https://theses.gla.ac.uk/

Theses Digitisation:

https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

A Mutational Analysis of the DNA Binding Domain of the Herpes Simplex Virus Immediate Early Protein Vmw175.

by

Kathryn Elizabeth Allen

A thesis presented for the Degree of Doctor of Philosophy

in

The Faculty of Science at the University of Glasgow

Institute of Virology,
Church Street,
Glasgow.
G11 5JR.

October, 1993.

ProQuest Number: 10992178

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10992178

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

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

Thesis
9719
copy 1

GLASGOW
UNIVERSITY
LIBRARY

For Mum and Dad.

TABLE OF CONTENTS

Summary

Upon infection, HSV-1 genes are transcribed by the host cell RNA polymerase II. The analysis of polypeptides synthesized during the course of HSV-1 lytic infection led to the identification of three co-ordinately regulated groups of genes, characterized by their specific temporal expression. The immediate-early or α genes are synthesized immediately following infection. Synthesis of immediate-early polypeptides is followed by the expression of early or β genes, whilst the synthesis of late or γ genes occurs subsequent to the onset of viral DNA synthesis.

The expression of these temporal classes of genes is controlled by various virally-encoded transactivators, one of which is the immediate-early polypeptide Vmw175. Vmw175 is encoded by the IE-3 gene, an unspliced open reading frame situated within the short repeat region of the HSV-1 genome. The isolation of a number of temperature-sensitive viruses containing lesions in the IE-3 gene led to the elucidation of the role of the Vmw175 polypeptide during the viral lytic cycle. The majority of these viruses showed a lack of early and late gene expression and an over-production of immediate-early products at the non-permissive temperature. These observations illustrated the importance of Vmw175 in early and late gene expression, and suggested that it may also play a role in the repression of immediate-early genes.

Whilst *in vitro* studies have demonstrated the possible involvement of Vmw175 DNA binding activity in the mechanism of viral gene transactivation, mutational analysis of early gene promoters within HSV-1 has demonstrated that no additional sequences other than those shown to bind cellular transcription factors are required for early gene transactivation during the normal course of infection. It has been proposed that such studies do not provide a true representation of the mechanism of Vmw175 transactivation, as viral infection supplies other virally-encoded transactivating proteins. However, it seems clear that for the transactivation of viral genes during a normal course of infection, Vmw175 does not require the presence of specific DNA binding sites within the viral genome. This suggests two possibilities: either Vmw175 mediates transactivation not through the interaction with viral DNA but through the modification of components of the cellular transcription machinery; or that Vmw175 is able to bind to a divergent range of DNA sequences, so that removal of a specific region of DNA does not abolish the binding activity of the protein *in vivo*.

The first scenario seems unlikely, as a wide range of mutational analyses have demonstrated that the DNA binding activity of Vmw175 is essential for gene

transactivation both *in vitro* and *in vivo*. In addition, statistical analysis has revealed that putative Vmw175 binding sites are distributed throughout the viral genome, with an average occurrence of 1 site to 250 genic bases and 1 site per 760 non-genic bases. This data supports the hypothesis that removal of specific regions of gene promoters will not remove all available sites for the interaction of Vmw175 with promoter-regulatory DNA sequences.

A wide range of mutational studies have illustrated that the abolition of the DNA binding activity of Vmw175 also led to a loss of transactivation and autoregulation activity both *in vitro* and *in vivo* (DeLuca and Schaffer, 1988; Paterson and Everett, 1988a, b; Shepard *et al.*, 1989; this thesis). However, certain mutant Vmw175 molecules do not fit this general pattern. In particular, a mutant Vmw175 molecule, ri12, which is described in greater detail below, appears to support the viral replication cycle in tissue culture whilst exhibiting no *in vitro* DNA binding properties.

The aim of this project was to identify regions of the Vmw175 molecule involved in contacting the DNA helix through mutational analysis, and to investigate whether the DNA binding properties of Vmw175 as measured *in vitro* correlate with the ability of the protein to support the viral life cycle in cell culture.

Vmw175 has been separated into 5 domains based on its homology with the VZV regulatory protein, 140K. Previous mutational analyses have shown that region 2 of Vmw175 is essential to its DNA binding and transactivational properties, and the DNA binding domain of the protein has also been mapped to this region. Analysis of the primary structure of the protein has shown that a helix-turn-helix predicted secondary structure motif is present within region 2, encompassing amino acids 445-487. This type of motif is present in many other DNA binding proteins, and has been shown in several studies to be a motif essential for DNA-protein interactions. The region spanning residues 445-487 of Vmw175 is also extremely highly conserved between regulatory proteins of alpha herpesviruses, and insertion mutations constructed within this region eliminate the *in vitro* DNA binding properties of the protein. In order to study the function of this motif within Vmw175, a detailed mutational analysis of the region was carried out.

16 single point amino acid mutations and 2 double amino acid mutations were introduced within the helix-turn-helix motif. In addition, a point base pair deletion mutation was isolated which resulted in an in-frame stop codon at amino acid 454, resulting in the production of a truncated polypeptide containing only the first helix of the proposed motif. Internal deletions of both proposed helices of the motif were also constructed. All mutations were expressed as non-fusion DNA binding domain polypeptides in bacteria using the pET vector system. The mutant polypeptides were analysed for their ability to bind to DNA probes containing a consensus Vmw175 DNA binding sequence by gel retardation analysis. It was found that the mutations resulted in

three different DNA binding phenotypes. Several point mutant polypeptides showed no change in DNA binding ability, as compared to a wild type DNA binding domain polypeptide. A second group, including several point mutant polypeptides and one double mutant polypeptide, exhibited a ts defect in DNA binding; they bound to DNA at room temperature but failed to bind at 37°C. Finally, three single point mutant polypeptides, one double mutant polypeptide, the truncated polypeptide and the polypeptides containing internal deletions were unable to bind to DNA at any temperature. The single amino acid mutations which eliminated the DNA binding ability of the polypeptide were all situated around the turn of the motif.

It has been shown that Vmw175 binds to DNA as a dimer and that the expressed DNA binding domain of the protein exists as a dimer in solution (Everett *et al.*, 1991b). Although it has not been shown conclusively that the protein monomers cannot bind to DNA, it has been suggested that mutations which eliminate the ability of the protein to dimerize will also eliminate the ability of the protein to bind to DNA. As a dimerization motif has not been identified within Vmw175, the mutant polypeptides were analysed for dimerization activity.

Glutaraldehyde cross-linking assays were used to investigate the dimerization properties of some mutant polypeptides which exhibited a ts DNA binding defect, both at room temperature and at 37°C. For those ts mutants that were investigated, the DNA binding defect was not due to a defect in dimerization. The dimerization properties of two point mutant polypeptides which do not bind to DNA at any temperature were also investigated. By gel filtration chromatography, it was shown that both mutant polypeptides exist as dimers in solution. These results suggest that the proposed helix-turn-helix motif is important for the DNA binding but not the dimerization properties of Vmw175.

The virus vri12 constructed by DeLuca *et al.* (Shepard and DeLuca, 1991b) contains a six amino acid insertional mutation and a base pair substitution within the IE-3 reading frame. This virus is able to grow in cell lines not expressing Vmw175, but Vmw175 isolated from cells infected with vri12 was unable to bind to DNA upon gel retardation analysis. This has been the only example of a Vmw175 mutant which cannot bind to DNA *in vitro* but which can support the growth of virus in tissue culture. A portion of the Bam 'y' fragment from vri12 was sequenced, and found to contain a single base pair mutation resulting in an ala-val mutation at codon 345. The DNA binding domain of Vmw175 containing the ri12 mutation was expressed in bacteria and gel retardation analysis demonstrated that this mutant polypeptide has a *ts* defect in DNA binding activity. Glutaraldehyde cross-linking assays demonstrated that this defect was not due to a lack of dimerization activity at higher temperatures.

A series of three amino acid insertional mutations have been introduced into various positions within the IE-3 coding region. Several of these mutations have been

introduced into the DNA binding domain of the protein, and it has been shown that mutations that affected the *in vitro* DNA binding properties of Vmw175 also affected the ability of the protein to transactivate and repress HSV-1 promoters in transient transfection assays. In order to also determine whether the DNA binding properties of Vmw175 are essential for its transactivation and repression properties *in vivo*, the insertional mutations have been introduced into the DNA binding domain of Vmw175 within the viral genome. The mutant viruses were constructed by co-transfection of the plasmids carrying the insertional mutations within the IE-3 coding region with genomic DNA from the virus D30EBA. The virus D30EBA contains two large deletions which remove both copies of the IE-3 gene. As D30EBA cannot express Vmw175, it can only replicate in cells which express complementary Vmw175 from the cellular genome.

The resultant recombinant viruses were tested for their ability to grow in BHK cells which do not express Vmw175. It was found that viruses carrying mutations within the IE-3 gene which eliminate the ability of Vmw175 to bind to DNA *in vitro* cannot support viral growth in tissue culture. These results support the proposal that the DNA binding properties of Vmw175 are essential for its role in the regulation of viral gene expression.

Finally, certain point and deletion mutations whose construction is described above were introduced into the viral genome. The growth of such recombinant viruses in non-complementing cell lines was analysed to determine the *in vivo* activity of the mutant Vmw175 molecules. Again, it was found that mutations which eliminated the DNA binding activity of a Vmw175 DNA binding domain polypeptide *in vitro* also abolished the ability of the protein to support viral replication. However, interesting results were obtained with two mutations previously found to result in a ts defect in DNA binding activity *in vitro*. Whilst one mutation resulted in a Vmw175 molecule unable to support viral growth, the second mutant protein resulted in a recombinant virus viable in BHK cells. The ts defect in DNA binding as measured *in vitro* represents an impairment in the function of the protein molecule, and it is possible that whilst interaction with other cellular factors may restore the activity of certain ts mutant proteins, other mutations disrupt the protein structure sufficiently to abolish its transactivation and repression functions *in vivo*.

In conclusion, the mutational analysis of a proposed helix-turn-helix motif within Vmw175 has illustrated that this region plays a role in the DNA binding but not the dimerization activity of the protein. In addition, the introduction of several mutations into the viral genome has illustrated a link between the DNA binding activity of Vmw175 as measured *in vitro*, and the ability of the protein to support the viral lytic cycle in cell culture.

Acknowledgements.

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

I would like to thank my project supervisor, Dr. Roger D. Everett for his unfailing support both during and after my time in the Institute of Virology, and for his continued enthusiasm which has provided motivation during difficult times. I am extremely grateful for all the help and friendship, cakes and chocolate biscuits that I have received from my co-workers - Anne Orr, Maggie Elliott and Jess Tyler. In particular, thanks to Maggie for her help in constructing the pGL series of plasmid vectors.

My years of study have been long and sometimes hard, but made all the more bearable by the great friends I have made, and the old friends that are luckily still around. Thankyou to Anne and Julia, for all the support you have shown me over the past years - you've put up with a lot and I am grateful! And to Cathi - hope we're still sharing each others company and music taste in another 12 years, even if we're not under the same roof. Thanks to my sister, Mags, for continual wise words and general kicks up the backside, which I have needed. And last, but not least, to Lisa, for all the good times and holidays, it wouldn't have been the same without you!

Finally, I would like to dedicate this thesis to my parents, Nancy and Brian Allen, who have provided complete and ungrudging moral and financial support through all my years of study. Thankyou for everything.

During this work the author was supported by a Medical Research Council Training Award. Unless otherwise stated, all the results described in this thesis were obtained by the author's own efforts.

Kathryn Elizabeth Allen October, 1993.

Abbreviations.

^{32}P: phosphorus-32 radioisotope
α-TIF: alpha trans-inducing factor
A: adenine
ADP: adenosine-5'-diphosphate
APS: ammonium persulphate
ATP: adenosine-5'-triphosphate
bFGF: basic fibroblast growth factor
BHK: baby hamster kidney
bHLH: basic helix loop helix
bp: base pairs
BSA: bovine serum albumin
C-terminal: carboxy-terminal
C: cytosine
CA: carbonic anhydrase
cAMP: cyclic adenosine monophosphate
CAV: cell associated virus
Ci: curie
cm: centimetre
cpe: cytopathic effect
cpm: counts per minute
CRV: cell released virus
CTD: carboxy terminal domain
Cyt C: cytochrome C
Da: dalton
dC: deoxy cytosine
ddNTP's: dideoxy nucleoside-5'-triphosphates
dI: deoxy inosine
DNA: deoxyribonucleic acid
DNAase: deoxyribonuclease
dNTP: 2'-deoxyribonucleoside-5'-triphosphates
DTT-dithiothreitol
E. coli: Eschericia coli
E: early
EBV: Epstein-Barr virus

EDTA: sodium ethylenediamine tetra-acetic acid
EHV-1 equine herpesvirus type 1
ELISA: enzyme-linked immuno-sorbent assay
EtBR: ethidium bromide
G-O-R: Garnier-Osguthorpe-Robson
G: guanine
g: grams
GCG: genetics computer group
gD: glycoprotein D
GMEM: Glasgow's modified Eagles mediuM
h: hour(s)
HCMV: human cytomegalovirus
HEPES: N-2 hydroxyethylpiperazine-N'-2-ethanesulphonic acid
HHV-6: human herpesvirus-6
HRP: horseradish peroxidase
HSV-1: herpes simplex virus type 1
HSV-2: herpes simplex virus type 2
HVT: herpesvirus of turkeys
IE: immediate-early
IR_L: long internal repeat
IR_S: short internal repeat
K: kilo
Kb: kilobases
L: late
l: litre
LAT's: latency associated transcripts
LFP: large fragment polymerase
LPF: late processing factor
m. o. i.: multiplicity of infection
M: molar
µ: micro
mA: milliamps
MDV: Marek's disease virus
mg: milligrams
mi: mock infected
min: minutes
ml: millilitres
mm: millimetres
MOPS: 3-(N-morpholine)propanesulphonic acid

mRNA: messenger ribonucleic acid
N-terminal: amino-terminal
n: nano
nm: nanometre
NP40: Nonidet P40
OD_{260}: optical density at wavelength 260nm
Ori: origin of replication
p: pico
PAGE: polyacrylamide gel electrophoresis
PEG: polyethylene glycol
pfu: plaque forming units
PMSF:phenylmethylsulphonylfluoride
poly(A): poly adenylic acid
PRV: pseudorabies virus
RE: restriction enzyme
RF: replicative form
RNA: ribonucleic acid
RNase: ribonuclease
rpm: revolutions per minute
rRNA: ribosomal ribonucleic acid
RT: room temperature
SDS: sodium dodecyl sulphate
sec: seconds
snRNP: small nuclear ribonuclear protein
ss: single stranded
SV40: simian virus 40
syn: syncitial plaque morphology locus (syn^+= non-syncitial; syn^-= syncitial)
T: thymine
TEMED: N,N,N',N'-tetramethylethylene diamine
tk: thymidine kinase
TLC: thin layer chromatography
TR_L: long terminal repeat
tRNA: transfer ribonucleic acid
TR_S: short terminal repeat
ts: temperature sensitive
u: unit
U_L: long unique region
U_S: short unique region
uv: ultraviolet

V: volts
v: volume
vol: volumes
VZV: varicella zoster virus
W: watts
w: weight
WT: wild type

AMINO ACID SYMBOLS

A alanine	G glycine	M methionine	S serine
C cysteine	H histidine	N asparagine	T threonine
D aspartate	I isoleucine	P proline	V valine
E glutamate	K lysine	Q glutamine	W tryptophan
F phenylalanine	L leucine	R arginine	Y tyrosine

(The three letter code abbreviations are underlined).

Chapter 1. Introduction.

1A. Biology of Herpes Simplex Virus Type 1.

1A.1. Herpesvirus Characterization.

Viruses are classified as members of the family *Herpesviridae* on the basis of the structure of the viral particle. Herpesvirus particles contain four major architectural elements; a central core containing the genomic DNA, an icosahedral capsid, approximately 100-120nm in diameter, an amorphous proteinaceous tegument, and an outer lipid envelope containing glycoprotein spikes.

The core of the virus is an electron dense mass within the capsid containing the genomic DNA, which exists as a linear double stranded molecule of 80-150x10^6Da molecular weight (reviewed by Dargan, 1986). The viral DNA was thought to form a toroid structure wound around a central cylindrical plug (Furlong *et al.*, 1972; Nazerian, 1974). However, recent cryoelectron microscopy analysis of herpesvirus particles has suggested that this structure was an artifact caused by the method of specimen preparation or was an intermediate structure in herpesvirus particle maturation. It has now been proposed that the DNA is packaged in a liquid-crystalline configuration, forming a universally dense ball which extends to the inner surface of the viral capsid (Booy *et al.*, 1991).

The morphology of the viral capsid is characteristic of all members of the family *Herpesviridae*. It is an icosahedral structure, 100-120nm in diameter, containing 162 hollow, elongated capsids, 150 of which are hexameric and 12 of which are pentameric (Wildy *et al.*, 1960).

The tegument layer is situated between the viral capsid and outer lipid membrane and varies in thickness between members of the herpesvirus family. It consists of an ill-defined proteinaceous mass which occasionally appears to be fibrous in negatively stained virus particles (Roizman and Furlong, 1974).

The outer envelope of herpesvirus particles has a trilaminar appearance, typical of a plasma membrane (Epstein, 1962). This envelope contains numerous glycoprotein spikes, estimated to be 8nm long (Wildy *et al.*, 1960; Wildy and Watson, 1963; Heine *et*

al., 1972). The stage at which a herpesvirus capsid acquires its envelope has been the subject of many studies, and is discussed in more detail in Section 1A.5.4.

Nearly 100 herpesviruses have been partially characterized (Roizman, 1990), and the members of this family infect a wide variety of hosts. At present, seven herpesviruses have been isolated from humans, these are discussed in further detail in Section 1A.4.

In addition to sharing conserved structural features, members of the *Herpesviridae* family also share a number of common biological properties. The replication cycle of the virus, including viral DNA synthesis and capsid assembly, occurs within the nucleus. Infection of host cells with herpesviruses and the subsequent replication and production of infectious particles results in the destruction of infected cells. Finally, all herpes viruses examined share the ability to establish a latent infection in the host. This is discussed further in Section 1A.7.

1A.2. Herpesvirus Classification.

Herpesviruses have been classified on the basis of their biological properties into three subfamilies - the alpha-, beta- and gamma-herpesviruses (Roizman, 1982; Matthews, 1982; Roizman, 1990).

Alphaherpesviruses have a variable host range in tissue culture, and exhibit a short reproductive cycle followed by destruction of infected cells and rapid spread in cell culture. These viruses readily establish latent infection in the ganglia. This subfamily includes the human viruses herpes simplex virus subtypes 1 and 2 and varicella zoster virus. Betaherpesviruses are characterized by a longer reproductive cycle and slow development of cytopathic effect in cell culture. Infection with betaherpesviruses frequently results in enlargement of the infected cell (cytomegalia) both in tissue culture and *in vivo*. This subfamily includes the human and murine cytomegaloviruses.

Gammaherpesviruses are identified mainly on the basis of the host range of the virus and characteristics of the latent infection. *In vivo*, the host range of the virus is limited to the family or order to which the natural host belongs. *In vitro*, members of this subfamily are restricted to replication in lymphoblastic cell cultures and occasionally may infect some epithelial and fibroblastic cell lines. These viruses frequently establish latent infections within lymphoid tissue. Two members of this subfamily are Epstein-Barr virus and herpesvirus saimiri.

Such biological classification does not take into account the evolutionary relatedness of the herpesviruses. As more sequence data becomes available, herpesviruses can also be classified on the basis of the conservation of genes and gene

clusters, as well as the conservation of reiterated sequences within the genome. At present, the analysis of the genome content of many herpesviruses has led to the confirmation of classification of the viruses into the three subfamilies described above (see, for example, McGeoch *et al.*, 1988a; Davison and Taylor, 1987). Two exceptions to this method of classification are the Marek's disease virus and human herpesvirus 6.

Marek's disease virus (MDV) and the closely related herpesvirus of turkeys (HVT) have been classified as gammaherpesviruses on the basis of their ability to establish infection in lymphoid tissue and, in the case of Marek's disease virus, to result in the formation of T-cell lymphomas in the infected host (Roizman, 1982). Human herpesvirus 6 was also provisionally classified as a gammaherpesvirus on the basis of its ability to infect lymphoid tissue both *in vitro* and *in vivo* and the isolation of HHV-6 genome sequences in some B-cell tumours. However, partial sequence analysis of the three viruses and comparison of predicted reading frames with those of other viruses has shown that MDV and HVT exhibit high homology to the alphaherpesvirus VZV, whilst HHV-6 exhibits high homology to the betaherpesvirus HCMV (Buckmaster *et al.*, 1988; Lawrence *et al.*, 1990). Both studies have shown that the viruses in question display a lower genetic homology to the gammaherpesvirus Epstein-Barr virus. Therefore, on the basis of sequence and mapping data, it appears that MDV and HVT belong to the alphaherpesvirus subfamily and HHV-6 to the betaherpesvirus subfamily, rather than the gammaherpesvirus subfamily .

Although these studies appear to present discrepancies between the biological and genomic methods for classification of the herpesviruses, genomic information may not directly reflect the fundamental properties of the virus *in vivo*. For example, in Marek's disease virus, several open reading frames identified as conserved and colinear with those of varicella zoster virus are not expressed during latency or in transformed cells (Buckmaster *et al.*, 1988). Therefore, several viruses may have evolutionarily adapted in order to lose or gain biological properties, leading to the expression of the biological properties of one herpesvirus subfamily whilst retaining genetic similarity to members of a second herpesvirus subfamily.

1A.3. Pathogenicity of Human Herpesviruses.

Seven herpesviruses are known to establish infection in man - the herpes simplex viruses subtypes 1 and 2; varicella zoster virus; human cytomegalovirus; Epstein-Barr virus; human herpesvirus 6; and human herpesvirus 7.

