,T. (1994) Getting a grip on DNA recognition: Structures of the basic region leucine zipper, and the basic region helix-loop-helix DNA-binding domains. Curr. Opin. Struct. Biol., **4**, 12-21.

Ellenberger,T.E., Brandl,C.J., Struhl,K. and Harrison,S.C. (1992) The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: Crystal structure of the protein-DNA complex. Cell, **71**, 1223-1237.

Ellis,R.W., Keller,P.M., Lowe,R.S. and Zivin,R.A. (1985) Use of a bacterial expression vector to map the varicella-zoster virus major glycoprotein gene, gC. J. Virol., **53**, 81-88.

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

Epstein,M.A. (1962) Observations on the fine structure of mature herpes simplex virus and on the composition of its nucleoid. J. Experim. Med., **115**, 1-9.

Everett,R.D. (1983) DNA sequence elements required for regulated expression of the HSV-1 glycoprotein D gene lie within 83bp of the RNA capsites. Nucleic Acids Res., **11**, 6647-6666.

Everett,R.D. (1984a) A detailed analysis of an HSV-1 early promoter: sequences involved in trans-activation by viral immediate-early gene products are not early-gene specific. Nucleic Acids Res., **12**, 3037-3056.

Everett,R.D. (1984b) *Trans* activation of transcription by herpes virus products: requirement for two HSV-1 immediate early polypeptides for maximum activity. EMBO J., **3**, 3135-3141.

Everett,R.D. (1985) Activation of cellular promoters during herpes virus infection of biochemically transformed cells. EMBO J., **4**, 1973-1980.

Everett,R.D. (1986) The products of herpes simplex virus type 1 (HSV-1) immediate early genes 1, 2 and 3 can activate HSV-1 gene expression in *trans*. J. Gen. Virol., **67**, 2507-2513.

Everett,R.D. (1987a) The regulation of transcription of viral and cellular genes by herpesvirus immediate-early gene products. Anticancer Research, **7**, 589-604.

Everett,R.D. (1987b) A detailed mutational analysis of Vmw110, a *trans*-acting transcriptional activator encoded by herpes simplex virus type 1. EMBO J., **6**, 2069-2076.

Everett,R.D. (1988a) Analysis of the functional domains of herpes simplex virus type 1 immediate-early polypeptide Vmw110. J. Mol., Biol., **202**, 87-96.

Everett,R.D. (1988b) Promoter sequence and cell type can dramatically affect the efficiency of transcriptional activation induced by herpes simplex virus type 1 and its immediate-early gene products Vmw175 and Vmw110. J. Mol. Biol., **203**, 739-751.

Everett,R.D. (1989) Construction and characterization of herpes simplex virus type 1 mutants with defined lesions in immediate early gene 1. J. Gen. Virol., **70**, 1185-1202.

Everett,R.D. and Dunlop,M. (1984) *Trans* activation of plasmid-borne promoters by adenovirus and several herpes group viruses. Nucleic Acids Res., **12**, 5969-5978.

Everett,R.D. and Orr,A. (1991) The Vmw175 binding site in the IE-1 promoter has no apparent role in the expression of Vmw110 during herpes simplex virus type 1 infection. Virology, **180**, 509-517.

Everett,R.D., Paterson,T. and Elliott,M. (1990) The major transcriptional regulatory protein of herpes simplex virus type 1 includes a protease resistant DNA binding domain. Nucleic Acids Res., **18**, 4579-4585.

Everett,R.D., Elliott,M., Hope,G. and Orr,A. (1991a) Purification of the DNA binding domain of herpes simplex virus type 1 immediate-early protein Vmw175 as a homodimer and extensive mutagenesis of its DNA recognition site. Nucleic Acids Res., **19**, 4901-4908.

Everett,R.D., Orr,A. and Elliott,M. (1991b) High level expression and purification of herpes simplex virus type 1 immediate early polypeptide Vmw110. Nucleic Acids Res., **19**, 6155-6161.

Everett,R.D., Preston,C.M. and Stow,N.D. (1991c) Functional and genetic analysis of the role of Vmw110 in herpes simplex virus replication. In: Herpesvirus transcription and its regulation. pp49-76, Ed. Wagner. E.K., LRC Press, Boca Raton, USA.

Everett,R.D., DiDonato,J., Elliott,M. and Muller,M. (1992) Herpes simplex virus type 1 polypeptide ICP4 bends DNA. Nucleic Acids Res., **20**, 1229-1233.

Everett,R.D., Barlow,P., Milner,A., Luisi,B., Orr,A., Hope,G. and Lyon,D. (1993a) A novel arrangement of zinc-binding residues and secondary structure in the C_3HC_4 motif of an alpha herpes virus protein family. J. Mol. Biol., **234**, 1038-1047.

Everett,R.D., Cross,A. and Orr,A. (1993b) A truncated form of herpes simplex virus type 1 immediate-early protein Vmw110 is expressed in a cell type depenent manner. Virology, **197**, 751-756.

Everett,R., Cross,A., Tyler,J. and Orr,A. (1993c) An epitope within the DNA-binding domain of herpes simplex virus immediate early protein Vmw175 is conserved in the varicella-zoster virus gene 62 protein. J. Gen. Virol., **74**, 1955-1958.

Faber,S.W. and Wilcox,K.W. (1986a) Characterization of a herpes simplex virus regulatory protein: Aggregation and phosphorylation of a temperature sensitive variant of ICP 4. Arch. Virol., **91**, 297-312.

Faber,S.W. and Wilcox,K.W. (1986b) Association of the herpes simplex virus regulatory protein ICP4 with specific nucleotide sequences in DNA. Nucleic Acids Res., **14**, 6067-6083.

Faber,S.W. and Wilcox,K.W. (1988) Association of herpes simplex virus regulatory protein ICP4 with sequences spanning the ICP4 gene transcription initiation site. Nucleic Acids Res., **16**, 555-570.

Feldman,L.T., Imperiale,M.J. and Nevins,J.R. (1982) Activation of early adenovirus transcription by a herpesvirus immediate early gene: Evidence for a common cellular control factor. Proc. Natl. Acad. Sci. USA, **79**, 4952-4956.

Felgner,P.L., Gadek,T.R., Holm,M., Roman,R., Chan,H.W., Wenz,M., Northrop,J.P., Ringold,G.M. and Danielsen,M. (1987) Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. Proc. Natl. Acad. Sci. USA, **84**, 7413-7417.

Felser,J.M., Straus,S.E. and Ostrove,J.M. (1987) Varicella-zoster virus complements herpes simplex virus type 1 temperature-sensitive mutants. J. Virol., **61**, 225-228.

Felser,J.M., Kinchington,P.R., Inchauspe,G., Straus,S.E. and Ostrove,J.M. (1988) Cell lines containing varicella-zoster virus open reading frame 62 and expressing the "IE" 175 protein complement ICP4 mutants of herpes simplex virus type 1. J. Virol., **62**, 2076-2082.

Fenwick,M.L. (1984) The effects of herpes viruses on cellular macromolecular synthesis. In: Comprehensive Virology. **19**, 359-390, Eds Fraenkel-Conrat,H. and Wagner,R.R., Plenum Press, NY.

Fenwick,M., Walker,M. and Marshall,L. (1980) Some characteristics of an early protein (ICP 22) synthesized in cells infected with herpes simplex virus. J. Gen. Virol., **47**, 333-341.

Ferre-D'Amare,A.R., Prendergast,G.C., Ziff,E.B. and Burley,S.K. (1993) Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. Nature, **363**, 38-45.

Ferre-D'Amare,A.R., Pognonec,P., Roeder,R.G. and Burley,S.K. (1994) Structure and function of the b/HLH/Z domain of USF. EMBO J., **13**, 180-189.

Flanagan,W.M., Papavassiliou,A.G., Rice,M., Hecht,L.B., Silverstein,S. and Wagner,E.K. (1991a) Analysis of the herpes simplex virus type 1 promoter controlling the expression of U_L38 a true late gene involved in capsid assembly. J. Virol., **65**, 769-786.

Flanagan,P.M., Kelleher,R.J.III, Sayre,M.H., Tschochner,H. and Kornberg,R.D. (1991b) A mediator required for activation of RNA polymerase II transcription *in vitro*. Nature, **350**, 436-438.

Flores,O., Lu,H., Killeen,M., Greenblatt,J., Burton,Z.F. and Reinberg,D. (1991) The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. Proc. Natl. Acad. Sci. USA, **88**, 9999-10003.

Forghani,B., Mahalingam,R., Vafai,A., Hurst,J.W. and DuPuis,K.W. (1990) Monoclonal antibody to immediate early protein encoded by varicella-zoster virus gene 62. Virus Res., **16**, 195-210.

Fortini,M.E., Lai,Z. and Rubin,G.M. (1991) The *Drosophila* zfh-1 and zfh-2 genes encode novel proteins containing both zinc-finger and homeodomain motifs. Mechan. Dev., **34**, 113-122.

Foulkes,N.S. and Sassone-Corsi,P. (1992) More is better: activators and repressors from the same gene. Cell, **68**, 411-414.

Frankel,A.D. and Kim,P.S. (1991) Modular structure of transcription factors: Implications for gene regulation. Cell, **65**, 717-719.

Fraser,N.W., Spivack,J.G., Wroblewska,Z., Block,T., Deshmane,S.L., Valyi-Nagy,T., Natarajan,R. and Gesser,R.M. (1990) A review of the molecular mechanisms of HSV-1 latency. Curr. Eye Res., **10** (Suppl.), 1-13.

Freedman,L.P. and Luisi,B.F. (1993) On the mechanism of DNA binding by nuclear hormone receptors: A structural and functional perspective. J. Cell. Biochem., **51**, 140-150.

Freeman,M.J. and Powell,K.L. (1982) DNA-binding properties of a herpes simplex virus immediate early protein. J. Virol., **44**, 1084-1087.

Freemont,P.S. (1993) The RING finger. A novel protein sequence motif related to the zinc finger. Anals NY Acad. Sci., **684**, 174-192.

Freemont,P.S., Hanson,I.M. and Trowsdale,J. (1991) A novel cysteine-rich sequence motif. Cell, **64**, 483-484.

Gaffney,D.F., McLaughlan,J., Whitton,J.L. and Clements,J.B. (1985) A modular system for the assay of transcription regulatory signals: the sequence TAATGARAT is required for herpes simplex virus immediate early gene activation. Nucleic Acids Res., **13**, 7847-7863.

Galas,D.J. and Schmitz,A. (1978) DNAase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res., **5**, 3157-3170.

Garnier,J., Osguthorpe,D.J. and 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.

Gelman,I.H. and Silverstein,S. (1985) Identification of immediate early genes from herpes simplex virus that transactivate the virus thymidine kinase gene. Proc. Natl. Acad. Sci. USA, **82**, 5265-5269.

Gelman,I.H. and Silverstein,S. (1986) Co-ordinate regulation of herpes simplex virus gene expression is mediated by the functional interaction of two immediate early gene products. J. Mol. Biol., **191**, 395-409.

Gershon,A.A., Cosio,L. and Brunell,P.A. (1973) Observations on the growth of varicella-zoster virus in human diploid cells. J. Gen. Virol., **18**, 21-31.

Gershon,A., LaRussa,P., Hardy,I., Steinberg,S. and Silverstein,S. (1992) Varicella vaccine: The American experience. J. Infect. Dis., **166** (Suppl. 1), 63-68.

Gerster,T. and Roeder,R.G. (1988) A herpes virus trans-acting protein interacts with transcription factor OTF-1 and other cellular proteins. Proc. Natl. Acad. Sci. USA, **85**, 6347-6351.

Gilden,D.H., Shtram,Y., Friedmann,A., Wellish,M., Devlin,M., Fraser,N. and Becker,Y. (1982) The internal organization of the varicella-zoster virus genome. J. Gen. Virol., **60**, 371-374.