The **herpes simplex viruses** primarily infect epithelial tissue through mucosal surfaces or abraded skin, followed by transport of virions or capsids to neuronal ganglia where the viruses establish a latent infection. Viruses can be reactivated from latency by a variety of stimuli. Primary HSV-1 infections are generally asymptomatic, with the occasional occurrence of fever and sore throat. Reactivation results in the formation of vesicular lesions at the site of primary infection. Infection with HSV-2 results in formation of genital lesions (reviewed in Wildy, 1990).

Immunocompromised patients have an increased susceptibility for HSV infection. The latent herpes simplex viruses may reactivate at multiple sites, and patients can develop a progressive disease involving the respiratory or gastrointestinal tracts. The increase in occurrence of acquired immune deficiency syndrome has resulted in widespread outbreaks of such opportunistic HSV infections, together with the isolation of many acyclovir-resistant strains of HSV (reviewed in Masur, 1992). Recent studies have suggested that HSV infection can trigger replication of proviral forms of the human immunodeficiency virus through a transcriptional mechanism involving the transcription factor NFκB (Vlack and Pitha, 1992).

Occasionally, infection with HSV can result in herpes simplex encephalitis, a sporadic, fatal disease involving fever and progressively deteriorating levels of consciousness (Olson *et al.*, 1967).

Primary infection with **varicella-zoster virus** usually occurs in childhood and is initially asymptomatic, with appearance of a rash after a 14-15 day incubation. Infection in older patients produces the initial symptoms of a fever accompanied by backache and headache. After the incubation period, a vesicular rash appears, predominantly concentrated on the scalp and trunk (reviewed in Gelb, 1990). Herpes zoster results from reactivation of latent virus from neuronal ganglia in older or immunocompromised hosts. Reactivation results in the formation of skin lesions situated around the infected nerve endings (Hope-Simpson, 1965).

Human cytomegalovirus infection is widespread, with a generally asymptomatic primary infection resulting in mononucleosis, followed by recurrent infection caused by chronic infection, reinfection by different genetic strains of the virus; or reactivation of latent virus (reviewed in Alford and Britt, 1984). Again, HCMV infection is increased in occurrence and severity in immunocompromised hosts. HCMV is one of the more common opportunistic infective agents associated with acquired immune deficiency syndrome, causing symptoms such as gastroenteritis and retinitis. It has been suggested that HCMV potentiates human immunodeficiency virus infectivity by a transactivational mechanism (Davis *et al.*, 1987).

Primary infection with **Epstein-Barr virus** is generally asymptomatic, although in older children and adults, EBV infection results in infectious mononucleosis. EBV has been implicated in the formation of Burkitt's lymphoma (Andiman *et al.*, 1985; de The *et al.*, 1987) and nasopharyngeal carcinoma (Raab-Traub *et al.*, 1987).

Human herpesvirus 6 was first isolated from the peripheral lymphocytes of patients with acquired immune deficiency syndrome and T-cell leukaemia (Salahuddin *et al.*, 1986). It is now thought that HHV-6 infection is widespread in the adult population (reviewed in Lopez and Honess, 1990). The virus is strongly linked to the childhood disease exanthum subitum (Roseala infantum; Yamanishi *et al.*, 1988). Most other cases involve isolation of HHV-6 from patients exhibiting some form of immunosuppression, such as patients with acquired immune deficiency syndrome or transplant recipients. Initial reports that HHV-6 may be involved in the progression of AIDS have not been supported by recent studies (Lopez and Honess, 1990).

Human herpesvirus-7 was isolated from activated T-cells of a healthy individual, and was found to exhibit typical herpesvirus morphology upon electron micrograph analysis. The virus exhibited cytopathic effect in cell culture and limited homology to HHV-6 by Southern hybridization (Frenkel *et al.*, 1990). As yet, no disease associated with HHV-7 has been described.

1A.4. HSV-1 Genome Structure.

The initial purification of DNA from nucleocapsids of HSV-1 and -2 showed that the viral DNA existed as a double stranded linear molecule with a molecular weight of 95+/-5 x 10^6 Da. This DNA molecule exhibited evidence of random single stranded nicks throughout the genome (Kieff *et al.*, 1971). Comparison of DNA from HSV-1 and -2 revealed that the genomes were of similar size, similar G+C content (67% and 69% respectively), and exhibited 50% homology (Kieff *et al.*, 1971; Wadsworth *et al.*, 1975).

Electron micrograph analysis of self-annealed DNA molecules from HSV-1 strain F suggested that the genome contained terminally redundant sequences that were inverted and repeated internally (Wadsworth *et al.*, 1975; Sheldrick and Berthelot, 1974). These repeated sequences flank unique long and short regions of the HSV genome (see Figure 1A.1). Partial denaturation analysis revealed that preparations of HSV-1 DNA contain an equimolar mixture of four isomers of the HSV-1 genome (Wadsworth *et al.*, 1975; reviewed in Roizman, 1979). One of these four isomers, the P isomer, is used as the prototype of the HSV-1 genome and is illustrated in Figure 1A.1.

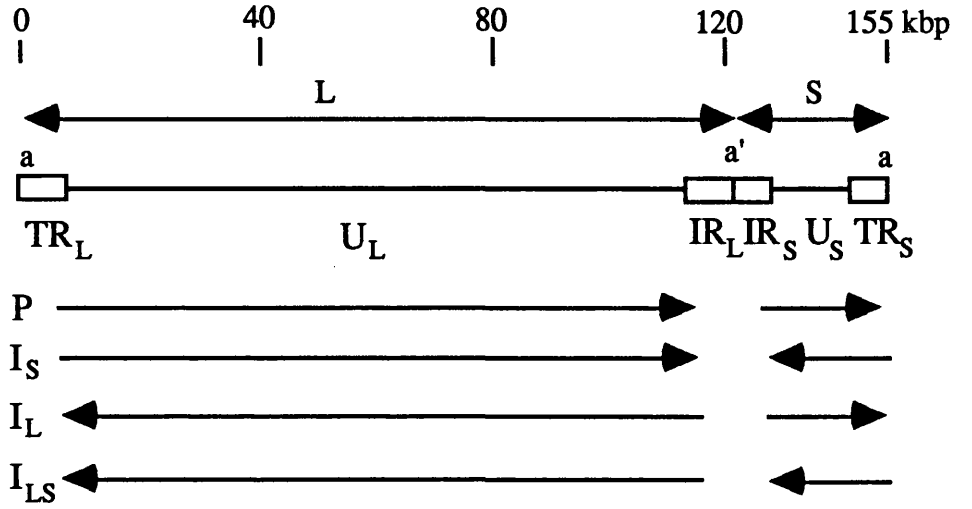


Figure 1A.1. Structure of the HSV-1 Genome.

The HSV-1 genome is represented as two covalently linked linear unique segments, U_L and U_S, spanned by major repeat elements, TR_L, TR_S, IR_L and IR_S. The a sequence and the internal inverted repeat a' sequence are also shown.

Isomers of the HSV-1 genome are formed by inversion of the covalently linked L and S segments, and are illustrated as the molecules P, I_L, I_S and I_{LS}

Restriction enzyme mapping illustrated the existence of the four isomers of HSV DNA. These studies also showed that the genome contained multiple copies of a sequence present at the joint of the long and short sequences, and also at the termini of the molecule (Wagner and Summers, 1978). In the simplest form of this model, only one copy of the repeated sequence was present at the joint of the segments.

This repeat sequence, termed the "a" sequence, is 401bp in size in the HSV-1 strain 17 syn$^+$ (McGeoch *et al.*, 1988a). Sequence analysis of the HSV-1 strain F DNA led to the elucidation of the structure of the a sequence, which can be represented as:

DR1—Ub—DR2$_n$—DR4$_m$—Uc—DR1

Where DR1, DR2 and DR4 = direct repeats of 20bp, 12bp and 37bp

Uc and Ub = "unique" segments of 58bp and 64bp

n = 19-22 in HSV-1 strain F

m = 2-3 in HSV-1 strain F.

The structure and size of the a sequence varies from strain to strain due to the differing numbers of DR2 and DR4 repeats (Mocarski and Roizman, 1982). Sequence analysis of the termini of a HSV-1 strain F DNA molecule showed that viral genomes were produced from concatameric molecules by cleavage within the DR1 element of the a sequence, resulting in overhanging 3' residues at the termini of the molecules (Mocarski and Roizman, 1982).

Complete genomic sequence data has now been obtained for four herpesviruses of man - HSV-1 (McGeoch *et al.*, 1986, 1988a); VZV (Davison and Scott, 1986);

HCMV (Chee *et al.*, 1990); and EBV (Baer *et al.*, 1984); and also for the channel catfish virus (Davison, 1992). The HSV-1 strain 17 syn^+ has a genome length of approximately 152000 residues and base composition of 68.3% G+C content. Analysis of the sequence of this genome led to the prediction of 72 open reading frames, as illustrated in Figure 1A.2, 52 of which corresponded to previous mRNA transcript mapping data. The function of the protein products of 33 of these genes were proposed (McGeoch *et al.*, 1988a). The identified proteins fell into three general categories - immediate early proteins; proteins involved in DNA replication or metabolism; and proteins involved in the structure or assembly of viral particles. Certain genes within the long unique region of the genome were shown to be non-essential for viral growth in tissue culture, including UL23, UL24, UL39, UL44, UL50 and UL56 (McGeoch *et al.*, 1988a).

Since this study, a further three genes of HSV-1 have been described, namely UL26.5 (Liu and Roizman, 1991a, b); RL1 (Chou and Roizman, 1990; Dolan *et al.*, 1992); and UL49.5 (Barker and Roizman, 1992) or UL49A (Barnett *et al.*, 1992). Of these three genes, only one has been assigned a function - the gene UL26.5, which encodes the ICP35 family of capsid proteins (Liu and Roizman, 1991a).

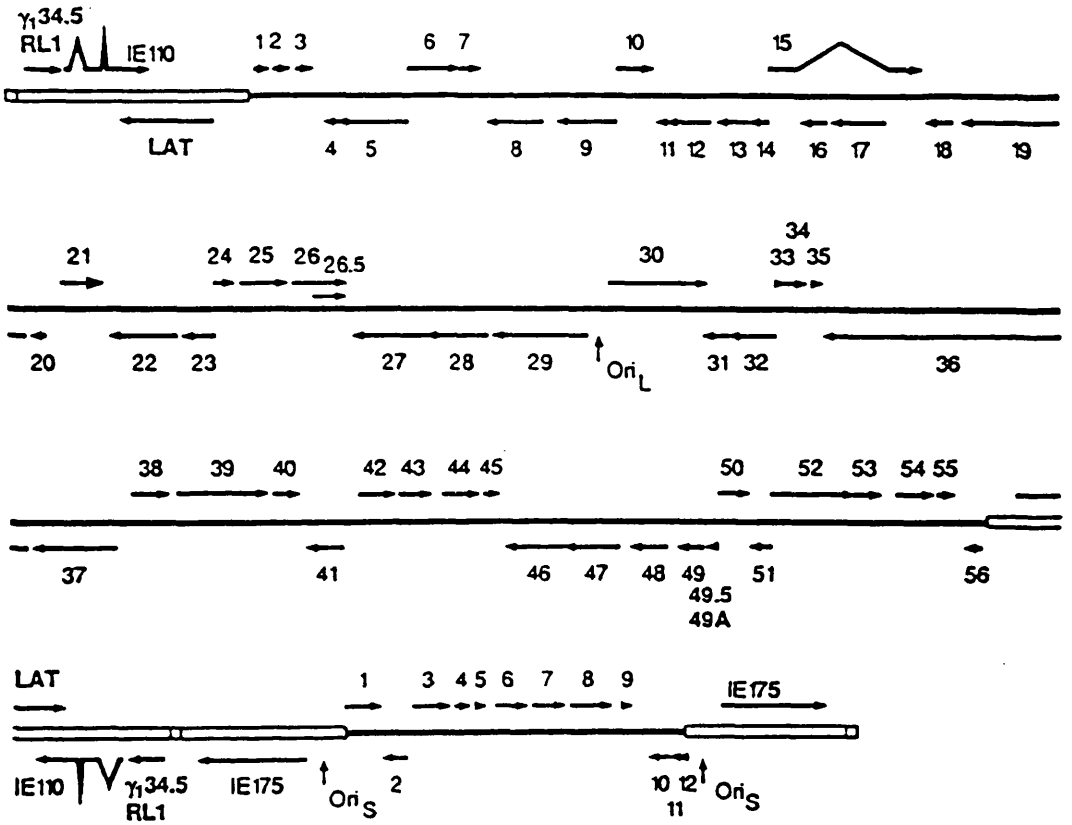


Figure 1A.2. The Coding Content of the HSV-1 Genome.

The HSV-1 genome is represented as the P isomer, showing the location of the open reading frames. Open reading frames are indicated by arrows. The genes UL1-UL56 are indicated by genes 1-56 on the top three lines, whilst genes US1-US12 are indicated by genes 1-12 on the bottom line. Ori$_L$ and Ori$_S$ indicate the viral origins of replication.

Reproduced from McGeoch, 1989, with the incorporation of three additional genes; UL26.5, UL49.5 and γ1 34.5/RL1 (see text).

1A.5. The Lytic Cycle of HSV-1.

1A.5.1. Adsorption, Penetration and Uncoating.

HSV-1 adsorption and penetration occurs by a pH independent pathway, involving fusion of the viral and cellular membranes (Fuller and Spear, 1987). This is a function of the viral glycoproteins. At least ten glycoproteins are encoded by HSV-1, designated gB, gC, gD, gE, gG, gH, gJ, gK and gL (Spear, 1985; Roizman and Sears, 1990; Fuller and Lee, 1992). Neutralization studies (Fuller and Spear, 1987; Highlander *et al.*, 1988; Fuller *et al.*, 1989) and characterization of mutant viruses (Ligas and

Johnson, 1988; Desai *et al.*, 1988; Cai *et al.*, 1988; Forrester *et al.*, 1992) have shown that glycoproteins B, D and H are essential for viral replication in cell culture, and block viral penetration at a step following cellular attachment.

Viral attachment to the cell surface has been proposed to occur via two stages. The initial attachment to the cell surface is mediated through heparan sulphate (WuDunn and Spear, 1989), and this attachment may involve the viral glycoproteins B or C (Herold *et al.*, 1991). Analysis of a deletion mutant virus lacking glycoprotein C coding sequences has led to the proposal that although the protein is dispensable for the viral lytic cycle, it facilitates adsorption by its interaction with heparan sulphate (Herold *et al.*, 1991).

The second stage of viral penetration is postulated to involve the interaction of viral glycoproteins with a specific, high affinity cellular receptor. Kaner *et al.* (1990) have proposed that the basic fibroblast growth factor (bFGF) receptor acts as the specific high-affinity portal of entry for HSV-1. However, later studies have led to the suggestion that the inhibition of viral attachment to the cellular surface by bFGF as observed by Kaner *et al.* was due to the sequestering and saturation of heparan sulphate molecules on the cellular surface by bFGF (Muggeridge *et al.*, 1992; Mirda *et al.*, 1992).

The mechanism by which viral glycoproteins act to initiate fusion with and penetration of the cell membrane has not been clearly defined. A wide variety of evidence has emerged to suggest that glycoprotein D binds to a cellular receptor (Johnson and Ligas, 1988; Johnson *et al.*, 1990). The fact that gD expressing cell lines are resistant to superinfection by HSV-1 has led to the suggestion that gD is sequestering a component of the cell membrane, but later work has shown that this interference is mediated through gD molecules present in the virion envelope (Campadelli-Fiume *et al.*, 1990; Dean *et al.*, 1992). Additionally, studies with a HSV-1 virus lacking glycoprotein H coding sequences has led to the suggestion that gH acts in viral-cellular penetration at a step subsequent to the action of gD (Forrester *et al.*, 1992).

A recent electron micrograph analysis of infectious HSV-1 particles and particles neutralized with anti-gD or -gH antibodies (Fuller and Lee, 1992) has resulted in the development of a model for HSV-1 adsorption and penetration, illustrated in Figure 1A.3. The authors proposed that viral adsorption and penetration occurs in five distinct stages, each of which can be observed during electron micrograph analysis of HSV-1 infection.

The five stages described by the authors involve the initial attachment of the virus to cell surface heparan sulphate molecules as described previously, followed by a stable or secondary attachment probably mediated by gD. Following attachment, formation of a fusion bridge is initiated, possibly through conformational changes of viral glycoproteins D and H. Expansion of the bridge follows, resulting in the condensation of the viral envelope and tegument and release of some tegument proteins into the cellular cytoplasm.

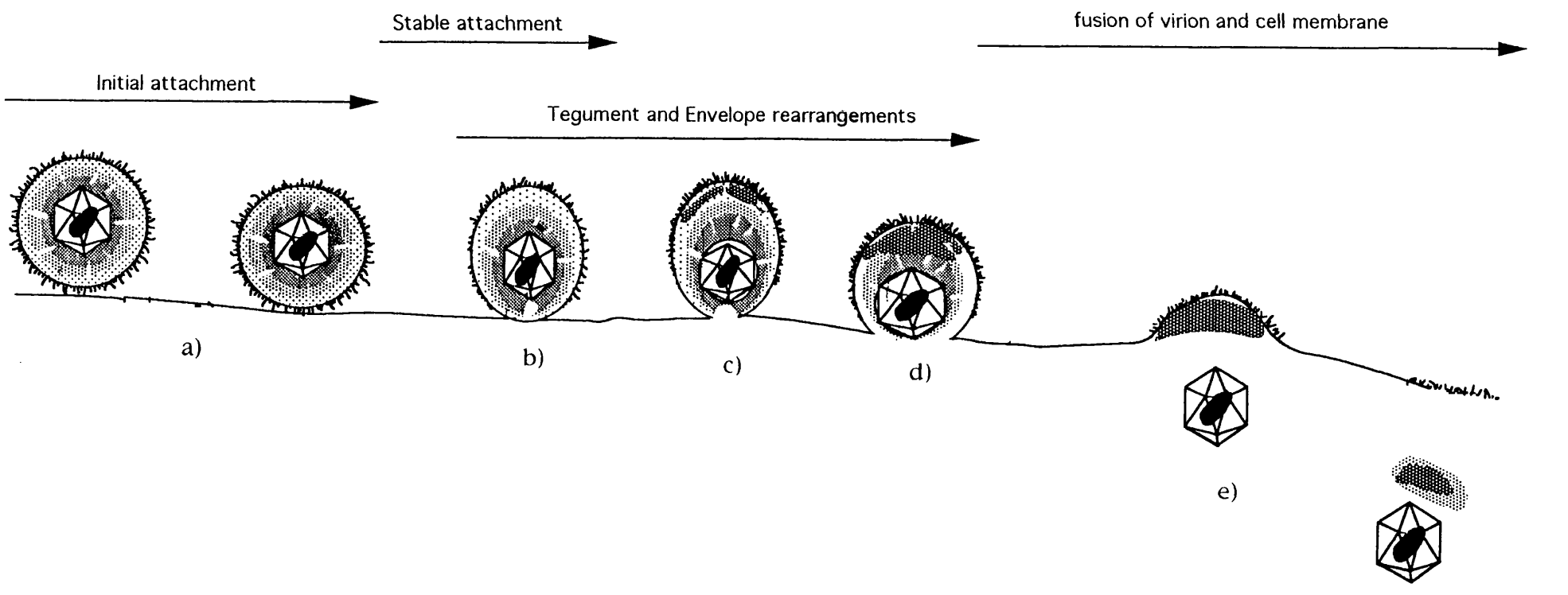


Figure 1A.3. Model of HSV-1 Entry.

The diagram illustrates the five distinct stages of HSV-1 entry. (a) Initial attachment of the virus to the cell surface. (b) Stable attachment. (c) Initiation of a fusion bridge. (d) Expansion of the fusion bridge. (e) Release of the nucleocapsid into the cytoplasm.

The proposed rearrangements of the virion envelope, tegument and nucleocapsid in relation to the cell plasma membrane are diagrammed. Reproduced from Fuller and Lee (1992).

Finally, viral adsorption and penetration is completed by movement of the viral nucleocapsid into the cellular cytoplasm.

Following penetration of the cellular membrane, viral nucleocapsids are transported along microfilaments to the nuclear membrane (Lycke *et al.*, 1984). From electron micrograph analysis of a temperature sensitive mutant of HSV-1, it appears that the viral DNA dissociates from the nucleocapsid and enters the nucleus through nuclear pores (Batterson *et al.*, 1983). A HSV-1 mutant with a temperature-sensitive defect in uncoating and late gene expression has been isolated (Knipe *et al.*, 1981; Batterson *et al.*, 1983). This mutation has been mapped to a 300bp region within the UL36 gene (McGeoch *et al.*, 1988a).

1A.5.2. The Effect of HSV-1 Infection on Host Cell Macromolecular Synthesis.

Upon infection, herpes simplex viruses redirect the host macromolecular synthesis machinery in order to allow replication of the viral genome (reviewed in Fenwick, 1984). Initial studies illustrated that as viral infection proceeded, synthesis of cellular DNA declined, whilst synthesis of viral DNA increased (Roizman and Roane, 1964). Host cell protein synthesis (Sydiskis and Roizman, 1966, 1967) and cellular RNA levels (Hay *et al.*, 1966; Pizer and Beard, 1976) were also found to decline.

Whilst the prevalent effect of HSV infection is to inhibit host cell macromolecular synthesis, the activation of a subset of cellular genes has also been observed. Latchman and co-workers have illustrated that activation of certain cellular genes is dependent upon the presence of active Vmw175 (Latchman *et al.*, 1987; Kemp and Latchman, 1988), whilst cellular promoters introduced into the cell by transfection are activated by HSV infection in a mechanism involving the HSV-1 immediate-early proteins Vmw110 and Vmw175 (Everett, 1985). The mechanism by which Vmw175 activates certain cellular promoters is still not fully understood, and is dealt with in greater detail in Section 1B.5. Interestingly, it appears that induction of certain cellular genes involves the binding of HSV-1 to the cell surface, but is independent of viral gene expression (Kemp *et al.*, 1986; Preston, 1990) and that cellular transformation by herpes simplex viruses also results in the overexpression of a number of cellular genes (Macnab *et al.*, 1985).

The HSV-1 mutant tsK contains a temperature-sensitive defect in the gene encoding Vmw175. Infections with this virus at the non-permissive temperature results in the induction of cellular heat shock proteins (Notariani and Preston, 1982; Estridge *et al.*, 1989). These proteins are also induced upon HSV-2 infection (LaThangue *et al.*, 1984). Further analysis of this cellular response has suggested that the heat shock proteins are expressed in response to high levels of an abnormal form of Vmw175 (Russell *et al.*, 1987a).

Host cell macromolecular synthesis shut-off is achieved through two clearly defined steps - the disassociaton of host cellular mRNA from polyribosomes and disaggregation of polyribosomes (Sydiskis and Roizman, 1967; Nishioka and Silverstein, 1978); and the degradation of mRNA (Nishioka and Silverstein, 1978). This degradation is not confined to host mRNA transcripts - HSV infection results in the degradation of virally-encoded RNA (Kwong and Frenkel, 1987; Oraskar and Read, 1987). In addition to the degradation of pre-existing mRNA, several groups have reported the repression of host RNA transcription by HSV infection (Kemp and Latchman, 1988; Smiley *et al.*, 1991). Whilst Kemp and Latchman (1988) have suggested that this transcriptional repression is caused by the immediate-early gene product ICP22, Smiley *et al.* (1991) have presented evidence that repression of cellular transcription only occurs in the presence of functional Vmw175. It is possible that the repression observed is not caused by the action of a specific virally-encoded protein, rather, it is a indirect result of the redirection of the cellular transcriptional machinery by HSV-1, resulting in the "hijacking" of factors essential for the transcription of cellular genes.

The shut-off of host macromolecular synthesis occurs in two distinct stages - "early" and "delayed" shut-off. "Early" host cell shut-off is mediated by a virion component, whilst "delayed" shut-off requires viral gene expression (Nishioka and Silverstein, 1978; Fenwick and Clark, 1982; Fenwick, 1984). Further analyses demonstrated that different strains of virus may have differing mechanisms for host cell shut-off, HSV-1 strain KOS demonstrating no requirement for "delayed" shut-off (Schek and Bachenheimer, 1985; Smiley *et al.*, 1991). In addition, HSV-2 strain G exhibits stronger host cell shut-off, which again does not require viral gene expression (Pereira *et al.*, 1977; Fenwick and Walker, 1978).

Read and Frenkel (1983) isolated a series of HSV-1 strain KOS mutants with defects in early host cell shut-off - designated *vhs* mutants. Analyses of these mutants demonstrated that the early and delayed shut-off phenotypes were functions of separate genes, that the early shut-off function was non-essential for viral replication in cell culture, and that a defect in host cell shut-off resulted in over-expression of viral genes. This over-expression of viral gene products by *vhs* mutants was shown to be due to the stabilization of viral RNA (Kwong and Frenkel, 1987; Oraskar and Read, 1987). Although non-essential for viral growth in tissue culture, the early host cell shut-off function may be advantageous under limiting growth conditions *in vivo*. The destabilization of viral as well as cellular mRNAs may prevent the over-expression of cellular promoters by Vmw175, halt disturbance of nucleotide metabolism by over-expressed early HSV gene products, and facilitate rapid changeover during the temporal cascade of viral gene expression (described by Honess and Roizman, 1975).

These *vhs* mutations have been mapped to the HSV-1 gene UL41 (McGeoch *et al.*, 1988a). Fenwick and Everett (1990) designed a series of experiments to determine whether the differences in host cell shut-off exhibited by various HSV strains were attributable to the UL41 gene product. Construction and analysis of recombinant viruses showed that the HSV-2 UL41 gene was solely responsible for the strong early shut-off phenotype. In contrast to these results, super-infection of HSV-1 (17^+) infected cells with HSV-2 (G) resulted in the domination of the weaker HSV-1 phenotype (Hill *et al.*, 1985). Fenwick and Everett (1990) explained this paradox by suggesting that UL41 functions as a dimeric molecule, acting on a cellular target (possibly a nuclease). Therefore, coinfection with HSV-1 and HSV-2 could allow the HSV-1 UL41 dimer which exhibits weaker shut-off characteristics to occupy the target site, thus inhibiting the action of the HSV-2 UL41 gene product. However, expression of the two genes from one viral genome could result in the occurrence of a heterodimeric molecule, in which the stronger shut-off phenotype is dominant.

Analysis of the UL41 coding sequences from various HSVstrains has shown that differences in viral host cell shut-off phenotype are reflected in their genetic content (Everett and Fenwick, 1990). The development of an antiserum to the UL41 gene product has allowed the identification of the 58KDa protein product and demonstrated that the protein is, as predicted, a component of the virion particle (Smibert *et al.*, 1992).

1A.5.3. HSV DNA Replication.

HSV DNA replication is thought to involve a rolling-circle method of replication, due to the identification of concatamers of the HSV genome within infected cells (Jacob *et al.*, 1979). Biochemical analysis of the *cis*-acting sequences and viral genes required for HSV DNA replication has resulted in the elucidation of some aspects of the replication process (reviewed in Challberg, 1991).

The viral origins of DNA replication were first identified from analysis of defective genomes obtained from propagation of HSV in tissue culture (Frenkel *et al.*, 1975; 1976; Schroder *et al.*, 1975), and were shown to direct viral replication in transient transfection assays (Stow and McMonagle, 1983; Weller *et al.*, 1985). HSV-1 contains three origins of replication: two copies of oriS, situated within the short repeat region of the genome and one copy of oriL, situated within the long unique region of the genome. It seems that the two origins of replication are functionally equivalent; only mutations affecting both oriS and oriL have a deleterious effect on viral replication (Longnecker and Roizman, 1986; Polvino-Bodnar *et al.*, 1987; Challberg, 1991). In addition, mutational analysis has shown that sequences essential for the function of oriS are also conserved in oriL (Weller *et al.*, 1985; Lockshon and Galloway, 1988; Deb and Doelberg, 1988; Hernandez *et al.*, 1991; Weir and Stow, 1990).

The essential regions of the origin of replication are illustrated in Figure 1A.4. The minimal core region of the origin contains the sequences essential for viral DNA replication, whilst addition of flanking sequences increases the efficiency of replication, an aspect of viral DNA replication also observed in adenovirus and SV40. The minimal core contains three binding sites for the UL9 origin-binding protein, discussed in greater detail below, and a central AT-rich region. The flanking regions of the oriS contain the promoter/regulatory regions of the IE-3 and IE-4/5 promoters. These sequences contain binding sites for a variety of cellular and viral transactivating proteins, such as Vmw65, Sp1, and NF-III. Binding of these proteins to the regulatory sequences increases the efficiency of replication. The mechanism of this enhancement is not fully understood, and may involve inhibition of nucleosome formation, local unwinding of the duplex DNA or interaction between transcription factors and proteins involved in DNA replication (Wong and Schaffer, 1991).

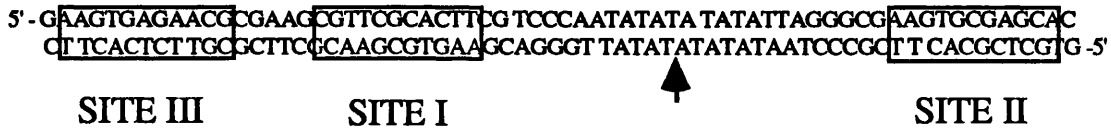


Figure 1A.4. Sequence of the HSV-1 Ori$_S$ Core Domain.
The sequence of the HSV-1 Ori$_S$ core domain, showing the 11bp consensus recognition elements of the three UL9 binding sites as open boxes. The arrow indicates the centre of a 53bp hyphenated palindrome. Reproduced from Challberg, 1991.

The elucidation of the genomic sequence of HSV-1 led to the identification of several genes proposed to be involved in viral DNA metabolism, including a thymidine kinase, a ribonucleotide reductase, a nuclease, a single-stranded DNA binding protein, an origin binding protein, a primase and a helicase (McGeoch *et al.*, 1988a). Subsequent studies have led to the identification of seven genes which are necessary and sufficient for viral DNA replication (Challberg, 1986; Wu *et al.*, 1988; McGeoch *et al.*, 1988b). The genes which are essential for viral DNA replication are UL5, UL8, UL9, UL29, UL30, UL42 and UL52. Expression of these seven genes from baculovirus vectors allows the replication of viral origin-containing DNA in insect cells and may provide a basis for the biochemical analysis of the gene products (Stow, 1992).

The HSV-1 UL9 gene encodes the origin-binding protein (Elias *et al.*, 1986; McGeoch *et al.*, 1988a). Footprinting and mutational analyses have shown that the protein binds to three sites within the oriS minimal region, designated sites I-III (Weir *et al.*, 1989; Olivo *et al.*, 1988; Elias *et al.*, 1990). These three binding sites, together with the AT-rich sequence, are essential for the full function of the origin of replication.

Binding of UL9 to sites I and II is co-operative (Elias *et al.*, 1990) and binding of the protein to the two sites induces conformational changes in the central AT rich region

(Lockshon and Galloway, 1988; Koff *et al.*, 1991; Challberg, 1991). Distortion of the central AT tract requires the situation of two adjacent UL9 binding sites on the same side of the DNA helix (Lockshon and Galloway, 1988; Challberg, 1991), and may involve interaction between two UL9 molecules bound to these sites (Fierer and Challberg, 1992a).

The *in vivo* function of the UL 9 protein is still unclear. The DNA binding function of the protein has been mapped to the carboxy-terminal 317 amino acids of the protein (Weir *et al.*, 1989; Deb and Deb, 1991). The protein exists as a dimer in solution and has been shown to exhibit limited helicase activity, which is enhanced by the addition of ICP8 (Fierer and Challberg, 1992b). The amino-terminal two-thirds of the protein contains six sequence motifs conserved within a superfamily of DNA and RNA helicases. Mutational analysis of the protein has shown that five of these conserved motifs are essential for the full function of the protein in transient transfection assays (Martinez *et al.*, 1992).

The products of the UL30 and UL42 genes form the functional viral DNA polymerase. Although the UL30 protein (the *pol* protein) is able to support DNA synthesis in the absence of UL42, the protein is isolated predominantly from infected cells as a heterodimer with the UL42 polypeptide, and UL42 acts to increase the processivity of the HSV-1 DNA polymerase in *in vitro* assays (Crute and Lehman, 1989; Gottlieb *et al.*, 1990; Hernandez and Lehman, 1990). The presence of the UL42 subunit increases the affinity of the UL30 subunit for DNA by 5 to 10-fold, decreasing the rate at which DNA polymerase disassociates from the template (Challberg, 1991).

The products of the UL5, UL8 and UL52 genes are subunits of the viral helicase-primase complex (Crute *et al.*, 1989; Dodson *et al.*, 1989; Crute and Lehman, 1991). A complex of the UL5 and UL52 gene products contains the helicase and primase activities (Calder and Stow, 1990; Dodson and Lehman, 1991). The UL5 gene product contains six motifs which are conserved within a superfamily of DNA and RNA helicases. Mutations within these motifs inhibit the helicase properties of UL5, both *in vitro* and *in vivo* (Zhu and Weller, 1992). However, the UL5 protein does not express helicase activity in the absence of the UL52 protein (Calder and Stow, 1990; Dodson and Lehman, 1991). It seems that UL8 acts to increase the efficiency of primer utilization by increasing the stability of the primer-template complex.

The final gene involved in viral DNA replication, the UL29 gene, has long been recognised to encode the viral major DNA binding protein (*dbp*), ICP8 (Bayliss *et al.*, 1975; Purifoy and Powell, 1976). Binding studies with this protein have illustrated that it exhibits high affinity for single-stranded DNA and that binding is non-sequence specific (Ruyechan and Weir, 1984). The protein acts to stimulate the activity of viral DNA polymerase (Ruyechan and Weir, 1984), is required for the synthesis of long stretches of DNA (Hernandez and Lehman, 1990) and specifically increases the DNA helicase

activity of the UL9 protein (Fierer and Challberg, 1992b). ICP8 may form multiple interactions with other replication proteins. Immunofluorescence studies have illustrated that functional ICP8 is required for the reorganisation of the cellular DNA replication machinery into viral replicative compartments within the nucleus (deBruyn Kops and Knipe, 1988), and directs DNA polymerase (UL30) to specific pre-replicative sites within the nucleus (Bush *et al.*, 1991).

1A.5.4. Maturation and Egress of HSV-1.

The replication of HSV-1 DNA results in the production of concatemeric DNA molecules consisting of head to tail repeats of the viral genome. The packaging of viral DNA involves the cleavage of concatemeric molecules into viral genomic molecules, a function of the viral 'a' sequences (Stow *et al.*, 1983; Deiss *et al.*, 1986; Varmuza and Smiley, 1985; Smiley *et al.*, 1992). Once cleaved, the viral genome is inserted into preformed nucleocapsids (Perdue *et al.*, 1976; Ladin *et al.*, 1980; Preston *et al.*, 1983). Analysis of defective HSV-1 particles suggested that only capsids containing full genomic length DNA molecules were processed further to form fully mature virion particles (Vlazny *et al.*, 1982).

The elucidation of the process of viral capsid maturation has been promoted by the identification of three forms of viral capsids within the infected cell nucleus. These forms have been termed light or A, intermediate or B and heavy or C capsids (Perdue *et al.*, 1976; Rixon, 1992). A capsids and B capsids contain no viral DNA. C capsids contain viral DNA and represent the fully mature virion nucleocapsid, whilst B capsids are intermediate structures formed during assembly.

Initial studies showed that B nucleocapsids contained large amounts of a structural or scaffolding protein termed VP22a (also termed p40 or ICP35) which is lost upon maturation to C capsids (Preston *et al.*, 1983; Rixon *et al.*, 1988). Analysis of a temperature sensitive mutant of HSV-1 mapped the VP22a open reading frame to the gene UL26 (Preston *et al.*, 1983). Further analysis has identified a gene designated UL26.5 which is present within, and in frame with, the UL26 open reading frame. The two reading frames have unique 5' termini but share a common 3' terminus, resulting in the expression of two proteins with identical carboxy-terminal portions (Liu and Roizman, 1991a).

The product of the UL26 gene has been shown to exhibit proteolytic properties. This protease will cleave the UL26.5 gene product at a region close to the carboxy-terminal of the protein, yielding the VP22a polypeptide. It will also catalyse a proteolytic cleavage within the UL26 protein, resulting in the formation of the VP24 and VP21 proteins (Liu and Roizman, 1991b; Liu and Roizman, 1992; Preston *et al.*, 1992; Deckman *et al.*, 1992). The cleavage sites are illustrated in Figure 1A.5. Both VP21 and

VP24 have been shown to be components of HSV virions (Davison *et al.*, 1992). However, only VP24 remains in fully mature nucleocapsids.

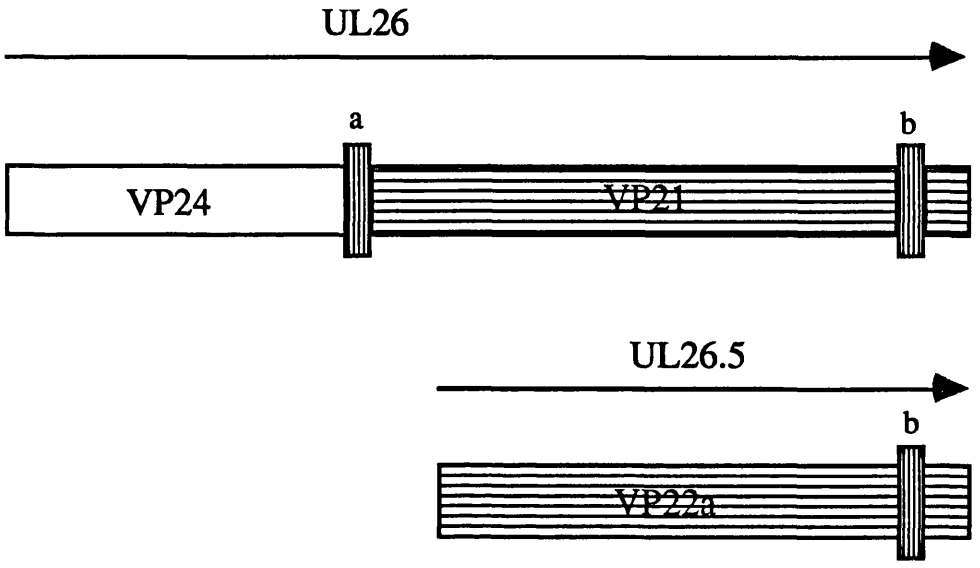


Figure 1A.5. The Organization of the UL26 and UL26.5 genes.
The UL26 and UL26.5 open reading frames are indicated by arrowed lines. Proteolytic cleavage sites of the gene products are ishown as the boxes a and b. Cleavage of the UL26 gene product results in the proteins VP21 and VP24, whilst VP22a is the result of proteolytic cleavage of the UL26.5 gene product.

The assembly of nucleocapsids appears to follow a pathway beginning with the formation of large-cored B capsids containing viral structural proteins. These capsids are processed to small-cored B capsids in a process that requires the proteolytic activity of UL26 and production of VP22a (Preston *et al.*, 1983). DNA is then packaged into these B capsids, triggering the removal of VP21 and VP22a proteins and resulting in the formation of fully mature C capsids, which then acquire a tegument and envelope during egress from the cell (Rixon, 1992).

The site at which virion nucleocapsids acquire their tegument and envelope is the subject of much speculation. Roffman *et al.*, (1990) have suggested that HHV-6 nucleocapsids acquire their tegument at a structure called the tegusome, membrane bounded cytoplasmic structures present at the outer nuclear membrane. Passage of HSV-1 in cell culture results in the production of L particles, consisting of enveloped tegument structures (Szilagyi and Cunningham, 1991). These L particles have been observed to be formed in the absence of capsid assembly (Rixon *et al.*, 1992), illustrating that tegumentation can occur independently of other assembly processes.

Initial analyses of lipid metabolism in infected cells suggested that herpesvirus capsids acquired their envelope at the nuclear membrane (Asher *et al.*, 1969). However, during multiple electron micrography studies, unenveloped capsids were observed within

the infected cell cytoplasm, often juxtaposed to cytoplasmic membranes (reviewed in Roizman and Sears, 1990). Two explanations have been proposed for the occurrence of unenveloped nucleocapsids within the cell cytoplasm. Firstly, that the viral capsid undergoes a series of envelopments and de-envelopments during its egress from the cell (Stackpole, 1969; Roffman, *et al.*, 1990). Secondly, that the unenveloped cytoplasmic capsids were products of fusion between the viral envelope, acquired at the nuclear membrane, and the membrane of cytoplasmic vesicles in which viral particles were transported to the cell surface (Campadelli-Fiume *et al.*, 1990; 1991). The second suggestion is based upon the observation that viral gD interferes with fusion of plasma membranes, viruses containing mutant forms of gD resulting in a higher proportion of unenveloped cytoplasmic capsids. However, previous analysis of a gD^- mutant virus suggested that the lack of gD did not interfere with the egress of virus from the cell (Ligas and Johnson, 1988).

The majority of evidence contained in these studies involves electron micrograph analyses of infected cells. These photographs provide frozen images of viral infection, but yield no indication of the direction of nucleocapsid transport. This presents difficulties in the interpretation of such images.

1A.6. The Latent Cycle of HSV-1.

HSV infection is characterized by the the ability of the virus to establish a latent infection in host cells, predominantly in neuronal ganglia. This latent infection was initially studied using mouse latency models, in which latent virus was reactivated by transplantation of ganglia. The use of this technique has also illustrated the existence of latent virus in human ganglia (reviewed in Ho, 1992). More recently, the development of *in vitro* models of latency have provided a means for the biochemical analysis of the latent state of the virus (Wighdal *et al.*, 1984; Russell and Preston, 1986). Studies of the activity of mutant HSV-1 viruses in these *in vitro* models have suggested that Vmw110 plays a part in the reactivation from latency (Russell *et al.*, 1987b; Harris *et al.*, 1989).

The HSV genome has been shown to exist in an endless or concatemeric state in latently infected cells (Rock and Fraser, 1983; 1985; Efstathiou *et al.*, 1986). Analysis of cellular DNA by CsCl gradient centrifugation has illustrated that the viral DNA exists as an extrachromosomal molecule (Mellerick and Fraser, 1987). The latent viral genome is packaged into nucleosomes, and this has been suggested as a mechanism by which viral transcription is limited (Deshmane and Fraser, 1989).

During latency, viral transcription is limited to a region within the long repeat of the genome which overlaps with the 3' portion of the open reading frame for the viral

immediate-early protein, Vmw110. Two transcripts from this region were originally detected, termed the latency associated transcripts or LAT's (Stevens *et al.*, 1987; Deatly *et al.*, 1987; see Figure 1A.2). They are colinear transcripts of approximately 2.0Kb (the major LAT) and 1.5Kb (the minor LAT). The minor LAT is produced from the major LAT by a splicing reaction (Wagner *et al.*, 1988b; Wechsler *et al.*, 1988).

Initially, it was proposed that the two transcripts contained a conserved open reading frame. However, the transcripts exhibit no evidence of polyadenylation (Wagner *et al.*, 1988a, b), and although a protein product has been detected, it is not of the predicted molecular weight (Doerig *et al.*, 1991).