Gill,G. and Ptashne,M. (1987) Mutants of GAL4 protein altered in an activation function. Cell, **51**, 121-126.

Gill,G., Pascal,E., Tseng,Z.H. and Tjian,R. (1994) A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAF$_{II}$110 component of the *Drosophila* TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. USA, **91**, 192-196.

Godowski,P.J. and Knipe,D.M. (1983) Mutations in the major DNA-binding protein gene of herpes simplex virus type 1 result in increased levels of viral gene expression. J. Virol., **47**, 478-486.

Godowski,P.J. and Knipe,D.M. (1985) Identification of a herpes simplex virus function that represses late gene expression from parental viral genomes. J. Virol., **55**, 357-365.

Godowski,P.J. and Knipe,D.M. (1986) Transcriptional control of herpesvirus gene expression: Gene functions required for positive and negative regulation. Proc. Natl. Acad. Sci. USA, **83**, 256-260.

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

Gordon,J.E. (1962) Chickenpox: an epidemiological review. Am. J. Med. Sci., **244**, 362-389.

Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol., **2**, 1044-1051.

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

Gray,W.L., Baumann,R.P., Robertson.A.T., Caughmann,G.B., O'Callaghan,D.J. and Staczek,J. (1987a) Regulation of equine herpesvirus type 1 gene expression: characterization of immediate early, early, and late transcription. Virology, **158**, 79-87.

Gray,W.L., Baumann,R.P., Robertson.A.T., O'Callaghan,D.J. and Staczek,J. (1987b) Characterization and mapping of equine herpesvirus type 1 immediate early, early and late transcripts. Virus Res., **8**, 233-244.

Gray,W.L, Pumphrey,C.Y., Ruyechan,W.T. and Fletcher,T.M. (1992) The simian varicella virus and varicella zoster virus genomes are similar in size and structure. Virology, **186**, 562-572.

Green,M.R., Treisman,R. and Maniatis,T. (1983) Transcriptional activation of cloned human β-globin genes by viral immediate-early products. Cell, **35**, 137-148.

Grose,C. and Ng,T.I. (1992) Intracellular synthesis of varicella-zoster virus. J. Infect. Dis., **166** (Suppl. 1), 7-12.

Grundy,F.J., Baumann,R.P. and O'Callaghan,D.J. (1989) DNA sequence and comparative analyses of the equine herpesvirus type 1 immediate early gene. Virology, **172**, 223-236.

Gu,B., Rivera-Gonzalez,R., Smith,C.A. and DeLuca,N.A. (1993) Herpes simplex virus infected cell polypeptide 4 preferentially represses Sp1-activated over basal transcription from its own promoter. Proc. Natl. Acad. Sci. USA, **90**, 9528-9532.

Guess,H.A., Broughton,D.D., Melton,L.J.III and Kurland,L.T. (1986) Population-based studies of varicella complications. Pediatrics, **78**, 723-727.

Hahn,S. (1993a) Efficiency in activation. Nature, **363**, 672-673.

Hahn,S. (1993b) Structure(?) and function of acidic transcription activators. Cell, **72**, 481-483.

Hampl,H., Ben-Porat,T., Erlicher,L., Habermehl,K.-O. and Kaplan,A.S. (1984) Characterization of the envelope proteins of pseudorabies virus. J. Virol., **52**, 583-590.

Han,K. and Manley,J.L. (1993) Transcriptional repression by the *Drosophila* even-*skipped* protein: definition of a minimal repression domain. Genes Dev., **7**, 491-503.

Hanes,S.D. and Brent,R. (1989) DNA specificity of the Bicoid activator protein is determined by homeodomain recognition helix residue 9. Cell, **57**, 1275-1283.

Hanes,S.D. and Brent,R. (1991) A genetic model for interaction of the homeodomain recognition helix with DNA. Science, **251**, 426-430.

Hardwicke,M.A., Vaughan,P.J., Sekulovich,R.E., O'Conner,R. and Sandri-Goldin,R.M. (1989) The regions important for the activator and repressor functions of the herpes simplex virus type 1 α protein ICP27 map to the C-terminal half of the molecule. J. Virol., **63**, 4590-4602.

Hardy,I., Gershon,A.A., Steinberg,S.P., LaRussa,P. and the Varicella Vaccine Collaborative Study Group. (1991) The incidence of zoster after immunization with live attenuated varicella vaccine: A study in children with leukemia. N. Engl. J. Med., **325**, 1545-1550.

Harris,R.A., Everett,R.D., Zhu,X., Silverstein,S. and Preston,C.M. (1989) Herpes simplex virus type 1 immediate-early protein Vmw110 reactivates latent herpes simplex virus type 2 in an in vitro latency system. J. Virol., **63**, 3513-3515.

Harris-Hamilton,E. and Bachenheimer,S.L. (1985) Accumulation of herpes simplex virus type 1 RNAs of different kinetic classes in the cytoplasm of infected cells. J. Virol., **53**, 144-151.

Harrison,S.C. (1991) A structural taxonomy of DNA binding domains. Nature, **353**, 715-719.

Harrison,S.C. and Aggarwal,A.K. (1990) DNA recognition by proteins with the helix-turn-helix motif. Annu. Rev. Biochem., **59**, 933-969.

Harrison,C.J., Bohm,A.A. and Nelson,H.C.M. (1994) Crystal structure of the DNA binding domain of the heat shock transcription factor. Science, **263**, 224-227.

Harty,R.N., Colle,C.F., Grundy,F.J. and O'Callaghan,D.J. (1989) Mapping the termini and intron of the spliced immediate-early transcript of equine herpesvirus 1. J. Virol., **63**, 5101-5110.

Harty,R.N., Colle,C.F. and O'Callaghan,D.J. (1991) Equine herpesvirus type 1 gene regulation: Characterization of transcription from the immediate-early gene region in a productive infection. In: Herpesvirus transcription and its regulation. pp319-337, Ed. Wagner,E.K., LRC Press, Boca Raton, USA.

Hay,R.T. and Hay,J. (1980) Properties of herpesvirus-induced "immediate-early" polypeptides. Virology, **104**, 230-234.

Hayakawa,Y. and Hyman,R.W. (1987) Isomerization of the UL region of varicella-zoster virus DNA. Virus Res., **8**, 25-31.

Hayakawa,Y., Yamamoto,T., Yamanishi,K. and Takahashi,M. (1986) Analysis of varicella-zoster virus DNAs of clinical isolates by endonuclease *Hpa*I. J. Gen. Virol., **67**, 1817-1829.

Hedge,R.S., Grossman,S.R., Laimins,L.A. and Sigler,P.B. (1992) Crystal structure at 1.7 Å of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. Nature, **359**, 505-512.

Hernandez,N. (1993) TBP, a universal eukaryotic transcription factor. Genes Dev., **7**, 1291-1308.

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

Herschlag,D. and Johnson,F.B. (1993) Synergism in transcriptional activation: a kinetic view. Genes Dev., **7**, 173-179.

Hoey,T., Weinzievl,R.O.J., Grace,G., Chen,J.L., Dynlacht,B.D. and Tjian,R. (1993) Molecular cloning and functional analysis of *Drosophila* AF110 reveal properties expected of coactivators. Cell, **72**, 247-260.

Hoffman,A., Sinn,E., Yamamoto,T., Wang,J., Roy,A., Horikoshi,M. and Roeder,R.G. (1990) Highly conserved core domain and unique N terminus with presumptive regulatory motifs in a human TATA factor (TFIID). Nature, **346**, 387-390.

Holland,L.E., Anderson,K.P., Shipman,C.Jr., Wagner,E.K. (1980) Viral DNA synthesis is required for the efficient expression of specific herpes simplex virus type 1 mRNA species. Virology, **101**, 10-24.

Holmes,D.S. and Quigley,M. (1981) A rapid boiling method for the preparation of bacterial plasmids. Analyt. Biochem., **114**, 193-197.

Homa,F.L., Otal,T.M., Glorioso,J.C. and Levine,M. (1986) Transcription control signals of a herpes simpex virus type 1 late (γ_2) gene lie within bases -34 to +124 relative to the 5' terminus of the mRNA. Mol. Cell. Biol., **6**, 3652-3666.

Homa,F.L., Glorioso,J.C. and Levine,M. (1988) A specific 15-bp TATA box promoter element is required for expression of a herpes simplex virus type 1 late gene. Genes Dev., **2**, 40-53.

Hondo,R. and Yogo,Y. (1988) Strain variation of R5 direct repeats in the right-hand portion of the long unique segment of varicella-zoster virus DNA. J. Virol., **62**, 2916-2921.

Hondo,R., Yogo,Y., Kurata,T. and Aoyama,Y. (1987) Genome variation among varicella-zoster virus isolates derived from different individuals and from the same individuals. Arch. Virol., **93**, 1-12.

Honess,R.W. and Roizman,B. (1973) Proteins specified by herpes simplex virus. XI. Identification and relative molar rates of synthesis of structural and non structural herpes virus polypeptides in the infected cell. J. Virol., **12**, 1347-1365.

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

Honess,R.W. and Roizman,B. (1975) Regulation of herpesvirus macromolecular synthesis: sequential transition of polypeptide synthesis requires functional viral polypeptides. Proc. Natl. Acad. Sci. USA. **72**, 1276-1280.

Hope,I.A. and Struhl,K. (1987) GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. EMBO J., **6**, 2781-2784.

Hope,I.A., Mahadevan,S. and Struhl,K. (1988) Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. Nature, **333**, 635-640.

Hope-Simpson,R.E. (1965) The nature of herpes zoster: a long-term study and a new hypothesis. Proc. R. Soc. Med., **58**, 9-20.

Ihara,S., Feldman,L., Watanabe,S. and Ben-Porat,T. (1983) Characterization of the immediate-early functions of pseudorabies virus. Virology, **131**, 437-454.

Imbalzano,A.N. and DeLuca,N.A. (1992) Substitution of a TATA box from a herpes simplex virus late gene in the viral thymidine kinase promoter alters ICP4 inducibility but not temporal expression. J. Virol., **66**, 5453-5463.

Imbalzano,A.N., Shepard,A.A. and DeLuca,N.A. (1990) Functional relevance of specific interactions between herpes simplex virus type 1 ICP4 and sequences from the promoter-regulatory domain of the viral thymidine kinase gene. J. Virol., **64**, 2620-2631.

Imbalzano,A.N., Coen,D.M. and DeLuca,N.A. (1991) Herpes simplex virus transactivator ICP4 operationally substitutes for the cellular transcription factor Sp1 for efficient expression of the viral thymidine kinase gene. J. Virol., **65**, 565-574.

Imperiale,M.J., Feldman,L.T. and Nevins,J.R. (1983) Activation of gene expression by adenovirus and herpesvirus regulatory genes acting in *trans* and by a *cis*-acting adenovirus enhancer element. Cell, **35**, 127-136.

Inchauspe,G. and Ostrove,J.M. (1989) Differential regulation by varicella-zoster virus (VZV) and herpes simplex virus type-1 *trans*-activating genes. Virology, **173**, 710-714.

Inchauspe,G., Nagpal,S. and Ostrove,J.M. (1989) Mapping of two varicella-zoster virus-encoded genes that activate the expression of viral early and late genes. Virology, **173**, 700-709.

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

Inostroza,J.A., Mermelstein,F.H., Ha,I., Lane,W.S. and Reinberg,D. (1992) Dr1, a TATA-binding protein-associated phosphoprotein and inhibitor of class II gene transcription. Cell, **70**, 477-489.

Jackers,P., Defechereux,P., Baudoux.L., Lambert,C., Massaer,M., Merville-Louis,M. -P., Rentier,B. and Piette,J. (1992) Characterization of regulatory functions of the varicella-zoster virus gene 63-encoded protein. J. Virol., **66**, 3899-3903.