Recent *in situ* hybridisation studies have allowed the detection of an RNA molecule which extends upstream of the major LAT. region. Further analysis of this transcript has shown that it is a 8.3Kb polyadenylated species. As the LAT's contain putative splice signals, and do not appear to be polyadenylated, it has been suggested that they are stable introns formed from the splicing of this large primary transcript (Zwaagstra *et al.*, 1990).

The LAT promoter has been mapped by a series of mutational analyses and transient transfection assays to a region approximately 30bp upstream from the large primary transcript (Dobson *et al.*, 1989; Batchelor and O'Hare, 1990; 1992; Zwaagstra *et al.*, 1990; 1991). Deletion of this region results in the loss of latency-associated transcription. The promoter contains many elements common to RNA polymerase II promoters and also additional elements with a possible involvement in transcriptional regulation. These include a TATA box, potential CAAT box, several Sp1 sites, a Vmw175 binding site (Batchelor and O'Hare, 1990), a binding site for a LAT promoter binding factor (LPBP, Zwaagstra *et al.*, 1990), and a cAMP resonse element (Leib *et al.*, 1991). The interaction of Vmw175 with its proposed binding site has been shown to down-regulate transcription from the promoter, and may be involved in repression of latency during the lytic cycle (Batchelor and O'Hare, 1990). Binding of the LPBF seems to up-regulate transcription in neuronal and non-neuronal cell lines (Zwaagstra *et al.*, 1990).

A region upstream of the proposed minimal promoter has been suggested to confer neuronal specificity on the transcription unit (Zwaagstra *et al.*, 1990; Batchelor and O'Hare, 1990; 1992). Batchelor and O'Hare (1992) have mapped a binding site for a neuronal specific factor, the IC-1 factor, to a 70bp region within this upstream element.

In addition, the cAMP response element mapped by Leib *et al.*, (1991) may play a role in the mechanism of reactivation. Deletion of this element results in a significant delay in reactivation *in vivo*, whilst induction of intracellular cAMP increased the activity of the latency-associated promoter. Interestingly, many physiological and external stimuli which result in the reactivation of latent HSV also lead to an increase in intracellular cAMP levels.

The function of the LAT's in the mechanism of latency establishment, maintenance and reactivation is still unclear. It is possible that they function as antisense messengers to the Vmw110 open reading frame, thus inhibiting its function in reactivation (Stevens *et al.*, 1987); or else they represent a species of RNA molecule which are transcribed upon reactivation (Steiner *et al.*, 1989). The analysis of mutations within the LAT promoter region which disrupt transcription has demonstrated that the LAT's do not appear to function in latency establishment or maintenance. However, mutant viruses lacking LAT's displayed delayed reactivation kinetics (Steiner *et al.*, 1989; Sedarati *et al.*, 1989; Hill *et al.*, 1990). This observation does not correlate with the proposed role of the LAT's as antisense messengers. It has been suggested that the LAT's act to express a protein that is essential for reactivation (Hill *et al.*, 1990).

Establishment of HSV latency may involve a reduction in activity of the viral transactivating protein Vmw65 (Roizman and Sears, 1990). This hypothesis is supported by the observation that a viral mutant defective in Vmw65 activity can establish latency *in vitro* and is reactivated by Vmw110 but not Vmw65 (Harris and Preston, 1991). This mutant virus is also impaired in lytic infection and is predisposed to latency *in vivo* (Steiner *et al.*, 1990). Kemp *et al.* (1990) have reported the existence of an octamer binding protein distinct from the protein Oct-1. This protein is present predominately in neuronal cells, and appears to interfere with the function of Vmw65. It is possible that this protein plays a role in the establishment of latency.

1B. An Overview of HSV-1 Gene Expression.

1B.1. Temporal Regulation of HSV-1 Gene Expression.

Upon infection, HSV genes are transcribed by the host cell RNA polymerase II (Costanzo *et al.*, 1977; Ben-Zeev and Becker, 1977). The analysis of polypeptides synthesized during the course of HSV-1 lytic infection led to the identification of three co-ordinately regulated groups of genes, characterized by their specific temporal expression (Honess and Roizman, 1974; 1975; Clements *et al.*, 1977). The α or immediate-early polypeptides are synthesized immediately upon infection, whereas β or early proteins are synthesized following the expression of immediate-early genes, reaching a maximum at 5-7 hours post infection. The final class of genes, the γ_2 or true late genes are synthesized following the onset of viral DNA replication at approximately

2 hours post infection. Late gene expression increases to a maximum at approximately 12 hours post infection. An additional subclass, termed γ_1 or leaky-late genes, are expressed at low levels prior to viral DNA synthesis, but require the onset of viral DNA replication for maximal expression (Wagner, 1985).

The use of metabolic inhibitors has led to the elucidation of certain biochemical aspects of the temporal cascade of HSV-1 gene expression (Honess and Roizman, 1974; 1975; Clements *et al.*, 1977; Jones and Roizman, 1979). The immediate-early viral proteins are the only gene products expressed in the absence of other viral protein synthesis. These genes are controlled in part by a *trans*-activating protein present in the virion particle (see Section 1B.2 below). The transcription of early genes requires the presence of viral immediate-early proteins, as illustrated by the lack of early gene expression during inhibition of viral protein synthesis. Finally, expression of the true late genes occurs upon the induction of DNA synthesis, and also requires the presence of immediate-early viral proteins (Sacks *et al.*, 1985; McCarthy *et al.*, 1989; Preston, 1979; Watson and Clements. 1980; Dixon and Schaffer, 1980; Everett, 1986). Figure 1B.1 illustrates the viral proteins known to be involved in the control of HSV gene expression, and indicates their role in the temporal cascade.

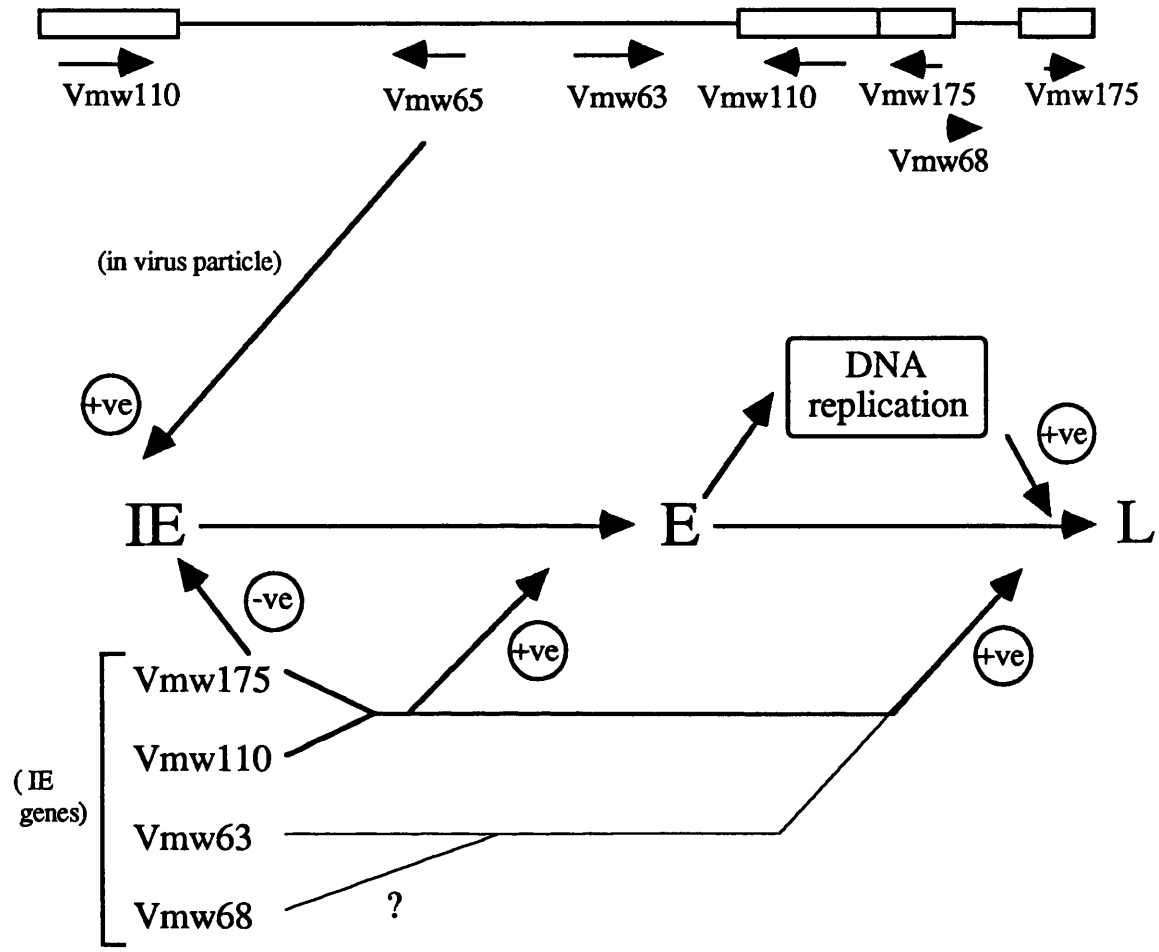


Figure 1B.1. HSV Gene Products Which Regulate Transcription.
The top line represents the HSV-1 genome, indicating the positions of the genes encoding transcriptional regulatory proteins (not drawn to scale). The role of each protein in the regulation of viral transcription is indicated. Adapted from an original diagram by R. D. Everett.

1B.2. Regulation of Immediate-Early Gene Expression.

1B.2.1. Sequences Within the Immediate-Early Genes Required for Viral-Induced Transactivation.

The positions of all five IE gene transcripts within the HSV-1 genome have now been mapped (Clements *et al.*, 1979; Rixon *et al.*, 1982; Murchie and McGeoch, 1982; McGeoch *et al.*, 1985; 1986; 1988b; Perry *et al.*, 1986). The IE genes are concentrated within the two repeat regions of the genome, with only the IE-2 gene located within the

unique long segment. Although HSV RNA is rarely spliced (Wagner, 1985), three of the five IE proteins are translated from spliced mRNAs. The RNAs for the IE4 and IE5 genes are transcribed from a common promoter situated within the short repeat region of the genome. The common leader sequences of the genes are spliced to individual open reading frames situated at either end of the short unique segment.

Analysis of the four HSV-1 IE promoters has defined two separable regions (Mackem and Roizman, 1982a; 1982b; Cordingley *et al.*, 1983; Preston *et al.*, 1984; Kristie and Roizman, 1984; Bzik and Preston, 1986): (i) the minimal promoter region consisting of a functional TATA box and proximal upstream sequences which directs basal transcription, and (ii) an upstream enhancer element which mediates the transactivation of IE gene expression by a component of the virion particle, later shown to be the virion protein Vmw65 (see Section 1B.2.2 below).

Sequence analysis of the IE promoters initially defined two conserved motifs within the IE enhancer regions - the TAATGARAT motif, and a GC rich motif (Mackem and Roizman, 1982b; Murchie and McGeoch, 1982; Preston *et al.*, 1984). Transfection experiments with mutant and chimaeric forms of HSV-1 IE and E promoters illustrated that these and possibly additional regulatory elements conferred Vmw65-mediated transactivation on both minimal IE and E promoters (Preston *et al.*, 1984; Kristie and Roizman, 1984). The linkage of an oligonucleotide containing the TAATGARAT motif to the IE4/5 minimal promoter sequences restored Vmw65-mediated transactivation (Gaffney *et al.*, 1985).

Sequences other than the defined TAATGARAT motif are also required for optimal expression of the IE genes. A GA-rich sequence upstream of the TAATGARAT motif within the IE3 promoter was shown to be necessary for the full activation of the promoter by Vmw65 (Triezenberg *et al.*, 1988a; 1988b; Bzik and Preston, 1986). The IE4/5 and IE2 promoters also contained motifs with homology to this region. Proteins which bind specifically to this region, termed the $(IEF)_{ga}$ or IE facilitator proteins have been purified. However, these proteins do not seem to mediate transactivation of IE genes by the Vmw65 protein (LaMarco and McKnight, 1989).

Analysis of the GC rich sequences illustrated the presence of GC box motifs, which had previously been identified as the binding site for the cellular transcription factor Sp-1. DNase I footprinting analysis of the HSV-1 short repeat region showed that Sp-1 bound to the GC boxes within the IE3 and IE4/5 promoters (Jones and Tjian, 1985). These Sp-1 sites increase the efficiency of expression of both IE and E promoters (see also Section 1B.3).

 origin of DNA replication

1B.2.2. Identification and Characterization of the Proteins Interacting with the Regulatory Sequences of IE Genes and Mediating IE Transactivation.

Early experiments showed that a protein component of the virion particle mediated IE gene transactivation, whilst further transfection assays illustrated that the mediator of viral IE transactivation was the virion protein Vmw65, also termed VP16 or α-TIF (Post *et al.*, 1981; Campbell *et al.*, 1984). This regulatory protein is located within the tegument of the virion particle (Batterson and Roizman, 1983). The observation that this protein exhibited no DNA binding properties raised questions about the function of the TAATGARAT motif within the IE promoter enhancer region (Marsden *et al.*, 1987; Preston *et al.*, 1988; O'Hare *et al.*, 1988).

The discovery that the cellular protein NF-III bound to a TAATGARAT motif within an adenovirus [illegible] (Pruijn *et al.*, 1986) suggested that HSV-1 IE transactivation by Vmw65 may be mediated by a cellular DNA-binding factor. In fact, it was soon discovered that a cellular factor did bind to the HSV-1 IE TAATGARAT motif. This factor was termed α-H1 (Kristie and Roizman, 1987); H3 (Preston *et al.*, 1988) and TRF (O'Hare and Goding, 1988). By the analysis of the sequences required for the formation of this complex, O'Hare and Goding (1988) identified the cellular factor as the factor NF-III, also termed Oct-1 or OTF-1.

The addition of an infected cell extract to this binding reaction resulted in the formation of a specific infected cell complex (Preston *et al.*, 1988; O'Hare *et al.*, 1988). The use of an anti-Vmw65 antibody and purified Vmw65 demonstrated that the specific infected cell complex contained Vmw65. Mutational analysis of the TAATGARAT motif illustrated that the first four base pairs of the motif mediated the interaction with OTF-1, whilst mutations within the GARAT portion of the motif interfered with the formation of the infected cell specific complex without inhibiting the DNA-binding activity of OTF-1. These mutations also inhibit the Vmw65 mediated activation of IE genes, illustrating the involvement of the infected cell complex in the mechanism of transactivation of IE genes (O'Hare *et al.*, 1988; O'Hare and Goding, 1988).

The mechanism by which the OTF-1 factor mediates the interaction of Vmw65 with DNA is poorly understood. It is possible that the interaction of the two proteins creates a conformational change within Vmw65 which facilitates its interaction with DNA. Alternatively, the binding of OTF-1 to the TAATGARAT motif may induce a conformational change in the DNA helix, again allowing the Vmw65 protein to bind to DNA. Finally, an additional cellular protein may interact with the 3' portion of the TAATGARAT motif, allowing the recruitment of Vmw65 to the complex.

Some experimental evidence exists to support the latter hypothesis. The analysis of complex formation in the absence of DNA has suggested that Vmw65 interacts with a cellular protein distinct from OTF-1 (Goding and O'Hare, 1989). In addition, several

groups demonstrated that cellular factors present in partially purified cellular extracts were required for formation of the infected cell complex with affinity purified Vmw65 and OTF-1 (Gerster and Roeder, 1988; Kristie *et al.*, 1989; Katan *et al.*, 1990). Katan *et al.* (1990) identified and purified a complex forming factor (CFF) from HeLa cells. The CFF was found to be a protein of approximately 70kDa molecular weight, which exhibited no DNA binding activity in isolation, but which allowed the interaction of OTF-1 and Vmw65 with the full TAATGARAT sequence.

A HSV-1 virus (*in*1814) has been constructed which encodes a mutant Vmw65 protein (Ace *et al.*, 1989).This Vmw65 molecule is fully functional in its role in virion structure, but is unable to interact with OTF-1. The virus has a high particle:pfu ratio, and is unable to replicate at low multiplicities of infection *in vitro*, whilst it is essentially avirulent *in vivo*. Loss of Vmw65-OTF-1 interaction results in a 5-10 fold reduction in expression of IE1 and IE4/5, whilst the IE2 and IE3 genes show only a small or negligible reduction in expression. The properties of this virus mirror those of a Vmw110 deletion mutant virus (Stow and Stow, 1986; Sacks and Scaffer, 1987; Everett, 1989), and may reflect the inability of *in*1814 to efficiently express Vmw110.

1B.2.3. Regions of the Vmw65 Protein Involved in Transactivation.

Mutational analysis of the Vmw65 protein has identified domains involved in both infected cell complex formation and IE gene transactivation, again suggesting that complex formation is an essential step in the mechanism of IE gene transactivation. Certain mutant Vmw65 proteins able to support IE gene expression were unable to rescue an HSV-2 assembly mutant virus, suggesting that another domain of Vmw65 is involved in the structural function of the protein. These results may represent a multiplicity of interactions between Vmw65 and cellular or viral proteins - one set of interactions mediating IE gene transactivation, whilst another set of interactions mediate virion assembly (Ace *et al.*, 1988).

Deletion of the C-terminal 80 amino acids of Vmw65 results in the loss of transactivating function of the protein, whilst the N-terminal 403 amino acids of the protein still retain the ability to form a cell specific complex (Greaves and O'Hare, 1989). The C-terminal 80 amino acid domain of Vmw65 contains a putative highly acidic amphipathic α-helix, which exhibits homology to the transcriptional domains of the yeast transcription factors GAL4 and GCN4 (Sadowski *et al.*, 1988; Goding and O'Hare, 1989). Deletion of this transcriptional domain results in a mutant form of Vmw65 which interferes with the transcriptional ability of wild type Vmw65, presumably by competing with wild type Vmw65 for the formation of a transactivation competent infected cell complex (Triezenberg *et al.*, 1988a; Friedman *et al.*, 1988).

It is believed that the acidic activator domain of Vmw65 acts through the interaction with one or more components of the basal transcription apparatus, although the identity of the interacting cellular factor(s) is not clearly defined. The mechanism of transactivation by Vmw65 is discussed in further detail in Section 1D.2.1.

1B.3. Regulation of HSV-1 Early Genes.

Analysis of temperature-sensitive mutant viruses has illustrated that Vmw175 function is essential for the activation of early and late genes (Marsden *et al.*, 1976; Courtney *et al.*, 1976; Watson and Clements, 1978; 1980; Preston, 1979a; Dixon and Schaffer, 1980), whilst further transient transfection assays illustrated that two out of the five IE gene products, Vmw175 and Vmw110, were sufficient for the transactivation of several early promoters. These two proteins were able to transactivate early gene expression independently. In addition, several studies illustrated that the two proteins acted synergistically in early gene transactivation whilst other studies did not observe this synergistic cooperation (Everett, 1984b; 1986; O'Hare and Hayward, 1985a, b; Gelman and Silverstein, 1985; Quinlan and Knipe, 1985). These differences in the *in vitro* properties of Vmw175 and Vmw110 may be explained by the observation of Everett (1988b) that experimental details such as cell type and transfection protocol can affect the transactivation properties of viral proteins in this assay. Although these experiments illustrated that the two proteins may interact to increase early gene expression, the use of a cell line stably expressing Vmw175 from the cellular genome illustrated that Vmw175 was sufficient to activate certain early promoters within a superinfecting viral genome (Persson *et al.*, 1985).

Some aspects of the mechanism of early gene transactivation by the two proteins are still unclear. Although Vmw110 activates early gene expression independently of Vmw175 in transfection assays, viruses containing temperature sensitive mutations in Vmw175 exhibit no early gene transcription at the non-permissive temperature (see for example, Dixon and Schaffer, 1980; Watson and Clements, 1980; DeLuca and Schaffer, 1985). It has been suggested that this property is due to the trans-dominance of the mutant Vmw175 molecule over the wild type Vmw110 (Gelman and Silverstein, 1986; Everett, 1987b). In fact, some temperature-sensitive mutant forms of Vmw175 accumulate in the cytoplasm and seem to interfere with the transport of Vmw110 to the nucleus (Knipe and Smith, 1986), whilst similar mutant forms of Vmw175 inhibit the transactivation activity of Vmw110 in transient expression assays (Gelman and Silverstein, 1986). However, this hypothesis cannot account for the fact that Vmw110 is not sufficient to support the viral lytic cycle in the absence of functional Vmw175

(DeLuca *et al.*, 1985; DeLuca and Schaffer, 1988; Shepard and DeLuca, 1989; this thesis).

Initial studies of HSV-1 early gene expression concentrated on the elucidation of sequences within early viral promoters specific for the viral trans-inducing activity. Such sequences have been discovered within viral IE promoters; these mediate the trans-inducing effect of the viral protein Vmw65 (see Section 1B.2.1).

Several groups constructed cell lines biochemically transformed with the viral thymidine kinase (tk) gene. This technique allowed the mutational analysis of the viral tk promoter, which could be assayed for activity both constitutively within the cellular genome, and also following transactivation by super-infecting virus. These studies illustrated the importance of the integrity of the 109bp fragment upstream of the tk capsite for both cellular and viral trans-induced expression of the gene (Smiley *et al.*, 1983; El Kareh *et al.*, 1985). However, such studies were limited by the integration of multiple copies of the tk gene in the cellular genome, and by the selection of tk expression, which may force the promoter into activity.

The development of transient expression techniques allowed a more sensitive analysis of the essential sequences of viral promoters. Studies have concentrated on two early HSV-1 promoters - the gD and tk promoters. For clarity, studies involving each promoter will be dealt with separately.

1B.3.1. Analysis of the Viral Glycoprotein D (gD) Promoter.

Transfection of reporter constructs containing gD promoter sequences followed by superinfection with HSV-1 resulted in the activation of the transfected promoter by viral proteins. Deletion mutational analysis of the gD promoter in this assay system illustrated that a fragment 83bp upstream of the gD capsite was required for efficient viral transactivation (Everett, 1983). The linkage of the SV40 enhancer region upstream of the gD minimal promoter region also allowed for the analysis of cellular mediated cis-activation of the promoter. Further mutational analyses demonstrated that transactivation by viral proteins required no additional sequences than those already found to be essential for cellular-mediated cis-activation (Everett, 1984a). These sequences included a G-rich sequence, an inverted CAAT box (described below), a TATA box and a functional cap-site. Therefore, it appeared that transactivation of the HSV-1 gD promoter required no specific sequences other than those previously utilized by the cellular transcription machinery.

The discovery that Vmw175 exhibited DNA binding properties (Freeman and Powell, 1982; Faber and Wilcox, 1986) led to the investigation of its ability to bind to viral promoter sequences. These studies identified three specific Vmw175 binding sites within the gD promoter region - two upstream and one downstream of the capsite (Faber

and Wilcox, 1986; Tedder and Pizer, 1988; Tedder *et al.*, 1989). The removal of the proximal upstream binding site reduced the transactivation of the promoter in a transient transfection assay, whilst the insertion of multiple Vmw175 binding sites increased the activation by Vmw175 (Tedder and Pizer, 1988). Further analysis illustrated that deletion of any one of the three Vmw175 binding sites reduced the transactivation of the promoter by Vmw175 in an *in vitro* transcription assay (Tedder *et al.*, 1989). Although these results implied that the specific Vmw175 binding sites may play a role in the Vmw175-mediated transactivation of the gD promoter, the removal of all three binding sites within the viral genome had little or no effect on gD expression during HSV-1 infection in cell culture (Smiley *et al.*, 1992). In addition, the Vmw175 binding site investigated above is not contained within the sequence shown by Everett (1983, 1984a) to be necessary and sufficient for gD transactivation. Therefore, as concluded by Everett (1983, 1984a), it seems that no specific sequences within the gD promoter mediate transactivation by Vmw175.

1B.3.2. Analysis of the Viral Thymidine Kinase (tk) Promoter.

Microinjection of a reporter construct containing HSV-1 tk promoter sequences demonstrated that this promoter was expressed constitutively in *Xenopus laevis* oocytes (McKnight *et al.*, 1981). Mutational analysis of the promoter region demonstrated that the constitutive expression required the integrity of four sites within the promoter - the TATA box and three upstream sequences - two GC boxes and one CAAT motif (McKnight and Kingsbury, 1982). These upstream elements were shown to bind two cellular factors, a CAAT-binding transcription factor (CBP or CTF) and the cellular transcription factor Sp-1, which has also been demonstrated to bind to and transactivate HSV-1 IE promoters (Jones *et al.*, 1985; Jones and Tjian, 1985).

Transient assays of the viral-mediated transactivation of the tk promoter, as described for the gD promoter in Section 1B.3.1, have allowed the delineation of promoter sequences essential for viral transactivation of the tk gene. These studies demonstrated that viral-mediated transactivation required only the sequences defined by McKnight and Kingsbury (1982). This again suggested that transactivation of the HSV-1 tk promoter required no specific sequences other than those previously utilized by the cellular transcription machinery. The introduction of mutations into the tk promoter within the HSV-1 genome illustrated that deletions within the upstream elements reduced tk expression 2-5 fold, whilst deletion of the tk TATA box virtually abolished expression from the promoter (Coen *et al.*, 1986).

The combined studies described above have suggested that Vmw175-mediated transactivation of early viral promoters requires a functional TATA box together with upstream binding sites for cellular factors such as CBP or Sp-1. However, Imbalzano *et*

al (1991) have carried out an analysis of the effects of mutations within the viral tk promoter in the presence and absence of Vmw175. They discovered that whilst the removal of the CAAT motif had a similar effect both in the presence and absence of Vmw175, deletion of the Sp-1 binding sites resulted in a 20-fold reduction in expression in the absence of Vmw175. The addition of Vmw175 in *trans* limited this effect to a modest 3-fold reduction in activity, similar to the results obtained by Coen *et al.* (1986). The authors speculated that these results illustrated the ability of Vmw175 to functionally substitute for Sp-1 in transactivation of viral early genes. As Sp-1 acts to increase cellular transcription through interaction with an adaptor or coactivator molecule (Pugh and Tjian, 1990), the authors postulated that Vmw175 may interact with this same adaptor protein. Alternatively, the interaction of Vmw175 with the cellular transcriptional machinery may bypass requirements for an upstream factor. It is possible that although the upstream sites are not required for Vmw175 transactivation and regulation of the tk gene with early kinetics, functional Sp-1 may be required to activate tk expression to levels required for the viral replicative cycle.

In conclusion, although recent studies have cast doubt on the integral requirements of functional Sp-1 binding sites, transactivation of viral early promoters by Vmw175 requires certain upstream elements together with a functional TATA box. The absolute requirement of a functional TATA box supports the hypothesis that Vmw175 mediates viral gene activation through an interaction with the cellular TFIID transcription factor. This proposed mechanism for Vmw175 transactivation is described in fuller detail in Section 1D.

1B.4. Regulation of HSV-1 Late Genes.

Transactivation of the late genes of HSV-1 involves three out of the five immediate-early proteins. Analyses of temperature-sensitive mutant viruses illustrated that functional Vmw175 was essential for the activation of late genes (Marsden *et al.*, 1976; Courtney *et al.*, 1976; Watson and Clements, 1978; 1980; Preston, 1979a; Dixon and Schaffer, 1980), and that the lack of a functional Vmw63 molecule led to over-expression of early gene products and no detectable late gene expression (Sacks *et al.*, 1985). The analysis of Vmw63 deletion mutant viruses also suggested that the protein was essential for late gene expression and possibly involved in the down-regulation of early genes (McCarthy *et al.*, 1989). Transfection assays illustrated that Vmw175 and Vmw110 transactivated late gene promoters, whilst the effect of Vmw63 was varied (for example, see Everett, 1984b, 1986; Shapira *et al.*, 1987; Sekulovich *et al.*, 1988).

Recent data has suggested that Vmw63 enhances late gene expression through a post-transcriptional mechanism (see Section 1B.7.2).

The observations made about the expression of late genes within the viral genome suggest two possible modes of transactivation - either the promoters are transactivated by a virally-encoded protein synthesized after the onset of viral DNA replication, or transactivation is mediated by an event associated with viral DNA synthesis. Mavromara-Nazos and Roizman (1987) demonstrated that a *cis*-acting element mediated by viral DNA replication allowed the temporal expression of late genes. Cells were infected with HSV-1, and infection allow to proceed until late times were reached. The cells were then superinfected with HSV-1 in the presence of an inhibitor of DNA synthesis. It was found that late genes in the superinfecting and unreplicated template were not expressed. As the initial infection provided all virally-encoded factors, it was concluded that DNA replication was essential for late gene expression.

1B.4.1. *Cis*-acting Signals Within Late Gene Promoters.

A number of studies have been carried out in order to determine the specific *cis*-acting control signals within late promoters which control their kinetics of expression and mediate their requirement for template replication before the induction of transcription. Initial analyses involved the transformation of cells with reporter constructs containing late promoters fused to tk structural sequences. Superinfection of these transformed cell lines enabled the investigation of the activation of late gene promoters by viral proteins. All reporter constructs tested were found to be activated by HSV-1 superinfection. However, late promoters resident in the cellular genome were expressed with the kinetics of early promoters (Dennis and Smiley, 1984), and were expressed in the presence of DNA replication inhibitors (Silver and Roizman, 1985). Therefore, it appeared that the local environment of the promoter controlled its requirement for template replication prior to induction.

A comprehensive series of mutational studies have defined the region of the true late gC promoter required for temporal control of expression. Initially, a series of deletions within the promoter region were introduced into the gC locus of HSV-1. These studies defined a region between -34 and +124 of the gC promoter which was solely required for expression following DNA replication. Unlike the studies of early gene promoters (described in Section 1B.3), these studies identified no upstream sequences other than a functional TATA box required for late gene expression (Homa *et al.*, 1986). Replacement of the tk TATA box with that of the gC promoter conferred late kinetics of expression on the tk gene, whilst the tk TATA box alone was unable to support gene expression during viral infection. Further mutational analyses narrowed down the region required for late gene expression to a 15bp fragment encompassing the TATA box of the

gC promoter (Homa *et al.*, 1988; 1991). Although results obtained by Imbalzano and DeLuca (1992) did not correlate with this observation, it was suggested that these discrepancies were due to sequences surrounding the TATA box.

Transient expression experiments also confirmed that sequences 34bp upstream of the gC cap site were sufficient for the HSV-induced activation of the promoter (Shapira *et al.*,1987). The expression of the gC promoter from a plasmid template in this transient expression assay did not require the replication of template DNA. This mirrored the results obtained with transformed cell lines - that removal of late genes from the viral genome resulted in a loss of the requirement for template replication prior to transcription.

The conclusion that the gC promoter required only a functional TATA box for expression was supported by results obtained by Johnson and Everett (1986a, b). Deletion of the upstream sequences of the true late US11 promoter illustrated that a functional TATA box was sufficient for viral mediated-transactivation in a transient expression assay. This promoter required the replication of template DNA for expression. In addition, the deletion of sequences upstream of the gD TATA box and linkage of the promoter to a viral origin of replication allowed for the expression of gD at late times in infection (Johnson and Everett, 1986).

Figure 1B.2 is a diagrammatical representation of the promoter requirements of the three classes of HSV-1 genes. This demonstrates that, as viral infection proceeds, each class of viral promoter exhibits a lower complexity of upstream signals required for activation of expression. It seems that early genes require the interaction of certain cellular factors with upstream sequences to boost expression early in infection, whilst late classes of genes are capable of expression from a TATA box alone. The dilution of cellular factors during viral infection mediated by the shut-off of host cell protein synthesis could result in the shut-off of early gene expression.

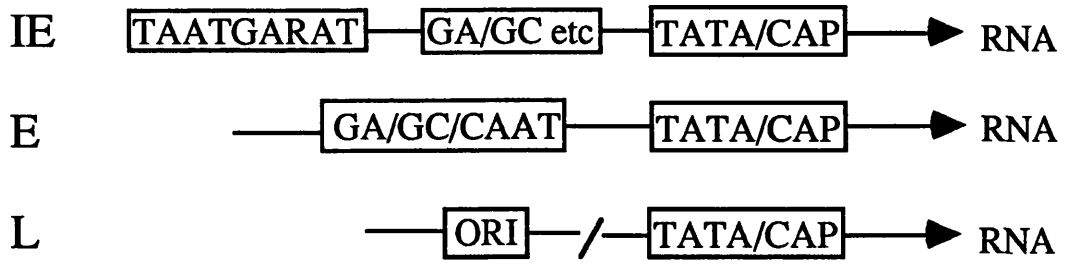


Figure 1B.2. The Essential Regions of the Three Classes of HSV-1 Gene Promoters.

The diagram shows representations of the three classes of HSV-1 gene promoters. All classes of promoters contain a TATA box and RNA CAP site. IE and E promoters in addition contain upstream regulatory regions, which bind a variety of cellular factors. In addition, IE genes also contain an upstream TAATGARAT regulatory region required for activation by the viral protein Vmw65. Late genes, in contrast, require only linkage to a functional origin of replication.

This theory implies that the TATA boxes incorporated within late gene promoters are stronger initiators of transcription, and is supported by several observations. Firstly, that the isolated tk TATA box is insufficient to direct transcription (Homa *et al.*, 1986; 1991), whilst the addition of the tk upstream elements to the gC TATA box resulted in the expression of gC early in infection (Homa *et al.*, 1988; 1991). In addition, gel shift experiments have illustrated that the gC TATA box has a stronger affinity for the TATA binding protein than the tk TATA box (Imbalzano and DeLuca, 1992).

However, this simplistic model of the sequence requirements within late gene promoters cannot explain the observations made during the mutational analysis of the US11 promoter sequences. This study illustrated that expression of the US11 gene late in infection required the incorporation of US11 cap/leader sequences (Kibler *et al.*, 1991). These observations are supported by other reports that non-coding 5' leader sequences increase the expression of certain late genes (Blair *et al.*, 1987; Flanagan *et al.*, 1991; Weir and Narayanan, 1990; Pederson *et al.*, 1992). These leader sequences may act through the stabilization of mRNA molecules, or through the binding of cellular or viral factors which activate initiation at the TATA box. Interestingly, Romanelli *et al.* (1992) have shown that incorporation of γ42 leader sequences into the tk promoter of HSV-1 increases the expression of the tk gene; whilst the same group has shown that these leader sequences contain a binding site for Vmw175 (Mavromara-Nazos *et al.*, 1987; Michael *et al.*, 1988). Mutation of the Vmw175 binding sites within the recombinant virus results in a loss of activation of the gene. This is the only reported example of the direct involvement of a Vmw175 binding site in the transactivation of a viral promoter (Romanelli *et al.*, 1992). The requirement of additional cap/leader

sequences for efficient expression may be an indication of the relative efficiency of the TATA box in directing transcription (Kibler *et al.*, 1991).

1B.4.2. A Model for the Role of Viral DNA Replication in Late Gene Expression.

Two hypotheses have been proposed to explain the requirement for viral DNA replication in late gene expression (Homa *et al.*, 1991). Both scenarios assume that all factors required for late gene transactivation are present in infected cells prior to the initiation of viral DNA replication.

The first model suggests that late gene expression is blocked by an inhibitory factor, and that the replication of viral DNA removes the inhibitor, allowing expression to proceed. This inhibitory block could be the chromatin structure of the viral DNA, or a bound cellular or viral protein. The hypothesis that chromatin structure plays a role in transactivation by viral gene products could also account for the expression of selected cellular genes during viral infection (see Section 1B.5). The cellular factors binding to the upstream elements of early genes could render the genes insensitive to this inhibition of transcription. Interestingly, Sp-1 has been shown to interfere with the chromatin-mediated repression of cellular transcription (Croston *et al.*, 1991).

Alternatively, activation of late genes may require the organization of the DNA into specific nuclear compartments. Knipe and co-workers have demonstrated that during viral DNA replication, certain viral proteins are co-localized with DNA into globular compartments termed replicative compartments (Knipe *et al.*, 1987; de Bruyn Kops and Knipe, 1988; Bush *et al.*, 1991). This localization may facilitate the expression of late genes.

1B.5. The Activation of Cellular Promoters by HSV-1 Products.

Latchman and colleagues have demonstrated that a subset of cellular genes are activated during HSV infection, and that expression of some of these genes requires the presence of functional Vmw175 (Patel *et al.*, 1986; Kemp and Latchman, 1988). The incorporation of viral promoters into the cellular genome has illustrated that they too can be activated by superinfection with HSV (Sandri-Goldin *et al.*, 1983; Dennis and Smiley, 1984; Silver and Roizman, 1985). This observation led to the suggestion that promoters within the cellular genome which are activated by HSV products share specific *cis*-regulatory sequences with viral promoters. However, it was found that the β–globin gene, which is not expressed during normal viral infection, was expressed by HSV products when it was introduced by transfection or incorporated into the genome of

transformed cells (Everett, 1985). The conclusions that can be drawn from such an experiment are limited by the possible incorporation of multiple promoter sequences into the cellular genome, and by the rather artificial conditions of transient expression assays, in which high numbers of promoter sequences are exposed to viral regulatory proteins for longer periods of time than the usual length of a viral replicative cycle. In order to reduce such experimental artifacts, the regulation of cellular promoters was tested within the environment of the viral genome.

Smiley *et al.* (1987) introduced the β-globin gene under the control of its own promoter or the early gD promoter into the tk locus of HSV-1. β-globin was expressed from both constructs with early kinetics, and required Vmw175 for expression. The transfer of the reporter construct to the gC locus had no effect on the kinetics of expression. These results support other conclusions that the expression of promoters by Vmw175 requires no additional *cis*-acting sequences other than those previously used by the cellular transcription machinery (see Section 1B.3).

In contrast, the transfer of the adenine phosphoribosyl transferase (*aprt*) gene into the viral genome results in the loss of transcription (Tackney *et al.*, 1984). This gene does not contain functional TATA box sequences. The analysis of viral early and late promoters has suggested that viral transactivation is mediated through the TATA box, and it seems likely that Vmw175 acts to increase transcription through an interaction with proteins that bind to the TATA box (see discussion in Section 1D.4). This would account for the inability of HSV transactivators to direct the expression of the *aprt* gene.

Further analysis has indicated that the α-globin gene which is normally expressed constitutively during transient expression assays in the absence of viral proteins is expressed with early kinetics when introduced into the viral genome (Smiley *et al.*, 1991). It is possible that the HSV genome exerts an inhibitory effect on the expression of viral genes, which is overcome by the action of viral or cellular transactivating proteins, or by DNA replication. Such a mechanism would account for the limited viral gene expression in the absence of IE gene products.

In conclusion, studies on the activation of cellular promoters by HSV products support the hypothesis that viral-mediated transactivation does not require the presence of specific *cis*-acting sequences within the viral promoter. Viral transactivators therefore act through sequences previously required for the cellular transcription machinery, probably by the modification of or interaction with pre-existing factors. As the major HSV transactivating protein is Vmw175, it is probable that the activation of cellular promoters within the viral genome is mediated through Vmw175. The absolute requirement of a functional TATA box demonstrated by the results of Tackney *et al.* (1984) supports the hypothesis that Vmw175 mediates viral gene activation through an interaction with the cellular TFIID transcription factor. This proposed mechanism for Vmw175 transactivation is described in fuller detail in Section 1D.4.

1B.6. Repression of Immediate-Early Gene Expression.

1B.6.1. Early Observations of Herpes Simplex Virus Immediate-Early Gene Expression.

Initial studies of protein synthesis during HSV-1 infection defined the immediate-early genes as those which were expressed during HSV-1 infection in the absence of *de novo* protein synthesis. These studies showed that IE protein synthesis began at 1h post infection, and suggested that synthesis was shut-off at 3-4h post infection (Honess and Roizman, 1974, 1975)

However, later studies questioned some of the conclusions made during these initial analyses. A detailed analysis of mRNA accumulation during HSV-1 lytic infection showed that the mRNA for the IE1 and IE3 genes could be detected at late times in infection. This analysis also illustrated that in the absence of protein synthesis, IE1 and IE3 mRNA accumulated to high levels, suggesting that some virally-encoded protein controlled the reduction of RNA transcription for these two genes at least (Harris-Hamilton and Bachenheimer, 1985).

Weinheimer and McKnight (1987) studied the rate of RNA transcription and mRNA accumulation for three IE genes. This study found that although IE1 and IE3 mRNA accumulated with similar kinetics, increasing from 1-6h post infection, the rate of transcription of only the IE3 gene was reduced after 1h post infection. In addition, the accumulation of the RNA from the IE4/5 promoter followed a similar pattern to the accumulation of early transcripts. During the inhibition of protein synthesis, whilst IE1 and IE3 mRNA accumulated, no accumulation of IE4/5 RNA was observed. These observations led the authors to suggest that HSV-1 IE expression can be regulated independently of transcriptional control.

1B.6.2. Transient Transfection Assays of Immediate-Early Gene Expression.

Several groups have demonstrated the transcriptional repression of certain IE promoters during transient expression assays (O'Hare and Hayward, 1985a, b, 1987; Gelman and Silverstein, 1985, 1986, 1987a, b; Resnick *et al.*, 1989; Everett, 1986; Everett and Orr, 1991). The repression of the promoters appears to be dependent on the relative ratio of Vmw175, Vmw110 and Vmw65 in the transfected cell (O'Hare and Hayward, 1985a, b, 1987; Gelman and Silverstein, 1986, 1987a, b). The repression of the promoters by Vmw175 alone differed - whilst the IE1 and IE3 promoters were strongly repressed, the IE2 promoter was weakly repressed and the IE4/5 promoter was not repressed at all (Gelman and Silverstein, 1987a, b). The interpretation of these assays was complicated by the widely varying results produced by duplicate assays

*see Section 1C.4.

carried out in differing cell lines. This again demonstrates the problems of short-term transfection assays, as illustrated by Everett (1988b).

It was found that the ability of Vmw175 to bind to IE promoter sequences mirrored its ability to repress transcription from the promoter. The IE3 promoter was shown to contain a strong Vmw175 binding site at its capsite which exhibited homology to the consensus sequence published by Faber and Wilcox (1986). This consensus sequence was also contained in a site within the upstream regulatory sequences of the IE1 promoter. The association of Vmw175 with the IE2 and IE4/5 promoters, in contrast, was found to be weak (Faber and Wilcox, 1988; Kristie and Roizman, 1986a, b). Whilst it has been proposed that Vmw175 represses IE3 transcription by associating strongly with a site downstream to the TATA box, thus hindering the movement of RNA polymerase (Muller, 1987; Roberts *et al.*, 1988; Faber and Wilcox, 1988), the position of the Vmw175 binding site upstream from the IE1 TATA box suggests that the IE1 promoter cannot be repressed by a similar mechanism. Faber and Wilcox (1988) observed that the Vmw175 binding site within the IE1 promoter overlapped with the binding sites for the cellular transcription factors Sp-1 and CBP, and suggested that the binding of Vmw175 resulted in the dissociation of the cellular proteins, thus reducing transcription from the promoter. Alternatively, DiDonato and Muller (1989) reported that the Vmw175 binding sites within the IE1 and IE3 promoters are equidistant from the TATA box, and presented data to show that binding of Vmw175 to either site resulted in a conformational change at the TATA box. This may then promote the disassociation of preformed transcription complexes, or prevent transcription complexes from assembling on the promoter.