Jako,G.J. and Jako,R.A. (1986) Short historical note. Connection between varicella and herpes zoster. J. Med., **17**, 267-269.

Jaynes,J.B. and O'Farrell,P.H. (1991) Active repression of transcription by the engrailed homeodomain protein. EMBO J., **10**, 1427-1433.

Johnson,P.A. and Everett,R.D. (1986a) DNA replication is required for abundant expression of a plasmid-borne late US11 gene of herpes simplex virus type 1. Nucleic Acids Res., **14**, 3609-3625.

Johnson,P.A. and Everett,R.D. (1986b) The control of herpes simplex virus type 1 late gene transcription: a 'TATA-box'/capsite region is sufficient for fully efficient regulated activity. Nucleic Acids Res., **14**, 8247-8264.

Jones,P.C. and Roizman,B. (1979) Regulation of herpesvirus macromolecular synthesis. VIII. The transcription program consists of three phases during which both extent of transcription and accumulation of RNA in the cytoplasm are regulated. J. Virol., **31**, 299-314.

Jones,K.A. and Tjian,R. (1985) Sp1 binds to promoter sequences and activates herpes simplex virus immediate-early gene transcription *in vitro*. Nature, **317**, 179-182.

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

Jordan,S.R. and Pabo,C.O. (1988) Structure of the lambda complex at 2.5 Å resolution: Details of the repressor-operator interactions. Science, **242**, 893-899.

Kalderon,D., Roberts,B.L., Richardson,W.D. and Smith,A.E. (1984) A short amino acid sequence able to specify nuclear location. Cell, **39**, 499-509.

Kamakaka,R.T., Bulger,M. and Kadonaga,J.T. (1993) Potentiation of RNA polymerase II transcription by GAL4-VP16 during but not after DNA replication and chromatin assembly. Genes Dev., **7**, 1779-1795.

Kao,C.C., Lieberman,P.M., Schmidt,M.C., Zhou,Q., Pei,R. and Berk,A.J. (1990) Cloning of a transcriptionally active human TATA binding factor. Science, **248**, 1646-1650.

Kaptein,R. (1992) Zinc-finger structues. Curr. Opin. Struct. Biol., **2**, 109-115.

Kattar-Cooley,P. and Wilcox,K.W. (1989) Characterization of the DNA-binding properities of herpes simplex virus regulatory protein ICP4. J. Virol., **63**, 696-704.

Kelleher,R.J.III, Flanagan,P.M. and Kornberg,R.D. (1990) A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. Cell, **61**, 1209-1215.

Keller,P.M., Davison,A.J., Lowe,R.S., Bennett,C.D. and Ellis,R.W. (1986) Identification and structures of the gene encoding gpII, a major glycoprotein of varicella-zoster virus. Virology, **152**, 181-191.

Keller,P.M., Davison,A.J., Lowe,R.S., Reimen,M.W. and Ellis,R.W. (1987) Identification and sequence of the gene encoding gpIII, a major glycoprotein of varicella-zoster virus. Virology, **157**, 526-533.

Kemp,L.M. and Latchman,L.M. (1988) Induction and repression of cellular gene transcription during herpes simplex virus infection are mediated by different viral immediate-early gene products. Eur. J. Biochem., **174**, 443-449.

Kibler,P.K., Duncan,J., Keith,B.D., Hupel,T. and Smiley,J.R. (1991) Regulation of herpes simplex virus true late gene expression: sequences downstream from the US11 TATA box inhibit expression from an unreplicated template. J. Virol., **65**, 6749-6760.

Kieran,M., Blank,V., Logeat,F., Vandekerckhove,J., Lottspeich,F., Le Bail,O., Urban,M.B., Kourilsky,P., Baeuerle,P.A. and Israel,A. (1990) The DNA binding subunit of NF-κB is identical to factor KBF1 and homolgous to the *rel* oncogene product. Cell, **62**, 1007-1018.

Kim,J.L., Nikolov,D.B. and Burley,S.K. (1993a) Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature, **365**, 520-527.

Kim,Y., Geiger,J.H., Hahn,S. and Sigler,P.B. (1993b) Crystal structure of a yeast TBP / TATA-box complex. Nature, **365**, 512-520.

Kinchington,P.R., Reinhold,W.C., Casey,T.A., Straus,S.E., Hay,J. and Ruyechan,W.T. (1985) Inversion and circularization of the varicella-zoster virus genome. J. Virol., **56**, 194-200.

Kinchington,P.R., Ling,P., Pensiero,M., Moss,B., Ruyechan,W.T. and Hay,J. (1990) The glycoprotein products of varicella-zoster virus gene 14 and their defective accumulation in a vaccine strain (Oka). J. Virol., **64**, 4540-4548.

Kinchington,P.R., Hougland,J.K., Arvin,A.M., Ruyechan,W.T. and Hay,J. (1992) The varicella-zoster virus immediate-early protein IE62 is a major component of virus particles. J. Virol., **66**, 359-366.

Kinoshita,H., Hondo,R., Taguchi,F. and Yogo,Y. (1988) Variation of R1 repeated sequence present in open reading frame 11 of varicella-zoster virus strains. J. Virol., **62**, 1097-1100.

Kissinger,C.R., Liu,B., Martin-Blanco,E., Kornberg,T.B. and Pabo,C.O. (1990) Crystal structure of an engrailed homeodomain-DNA complex at 2.8Å resolution: A framework for understanding homeodomain-DNA interactions. Cell, **63**, 579-590.

Klig,L.S., Crawford,I.P. and Yanofsky,C. (1987) Analysis of *trp* repressor-operator interaction by filter binding. Nucleic Acids Res., **15**, 5339-5351.

Klug,A. (1993) Opening the gateway. Nature, **365**, 486-487.

Knipe,D.M. and Smith,J.L. (1986) A mutant herpesvirus protein leads to a block in nuclear localization of other viral proteins. Mol. Cell. Biol., **6**, 2371-2381.

Knipe,D.M., Senechek,D., Rice,S.A. and Smith,J.L. (1987) Stages in the nuclear association of the herpes simplex virus transcriptional activator protein ICP4. J. Virol., **61**, 276-284.

Kobubo,T., Gong,D.-W., Wootton,J.C., Horikoshi,M., Roeder,R.G. and Nakatani,Y. (1994) Molecular cloning of Drosophila TFIID subunits. Nature, **367**, 484-487.

Konig,P. and Richmond,T.J. (1993) The X-ray structure of the GCN4-bZIP bound to ATF/CREB site DNA shows the complex depends on DNA flexibility. J. Mol. Biol., **233**, 139-154.

Koop,K.E., Duncan,J. and Smiley,J.R. (1993) Binding sites for the herpes simplex virus immediate-early protein ICP4 impose an increased dependence on viral DNA replication on simple model promoters located in the viral genome. J. Virol., **67**, 7254-7263.

Kornberg,T.B. (1993) Understanding the homeodomain. J. Biol. Chem., **268**, 26813-26816.

Krippl,B., Ferguson,B., Rosenberg,M. and Westphal,H. (1985) Functions of purified E1A protein microinjected into mammalian cells. Proc. Natl. Acad. Sci. USA, **81**, 6988-6992.

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

Kristie,T.M. and Roizman,B. (1986a) α4, the major regulatory protein of herpes simplex virus type1, is stably and specifically associated with promoter-regulatory domains of α genes and of selected other viral genes. Proc. Natl. Acad. Sci. USA, **83**, 3218-3222.

Kristie,T.M. and Roizman,B. (1986b) DNA-binding site of major regulatory protein α4 specifically associated with promoter-regulatory domains of α genes of herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA, **83**, 4700-4704.

Kristie,T.M. and Roizman,B. (1987) Host cell proteins bind to the *cis*-acting site required for virion-mediated induction of herpes simplex virus 1 α genes. Proc. Natl. Acad. Sci. USA, **84**, 71-75.

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

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

Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, **227**, 680-685.

LaMarco,K.L., Thompson,C.C., Byers,B.P., Walton,E.M. and McKnight,S.L. (1991) Identification of Ets- and notch-related subunits in GA binding protein. Science, **253**, 789-791.

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

Lang,J.C., Spandidos,D.A. and Wilkie,N.M. (1984) Transcriptional regulation of a herpes simplex virus immediate early gene is mediated through an enhancer-type sequence. EMBO J., **3**, 389-395.

Laughon,A. (1991) DNA binding specificity of homeodomains. Biochem., **30**, 11357-11367.

Laybourn,P.J. and Dahmus,M.E. (1989) Transcription-dependent structural changes in the C-terminal domain of mammalian RNA polymerase IIa/o. J. Biol. Chem., **264**, 6693-6698.

Laybourn,P.J. and Dahmus,M.E. (1990) Phosphorylation of RNA polymerase IIA occurs subsequent to interaction with the promoter and before the initiation of transcription. J. Biol. Chem., **265**, 13165-13173.

Lee,W.S., Kao,C.C., Bryant,G.O., Liu,X. and Berk,A.J. (1991) Adenovirus E1A activation domain binds the basic repeat in the TATA box transcription factor. Cell, **67**, 365-376.

Lee,D.K., DeJong,J., Hashimoto,S., Horikoshi,M. and Roeder,R.G. (1992) TFIIIA induces coformational changes in TFIID via intermediates with the basic repeat. Mol. Cell. Biol., **12**, 5189-5196.

Leinbach,S.S., and Summers,W.C. (1980) The structure of herpes simplex virus type 1 DNA as probed by micrococcal nuclease digestion. J. Gen. Virol., **51**, 45-49.

Leuther,K.K., Salmeron,J.M. and Johnston,S.A. (1993) Genetic evidence that an activation domain of GAL4 does not require acidity and may form a β sheet. Cell, **72**, 575-585.

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

Li,R., Knight,J.D., Jackson,S.P., Tjian,R. and Botchan,M.R. (1991) Direct interaction between Sp1 and the BPV enhancer E2 protein mediates synergistic activation of transcription. Cell, **65**, 493-505.

Li,Y., Flanagan,P.M., Tschochner,H. and Kornberg,R.D. (1994) RNA polymerase II initiation factor interactions and transcription start site selection. Science, **263**, 805-807.

Lieberman,P.M. and Berk,A.J. (1991) The Zta *trans*-activator stabilizes TFIID association with promoter DNA by direct protein-protein interaction. Genes Dev., **5**, 2441-2454.

Lin,Y.-S. and Green,M.R. (1991) Mechanism of action of an acidic transcriptional activator in vitro. Cell, **64**, 971-982.

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

Ling,P., Kinchington,P.R., Ruyechan,W.T. and Hay,J. (1991) A detailed anaysis of transcripts mapping to varicella-zoster virus gene 14 (glycoprotein V). Virology, **184**, 625-635.

Ling,P., Kinchington,P.R., Sadeghi-Zadeh,M., Ruyechan,W.T. and Hay,J. (1992) Transcription from varicella-zoster virus gene 67 (glycoprotein IV). J. Virol., **66**, 3690-3698.

Liu,F. and Green,M.R. (1990) A specific member of the ATF transcription factor family can mediate transcription activation by the adenovirus E1a protein. Cell, **61**, 1217-1224.

Longnecker,R. and Roizman,B. (1986) Generation of an inverting herpes simplex virus 1 mutant lacking the L-S junction *a* sequences, an origin of DNA synthesis, and several genes including those specifying glycoprotein E and the α47 gene. J. Virol., **58**, 583-591.

Looman,A.C., Bodlaender,J., Comstock,L.J., Eaton,D., Jhurani,P., de Boer,H.A. and van Kippenberg,P.H. (1987) Influence of the codon following the AUG initiation codon on the expression of a modified *lacZ* gene in *Escherichia coli*. EMBO J., **6**, 2489-2492.