Transient transfection assays have demonstrated that mutation or deletion of the Vmw175 binding sites within the IE1 and IE3 promoters abolishes the binding of Vmw175 to DNA and also abolishes the repression of transcription *in vitro* (Roberts *et al.*, 1988; Resnick *et al.*, 1989). In addition, with only one exception (Paterson *et al.*, 1989), mutations which result in a reduction or loss of the DNA binding activity of Vmw175 also result in the loss of autoregulatory activity. These results point to a strong correlation between the binding of Vmw175 to the IE1 and IE3 promoters and the repression of transcription.

1B.6.3. Analysis of Immediate-Early Gene Expression Through the Construction of Mutant Viral Strains.

Although transient transfection assays have illustrated the importance of the Vmw175 DNA binding activity in its transcriptional repression of two IE promoters, it is important to study the importance of the DNA binding activity of the protein during a normal course of infection. To this end, mutations have been introduced into the

Vmw175 binding site upstream of the IE1 TATA box within HSV-1 (Everett and Orr, 1991). The mutations introduced were shown to abolish binding of Vmw175 and repression of transcription during a transient expression assay. However, introduction of the mutations into the virus resulted in no detectable change in the accumulation of Vmw110. Whilst the effect of the mutations on the transcription of IE1 RNA was not demonstrated, the authors argued that the observed accumulation of Vmw110 was indicative of the expression of the IE1 gene, and that repression of IE1 transcription was of little importance during a productive viral infection. Although no mutational analysis of the binding site at the IE3 capsite within HSV-1 has been published, DeLuca and Schaffer (1988) implied that mutations within this site which interfere with repression during transient assays also reduced IE3 repression during a viral infection.

1B.6.4. Conclusions Regarding the Expression of Immediate-Early Genes.

Whilst initial studies of HSV-1 gene expression suggested that immediate-early gene expression as a whole is repressed during viral infection, it now appears that the transcriptional control of this subset of genes is rather more complex. Studies of RNA transcription and accumulation of transcripts have illustrated that the kinetics of expression of the various IE genes varies during infection, as discussed in Section 1B.6.1.

Studies of the levels of IE proteins during HSV-1 infection in cell culture have illustrated that both Vmw110 and Vmw175 are present in high quantities at late times during infection (Metzler and Wilcox, 1985; Everett and Orr, 1991). Metzler and Wilcox (1985) suggested that this represented the high stability of the Vmw175 protein. Whether this is true for the Vmw110 protein also is unknown. It is interesting to note that IE1 transcription rate and mRNA accumulation have been shown to be higher later during infection than that of the IE3 gene (Harris-Hamilton and Bachenheimer, 1985; Weinheimer and Mcknight, 1987). In fact, taken as a whole, all experimental evidence suggests that the IE3 gene is the only gene from which transcription is repressed by Vmw175 during viral infection.

Vmw175 is the most potent transactivating protein of HSV-1, as demonstrated by the inability of Vmw175 deletion mutant viruses to transactivate early and late genes (DeLuca and Schaffer, 1988). Therefore, overexpression of this protein would lead to abnormalities in both viral and possibly cellular gene expression. Avoidance of such a situation would require the fine control of Vmw175 expression, which appears to be mediated in part by the autoregulatory properties of the protein.

1B.7. IE Products of HSV-1 Involved in the Transcriptional Regulation of HSV Early and Late Genes.

HSV-1 encodes five immediate-early proteins. Table 1B.1 summarizes the immediate-early genes, and the terms used to identify their protein products.

Glasgow nomemclature			**Chicago nomenclature**	
Gene *	Gene §	Product	Gene	Product
IE-1	IE110	Vmw110	α0	ICP0
IE-2	UL54	Vmw63	α27	ICP27
IE-3	IE175	Vmw175	α4	ICP4
IE-4	US1	Vmw68	α22	ICP22
IE-5	US12	Vmw12	α47	ICP47
	UL39	R1		ICP6

* Original nomenclature (Clements *et al.*, 1979).
§ Later designation (McGeoch *et al.*, 1988a)

Table 1B.1. The IE Products of HSV-1.

As described in previous sections, biochemical analyses have illustrated the importance of three immediate-early proteins in the transcriptional regulation of HSV-1 genes, these are the Vmw175, Vmw110 and Vmw63 proteins. These proteins are dealt with in greater detail in later sections. The role of the other two IE proteins in HSV gene regulation is less clear. Viruses with deletions in the genes encoding Vmw12 and Vmw68 have been isolated and shown to be viable in tissue culture (Longnecker and Roizman, 1986; Post and Roizman, 1981). Although a virus containing a deletion within the carboxy-region of Vmw68 was shown to produce lower amounts of some late genes (Sears *et al.*, 1985), the role of this protein in the control of late gene expression has not been elucidated.

1B.7.1. The IE Protein Vmw110.

The properties of Vmw110 observed in transient transfection assays have been described in detail in Sections 1B.3 and 1B.4. The conclusions from these experiments are that Vmw110 in isolation can transactivate a variety of HSV-1 immediate-early, early and late gene promoters, and that it can also act synergistically with Vmw175.

The Vmw110 protein is a phosphorylated molecule which is located at specific sites within the nucleus both upon transfection and during viral infection (Everett, 1987a;

1988b, 1989; personal communication; Knipe and Smith, 1986). Cotransfection of the Vmw175 protein appears to result in the diffuse distribution of the Vmw110 throughout the nucleus (Gelman and Silverstein, 1986). Sequence analysis of the coding region of the IE1 gene has identified a cysteine-rich sequence between residues 99-156 (Perry *et al.*, 1986). This motif has been identified as a novel motif conserved within several eukaryotic proteins. The motif resembles the well-studied zinc finger motif, in that it contains a regular spacing of cysteine and histidine residues (Freemont *et al.*, 1991) and an isolated polypeptide encompassing the motif within Vmw110 does bind zinc, although the exact stoichiometry of this binding has not been elucidated (Everett, personal communication). Although the protein has been shown to interact non-specifically with DNA (Hay and Hay, 1980), no direct evidence has been shown that the cysteine-rich motif is involved in the interaction with DNA. In fact, the isolated zinc finger region of Vmw110 when expressed as a polypeptide in bacteria exhibits no DNA binding properties and none of the related proteins exhibit any sequence-specific DNA-binding activity (Freemont *et al.*, 1991; Everett *et al.*, 1991c).

The regions involved in the transactivatory properties of Vmw110 have been examined by a variety of mutational analyses (reviewed in Everett *et al.*, 1991d). Everett (1987b) found that mutations within five regions of the protein reduced its transactivatory ability in *in vitro* assays. Two of these regions were relatively large tracts within the molecule, region 5 located towards the carboxy terminal region of Vmw110, and region 1 located between amino acids 150-222, which incorporates the C-terminal portion of the cysteine-rich motif. Further analysis suggested that region 5 plays some role in mediating the interaction between Vmw110 and Vmw175. Several viruses containing insertional and deletion mutations within the IE1 open reading frame have also been constructed and characterized (Everett, 1989; Chen and Silverstein, 1992). The results mirrored those obtained in the transfection assays described above. One important point found during these mutational analyses was that mutations within the cysteine-rich domain of Vmw110 result in a differing ability to support viral growth in different cell lines. It is possible that this region of the protein mediates an interaction with cellular factors which are present in varying amounts in different cell types.

This hypothesis is also supported by the data of Weber and Wighdahl (1992), who found that truncated Vmw110 polypeptides encompassing the N-terminal portion of the protein act as dominant transcriptional repressors which interfere with the transactivatory properties of both Vmw110 and Vmw175 (Weber and Wighdahl, 1992). It is possible that the N-terminal domain interacts with a cellular factor essential for the transactivation activity of Vmw110, and the above data implies that both Vmw110 and Vmw175 interact with this cellular factor. It would be interesting to determine whether the cysteine-rich domain of Vmw110 is essential for this *trans*-dominant inhibitory effect. Cai and Schaffer (1989) have also suggested that a functional site of Vmw110

maps within its N-terminal domain, and have shown that a polypeptide consisting of the N-terminal 427 amino acids of Vmw110 is sufficient to transactivate a variety of viral promoters in a transient transfection assays.

The analysis of HSV-1 viruses containing non-functional Vmw110 molecules or deletions within the IE1 reading frame has assisted in the elucidation of the probable role of Vmw110 within the lytic cycle of the virus (Stow and Stow, 1986; Sacks and Scaffer, 1987; Everett, 1988a; Chen and Silverstein, 1992). The viruses exhibited no impairment in the production of viral particles, but differed in their growth efficiency as measured by one step growth curves. The analysis of infection in a range of cell lines also illustrated a difference in viral growth efficiency at low m.o.i., with different cell types exhibiting a variety in permissiveness. At low multiplicities of infection, all viruses exhibited an impairment in growth efficiency and infectious viral yield.

It appears that functional Vmw110, whilst not essential for viral growth in tissue culture, plays a role in the general transactivation of viral genes. A defect in this protein results in the impairment of viral replication and delay in early and late gene expression. These effects are most distinct during infections involving a low number of viral particles. The phenotype of Vmw110 negative viruses may be explained by an inability to efficiently transactivate the IE3 gene. At low multiplicities of infection, the reduced expression of Vmw175 may present a block to subsequent viral gene expression. Infection with a high number of viral particles providing more templates for IE3 expression may provide sufficient Vmw175 to trigger transactivation of viral early genes and enable the virus to enter the lytic cycle of infection.

The mechanism of the interaction (if any) between Vmw175 and Vmw110 is not clearly understood. Everett (1987b) suggested that the two proteins may not utilize similar mechanisms to transactivate viral promoters. For example, whilst Vmw175 may fumction through the direct enhancement of transcription, Vmw110 may act to increase the stability of viral transcripts. This proposal seems more valid in the light of the discovery of a family of proteins containing the distinctive cysteine-rich motif which is present within Vmw110 (Freemont *et al.*, 1991). No member of this family has been shown to direct enhancement of gene transcription. This mechanism may also account for the apparent requirement for a specific nuclear distribution of Vmw110 during viral infection (R. D. Everett, personal communication), and for the requirement of a functional Vmw175 molecule for Vmw110 activity during both transient expression assays (Gelman and Silverstein, 1986) and viral infection (DeLuca *et al.*, 1985; DeLuca and Schaffer, 1988; Shepard and DeLuca, 1989; this thesis).

1B.7.2. The IE Protein Vmw63.

The first evidence for a role for the viral IE product Vmw63 in the regulation of HSV-1 gene expression came from the study of temperature-sensitive viruses containing mutations within the IE2 gene. These viruses are non-viable in tissue culture at the non-permissive temperature and exhibit an over-expression of early gene products and no detectable late gene expression under restrictive conditions (Sacks *et al.*, 1985). Analysis of Vmw63 deletion mutant viruses also suggested that the protein was essential for late gene expression and possibly also involved in the down-regulation of early genes (McCarthy *et al.*, 1989).

Transient expression assays were then used to define more clearly the role of Vmw63 in the transactivation of late genes. Vmw63 in isolation had no effect upon gene expression. However, in conjunction with Vmw175 and Vmw110, Vmw63 either increased transactivation (with the γ_1 gB and VP5 promoters), repressed transactivation (with the β tk promoter), or had no effect (with the γ_2 gC promoter) (Everett, 1986; Sekulovich *et al.*, 1988; Shapira *et al.*, 1987; Rice and Knipe, 1988). The variable results in transient expression assays made it difficult to assign a specific transactivatory property to Vmw63. The observation of Everett (1988b) that differences in cell type and transfection protocol leads to a wide variance in the results of transient expression assays illustrates the limitations of the short term transfection assay in determining the *in vivo* transactivatory property of a viral protein.

In addition, transient transfection assays illustrated that Vmw63 may play a dual role in HSV-1 gene regulation, and may act post-transcriptionally to control mRNA 3' polyadenylation and reduce the expression of spliced genes (Rice and Knipe, 1990; Sandri-Goldin and Mendoza, 1992; Chapman *et al.*, 1992). McLauchlan *et al.* (1989) illustrated that HSV-1 encodes a factor, termed LPF, which selectively increases the 3' processing at the poly(A) signal of the UL38 transcript (a true late transcript). This factor does not affect the processing of additional immediate early transcripts. McLauchlan *et al.* (1992) illustrated that this process is independent of the late gene promoter and is mediated by an IE product. The analysis of mutant viral extracts and further transient expression assays indicated that Vmw63 was essential for this processing of late poly(A) signals,

Although it is now becoming apparent that Vmw63 may mediate its affect on viral gene expression through a post-transcriptional mechanism, it is still not clear how the protein utilizes this mechanism in the activation and repression of early and late genes.

* Marsden *et al.* (1976).

1B.7.3. The IE Protein Vmw175.

The early isolation of temperature-sensitive mutations in the gene encoding the viral protein Vmw175 enabled the elucidation of its role in viral gene expression. Vmw175 is a potent transactivator, and is essential for the viral replicative cycle. Loss of a functional Vmw175 molecule results in the lack of early and late gene expression, together with the over-expression of some immediate-early genes (Preston, 1979a; Dixon and Schaffer, 1980; Watson and Clements, 1980).

A great deal of evidence has been compiled concerning the *in vitro* and *in vivo* properties of Vmw175; this data may be of use in the determination of the mechanism by which the protein mediates viral gene activation and repression. The properties of the protein and their possible role in the mechanism of transactivation by Vmw175 are described in detail in Section 1C.

1C. A Detailed Analysis of the IE Protein Vmw175.

1C.1. Structural Properties of the Vmw175 Protein.

The immediate-early polypeptide Vmw175 is encoded by the IE-3 gene, an unspliced open reading frame situated within the short repeat region of the HSV-1 genome. This gene has an unusually high G+C content (81.5%), and the Vmw175 protein contains high amounts of the four amino acids lysine, arginine, isoleucine and methionine. The predicted molecular weight of the protein from amino acid composition is 132,835Da. However, the observed molecular weight of the protein upon SDS-polyacrylamide gel electrophoresis is nearer 175,000Da (McGeoch *et al.*, 1986).*

The difference in the observed and predicted molecular weights of the protein suggests that Vmw175 is post-translationally modified. In fact, it has been shown that Vmw175 is a nuclear-localized phosphoprotein (Pereira *et al.*, 1977; Metzler and Wilcox, 1985). Phosphorylation occurs at serine/threonine residues within the protein (Metzler and Wilcox, 1985), involving residues both within the serine-rich tract situted between amino acids 170-250 and also at sites carboxy-terminal to amino acid 309 (DeLuca and Schaffer, 1988). Analysis of [^{32}P] radiolabelled polypeptides synthesized during the HSV-1 replicative cycle illustrated that Vmw175 exists as three phosphorylated forms which exhibit different mobilities upon gel electrophoresis. The phosphate group cycles

✳

Vmw175 is guanylylated and adenylylated during infection, in a process which appears to involve a HSV-1 early or late gene product (Blaho *et al.*, 1993).

on and off two of these forms during a lytic infection, whereas the third form is stably phosphorylated (Wilcox *et al.*, 1980).

Further studies have illustrated that the phosphorylation state of Vmw175 affects its ability to bind to different gene leader sequences (Papavassiliou *et al.*, 1991). Dephosphorylation of the protein reduces its ability to bind to lower affinity, non-consensus binding sites. Phosphorylation was found to be required for the binding of the protein to sites within the leader sequences of early (tk) and true late (UL38) promoters, whilst the dephosphorylated protein could bind to the IE-3 gene cap site.

Vmw175 is also post-translationally poly(ADP ribosyl)ated during infection (Preston and Notarianni, 1983; Blaho *et al.*, 1992). However, the importance of poly(ADP ribosyl)ation to the transactivation and DNA binding functions of Vmw175 is unclear. ✳

Vmw175 is located within the nucleus during infection, possibly through a proposed nuclear localisation signal situated between amino acids 727-730 (DeLuca and Schaffer, 1987; Paterson, 1989). Studies of the nuclear localisation of several temperature sensitive Vmw175 mutants illustrated that certain mutations interfere with the nuclear localization of the protein. These mutant forms of Vmw175 also interfere with the nuclear localization of Vmw110, suggesting that the two proteins utilize the same nuclear localization pathway (Knipe and Smith, 1986). Varying results have been obtained with the tsK mutant form of Vmw175 (described in further detail in Section 1C.3). Whilst one group reported wild type nuclear distribution of this mutant protein (Knipe and Smith, 1986; Knipe *et al.*, 1987), Preston (1979b) reported that the protein did not associate with the nuclear fraction of infected cells. These observed differences could be due to differences in cell type or multiplicity of infection used in the experiments.

Knipe *et al* (1987) also demonstrated that Vmw175 co-localized with ICP8 in specific globular compartments within the nucleus late in infection. These compartments, termed replicative compartments, are the site of viral DNA replication. The authors suggested that these compartments are also the site of late gene transcription, and that Vmw175 is required to associate into these complexes in order to transactivate late viral genes. Several temperature sensitive mutants of Vmw175 have been isolated which transactivate early but not late genes at the non-permissive temperature (DeLuca *et al.*, 1984; Paterson *et al.*, 1990). It would be of interest to determine whether these mutations interfere with the association of Vmw175 into these replication compartments.

The sequence analysis of other α-herpesvirus regulatory proteins which are analogous to Vmw175 has illustrated that this family of proteins shares a specific pattern of sequence conservation. All the proteins can be divided into five regions, based upon their amino acid homology (McGeoch *et al.*, 1986; Davison and McGeoch, 1986; Grundy *et al.*, 1989; Cheung, 1989; Anderson *et al.*, 1992; see Figure 1C.1). The

regions 2 and 4 exhibit greatest amino acid conservation, and mutational analysis of Vmw175 has illustrated that these regions are important for the *in vitro* and *in vivo* properties of the protein (see Section 1C.3).

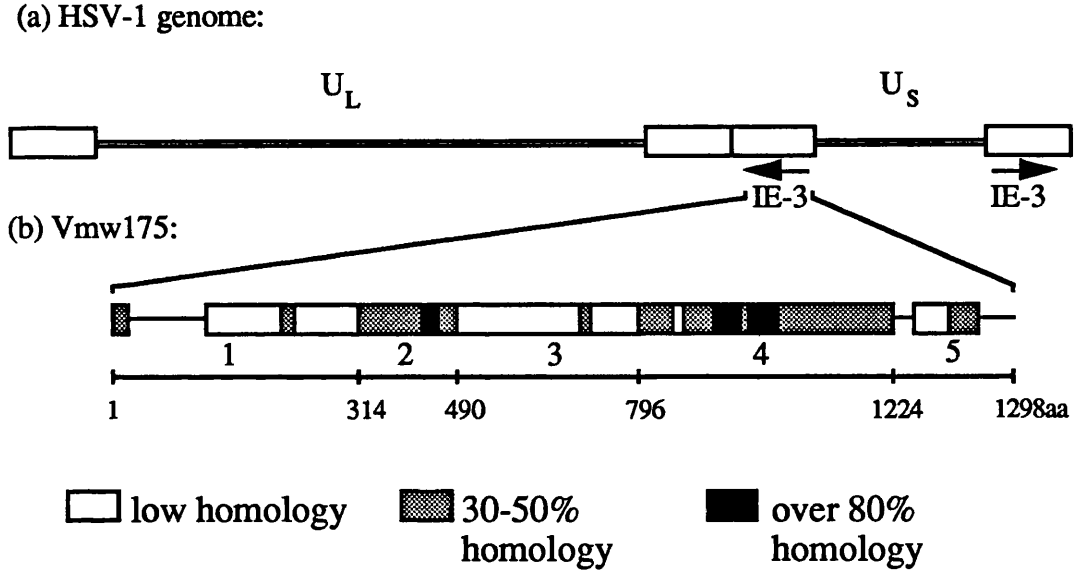


Figure 1C.1. Conserved Regions of Vmw175.
(a). A diagramatical representation of the HSV-1 genome, indicating the position of the gene IE-3, encoding Vmw175.
(b). The protein Vmw175. divided into five sections based upon its homology with the VZV regulatory protein 140K (McGeoch et al., 1986). The lower numbers indicate codon number within the IE-3 ORF. Analysis of the regulatory proteins of other α-herpesviruses indicates that they too share the conserved regions indicated (see text for details).

The ability of the HSV-2 and VZV regulatory proteins to substitute for Vmw175 in the replicative cycle of HSV-1 has been investigated. Whilst the HSV-2 equivalent of Vmw175 is functionally interchangeable with Vmw175 (Smith and Schaffer, 1987), the VZV regulatory protein (140K) exhibited some slight defects in its ability to replace Vmw175 during the viral life cycle. A HSV-1 virus with both copies of the IE-3 reading frame replaced by the 140K reading frame was able to propagate but unable to form plaques in tissue culture in the absence of complementing Vmw175. In addition, this virus exhibited low expression of certain late genes, and was unable to efficiently transactivate early and late genes at low multiplicities of infection (Disney and Everett, 1990a).

1C.2. The Early Genetic Characterization of Vmw175.

The investigation of the role of Vmw175 in the lytic cycle of HSV-1 has been enhanced by the isolation of a number of temperature-sensitive viruses containing lesions in the gene encoding this protein. The majority of these viruses showed a lack of early and late gene expression and an over-production of immediate-early products at the non-permissive temperature (Marsden *et al.*, 1976; Courtney *et al.*, 1976; Watson and Clements, 1978; 1980; ; Preston, 1979a; Dixon and Schaffer, 1980). These observations illustrated the importance of Vmw175 in early and late gene expression, and suggested that it may also play a role in the repression of immediate-early genes, a phenomenom described in greater detail in Section 1B.6. Temperature shift experiments illustrated that Vmw175 is synthesized but is non-functional at the non-permissive temperature (Preston, 1979a). In addition, temperature upshift experiments have illustrated that functional Vmw175 is required throughout the course of infection for the efficient expression of the full range of viral genes (Watson and Clements, 1980; Dixon and Scaffer, 1980). This observation is supported by the data of Metzler and Wilcox (1985) who illustrated the presence of stable Vmw175 protein in infected cells at late times during infection, after Vmw175 synthesis has decreased.

Certain of the mutations contained in these temperature sensitive viruses have been mapped to precise positions within the IE-3 open reading frame. One such mutant, *ts*K, which synthesizes solely immediate-early transcripts at the non-permissive temperature, has been mapped to within region 2 of the protein (Davison *et al.*, 1984). Additional mutations which do not exhibit such tight restrictions on gene expression have been mapped to positions towards the carboxy-terminus of the protein (Paterson *et al.*, 1990).

The construction of cell lines transformed with fragments of DNA containing HSV-1 promoter sequences has allowed for the further analysis of details of early and late promoter activation. Infection of such cell lines with the temperature sensitive viruses described above illustrated the requirement for Vmw175 in the transactivation of such promoters (Dennis and Smiley, 1984; Silver and Roizman, 1985).

The construction of cell lines stably expressing Vmw175 from the cellular genome demonstrated that Vmw175 was solely required for the activation of early genes in the context of the viral genome (Persson *et al.*, 1985). However, such cell lines could not facilitate the activation of late viral genes in the absence of additional protein synthesis, suggesting that some other factor was required for late gene activation. These cell lines have enabled the construction of mutant viruses containing engineered lesions and point mutations within the IE-3 open reading frame for further mutational analyses of Vmw175 (see Section 1C.3).

1C.3. Mutational Analysis of Vmw175.

As described in Section 1C.2, the isolation of a number of temperature-sensitive mutant forms of Vmw175 have facilitated the study of the function of Vmw175 during the viral life cycle. From the study of a number of temperature-sensitive lesions within Vmw175, it appeared that the more amino-terminal lesions exhibited a tighter phenotype, in that such lesions abolished the transactivation of all early and late viral genes. This type of temperature sensitive lesion of Vmw175 is represented by the widely studied mutation *ts*K (Preston, 1979a, b; Watson and Clements, 1978; 1980; Preston, 1981; Davison *et al.*, 1984). This mutation is situated within a highly conserved portion of region 2 of the protein, and the importance of this region of the protein to its *in vitro* and *in vivo* functions has been highlighted by a number of mutational studies outlined below.

Another class of temperature sensitive mutations within region 4 of Vmw175, which do not display such a tight restriction of viral gene expression have been described. DeLuca *et al.* (1984) studied two temperature-sensitive mutant viruses, *ts*48 and *ts*303 which contain lesions mapping within region 4 of Vmw175. These viruses overexpress immediate-early products and also produce certain early products at the non-permissive temperature. Although these viruses express early gene products at the non-permissive temperature, they replicate to 15-50% of wild type viral DNA levels. Transient transfection assays illustrated that these mutant proteins have a slight activatory effect on certain IE promoters, which may account for the observed over-production of immediate-early proteins (DeLuca and Schaffer, 1985).

Paterson *et al.* (1990) isolated and characterised five temperature sensitive viruses containing lesions within a similar region of Vmw175. Three of the temperature sensitive mutations, *ts*1219, *ts*1221 and *ts*1225, were mapped to conserved amino acids within region 4 of the protein, whilst the other two mutations, *ts*1211 and *ts*1223, were mapped to a domain further towards the carboxy-terminus of the protein, within the same region as the mutations *ts*48 and *ts*303. These viruses also overexpressed immediate-early products and expressed certain early proteins at the non-permissive temperature. Again, the phenotypes of these five mutant viruses suggest that the conserved region 4 of Vmw175 plays a role in the activation of late viral genes. These results are supported by the data of DeLuca and Schaffer (1988) described below.

No evidence has been obtained to explain why such mutants were defective for expression of the late class of viral genes. DeLuca *et al.* (1984) suggested that the defect in late gene activation may have been due to a lack of interaction between mutant Vmw175 and other regulatory proteins, or differences in phosphorylation of the mutant proteins. It would be interesting to study the intranuclear localization of these proteins, as other mutant Vmw175 molecules with lesions in the carboxy-terminal region of the

protein which exhibit similar activation properties fail to associate with the replicative compartments suggested to be sites of late gene activation (Knipe *et al.*, 1987; DeLuca and Schaffer, 1988).

Several workers have constructed and characterised insertional and deletion mutations within Vmw175, in order to attempt to define regions of the protein essential for its autoregulation, transactivation and DNA binding properties. (DeLuca and Schaffer, 1987, 1988; Paterson and Everett, 1988a, b; Shepard *et al.*, 1989). The construction of cell lines expressing Vmw175 from the cellular genome has enabled the introduction of such mutations into the IE-3 reading frame within HSV-1 (DeLuca *et al.*, 1985; DeLuca and Schaffer, 1988; Shepard *et al.*, 1989; Imbalzano *et al.*, 1990; Paterson *et al.*, 1990; Shepard and DeLuca, 1991a, b).

DeLuca and Schaffer (1987, 1988) created a series of nonsense mutations within Vmw175. In transient transfection assays, it was found that the N-terminal 60% of Vmw175 was sufficient for the transactivation of the early tk promoter and the autoregulation of the IE-3 promoter. However, this polypeptide was unable to complement a mutant virus containing a large deletion within the IE-3 open reading frame. A polypeptide lacking the N-terminal 90 amino acids of Vmw175 (the *n*6 polypeptide) exhibited wild type transactivation, autoregulation and complementation properties, illustrating that this region of the protein is inessential for its properties both *in vitro* and *in vivo*.

The same nonsense and additional deletion mutations were also analysed in the context of the viral genome. It was found that all the mutant Vmw175 molecules analysed, with the exception of the *n*6 polypeptide, were unable to support the growth of HSV-1 in tissue culture. From the analysis of the polypeptide profiles of cells infected with the mutant viruses, the Vmw175 proteins were separated into three classes. Class I, which consisted of the *n*6 polypeptide alone, exhibited wild type expression of early and late viral polypeptides. Class II comprised the deletion mutant and nonsense mutant polypeptides which lacked residues within the first 60% of Vmw175. These viruses exhibited very low levels of early and late gene expression, and overexpressed the immediate early genes. Interestingly, this class of mutants also expressed the ICP6 polypeptide, the product of a supposedly early gene. Deletion mutant Vmw175 viruses have also been observed to express this protein in addition to immediate-early products (DeLuca *et al.*, 1985; this thesis, Section 3F.1). The final group, Class III, comprised truncated polypeptides which contained 60% or more of the amino-terminal portion of Vmw175. These viruses expressed a higher level of early products than Class II viruses. However, they expressed no true late products and also exhibited no viral DNA synthesis.

These Class III Vmw175 molecules exhibit similar phenotypes to the temperature sensitive mutant viruses described previously (DeLuca *et al.*, 1984; Paterson *et al.*,

1990). Interestingly, one Class III polypeptide, the *n*208 polypeptide which consists of the N-terminal 774 amino acids of Vmw175, was not distributed into replicative compartments at late times during viral infection. This supports the hypothesis that Vmw175 is required to associate into such complexes in order to express late viral genes (Knipe *et al.*, 1987).

Further analysis of deletion polypeptides mapped the nuclear localisation signal of Vmw175 to a region between amino acids 590-773, and DeLuca and Schaffer located a potential nuclear localization signal centred around amino acid 728. These results are in agreement with those of Paterson (1989) who mapped the nuclear localization signal to a region encompassing amino acids 682-774.

The results obtained from this study of nonsense and deletion mutations of Vmw175 are summarized in Figure 1C.2. In conclusion, it was found that the N-terminal 774 amino acids of Vmw175 are sufficient for the DNA binding, autoregulation and early gene transactivation properties of Vmw175, whilst the N-terminal region is essential for late gene expression and possibly viral DNA replication, and may determine the intranuclear localisation of Vmw175.

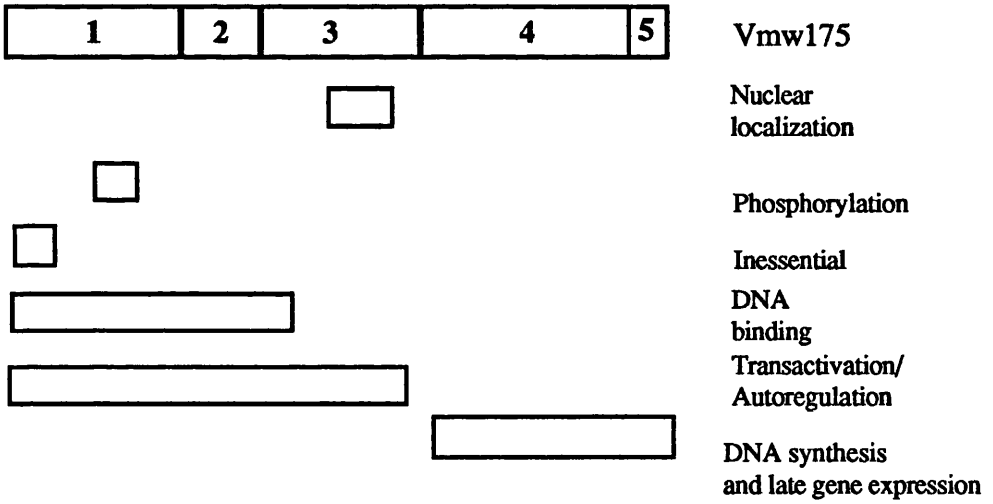


Figure 1C.2. Functional Domains of Vmw175 as Proposed by DeLuca and Schaffer, 1988.

Boxes indicate the possible functional domains decribed in the conclusions of DeLuca and Schaffer, 1988. The top line represents the Vmw175 protein, divided into five structural domains as described by McGeoch et al., 1986

Following this initial analysis of the functional domains of Vmw175, two more detailed mutational studies of the protein were carried out. Shepard *et al.* (1989) created a series of two amino acid insertional mutations throughout the whole of the *n*208 polypeptide. The parent *n*208 polypeptide consists of the N-terminal 774 amino acids of Vmw175, and is able to transactivate the early tk promoter, autoregulate the IE-3 promoter and bind to the IE-3 cap site *in vitro*. However, this polypeptide is defective in its ability to support viral growth in tissue culture (DeLuca and Schaffer, 1988). The

mutant polypeptides were tested for transactivation of the tk promoter, autoregulation of the IE-3 promoter, and DNA binding ability *in vitro*. Initially, the mutants were tested for their ability to bind to the consensus Vmw175 binding site situated at the IE-3 cap site. It was found that mutations which affected the DNA binding activity of *n*208 were situated proximal to or within region 2 of the protein. These mutations appeared to result in two DNA binding phenotypes. Mutations within the region encompassing amino acids 263-338 reduced but did not eliminate DNA binding activity, whereas mutations within the region encompassing amino acids 449-487 abolished the DNA binding activity of the protein. An insertion mutation at amino acid 386, which is situated between these two regions, exhibited wild type DNA binding activity. The mutant polypeptides were then tested for their ability to interact with a non-consensus Vmw175 binding site situated within the tk promoter. All mutations which reduced or abolished the binding of Vmw175 to a consensus DNA binding site eliminated the ability of the protein to interact with non-consensus DNA binding sites. Interestingly, the region defined by these mutations to be essential for the protein-DNA interaction, situated within amino acids 449-487 shows high conservation in other members of the α-herpesvirus regulatory protein family (see Section 3A). This region also contains the site of the *ts*K mutation described above. Therefore, it appears that the integrity of region 2, and particularly this carboxy-terminal portion of region 2 is essential for the DNA binding activity of Vmw175.

The mutant polypeptides were then tested in transient transfection assays for their ability to transactivate the tk promoter and to autoregulate the IE-3 promoter. The results mirrored those obtained in the DNA binding assays; that is, the mutations which affected DNA binding also affected the transactivation and autoregulation properties of Vmw175.

Two insertional mutations were introduced into the viral genome to test whether their *in vitro* properties were an indication of their ability to support viral growth in tissue culture. The mutations tested were representative of the two types of mutations which exhibited reduced DNA binding properties. Both mutant proteins were unable to induce early and late gene expression or to autoregulate the IE-3 gene, and viruses containing these mutations were non-viable in tissue culture in the absence of complementing wild type Vmw175.

In addition, a series of deletion mutations were introduced into the *n*208 polypeptide to test the functional importance of sequences N-terminal to amino acid 274. These deletions removed amino acids 31-210, 31-274 and 143-210. All deletions did not affect the ability of the polypeptide to transactivate early promoters or bind to a consensus DNA binding site, but did reduce the transactivation of the tk early promoter. These deletions remove a serine-rich region of the protein encompassing amino acids 176-206, which the authors proposed played a role in the transactivatory function of Vmw175. However, the authors cited data which showed that the deletion mutations

tested did not affect the transactivation properties of the whole Vmw175 protein. In addition, a virus containing a deletion of amino acids 162-229 is viable in tissue culture, further showing that deletion of the serine-rich tract does not have a deleterious effect on the *in vivo* properties of the protein (Paterson and Everett, 1990).

A second study involved the construction and characterization of a series of four amino acid insertional mutations throughout the whole of the IE-3 reading frame (Paterson and Everett, 1988a, b). These mutant proteins were tested for DNA binding, transactivation and autoregulation properties *in vitro*. The most clear-cut results were obtained with insertional mutations within region 2. Although these mutations were not in complete agreement with those of Shepard *et al.* (1989) (see Figure 1C.3), most mutations proximal to or within region 2 of the protein reduced or abolished the DNA binding activity of Vmw175. This was with the exception of the mutation I13, situated at amino acid 324, which exhibited near wild type DNA binding properties. However, a polypeptide containing this mutation exhibited no DNA binding activity in gel retardation assays (R. D. Everett, personal communication).

The transactivation and autoregulation properties of these mutant proteins mirrored their ability to bind to DNA; that is, those that reduced DNA binding activity also reduced transactivation and autoregulation activity.

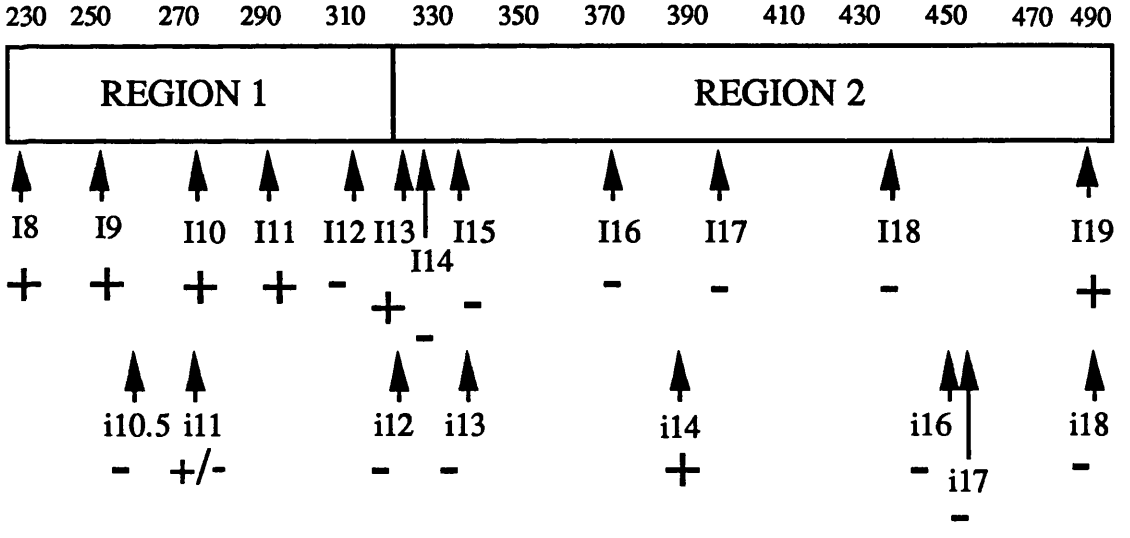


Figure 1C.3. Insertional Mutagenesis of Vmw175.
The open boxes represent regions 1 and 2 of Vmw175, as defined by McGeoch et al., 1986, with numbers above indicating codon number within the IE-3 ORF. Below are shown the positions of insertional mutations introduced by Paterson, 1989 (I series) and Shepard et al., 1989 (i series). The ability of mutant Vmw175 molecules to bind to DNA is indicated. + = wild type binding activity; - = below wild type binding activity.

The results obtained with mutations in region 4 of Vmw175 were more complicated. For example, several mutations which reduced DNA binding and transactivation activity did not affect the autoregulatory activity of the protein. These

mutations may have been exerting their effects through the destabilisation of protein conformation, an observation also made about temperature-sensitive mutations within this region (Paterson *et al.*, 1990).

The analysis of deletion mutants of Vmw175 in these assays also gave quite varied results. Many deletions within the IE-3 coding region may result in a deleterious effect on protein conformation and stability, and so the most informative evidence from such an analysis stems from the observation of deletions which have little negative effect on the activity of the protein. The main conclusion from this deletion mutational analysis was again that the integrity of region 2 is important for the DNA binding activity of Vmw175.

All the mutational analyses discussed so far have suggested that the DNA binding, transactivation and autoregulation properties of Vmw175 are properties of similar regions of the protein, and that no separable domain is responsible for any one function of the protein. These results support the hypothesis that the DNA binding activity of Vmw175 is essential for its full function *in vitro* and *in vivo*. However, a recent piece of evidence has arisen which appears to contradict this hypothesis.

The evidence arises from the characterization of the virus vi12 (Shepard and DeLuca, 1991b). This virus contains a two amino acid insertional mutation at residue 320 of the IE-3 open reading frame. The Vmw175 molecule containing this mutation was found to exhibit no detectable DNA binding activity *in vitro*, and the virus vi12 was non-viable in tissue culture in the absence of complementing levels of wild type Vmw175. During the propagation of vi12 on non-complementing cell lines, a pseudo-revertant virus, vri12 was isolated. Sequence analysis of the Vmw175 molecule encoded by this virus showed that the protein contained a second mutation at amino acid 342, although recent sequence data has suggested that the mutation may lie at residue 345 (this thesis). This pseudo-revertant virus was viable in tissue culture in the absence of complementing wild type Vmw175, and gave a polypeptide profile similar to that of wild type virus. The transcription rates of the Vmw175, Vmw63 and tk genes in both vri12 and vi12 were analysed, and it was found that whilst vi12 over expressed the two IE products and showed a reduced level of tk transcription, the rates of transcription of vi12 genes were identical to that of wild type virus. The ri12 Vmw175 molecule exhibited no DNA binding activity both in infected cell extracts and in isolation. This has been the sole example of a mutant Vmw175 molecule that exhibits no DNA binding function *in vitro* but is able to transactivate viral genes, and support the growth of HSV-1 in tissue culture. However, there are certain factors which must be taken into consideration when analysing these results, such as the relevance of the *in vitro* DNA binding assays in measuring the *in vivo* properties of Vmw175. The properties of this mutant protein are described in further detail in Section 1C.4.

1C.4. The DNA Binding Properties of Vmw175.

The early biochemical characterization of Vmw175 illustrated that the protein could interact non-specifically with calf thymus DNA (Freeman and Powell, 1982). Further experiments with partially purified Vmw175 from infected cell extracts illustrated that Vmw175 protected one site within the upstream regulatory sequences of the gD promoter and two sites within pBR322 vector sequences. Comparison of the DNA sequence of these three sites led to the identification of conserved residues within all three sites. The consensus sequence ATCGTCNNNNYCGRC (N = any base; Y = pyrimidine; R = purine) was proposed to represent an essential core sequence present within the binding sites of Vmw175 (Faber and Wilcox, 1986).

Further footprinting analyses identified other Vmw175 binding sites containing this consensus sequence in both the upstream regulatory sequences of the IE-1 promoter (Kristie and Roizman, 1986b; Kattar-Cooley and Wilcox, 1989) and the cap site of the IE-3 gene (Muller, 1987; Faber and Wilcox, 1988). The involvement of these sequences in the mechanism of IE gene repression by Vmw175 is discussed in Section 1B.6. Additional immunoprecipitation, gel retardation and South-Western experiments illustrated that Vmw175 directly and specifically interacted with sequences within many promoter sequences, including the upstream and leader sequences of the tk gene (Papavassiliou and Silverstein, 1990; Imbalzano *et al.*, 1990); the upstream sequences of the IE-2 gene (Kristie and Roizman, 1986a); the upstream sequences of the gD gene (Tedder *et al.*, 1989); the leader sequences of the Vmw65 gene (Michael *et al.*, 1988); the upstream regulatory and leader sequences of the $\gamma_2$42 gene (Michael *et al.*, 1988; Michael and Roizman, 1989); and the leader sequences of the UL38 gene (Flanagan *et al.*, 1991). Whilst certain of these binding sites contain the previously proposed Faber and Wilcox consensus sequence, many of the sites contain very little homology to this binding site. Therefore, it appears that Vmw175 can interact with a wide variety of divergent DNA sequences. This observation has been supported by a recent statistical analysis of Vmw175 DNA binding sites.

Statistical and missing contact analysis of several Vmw175 DNA binding sites led to the suggestion that Vmw175 preferentially binds to sequences which contain a more divergent consensus sequence RTCGTCNNYNYSG (N = any base; Y = pyrimidine; R = purine; S = G or C) (DiDonato *et al.*, 1991). From this data, the development of a matrix illustrating the probability of any one base occurring at each site within the 13-base core Vmw175 binding site allowed the accurate predicition of Vmw175 binding sites within a given DNA sequence. Analysis of the HSV-1 genome with this matrix mean model allowed the examination of the distribution of Vmw175 binding sites within the genome. This analysis revealed that Vmw175 binding sites are

distributed throughout the genome, with an average occurrence of 1 site to 250 genic bases and 1 site per 760 non-genic bases. The highest density of binding sites was found to occur within the IE-3 gene and several glycoprotein genes including the gD gene. This data supports the hypothesis that removal of specific regions of gene promoters will not remove all available sites for the interaction of Vmw175 with promoter-regulatory DNA sequences. The consensus sequence suggested by DiDonato *et al.* (1991) is in good agreement with a second consensus sequence suggested after extensive mutational analysis of a DNA binding site within the gD promoter (ATCGTNNNNNYSG, Everett *et al.*, 1991a, b).