Lopetegui,P., CampoVera,H. and Yamanishi,K. (1985) Varicella-zoster virus (VZV) specific polypeptides detected in cells treated with metabolic inhibitors. Micobiol. Immunol., **29**, 569-575.

Lowe,R.S., Keller,P.M., Keech,B.J., Davison,A.J., Whang,Y., Morgan,A.J., Kieff,E. and Ellis,R.W. (1987) Varicella-zoster virus as a live vector for the expression of foreign genes. Proc. Natl. Acad. Sci. USA, **84**, 3896-3900.

Lu,H., Flores,O., Weinmann,R. and Reinberg,D. (1991) The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. Proc. Natl. Acad. Sci. USA, **88**, 10004-10008.

Lu,H., Zawel,L., Fisher,L., Egly,J.M. and Reinberg,D. (1992) Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. Nature, **358**, 641-645.

Ludwig,H., Haines,H.G., Biswal,N. and Benyesh-Melnick,M. (1972) The characterization of varicella-zoster virus DNA. J. Gen. Virol., **14**, 111-114.

Luisi,B. (1994) The fittest species of the protein-DNA complexes. In: Frontiers in Molecular Biology. Ed. Lilley,D., Oxford University Press, Oxford. In press.

MacPherson,I. and Stoker,M. (1962) Polyoma transformation of hamster cell clones - an investigation of genetic factors affecting cell competence. Virology, **16**, 147-151.

McCarthy,A.M., McMahan,L. and Schaffer,P.A. (1989) Herpes simpex virus type 1 ICP27 deletion mutants exhibit altered patterns of transcription and are DNA deficient. J. Virol., **63**, 18-27.

McGeoch,D.J. and Davison,A.J. (1986) Alphaherpesviruses possess a gene homologous to the protein kinase gene family of eukaryotes and retroviruses. Nucleic Acids Res., **14**, 1765-1777.

McGeoch,D.J., Dolan,A., Donald,S. and Brauer,D.H.K. (1986) Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. Nucleic Acids Res., **14**, 1727-1764.

McGeoch,D.J., Dalrymple,M.A., Davison,A.J., Dolan,A., Frame,M.C., McNab,D., Perry,L.J., Scott,J.E. and Taylor,P. (1988) The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol., **69**, 1531-1574.

McGeoch,D.J., Barnett,B.C. and MacLean,C.A. (1993) Emerging functions of alphaherpesvirus genes. Semin. Virol., **4**, 125-134.

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

McKnight,S.L. and Kingsbury,R. (1982) Transcriptional control signals of a eukaryotic protein-coding gene. Science, **217**, 316-324.

McKnight,S.L., Gavis,E.R., Kingsbury,R. and Axel,R. (1981) Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identifictaion of an upstream control region. Cell, **25**, 385-398.

McKnight,J.L.C., Kristie,T.M. and Roizman,B. (1987) Binding of the virion protein mediating α gene induction in herpes simplex virus 1-infected cells to its *cis* site requires cellular proteins. Proc. Natl. Acad. Sci. USA, **84**, 7061-7065.

McLauchlan,J. and Rixon,F.J. (1992) Characterization of enveloped tegument structures (L particles) produced by alphaherpesviruses: integrity of the tegument does not depend on the presence of capsid or envelope. J. Gen. Virol., **73**, 269-276.

McLauchlan,J., Simpson,S. and Clements,J.B. (1989) Herpes simplex virus induces a processing factor that stimulates poly(A) site usage. Cell, **59**, 1093-1105.

McLauchlan,J., Phelan,A., Loney,C., Sandri-Goldin,R.M. and Clements,J.B. (1992) Herpes simplex virus IE63 acts at the post transcriptional level to stimulate viral mRNA 3' processing. J. Virol., **66**, 6939-6945.

McMahan,L. and Schaffer,P.A. (1990) The repressing and enhancing functions of the herpes simplex virus regulatory protein ICP27 map to the C-terminal regions and are required to modulate viral gene expression very early in infection. J. Virol., **64**, 3471-3485.

Mader,S., Kumar,V., de Verneuil,H. and Chambon,P. (1989) Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid-responsive element. Nature, **338**, 271-274.

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

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

Maguire,H.F. and Hyman,R.W. (1986) Polyadenylated, cytoplasmic transcripts of varicella-zoster virus. Intervirology, **26**, 181-191.

Marmorstein,R., Carey,M., Ptashne,M. and Harrison,S.C. (1992) DNA recognition by GAL4: Structure of a protein-DNA complex. Nature, **356**, 408-414.

Marsden,H.S., Crombie,I.K. and Subak-Sharpe,J.H. (1976) Control of protein synthesis in herpesvirus-infected cells: Analysis of the polypeptides induced by wild type and sixteen temperature-sensitive mutants of HSV strain 17. J. Gen. Virol., **31**, 347-372.

Marsden,H.S., Lang,J., Davison,A.J., Hope,R.G. and MacDonald,D.M. (1982) Genomic location and lack of phosphorylation of the HSV immediate-early polypeptide IE 12. J. Gen. Virol., **62**, 17-27.

Marsden,H.S., Campbell,M.E.M., Haarr,L., Frame,M.C., Parris,D.S., Murphy,M., Hope,R.G., Muller,M.T. and Preston,C.M. (1987) The 65,000-M_r DNA-binding and virion *trans*-inducing proteins of herpes simplex virus type 1. J. Virol., **61**, 2428-2437.

Martin,K.J., Lillie,J.W. and Green,M.R. (1990a) Evidence for interaction of different eukaryotic transcriptional activators with distinct cellular targets. Nature, **346**, 147-152.

Martin,K.J., Lillie,J.W. and Green,M.R. (1990b) Transcriptional activation by the pseudorabies virus immediate early protein. Genes Dev., **4**, 2376-2382.

Mastrangelo,I.A., Courey,A.J., Wall,J.S., Jackson,S.P. and Hough,P.V.C. (1991) DNA looping and Sp1 multimer links: A mechanism for transcriptional synergism and enhancement. Proc. Natl. Acad. Sci. USA, **88**, 5670-5674.

Mattaj,I., Lienhard,S., Jiricny,J. and De Robertis,E.M. (1985) An enhancer-like sequence within the *xenopus* U2 gene promoter facilitates the formation of stable transcription complexes. Nature, **316**, 163-167.

Matthews,R.E.F. (1982) Classification and nomenclature of viruses: Fourth report of the International Committee of Taxonomy of Viruses. Intervirology, **17**, 1-199.

Maul,G.G. and Everett,R.D. (1994) The nuclear location of PML, a cellular memeber of the C_3HC_4 zinc-binding domain protein family, is rearranged during herpes simplex virus infection by the C_3HC_4 viral protein ICP0. J. Gen. Virol., in press.

Maul,G.G., Guldner,H.H. and Spivack,J.G. (1993) Modification of discrete nuclear domains induced by herpes simplex virus type 1 immediate early gene 1 product (ICP0). J. Gen. Virol., **74**, 2679-2690.

Mavromara-Nazos,P. and Roizman,B. (1987) Activation of herpes simplex virus 1 γ_2 genes by viral DNA replication. Virology, **161**, 593-598.

Mavromara-Nazos,P. and Roizman,B. (1989) Delineation of regulatory domains of early (β) and late (γ_2) genes by construction of chimeric genes exprssed in herpes simplex virus 1 genomes. Proc. Natl. Acad. Sci. USA, **86**, 4071-4075.

Mavromara-Nazos,P., Ackermann,M. and Roizman,B. (1986a) Construction and properties of a viable herpes simplex virus 1 recombinant lacking coding sequences of the α47 gene. J. Virol., **60**, 807-812.

Mavromara-Nazos,P., Silver,S., Hubenthal-Voss,J., McKnight,J.L.C. and Roizman,B. (1986b) Regulation of herpes simplex virus 1 genes: α gene sequence requirements for transient induction of indicator genes regulated by β or late (γ_2) promoters. Virology, **149**, 152-164.

Maxam,A.M. and Gilbert,W. (1980) Sequencing end-labelled DNA with base-specific chemical cleavages. Methods Enzymol., **65**, 499-560.

Meier,J.L. and Straus,S.E. (1992) Comparative biology of latent varicella-zoster virus and herpes simplex virus infections. J. Infect. Dis., **166** (Suppl. 1), 13-23.

Meier,J.L., Holman,R.P., Croen,K.D., Smialek,J.E. and Straus,S.E. (1993) Varicella-zoster virus transcription in human trigeminal ganglia. Virology, **193**, 193-200.

Meisterernst,M. and Roeder,R.G. (1991) Family of proteins that interact with TFIID and regulate promoter activity. Cell, **67**, 557-567.

Meisterernst,M., Roy,A.L., Lieu,H.M. and Roeder,R.G. (1991) Activation of class II gene transcription by regulatory factors is potentiated by a novel activity. Cell, **66**, 981-993.

Meredith,M., Orr,A. and Everett,R. (1994) Herpes simplex virus type 1 immediate-early protein Vmw110 binds strongly and specifically to a 135kD cellular protein. Virology, in press.

Messing,J. (1979) A multipurpose cloning system based on a single stranded DNA bacteriophage M13. Recomb. DNA Technic. Bullet., **2**, 43-48.

Metzler,D.W. and Wilcox,K.W. (1985) Isolation of herpes simplex virus regulatory protein ICP4 as a homodimeric complex. J. Virol., **55**, 329-337.

Michael,N. and Roizman,B. (1989) Binding of the herpes simplex virus major regulatory protein to viral DNA. Proc. Natl. Acad. Sci. USA, **86**, 9808-9812.

Michael,N. and Roizman,B. (1993) Repression of the herpes simplex virus 1 α4 gene by its gene product occurs within the context of the viral genome and is associated with all three identified cognate sites. Proc. Natl. Acad. Sci. USA, **90**, 2286-2290.

Michael,N., Spector,D., Mavromara-Nazos,P., Kristie,T.M. and Roizman,B. (1988) The DNA-binding properties of the major regulatory protein α4 of herpes simplex viruses. Science, **239**, 1531-1534.

Miller,J., McLachlan,A.D. and Klug,A. (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J., **4**, 1609-1614.

Mills,L.K., Shi,Y. and Millette,R.L. (1994) YY1 is the cellular factor shown previously to bind to regulatory regions of several leaky-late ($\beta\gamma$, γ_1) genes of herpes simplex virus type 1. J. Virol., **68**, 1234-1238.

Mitchell,P.J. and Tjian,R. (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science, **245**, 371-378.

Moriuchi,H., Moriuchi,M., Smith,H.A., Straus,S.E. and Cohen,J.I. (1992) Varicella-zoster virus open reading frame 61 is functionally homologous to herpes simplex virus type 1 ICP0. J. Virol., **66**, 7303-7308.

Moriuchi,H., Moriuchi,M., Straus,S.E. and Cohen,J.I. (1993a) 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.

Moriuchi,H., Moriuchi,M., Straus,S.E. and Cohen,J.I. (1993b) Varicella-zoster virus (VZV) open reading frame 61 protein transactivates VZV gene promoter and enhances the infectivity of VZV DNA. J. Virol., **67**, 4290-4295.

Moriuchi,H., Moriuchi,M., Smith,H.M. and Cohen,J.I. (1994) Varicella-zoster virus open reading frame 4 protein is functionally destinct from and does not complement its herpes simplex virus type 1 homolog, ICP27. J. Virol., **68**, 1987-1992.

Muggeridge,M.I. and Fraser,N.W. (1986) Chromosomal organization of the herpes simplex virus genome during acute infection of the mouse central nervous system. J. Virol., **59**, 764-767.

Muller,M.T. (1987) Binding of the herpes simplex virus immediate-early gene product ICP4 to its own transcription start site. J. Virol., **61**, 858-865.

Muller,D., Ugi,I., Ballas,K., Reiser,P., Henning,R. and Montenarh,M. (1987) The AT-rich sequence of the SV40 control region influences the binding of SV40 T antigen to binging sites II and III. Virology, **161**, 81-90.

Murre,C., McCaw,P.S. and Baltimore,D. (1989) A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless, MyoD* and myc proteins. Cell, **56**, 777-783.

Myers,M.G., Duer,H.L. and Hausler,C.K. (1980) Experimental infection of guinea pigs with varicella-zoster virus. J. Infect. Dis., **142**, 414-420.

Nagashima,K., Nakazawa,M. and Endo,H. (1975) Pathology of the human spinal ganglia in varicella-zoster virus infection. Acta Neuropath. (Berl.), **33**, 105-117.

Nagpal,S. and Ostrove,J.M. (1991) Characterization of a potent varicella-zoster virus-encoded *trans*-repressor. J. Virol., **65**, 5289-5296.

Ng,T.I., Keenan,L., Kinchington,P.R. and Grose,C. (1994) Phosphorylation of varicella-zoster virus open reading frame (ORF) 62 regulatory product by viral ORF 47-associated protein kinase. J. Virol., **68**, 1350-1359.

Nikolov,D.B., Hu,S.-H., Lin,J., Gasch,A., Hoffmann,A., Horikoshi,M., Chau,N.-H., Roeder,R.G. and Burley,S.K. (1992) Crystal structure of TFIID TATA-box binding protein. Nature, **360**, 40-46.

Ogata,K., Hojo,H., Aimoto,S., Nakai,T., Nakamura,H., Sarai,A., Ishii,S. and Nishimura,Y. (1992) Solution structure of a DNA-binding unit of Myb: A helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. Proc. Natl. Acad. Sci. USA, **89**, 6428-6432.

Ogino,T., Otsuka,T. and Takahashi,M. (1977) Induction of deoxypyrimidine kinase activity in human embryonic lung cells infected with varicella-zoster virus. J. Virol., **21**, 1232-1235.

O'Hare,P. (1993) The virion transactivator of herpes simplex virus. Semin. Virol., **4**, 145-155.

O'Hare,P. and Goding,C.R. (1988) Herpes simplex virus regulatory elements and the immunoglobulin octamer domain bind a common factor and are both targets for virion transactivation. Cell, **52**, 435-445.

O'Hare,P. and Hayward,G.S. (1984) Expression of recombinant genes containing herpes simplex virus delayed-early and immediate early regulatory regions and *trans* activation by herpesvirus infection. J. Virol., **52**, 522-531.

O'Hare,P. and Hayward,G.S. (1985a) Evidence for a direct role for both the 175,000-and 110,000-molecular-weight immediate-early proteins of herpes simplex virus in the transactivation of delayed-early promoters. J. Virol., **53**, 751-760.

O'Hare,P. and Hayward,G.S. (1985b) Three *trans*-acting regulatory proteins of herpes simplex virus modulate immediate-early gene expression in a pathway involving positive and negative feedback regulation. J. Virol., **56**, 723-733.

O'Hare,P., Goding,C.R. and Haigh,A. (1988) Direct combinatorial interaction between a herpes simplex virus regulatory protein and a cellular octamer-binding factor mediates specific induction of virus immediate-early gene expression. EMBO J., **7**, 4231-4238.

Oliphant,A.R., Brandl,C.J. and Struhl,K. (1989) Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides, analysis of yeast GCN4 protein. Mol. Cell. Biol., **9**, 2944-2949.

Omichinski,J.G., Clore,G.M., Schaad,O., Felsenfeld,G., Trainor,C., Appella,E., Stahl,S.J. and Gronenborn,A.M. (1993) NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. Science, **261**, 438-446.

O'Shea,E.K., Klemm,J.D., Kim,P.S. and Alber.T. (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science, **254**, 539-544.

Ostrove,J.M. (1990) Molecular biology of varicella zoster virus. Advan. Virus Res., **38**, 45-98.

Ostrove,J.M., Reinhold,W., Fan,C.-M., Zorn,S., Hay,J. and Straus,S.E. (1985) Transcriptional mapping of the varicella-zoster virus genome. J. Virol., **56**, 600-606.

Otting,G., Qian,Y.Q., Billeter,M., Muller,M., Affolter,M., Gehring,W.J. and Wuthrich,K. (1990) Protein-DNA contacts in the structure of a homeodomain-DNA complex detemined by nuclear magnetic resonance spectroscopy in solution. EMBO J., **9**, 3085-3092.

Pabo,C.O. and Sauer,R.T. (1992) Transcription factors: Structural families and principles of DNA recognition. Annu. Rev. Biochem., **61**, 1053-1095.

Panning,B. and Smiley,J.R. (1989) Regulation of cellular genes transduced by herpes simplex virus. J. Virol., **63**, 1929-1937.

Papavassiliou,A.G. and Silverstein,S.J. (1990a) Characterization of DNA-protein complex formation in nuclear extracts with a sequence from the herpes simplex virus thymidine kinase gene. J. Biol. Chem., **265**, 1648-1657.

Papavassiliou,A.G. and Silverstein,S.J. (1990b) Interaction of cell and virus proteins with DNA sequences encompassing the promoter / regulatory and leader regions of the herpes simplex virus thymidine kinase gene. J. Biol. Chem., **265**, 9402-9412.

Papavassiliou,A.G., Wilcox,K.W. and Silverstein,S.J. (1991) The interaction of ICP4 with cell/infected-cell factors and its state of phosphorylation modulate differential recognition of leader sequences in herpes simplex virus DNA. EMBO J., **10**, 397-406.

Parvin,J.D. and Sharp,P.A. (1993) DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. Cell, **73**, 533-540.

Paterson,T. and Everett,R.D. (1988a) Mutational dissection of the HSV-1 immediate-early protein Vmw175 involved in transcriptional transactivation and repression. Virology, **166**, 186-196.

Paterson,T. and Everett,R.D. (1988b) The regions of the herpes simplex virus type 1 immediate early protein Vmw175 required for site specific DNA binding closely correspond to those involved in transcriptional regulation. Nucleic Acids Res., **16**, 11005-11025.

Paterson,T. and Everett,R.D. (1990) A prominent serine-rich region in Vmw175, the major transcriptional regulator protein of herpes simplex virus type 1, is not essential for virus growth in tissue culture. J. Gen. Virol., **71**, 1775-1783.

Paterson,T., Preston,V.G. and Everett,R.D. (1990) A mutant of herpes simplex virus type 1 immediate early polypeptide Vmw175 binds to the cap site of its own promoter *in vitro* but fails to autoregulate *in vivo*. J. Gen. Virol., **71**, 851-861.

Payne,J.M., Laybourn,P.J. and Dahmus,M.E. (1989) The transition of RNA polymerase II from initiation to elongation is associated with phosphorylation of the carboxyl-terminal domain of subunit IIa. J. Biol. Chem., **264**, 19621-19629.

Pedersden,M., Talley-Brown,S. and Millette,R.L. (1981) Gene expression of herpes simplex virus. III. Effect of arabinosyladenine on viral polypeptide synthesis. J. Virol., **38**, 712-719.

Pellet,P.E., McKnight,J.L.C., Jenkins,F.J. and Roizman,B. (1985) Nucleotide sequence and predicted amino acid sequence of a protein encoded in a small herpes simplex virus DNA fragment capable of *trans*-inducing α genes. Proc. Natl. Acad. Sci. USA, **82**, 5870-5874.

Pereira,L., Wolff,M.H., Fenwick,M. and Roizman,B. (1977) Regulation of herpesvirus macromolecular synthesis. V Properties of α polypeptides made in HSV-1 and HSV-2 infected cells. Virology, **77**, 733-749.

Perera,L.P., Mosca,J.D., Ruyechan,W.T. and Hay,J. (1992a) Regulation of varicella-zoster virus gene expression in human T lymphocytes. J. Virol., **66**, 5298-5304.

Perera,L.P., Mosca,J.D., Sadeghi-Zadeh,M., Ruyechan,W.T. and Hay,J. (1992b) The varicella-zoster virus immediate early protein, IE62, can positively regulate its cognate promoter. Virology, **191**, 346-354.

Perera,L.P., Mosca,J.D., Ruyechan,W.T., Hayward,G.S., Straus,S.E. and Hay,J. (1993) A major transactivator of varicella-zoster virus, the immediate-early protein IE62, contains a potent N-terminal activation domain. J. Virol., **67**, 4474-4483.

Perisic,O., Xiao,H. and Lis,J.T. (1989) Stable binding of Drosophila heat shock factor to head-to-head and tail-to-tail repeats of a conserved 5bp recognition unit. Cell, **59**, 797-806.

Persson,R.H., Bacchetti,S. and Smiley,J.R. (1985) Cells that constitutively express the herpes simplex virus immediate-early protein ICP4 allow efficient activation of delayed-early genes in *trans*. J. Virol., **54**, 414-421.

Peters,K. and Richards,F.M. (1977) Chemical cross-linking: reagents and problems in studies of membrane structure. Annu. Rev. Biochem., **46**, 523-551.

Peterson,M.G., Tanese,N., Pugh,B.F. and Tjian,R. (1990) Functional domains and upstream activation properties of cloned human TATA binding protein. Science, **248**, 1625-1630.

Peterson,C.L., Kruger,W. and Herskowitz,I. (1991) A functional interaction between the C-terminal domain of RNA polymerase II and the negative regulator SIN1. Cell, **64**, 1135-1143.

Phelan,A., Carmo-Fonseca,M., McLauchlan,J., Lamond,A.I. and Clements,J.B. (1993) A herpes simplex virus type 1 immediate-early gene product, IE63 regulates small nuclear ribonucleoprotein distribution. Proc. Natl. Acad. Sci. USA, **90**, 9056-9060.

Pizer,L.I., Tedder,D.G., Betz,J.L., Wilcox,K.W. and Beard,P. (1986) Regulation of transcription *in vitro* from herpes simplex virus genes, J. Virol., **60**, 950-959.

Pizer,L.I., Everett,R.D., Tedder,D.G., Elliott,M. and Litman,B. (1991) Nucleotides within both proximal and distal parts of the consensus sequence are important for specific DNA recognition by the herpes simplex virus regulatory protein ICP4. Nucleic Acids Res., **19**, 477-483.

Post,L.E. and Roizman,B. (1981) A generalized technique for deletion of specific genes in large genomes: α gene 22 of herpes simplx virus 1 is not essential for growth. Cell, **25**, 227-232.

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

Powell,K.L. and Purifoy,D.J.M. (1976) DNA-binding proteins of cells infected by herpes simplex virus type 1 and 2. Intervirology, **7**, 225-239.

Powell,K.L., Purifoy,D.J.M. and Courtney,R.J. (1975) The synthesis of herpes simplex virus proteins in the absence of virus DNA synthesis. Biochem. Biophys. Res. Comm., **66**, 262-271.

Preston,C.M. (1979a) Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant *ts*K. J. Virol., **29**, 275-285.

Preston,C.M. (1979b) Abnormal properties of an immediate early polypeptide in cells infected with the herpes simplex virus type 1 mutant *ts*K. J. Virol., **32**, 357-369.

Preston,V.G. (1981) Fine-structure mapping of herpes simplex virus type 1 temperature-sensitive mutations within the short repeat of the genome. J. Virol., **39**, 150-161.

Preston,C.M. and Notarianni,E.L. (1983) Poly(ADP-ribosyl)ation of a herpes simplex virus immediate early polypeptide. Virology, **131**, 492-501.

Preston,C.M. and Tannahill,D. (1984) Effects of orientation and position on the activity of a herpes simplex virus immediate early gene far-upstream region. Virology, **137**, 439-444.

Preston,C.M., Cordingley,M.G. and Stow,N.D. (1984) Analysis of DNA sequences which regulate the transcription of a herpes simplex virus immediate early gene. J. Virol., **50**, 708-716.

Preston,C.M., Frame,M.C. and 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.

Provost,P.J., Keller,P.M., Banker,F.S., Keech,B.J., Klein,H.J., Lowe,R.S., Morton,D.H., Phelps,A.H., McAleer,W.J. and Ellis,R.W. (1987) Successful infection of the common marmoset (*callithrix jacchus*) with human varicella-zoster virus. J. Virol., **61**, 2951-2955.

Pruijn,G.J.M., van Driel,W. and van der Vliet,P.C. (1986) Nuclear factor III, a novel sequence-specific DNA-binding protein from HeLa cells stimulating adenovirus DNA replication. Nature, **322**, 656-659.

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

Ptashne,M. (1988) How eukaryotic transcriptional activators work. Nature, **335**, 683-689.

Pugh,B.F. and Tjian,R. (1990) Mechanism of transcriptional activation by Sp1: evidence for coactivators. Cell, **61**, 1187-1197.

Pugh,B.F. and Tjian,R. (1991) Transcription from a TATA-less promoter requires a multisubunit TFIID complex. Genes Dev., **5**, 1935-1945.

Pugh,B.F. and Tjian,R. (1992) Diverse transcriptional functions of the multisubunit eukaryotic TFIID complex. J. Biol. Chem., **267**, 679-682.

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

Qian,Y.Q., Billeter,M., Otting,G., Muller,M., Gehring,W.J. and Wuthrich,K. (1989) The structure of the *Antennapedia* homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. Cell, **59**, 573-580.

Qian,X., Jeon,C.J., Yoon,H.S., Agarwal,K. and Weiss,M.A. (1993) Structure of a new nucleic-acid-binding motif in eukaryotic transcriptional elongation factor TFIIS. Nature, **365**, 277-279.

Quinlan,M.P. and Knipe,D.M. (1985) Stimulation of expression of a herpes simplex virus DNA-binding protein by two viral functions. Mol. Cell. Biol., **5**, 957-963.

Quinlan,M.P., Chen,L.B. and Knipe,D.M. (1984) The intranuclear location of a herpes simplex virus DNA-binding protein is determined by the status of viral DNA replication. Cell, **36**, 857-868.

Ramakrishnan,V., Finch,J.T., Graziano,V., Lee,P.L. and Sweet,R.M. (1993) Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. Nature, **362**, 219-223.

Randall,R.E. and Dinwoodie,N. (1986) Intranuclear localization of herpes simplex virus immediate-early and delayed-early proteins: Evidence that ICP4 is associated with progeny virus DNA. J. Gen. Virol., **67**, 2163-2177.

Raumann,B.E., Brown,B.M. and Sauer,R.T. (1994) Major groove DNA recognition by β-sheets: the ribbon-helix-helix family of gene regulatory proteins. Curr. Opin. Struct. Biol., **4**, 36-43.

Reinberg,D., Horikoshi,M. and Roeder,R.G. (1987) Factors involved in specific transcription in mammalian RNA polymerase II: Functional analysis of initiation factors IIA and IID and identification of a new factor operating at sequences downstream of the initiation site. J. Biol. Chem., **262**, 3322-3330.

Reinhold,W.C., Straus,S.E. and Ostrove,J.M. (1988) Directionality and further mapping of varicella zoster virus transcripts. Virus Res., **9**, 249-261.

Renkawitz,R. (1990) Transcriptional repression in eukaryotes. Trends Genet., **6**, 192-197.

Resnick,J., Boyd,B.A. and Haffey,M.L. (1989) DNA binding by herpes simplex type 1 ICP4 protein is necessary for efficient down regulation of the ICP0 promoter. J. Virol., **63**, 2497-2503.

Rice,S.A. and Knipe,D.M. (1988) Gene-specific transactivation by herpes simplex virus type 1 alpha protein ICP27. J. Virol., **62**, 3814-3823.

Rice,S.A. and Knipe,D.M. (1990) Genetic evidence for two distinct transactivation functions of the herpes simplex virus α protein ICP27. J. Virol., **64**, 1704-1715.

Rice,S.A., Su,L. and Knipe,D.M. (1989) Herpes simplex virus alpha protein ICP27 possesses separable positive and negative regulatory activities. J. Virol., **63**, 3399-3407.

Rigby,R.W.J. (1993) Three in one and one in three: it all depends on TBP. Cell, **72**, 7-10.

Rixon,F.J. (1993) Structure and assembly of herpesviruses. Semin. Virol., **4**, 135-144.

Rixon,F.J., Campbell,M.E. and Clements,J.B. (1982) The immediate-early mRNA that encodes the regulatory polypeptide Vmw175 of herpes simplex virus type 1 is unspliced. EMBO J., **1**, 1273-1277.

Roberts,M.S., Boundy,A., O'Hare,P., Pizzorno,M.C., Ciufo,D.M. and Hayward,G.S. (1988) Direct correlation between a negative autoregulatory response element at the cap site of the herpes simplex virus type 1 IE175 (α4) promoter and a specific binding site for the IE175 (ICP4) protein. J. Virol., **62**, 4307-4320.

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

Robertson,A.T., Caughman,G.B., Gray,W.L., Baumann,R.P., Staczek,J. and O'Callaghan,D.J. (1988) Analysis of the *in vitro* translation products of the equine herpesvirus type 1 immediate early mRNA. Virology, **166**, 451-462.

Roizman,B. (1982) The family herpesviridae: general description, taxonomy, and classification. In: The Herpesviruses. Vol **1**, 1-23, Ed. Roizman,B., Plenum press, New York.

Roizman,B. and Furlong,D. (1974) The replication of herpesviruses. In: Comprehensive Virology. Vol **3**, 229-403, Eds. Fraenkel-Conrat,H., Wagner,R.R., Plenum Press, New York.

Roizman,B. and Sears,A.E. (1993) Herpes simpex viruses and their replication. In: The Human Herpesviruses. pp11-68, Eds. Roizman,B., Whitley,R.J. and Lopez,C., Raven Press Ltd., New York.

Roizman,B., Carmichael,L.E., Deinhardt,F., de-The,G., Nahmias,A.J., Plowright,W., Rapp,F., Sheldrick,P., Takahashi,M. and Wolf,K. (1981) Herpesviridae. Definition, provisional nomenclature and taxonomy. Intervirology, **16**, 201-217.

Roizman,B., Desrosiers,R.C., Fleckenstein,B., Lopez,C., Minson,A.C. and Studdert,M.J. (1992) The family *Herpesviridae*: an update. Arch. Virol., **123**, 425-449.

Romanelli,M.G., Mavromara-Nazos,P., Spector,D. and Roizman,B. (1992) Mutational analysis of the ICP4 binding sites in the 5' transcribed noncoding domains of the herpes simplex virus 1 UL49.5 γ_2 gene. J. Virol., **66**, 4855-4863.

Ruppert,S., Wang,E.H. and Tjian,R. (1993) Cloning and expression of human TAF250: A TBP-associated factor implicated in cell cycle regulation. Nature, **362**, 175-179.

Russell,J., Stow,N.D., Stow,E.C. and Preston,C.M. (1987) Herpes simplex virus genes involved in latency *in vitro*. J. Gen. Virol., **68**, 3009-3018.

Ruvkun,G. and Finney,M. (1991) Regulation of transcription and cell identity by POU domain proteins. Cell, **64**, 475-478.

Ruyechan,W.T., Ling,P., Kinchington,P.R. and Hay,J. (1991) The correlation between varicella-zoster virus transcription and the sequence of the viral genome. In: Herpesvirus transcription and its regulation. pp301-316, Ed. Wagner,E.K., LRC Press, Boca Raton, USA.

Sabella,C., Lowry,P.W., Abbruzzi,G.M., Koropchak,C.M., Kinchington,P.R., SadeghZadeh,M., Hay,J., Ruyechan,W.T. and Arvin,A.M. (1993) Immunization with the immediate-early tegument protein (open reading frame 62) of varicella-zoster virus protects guinea pigs against virus challenge. J. Virol., **67**, 7673-7676.

Sacks,W.R. and Schaffer,P.A. (1987) Deletion mutants in the gene encoding the herpes simplex virus type 1 imediate-early protein ICP0 exhibit impaired growth in cell culture. J. Virol., **61**, 829-839.

Sacks,W.R., Greene,C.C., Aschman,D.P. and Schaffer,P.A. (1985) Herpes simplex virus type 1 ICP27 is an essential regulatory protein. J. Virol., **55**, 796-805.

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

Saltzmann,A.G. and Weinmann,R. (1989) Promoter specificity and modulation of RNA polymerase II transcription. FASEB J., **3**, 1723-1733.

Sandri-Goldin,R.M. and Mendoza,G.E. (1992) A herpesvirus regulatory protein appears to act post-transcriptionally by affecting mRNA processing. Genes Dev., **6**, 848-863.

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

Sauer,R.T., Smith,D.L. and Johnson,A.D. (1988) Flexibility of the yeast α2 repressor enables it to occupy the ends of its operator, leaving the centre free. Genes Dev., **2**, 807-816.

Sawadogo,M. and Roeder,R.G. (1985) Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell, **43**, 165-175.

Sawyer,M.H., Ostrove,J.M., Felser,J.M. and Straus,S.E. (1986) Mapping of the varicella zoster virus deoxypyrimidine kinase gene and preliminary identification of its transcript. Virology, **149**, 1-9.

Sayers,J.R., Schmidt,W., Wendler,A. and Eckstein,F. (1988) Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide. Nucleic Acids Res., **16**, 803-814.

Scafe,C., Chao,D., Lopes,J., Hirsch,J.P., Henry,S. and Young,R.A. (1990) RNA polymerase II C-terminal repeat influences response to transcriptional enhancer signals. Nature, **347**, 491-494.

Schmidt,N.J. and Lennette,E.H. (1976) Improved yields of cell-free varicella-zoster virus. Infect. Immun., **14**, 709-715.

Schmidt,M.C., Kao,C.C., Pei,R. and Berk,A.J. (1989) Yeast TATA box transcription factor gene. Proc. Natl. Acad. Sci. USA, **86**, 7785-7789.

Schmiedeskamp,M. and Kevit,R.E. (1994) Zinc finger diversity. Curr. Opin. Struct. Biol., **4**, 28-35.

Schroder,C.H., DeZazzo,J., Knopf,K.W., Kaerner,H.C., Levine,M. and Glorioso,J. (1985) A herpes simplex virus type 1 mutant with a deletion in the polypeptide-coding sequences of the ICP4 gene. J. Gen. Virol., **66**, 1589-1593.

Schwyzer,M., Vlcek,C., Menekse,O., Fraefel,C. and Paces,V. (1993) Promoter, spliced leader, and coding sequence for BICP4, the largest of the immediate-early proteins of bovine herpesvirus 1. Virology, **197**, 349-357.

Scott,M.P., Tamkun,J.W. and Hartzell,G.W.III (1989) The structure and function of the homeodomain. Biochim. Biophys. Acta, **989**, 25-48.

Seal,B.S., Martinez,J.D., Hall,M.R. and St.Joer,S.C. (1988) Occurance of bovine herpesvirus-1 DNA in nucleosomes and chromatin of bovine herpes virus-1-infected cells: identification of a virion-associated protein in chromatin of infected cells. Arch. Virol., **99**, 221-236.

Sears,A.E., Halliburton,I.W., Meignier,B., Silver,S. and Roizman,B. (1985) Herpes simplex virus 1 mutant deleted in the α22 gene: growth and gene expression in permissive and restrictive cells and establishment of latency in mice. J. Virol., **55**, 338-346.

Sekulovich,R.E., Leary,K. and Sandri-Goldin,R.M. (1988) The herpes simplex virus type 1 α protein ICP27 can act as a *trans*-repressor or a *trans*-activator in combination with ICP4 and ICP0. J. Virol., **62**, 4510-4522.

Seto,E., Shi,Y. and Shenk,T. (1991) YY1 is an initiator sequence-binding protein that directs and activates transcription *in vitro*. Nature, **354**, 241-245.

Shapira,M., Homa,F.L., Glorioso,J.C. and Levine,M. (1987) Regulation of the herpes simplex virus type 1 late (γ) glycoprotein C gene: sequences between base pairs -34 to +29 control transient expression and responsiveness to transactivation by the products of the immediate early (α) 4 and 0 genes. Nucleic Acids Res., **15**, 3097-3111.

Sharp,P.A. (1992) TATA-binding protein is a classless factor. Cell, **68**, 819-821.

Shepard,A.A. and DeLuca,N.A. (1989) Intragenic complementation among partial peptides of herpes simplex virus regulatory protein ICP4. J. Virol., **63**, 1203-1211.

Shepard,A.A. and DeLuca,N.A. (1991a) Activities of heterodimers composed of DNA-binding and transactivation-deficient subunits of the herpes simplex virus regulatory protein ICP4. J. Virol., **65**, 299-307.

Shepard,A.A. and DeLuca,N.A. (1991b) A second-site revertant of a defective herpes simplex virus ICP4 protein with restored regulatory activities and impaired DNA-binding properties. J. Virol., **65**, 787-795.

Shepard,A.A., Imbalzano,A.N. and DeLuca,N.A. (1989) Separation of primary structural components conferring autoregulation, transactivation, and DNA-binding properties to the herpes simplex virus transcriptional regulatory protein ICP4. J. Virol., **63**, 3714-3728.

Shepard,A.A., Tolentino,P. and DeLuca,N.A. (1990) *Trans*-dominant inhibition of herpes simplex virus transcriptional regulatory protein ICP4 by heterodimer formation. J. Virol., **64**, 3916-3926.

Shiraki,K. and Hyman,R.W. (1987) The immediate early proteins of varicella-zoster virus. Virology, **156**, 423-426.

Shiraki,K., Hayakawa,Y., Mori,H., Namazue,J., Takamizawa,A., Yoshida,I., Yamanishi,K. and Takahashi,M. (1991) Development of immunogenic recombinant Oka varicella vaccine expressing hepatitis B virus surface antigen. J. Gen. Virol., **72**, 1393-1399.

Silver,S. and Roizman,B.R. (1985) γ_2-thymidine kinase chimeras are identically transcribed but regulated as γ_2 genes in herpes simplex virus genomes and as β genes in cell genomes. Mol. Cell. Biol., **5**, 518-528.

Smiley,J.R. and Duncan,J. (1992) The herpes simplex virus type 1 immediate-early polypeptide ICP4 is required for expression of globin genes located in the viral genome. Virology, **190**, 538-541.

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

Smiley,J.R., Johnson,D.C., Pizer,L.I. and Everett,R.D. (1992) The ICP4 binding sites in the herpes simplex virus type 1 glycoprotein D (gD) promoter are not essential for efficient gD transcription during virus infection. J. Virol., **66**, 623-631.

Smith,C.A. and DeLuca,N.A. (1992) Transdominant inhibition of herpes simplex virus growth in transgenic mice. Virology, **191**, 581-588.

Smith,I.L. and Sandri-Goldin,R.M. (1988) Evidence that transcriptional control is the major mechanism of regulation for the glycoprotein D gene in herpes simplex virus type 1-infected cells. J. Virol., **62**, 1474-1477.

Smith,C.A. and Schaffer,P.A. (1987) Intertypic recombinants of herpes simplex virus types 1 and 2 infected-cell polypeptide 4. Virology, **160**, 176-182.

Smith,R.F. and Smith,T.F. (1989) Identification of new protein kinase-related genes in three herpesviruses, herpes simpex virus, varicella-zoster virus, and Epstein-Barr virus. J. Virol., **63**, 450-455.

Smith,R.H., Caughman,G.B. and O'Callaghan,D.J. (1992a) Characterization of the regulatory functions of the equine herpesvirus 1 immediate-early gene product. J. Virol., **66**, 936-945.

Smith,I.L., Hardwicke,M.A and Sandri-Goldin,R.M. (1992b) Evidence that the herpes simplex virus immediate early protein ICP27 acts post-transcriptionally during infection to regulate gene expression. Virology, **186**, 74-86.

Smith,C.A., Bates,P., Rivera-Gonzalez,R., Gu,B. and DeLuca,N. (1993) ICP4, the major transcriptional regulatory protein of herpes simplex virus type 1, forms a tripartite complex with TATA-binding protein and TFIIB. J. Virol., **67**, 4676-4687.

Somers,W.S. and Phillips,S.E.V. (1992) Crystal structure of the *met* repressor-operator complex at 2.8 Å resolution reveals DNA recognition by β-strands. Nature, **359**, 387-393.

Spear,P.G. and Roizman,B. (1972) Proteins specified by herpes simplex virus. V. Purification and structural proteins of the herpesvirion. J. Virol., **9**, 143-159.

Spector,D., Purves,F. and Roizman,B. (1990) Mutational analysis of the promoter region of the α27 gene of herpes simplex virus 1 within the context of the viral genome. Proc. Natl. Acad. Sci. USA, **87**, 5268-5272.

Stannard,L.M., Fuller,A.O. and Spear,P.G. (1987) Herpes simplex virus glycoproteins associated with different morphological entities projecting from the virion envelope. J. Gen. Virol., **68**, 715-725.

Steffy,K.R. and Weir,J.P. (1991) Mutational analysis of two herpes simplex virus type 1 late promoters. J. Virol., **65**, 6454-6460.

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

Stern,S., Tanaka,M. and Herr,W. (1989) The Oct-1 homeodomain directs formation of a multiprotein-DNA complex with the HSV transactivator VP16. Nature, **341**, 624-630.

Stevens,J.G.(1989) Human herpesviruses: a consideration of the latent state. Microbiol. Rev., **53**, 318-332.

Stevenson,D., Colman,K.L. and Davison,A.J. (1992) Characterization of the varicella-zoster virus gene 61 protein. J. Gen. Virol., **73**, 521-530.

Stow,N.D. and Davison,A.J. (1986) Identification of a varicella-zoster virus origin of DNA replication and its activation by herpes simplex virus type 1 gene products. J. Gen. Virol., **67**, 1613-1623.

Stow,N.D. and Stow,E.C. (1986) Isolation and characterization of a herpes simplex virus type 1 mutant containing a deletion within the gene encoding the immediate-early polypeptide Vmw110. J. Gen. Virol., **67**, 2571-2585.

Stow,N.D., Murray,M.D. and Stow,E.C. (1986) *Cis*-acting signals involved in the replication and packaging of herpes simplex virus type 1 DNA. In: CANCER CELLS **4** / DNA tumor viruses. pp497-507, Eds. Botchan,M., Grodzicker,T. and Sharp,P.A., Cold Spring Harbor Lab., New York.

Straus,S.E., Aulakh,H.S., Ruyechan,W.T., Hay,J., Casey,T.A., Vande Woude,G.F., Owens,J. and Smith,H.A. (1981) Structure of varicella-zoster virus DNA. J. Virol., **40**, 516-525.

Straus,S.E., Owens,J., Ruyechan,W.T., Takiff,H.E., Casey,T.A., Vande Woude,G.F. and Hay,J. (1982) Molecular cloning and physical mapping of varicella-zoster virus DNA. Proc. Natl. Acad. Sci. USA, **79**, 993-997.

Straus,S.E., Hay,J., Smith,H. and Owens,J. (1983) Genome differences among varicella-zoster virus isolates. J. Gen. Virol., **64**, 1031-1041.

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

Struhl,K. (1994) Duality of TBP, the universal transcription factor. Science, **263**, 1103-1104.

Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol., **185**, 60-89.

Su,L. and Knipe,D.M. (1989) Herpes simplex virus α protein ICP27 can inhibit or augment viral gene transactivation. Virology, **170**, 496-504.

Su,W., Jackson,S., Tjian,R. and Echols,H. (1991) DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. Genes Dev., **5**, 820-826.

Suzuki,M. (1990) The heptad repeat in the largest subunit of RNA polymerase II binds by intercalating into DNA. Nature, **344**, 562-565.

Szilagyi,J.F. and Cunningham,C. (1991) Identification and characterization of a novel non-infectious herpes simplex virus-related particle. J. Gen. Virol., **72**, 661-668.

Suzuki,M. (1993) Common features in DNA recognition helices of eukaryotic transcription factors. EMBO J., **12**, 3221-3226.

Tackney,C., Cachianes,G. and Silverstein,S. (1984) Transduction of a chinese hamster ovary *aprt* gene by herpes simplex virus. J. Virol., **52**, 606-614.

Tanaka,I., Appelt,K., Dijk,J., White,S.W. and Wilson,K.S. (1984) 3-Å resolution structure of a protein with histone-like properties in prokaryotes. Nature, **310**, 376-381.

Tanese,N., Pugh,B.F. and Tjian,R. (1991) Coactivators for a proline-rich activator purified from the multisubunit human TFIID. Genes Dev., **5**, 2212-2224.

Tasset,D., Tora,L., Fromental,C., Scheer,E. and Chambon,P. (1990) Distinct classes of transcriptional activating domains function by different mechanisms. Cell, **62**, 1177-1187.

Tedder,D.G. and Pizer,L.I. (1988) Role for DNA-protein interactions in activation of the herpes simplex virus glycoprotein D gene. J. Virol., **62**, 4661-4672.

Tedder,D.G., Everett,R.D., Wilcox,K.W., Beard,P. and Pizer,L.I. (1989) ICP4-binding sites in the promoter and coding regions of the herpes simplex virus gD gene contribute to activation of *in vitro* transcription by ICP4. J. Virol., **63**, 2510-2520.

Telford,E.A.R., Watson,M.S., McBride,K. and Davison,A.J. (1992) The DNA sequence of equine herpesvirus-1. Virology, **189**, 304-316.

Thali,M., Rusconi,S. and Schaffner,W. (1990) Immediate early protein of pseudorabies virus is a general transactivator but stimulates only suboptimally utilized promoters. A clue to specificity? J. Mol. Biol., **215**, 301-311.

Thompson,R., Honess,R.W., Taylor,L., Morran,J. and Davison,A.J. (1987) Varicella-zoster virus specifies a thymidylate synthetase. J. Gen. Virol., **68**, 1449-1455.

Timmers,H.Th.M. and Sharp,P.A. (1991) The mammalian TFIID protein is present in two functionally distinct complexes. Genes Dev., **5**, 1946-1956.

Timmers,H.Th.M., Meyers,R.E., and Sharp,P.A. (1992) Composition of transcription factor B-TFIIB. Proc. Natl. Acad. Sci. USA, **89**, 8140-8144.

Tjian,R. (1981) T antigen binding and the control of SV40 gene expression. Cell, **26**, 1-2.

Towbin,H., Staehelin,T. and Gordon,J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA., **76**, 4350-4354.

Treisman,J., Gonczy,P., Vashishtha,M., Harris,E. and Desplan,C. (1989) A single amino acid can determine the DNA binding secificity of homeodomain proteins. Cell, **59**, 553-562.

Treisman,J., Harris,E. and Desplan,C. (1991) The paired box encodes a second DNA-paired domain in the paired homeo domain protein. Genes Dev., **5**, 594-604.

Tremblay,M.L., Yee,S.-P., Persson,R.H., Bacchetti,S., Smiley,J.R. and Branton,P.E. (1985) Activation and inhibition of expression of the 72,000-Da early protein of adenovirus type 5 in mouse cells constitutively expressing an immediate early protein of herpes simplex virus type 1. Virology, **144**, 35-45.

Triezenberg,S.J., LaMarco,K.L. and McKnight,S.L. (1988) Evidence of DNA:protein interactions that mediate HSV-1 immediate early gene activation by VP16. Genes Dev., **2**, 730-742.

Tsukiyama,T., Becker,P.B. and Wu,C. (1994) ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. Nature, **367**, 525-532.

Tyler,J.K., Orr,A. and Everett,R.D. (1994) A hybrid HSV-1 ICP4 protein, with its DNA binding domain replaced by that of the VZV gene 62 transactivator, is functional during virus infection. Manuscript in preparation.

Tyree,C.M., George,C.P., Lira-DeVito,L.M., Wampler,S.L., Dahmus,M.E., Zawel,L. and Kadonaga,J.T. (1993) A minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymerase II. Genes Dev., **7**, 1254-1265.

Usheva,A., Maldonado,E., Goldring,A., Lu,H., Houbavi,C., Reinberg,D. and Aloni,Y. (1992) Specific interaction between the nonphosphorylated from of RNA polymerase II and the TATA-binding protein. Cell, **69**, 871-881.

Vafai,A., Wroblewska,Z., Wellish,M., Green,M. and Gilden,D. (1984) Analysis of three late varicella-zoster virus proteins, a 125,000 molecular weight protein and gp1 and gp3. J. Virol., **52**, 953-959.

Van Dyke,M.W., Roeder,R.G. and Sawadogo,M. (1988) Physical analysis of transcription preinitiation complex assembly on a class II gene promoter. Science, **241**, 1335-1338.

Van Hoy,M., Leuther,K.K., Kodadek,T. and 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.

van Zijl,M., Quint,W., Briaire,J., de Rover,T., Gielkens,A. and Berns,A. (1988) Regeneration of herpesviruses from molecularly cloned subgenomic fragments. J. Virol., **62**, 2191-2195.

Verrijzer,C.P. and van der Vliet,P.C. (1993) POU domain transcription factors. Biochim. Biophys. Acta., **1173**, 1-21.

Verrijzer,C.P., Kal,A.J. and van der Vliet,P.C. (1990) The Oct-1 homeo domain contacts only part of the octamer sequence and full oct-1 DNA binding activity requires the POU-specific domain. Genes Dev., **4**, 1964-1974.

Verrijzer,C.P., Alkema,M.J., van Weperen,W.W., van Leeuwen,H.C., Strating,M.J.J. and van der Vliet,P.C. (1992) The DNA binding specificity of the bipartitie POU domain and its subdomains. EMBO J., 11, 4493-5003.

Vlcek,C., Paces,V. and Schwyzer, M. (1989) Nucleotide sequence of the pseudorabies virus immediate early gene, encoding a strong trans-activator protein. Virus Genes, **2**, 335-346.

Vogelstein,B. and Gillespie, D. (1979) Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. USA, **76**, 615-619.

Wagner,E.K. (1985) Individual HSV transcripts: Characterization of specific genes. In: The Herpesviruses. **3**, 45-104, Ed. Roizman,B., Plenum Press, New York.

Wang,E.H. and Tjian,R. (1994) Promoter-selective transcriptional defect in cell cycle mutant ts13 rescued by $hTAF_{II}250$. Science, **263**, 811-814.

Wang,W., Gralla,J.D. and Carey,M. (1992a) The acidic activator GAL4-AH can stimulate polymerase II transcription by promoting assembly of a closed complex requiring TFIID and TFIIA. Genes Dev., **6**, 1716-1727.

Wang,W., Carey,M. and Gralla,J.D. (1992b) Polymerase II promoter activation: closed complex formation and ATP-driven start site opening. Science, **255**, 450-453.

Ward,R.L. and Stevens,J.G. (1975) Effect of cytosine arabinoside on viral-specific protein synthesis in cells infected with herpes simplex virus. J. Virol., **15**, 71-80.

Watson,R.J. and Clements,J.B. (1978) Characterization of transcription-deficient temperature-sensitive mutants of herpes simplex virus type 1. Virology, **91**, 364-379.

Watson,R.J. and Clements,J.B. (1980) A herpes simplex virus type 1 function continuously required for early and late virus RNA synthesis. Nature, **285**, 329-330.

Weinheimer,S.P. and McKnight,S.L. (1987) Transcriptional and post-transcriptional controls establish the cascade of herpes simplex virus protein synthesis. J. Mol. Biol., **195**, 819-833.

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

Weintraub,H. (1988) Formation of stable transcription complexes as assayed by analysis of individual templates. Proc. Natl. Acad. Sci. USA, **85**, 5819-5823.

Weir,H.M., Kraulis,P.J., Hill,C.S., Raine,A.R.C., Laue,E.D. and Thomas,J.O. (1993) Structure of the HMG box motif in the B-domain of HMG1. EMBO J., **12**, 1311.

Weiss,M.A., Ellenberger,T., Wobbe,C.R., Lee,J.P., Harrison,S.C. and Struhl,K. (1990) Folding transition in the DNA-binding domain of GCN4 on specific binding to DNA. Nature, **347**, 575-578.

Weller,T.H. (1953) Serial propagation *in vitro* of agents producing inclusion bodies derived from varicella and herpes zoster. Proc. Soc. Experim. Biol. Med., **83**, 340-346.

Weller,T.H. and Witton,H.M. (1958) The etiologic agents of varicella and herpes zoster. Serologic studies with the viruses as propagated in vitro. J. Exp. Med., **108**, 869-890.

Weller,T.H., Witton,H.M. and Bell,E.J. (1958) The etiological agents of varicella and herpes zoster. Isolation, propagation, and cultural characteristics in vitro. J. Exp. Med., **108**, 843-863.

White,R.J. and Jackson,S.P. (1992) The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III. Trends Genet., **8**, 284-288.

White,J.H., Brou,C., Wu,J., Burton,N., Egly,J.-M. and Chambon,P. (1991) Evidence for a factor required for transcriptional stimulation by the chimeric acidic activator GAL4-VP16 in HeLa cell extracts. Proc. Natl. Acad. Sci. USA, **88**, 7674-7678.

Whitley,R.J. (1992) Therapeutic approaches to varicella-zoster virus infections. J. Infect. Dis., **166** (Suppl. 1), 51-57.

Wilcox,K.W., Kohn,A., Sklyanskaya,E. and Roizman,B. (1980) Herpes simplex virus phosphoproteins. I. Phosphate cycles on and off some viral polypeptides and can alter their affinity for DNA. J. Virol., **33**, 167-182.

Wildy,P. and Watson,D.H.(1962) Electron microscopic studies on the architecture of animal viruses. Cold Spring Harbor Symp. Quant. Biol., **27**, 25-47.

Wildy,P., Russell,W.C. and Horne,R.W. (1960) The morphology of herpes virus. Virology, **12**, 204-222.

Wilson,A.C., LaMarco,K., Peterson,M.G. and Herr,W. (1993) The VP16 accessory protein HCF is a family of polypeptides processed from a large precursor protein. Cell, **74**, 115-125.

Winkler,F.K., Banner,D.W., Oefner,C., Tsernoglou,D., Brown,R.S., Heathman,S.P., Bryan,R.K., Martin,P.D., Petratos,K. and Wilson,K.S. (1993) The crystal structure of EcoRV enconuclease and of its complexes with cognate and non-cognate DNA fragments. EMBO J., **12**, 1781-1795.

Wirth,U.V., Vogt,B. and Schwyzer,M. (1991) The three immediate-early transcripts of bovine herpesvirus 1 arise from two divergent and spliced transcription units. J. Virol., **65**, 195-205.

Wolberger,C., Vershon,A.K., Liu,B., Johnson,A.D. and Pabo,C.O. (1991) Crystal structure of a Matα2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. Cell, **67**, 517-528.

Workman,J.L. and Buchman,A.R. (1993) Multiple functions of nucleosomes and regulatory factors in transcription. Trends. Biochem. Sci., **18**, 90-95.

Workman,J.L., Abmayr,S.M., Cromlish,W.A. and Roeder,R.G. (1988) Transcriptional regulation by the immediate early protein of pseudorabies virus during in vitro nucleosome assembly. Cell, **55**, 211-219.

Wright,P.E. (1994) POU domains and homeodomains. Curr. Opin. Struct. Biol., **4**, 22-27.

Wu,C.-L. and Wilcox,K.W. (1990) Codons 262 to 490 from the herpes simpex virus ICP4 gene are sufficient to encode a sequence-specific DNA binding protein. Nucleic Acids Res., **18**, 531-538.

Wu,C.-L. and Wilcox,K.W. (1991) The conserved DNA-binding domains encoded by the herpes simplex virus type 1 ICP4, pseudorabies virus IE180, and varicella-zoster virus ORF62 genes recognize similar sites in the corresponding promoters. J. Virol., **65**, 1149-1159.

Wu,C.A., Nelson,N.J., McGeoch,D.J. and Challberg,M.D. (1988) Identification of herpes simplex virus type 1 genes required for origin-dependent DNA synthesis. J. Virol., **62**, 435-443.

Xiao,P. and 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.

Yang,C.-C. and Nash,H.W. (1989) The interaction of E. coli IHF protein with its specific binding sites. Cell, **57**, 869-880.

Yao,F. and Courtney,R.J. (1989) A major transcriptional regulatory protein (ICP4) of herpes simplex virus type 1 is associated with purified virions. J. Virol., **63**, 3338-3344.

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

Zawel,L. and Reinberg,D. (1992) Advances in RNA polymerase II transcription. Curr. Opin. Cell. Biol., **4**, 488-495.

Zawel,L. and Reinberg,D. (1993) Initiation of transcription by RNA polymerase II: A multi-step process. Prog. Nucl. Acid Res. Mol. Biol., **44**, 67-108.

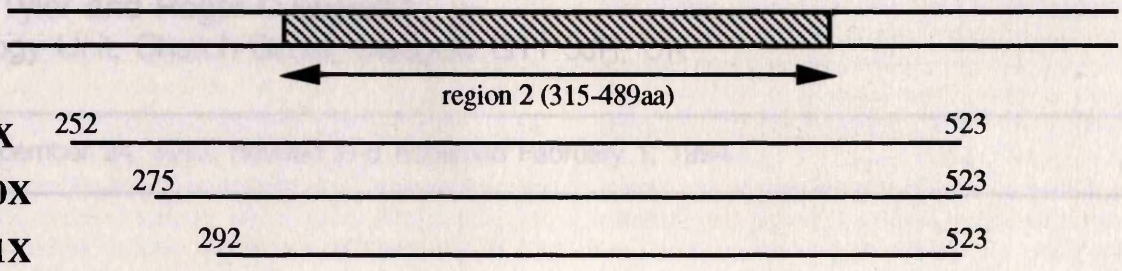
Zhang,Y.-F. and Wagner,E.K. (1987) The kinetics of individual herpes simplex virus type 1 transcripts. Virus Genes, **1**, 49-60.

Zhou,Q., Lieberman,P.M., Boeyer,T.G. and Berk.A.J. (1992) Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. Genes Dev., **6**, 1964-1974.

Zhou,Q., Boeyer,T.G. and Berk.A.J. (1992) Factors (TAFs) required for activated transcription interact with TATA box-binding protein conserved core domain. Genes Dev., **7**, 180-187.

Zweib,C., Kim,J. and Adhya,S. (1989) DNA bending by negative regulatory proteins: Gal and Lac repressors. Genes Dev., **3**, 606-611.

HSV-1 Vmw175



VZV 140k