1C.4.1. The Role of DNA Binding in Transactivation by Vmw175.

A wide range of mutational studies have illustrated that the abolition of the DNA binding activity of Vmw175 also led to a loss of transactivation and autoregulation activity both *in vitro* and *in vivo* (DeLuca and Schaffer, 1988; Paterson and Everett, 1988a, b; Shepard *et al.*, 1989; this thesis). However, certain mutant Vmw175 molecules do not fit this general pattern.

Several temperature-sensitive forms of Vmw175 (*ts*1211, *ts*1219, *ts*1221 and *ts*1223) do not exhibit DNA binding activity whilst activating certain early viral genes, whereas one mutant protein (*ts*1225) which can bind to DNA at the non-permissive temperature fails to autoregulate the IE-3 gene (Paterson *et al.*, 1990). These Vmw175 molecules gave complexes with an abberant mobility upon gel retardation analysis, suggesting that the mutant proteins were partially unfolded. In the case of the mutants *ts*1211, *ts*1219, *ts*1221and *t s*1223, the lack of autoregulatory activity is mirrored by the lack of the proteins to bind to the IE3 cap site. Interaction of the mutant proteins with additional cellular or viral factors may enable sufficient DNA binding activity for early gene transactivation. However, the fact that the *ts*1225 protein, whilst able to bind to the IE-3 cap site *in vitro*, is unable to autoregulate the IE-3 gene suggests that other properties of the protein are required for the efficient repression of transcription from the IE-3 promoter.

In addition, one double mutant Vmw175 protein, the ri12 protein, exhibits no DNA binding properties *in vitro* but is able to support the viral growth cycle in tissue culture (Shepard and DeLuca, 1991b). The interpretation of such data is complicated, as it is difficult to be sure that DNA binding activity as measured *in vitro* is a true reflection of the actual properties of the protein *in vivo*. In the case of the mutant ri12, a DNA binding polypeptide of Vmw175 containing both mutations exhibits DNA binding activity *in vitro* at lower temperatures (this thesis). Also, as mentioned above, the interaction of Vmw175 with other cellular factors during viral infection may mediate the binding of the protein to DNA under certain circumstances. With respect to this latter

suggestion, two pieces of evidence have been presented which demonstrate the interaction of Vmw175 with the cellular transcription factors TFIIB and TFIID (Smith and DeLuca, 1992; Pizer *et al.*, 1992). Whilst such *in vitro* experiments may identify protein-protein interactions which do not occur *in vivo*, it is interesting to note that addition of TFIIB to binding reactions increases the DNA binding activity of Vmw175 and also restores the DNA binding activity of several DNA binding defective mutants of Vmw175, including the ri12 protein (Smith and DeLuca, 1992).

The function of specific Vmw175 DNA binding sites in the transactivation of viral early genes has also been discussed in Section 1B.3. It seems clear that for the transactivation of viral genes during a normal course of infection, Vmw175 does not require the presence of specific DNA binding sites within the target promoter. The only Vmw175 binding site for which evidence exists to implicate its importance in the mechanism of transactivation by Vmw175 is the site present within the leader sequence of the γ42 gene (Romanelli *et al.*, 1992; see Section 1B.4). This suggests two possibilities: either Vmw175 mediates transactivation not through the interaction with viral DNA but through the modification of components of the cellular transcription machinery; or that Vmw175 is able to bind to a divergent range of DNA sequences, so that removal of a specific region of DNA does not abolish the binding activity of the protein *in vivo*. In light of the recent statistical analysis of Vmw175 binding sites discussed above, and the fact that a great number of mutations which abolish the DNA binding activity of Vmw175 also abolish its transactivatory activity, the latter hypothesis seems more likely.

In conclusion, it appears that the majority of evidence concerning the possible mechanism of transactivation by Vmw175 supports the hypothesis that the protein requires an interaction with DNA in order to efficiently transactivate the full range of viral genes. The size of the protein and its interaction with a wide range of divergent DNA sequences suggests that the protein may utilize a variety of mechanisms to activate different classes of viral genes, including the interaction with a variety of cellular proteins to mediate the transactivation of different promoters. However, the mechanism of transactivation by Vmw175 is insufficiently resolved to verify such speculations.

1C.5. The Dimerization Properties of Vmw175.

The partial purification of Vmw175 from infected cells under nondenaturing conditions revealed that the protein existed as a homodimeric complex (Metzler and Wilcox, 1985). Analysis of partial peptides of Vmw175 synthesized from the viral genome or through *in vitro* translation has illustrated that the protein binds to DNA as a

dimer (Shepard and DeLuca, 1989; Shepard *et al.*, 1990; Shepard and DeLuca, 1991a; Roberts and Hayward, 1991). In contrast to other dimeric DNA binding proteins, such as the leucine zipper proteins (Vinson *et al.*, 1989), the proposed consensus DNA binding site for Vmw175 as defined by DiDonato *et al.* (1991) exhibits no palindromic symmetry. How the Vmw175 dimer interacts with such an assymetrical site has still to be investigated.

Many studies illustrated that the dimerization of a DNA-binding proficient and a DNA-binding deficient Vmw175 peptide led to the formation of a heterodimeric complex which exhibited DNA binding activity upon gel retardation analysis (Shepard and DeLuca, 1989; Shepard *et al.*, 1990; Shepard and DeLuca, 1991a; Roberts and Hayward, 1991). Two possible explanations exist for such a scenario: (i) the heterodimer interacts with DNA through the DNA-binding proficient monomeric subunit, or (ii) the formation of the heterodimeric Vmw175 complex restores the DNA-binding activity of the DNA-binding deficient monomeric subunit. The data obtained so far with these heterodimeric complexes cannot illustrate which of the two possible explanations holds.

The co-expression of peptides of Vmw175 which retain specific properties of the whole protein molecule during mixed infections of mutant viruses led to the complementation of the partial peptides (Shepard and DeLuca, 1989). As described above, this complementation was shown to occur through the formation of heterodimeric molecules which retained the full range of activities expressed by a wild type Vmw175 molecule.

In contrast, cell lines which express a deletion mutant form of Vmw175 (X25) were shown to interfere with the normal growth characteristics of HSV-1, presumably through the formation of heterodimeric Vmw175 molecules. The mutant polypeptide X25 contains two deletions which remove amino acids 774-1298 and 32-275. These deletions remove a domain of the protein believed to be essential for the transactivational activity of Vmw175 and another which is a major site of phosphorylation. They result in a polypeptide which exhibits wild type multimerization, DNA binding and autoregulation properties, but which fails to transactivate early and late genes (Shepard *et al.*, 1989). Immunological analysis illustrated that the wild type Vmw175 monomeric subunit within the X25-Vmw175 heterodimeric molecule contained a conformational alteration within a carboxy-terminal epitope. Whilst the heterodimeric molecule exhibited wild type DNA-binding activity, this alteration presumably resulted in a reduced transactivation activity, which led to the reduced growth rate of HSV-1 in the X25-expressing cell line (Shepard *et al.*, 1990).

Both the above studies involved the formation of heterodimeric Vmw175 complexes which exhibited wild type DNA-binding activity *in vitro*. Shepard and DeLuca (1991a) illustrated that the formation of heterodimers between an insertional

mutant form of Vmw175 (i12) and the X25 polypeptide resulted in a complex which exhibited no DNA binding activity *in vitro*, but which was able to activate a subset of viral early and leaky-late (γ_1) genes. The lack of DNA-binding activity *in vitro* may not be a true representation of the situation during the course of a viral infection, nevertheless, it seemed paradoxical that a heterodimeric complex with greatly reduced DNA binding ability was able to transactivate a subset of viral genes. This situation is similar to that observed with the mutant virus vri12 (Shepard and DeLuca, 1991b; see Section 1C.4). In the case of the heterodimeric i12-X25 complex, whilst the Vmw175 complex was correctly distributed into nuclear replicative compartments, and viral DNA replication was not inhibited to a level which interfered with late gene expression, this heterodimer could not direct the activation of true late (γ_2) genes.

The analysis of heterodimer formation between partial peptides and deletion mutants of Vmw175 suggests that the sequences necessary for dimer formation must lie between amino acids 309-774 (Shepard and DeLuca, 1989; Shepard *et al.*, 1990; Shepard and DeLuca, 1991a). The observation that a peptide encompassing amino acids 275-523 of Vmw175 expressed from bacteria forms homodimeric molecules delineates the dimerization motif further, to a site situated between amino acids 309-523 (Everett *et al.*, 1991b). Evidence presented in this thesis shows that the dimerization motif does not lie between amino acids 445-487, suggesting that sequences within amino acids 309-445 or 487-523 contribute towards the dimerization of the protein.

1C.6. The DNA Binding Domain of Vmw175.

Mutational analyses of Vmw175 illustrated that alterations within region 2 of the protein had the most deleterious effect on its *in vitro* DNA binding activity (DeLuca and Schaffer, 1988; Paterson and Everett, 1988a, b; Shepard *et al.*, 1989). In order to more closely define the regions of the protein involved in DNA binding activity, Everett *et al.* (1990) digested Vmw175 with protease K to reveal a protease resistent DNA binding polypeptide. Analysis of the ability of mutant forms of Vmw175 to yield this protease resistant DNA binding domain illustrated that the N-terminus of the polypeptide was situated between amino acids 229 and 292 whilst its C-terminus was situated between amino acids 495 and 518. Therefore, this proposed protease resistant DNA binding domain consisted of a proximal portion of region 1, together with the whole of region 2 and a proximal portion of region three of Vmw175.

In a parallel study, Wu and Wilcox (1990) expressed portions of the IE-3 reading frame as *trpE* fusion polypeptides in bacteria, and tested the ability of these polypeptides to bind a consensus Vmw175 binding site *in vitro*. Analysis of these polypeptides

demonstrated that residues 186 to 256 were not required for DNA binding and that a polypeptide encompassing residues 262 to 490 formed a high affinity DNA binding polypeptide, whilst a polypeptide encompassing residues 275 to 490 formed a lower affinity DNA binding domain. These polypeptides encompass the whole of region 2 together with a proximal portion of region 1 of Vmw175. DNase footprinting experiments illustrated that the larger DNA binding polypeptide protected the Vmw175 consensus binding site situated within the upstream regulatory sequences of the IE-1 promoter. In addition, gel retardation experiments revealed that the DNA binding domain polypeptide bound to probes containing several consensus and non-consensus Vmw175 binding sites. These results illustrated that the expressed DNA binding domain exhibited similar DNA binding specificity as the whole Vmw175 protein.

The expression of similar regions of the open reading frames for the major regulatory proteins of VZV and PRV resulted in the definition of DNA binding polypeptides for both proteins. The DNA binding domain of each protein was found to lie within residues 417 to 647 (VZV 140K protein) and 448 to 696 (PRV IE180 protein) (Wu and Wilcox, 1991). The three DNA binding polypeptides described above were found to give a DNase footprint at the cap site of the corresponding gene, suggesting that each protein shared a similar mechanism for autoregulation of gene expression. Further mutational analysis has illustrated that the DNA binding domain of the VZV 140K protein lies within residues 472 to 633 (Tyler and Everett, 1993).

Expression of portions of the IE-3 reading frame as non-fusion polypeptides in a bacterial T7 expression system has demonstrated that polypeptides encompassing amino acids 252 to 273 and 276 to 523 of Vmw175 bind to a consensus Vmw175 DNA binding site *in vitro*. Whilst both polypeptides demonstrated DNA binding activity, analysis of the ability of the two peptides to bind to mutant probes revealed that the smaller polypeptide had a slightly relaxed DNA binding specificity (Pizer *et al.*, 1991). Further analysis of these DNA binding domain polypeptides demonstrated that both existed as homodimeric complexes in solution (Everett *et al.*, 1991b). This bacterial expression system provided a means by which a more detailed mutational analysis of the DNA binding domain could be carried out, and it is this expression system which provided a basis for the analysis presented in this thesis.

1D. RNA Polymerase II Transcription and the Mechanisms of Transactivation.

1D.1. RNA Polymerase II and the Cellular Transcription Machinery.

Eukaryotic organisms utilize three DNA-dependent RNA polymerases for the transcription of subsets of cellular genes. RNA polymerase I transcribes rRNA precursors, RNA polymerase III transcribes 5SRNA and tRNA genes, whilst RNA polymerase II transcribes the largest subset of cellular genes to produce pre-mRNA (reviewed in Young, 1991). Genes transcribed by RNA polymerase II contain two elements that direct basal transcription from the promoter - the TATA box, at which the preinitiation complex forms, and the initiator element (Inr), the site of RNA transcription initiation. Whilst some eukaryotic promoters do not contain an obvious TATA box, transcription from such promoters utilizes a similar basic transcription machinery (reviewed in Saltzman and Weinmann, 1989).

1D.1.1. RNA Polymerase II.

RNA polymerase II is a multi-protein complex, comprising 8-10 subunits, ranging in molecular weight from 240-10kDa. All 10 subunits of the yeast RNA polymerase II enzyme have been cloned and characterized. The two largest subunits, termed RPB1 and RPB 2 are analagous to the prokaryotic RNA polymerase subunits β' and β. The largest subunit of the complex has been cloned from several eukaryotic sources, including yeast, hamster, mouse, *Drosophila* and humans. The mouse subunit has been found to exist in three forms - $RNAPII_O$, a 240kDa phosphorylated protein, $RNAPII_A$, the 215kDa dephosphorylated protein, and $RNAPII_B$, a proteolytically degraded form found in *in vitro* preparations of the enzyme, but not detected *in vivo*. Photoaffinity labelling of an actively transcribing RNA polymerase II complex detected only the $RNAPII_O$ form, whereas the $RNAPII_A$ form preferentially associated with the preinitiation complex. Therefore, it appears that RNA polymerase II is phosphorylated coincidentally with or very soon after the initiation of transcription (Saltzman and Weinmann, 1989; Woychik and Young, 1990; Corden, 1990; Zawel and Reinberg, 1992).

The site of this phosphorylation reaction is the conserved carboxyl-terminal domain (CTD) of the large subunit. This domain consists of repeats of the heptapeptide sequence Pro-Thr-Ser-Pro-Ser-Tyr-Ser. The number of repeats within the CTD varies from species to species, with higher eukaryotes containing the highest number of

repeated sequences. Two kinases have been found to specifically phosphorylate this conserved CTD region. The first identified kinase was found to be homologous to the mammalian $p34^{cdc2/CDC28}$ M-phase promoting factor (Nurse, 1990). The second identified kinase activity was found to migrate with the basal transcription factor TFIIH (see below). The function of this phosphorylation reaction is unclear. Whilst deletions of the CTD of the large subunit of RNA polymerase II are lethal in mouse and yeast (Corden, 1990), the proteolysed form $RNAPII_B$ which lacks the majority of the CTD repeat sequences can still support *in vitro* reconstituted transcription systems (Zehring *et al.*, 1988; Buratowski and Sharp, 1990). It was found that the basal transcription factor TFIID specifically interacted with $RNAPII_A$ (but not $RNAPII_O$), via the heptapeptide repeat sequences (Zawel and Reinberg, 1992; Usheva *et al.*, 1992). This observation led to the suggestion that phosphorylation of the largest subunit of RNA polymerase II releases the enzyme from its interaction with the preinitiation complex, leaving it free to form elongated transcripts (Zawel and Reinberg, 1992). Other suggestions for the function of this phosphorylation reaction include the involvement of phosphorylation in promoter specificity; in the interaction with DNA; in the interaction with upstream factors; and in the removal of histones from the transcribed template (Corden, 1990).

1D.1.2. Formation of the Preinitiation Complex.

The chromatographic analysis of crude nuclear extracts demonstrated that *in vitro* transcription systems utilized several components, termed basal transcription factors. These factors associate with RNA polymerase II to form the preinitiation complex, which is assembled on the TATA box of transcribed promoters in an ordered fashion. The order of addition of certain individual transcription factors has been determined by the analysis of intermediate preinitiation complexes (Buratowski *et al.*, 1989). This assembly pathway is illustrated in Figure 1D.1.

The initial interaction with template DNA is mediated by the transcription factor TFIID, which binds to the TATA box sequence of the promoter. This interaction is stabilized by the binding of the TFIIA factor, which may remain or leave the preinitiation complex. TFIIB recognises the TFIIA-TFIID-DNA or TFIID-DNA complex and binds, forming a platform for the interaction of RNA polymerase II. RNA polymerase II enters the complex already bound by the TFIIF factor. The formation of the preinitiation complex is completed by the entry of TFIIE and H (Zawel and Reinberg, 1992).

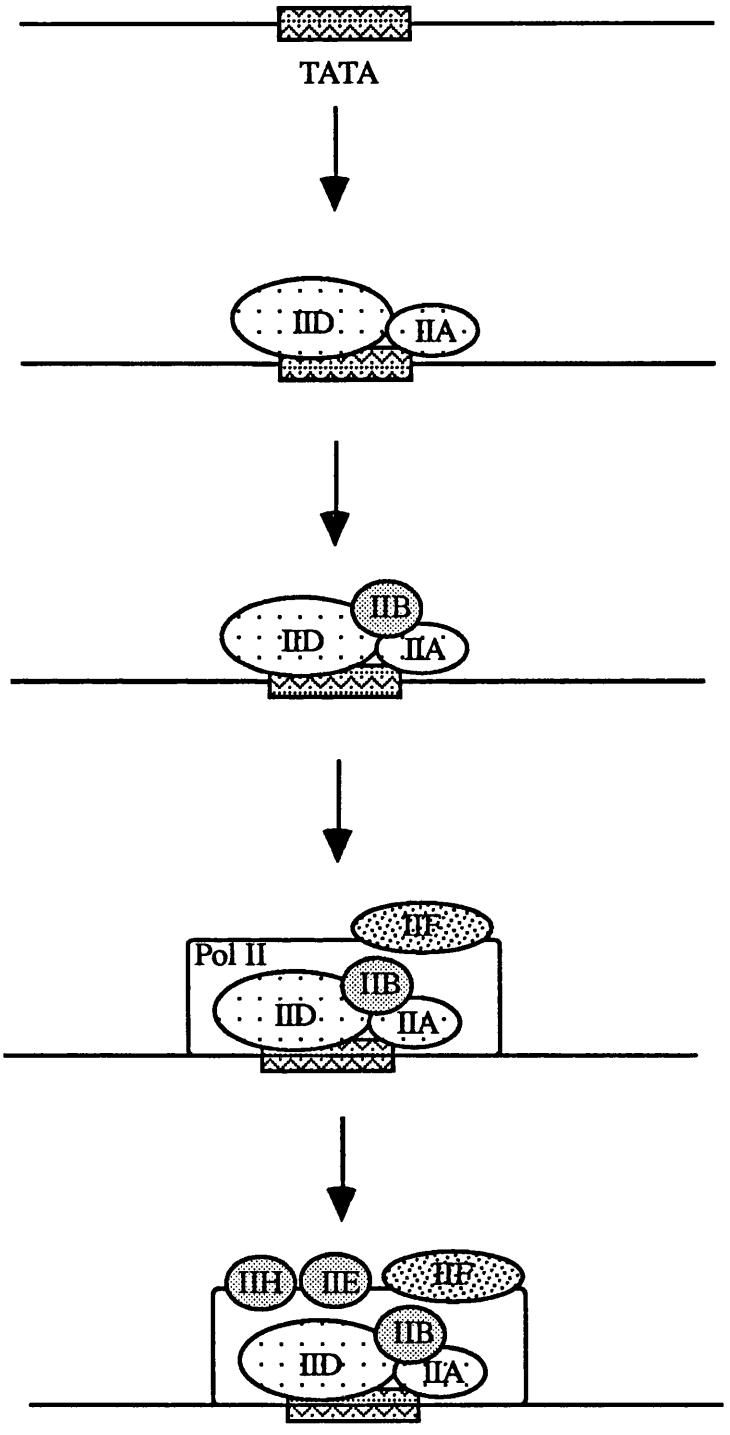


Figure 1D.1. Schematic Representation of the Assembly of a Transcription Complex.

The diagram illustrates the pathway of assembly of a cellular transcription complex on the TATA box. The recruitment of the major basal transcription factors and RNA polymerase II is shown. The exact protein - protein interactions between the sub-components of this complex are not yet clearly defined. Therefore, this representation of the transcription complex is purely diagrammatical.

1D.1.3. The Basal Transcription Factors.

Certain basal transcription factors have been cloned and characterized (Ha *et al.*, 1991; Kao *et al.*, 1990; Peterson *et al.*, 1990; Peterson *et al.*, 1991; Aso *et al.*, 1992; reviewed in Zawel and Reinberg, 1992). The majority of them show some homology to the prokaryotic σ factor, suggesting an evolutionary relationship between the two transcription systems. The functions of these individual factors as understood at present will now be dealt with in greater detail (for a review, see Zawel and Reinberg, 1992).

Initially, it was thought that the factor TFIIA was inessential for transcription by RNA polymerase II. Reinberg *et al.* (1987) found that TFIIA was required for *in vitro* transcription reactions which contained highly purified TFIID. The studies of Cortes *et al.* (1991) illustrated that TFIIA consisted of two activities, termed TFIIA and TFIIJ. The TFIIJ factor was required for transcription reactions with highly purified or recombinant TFIID, but the incorporation of crude preparations of TFIID eliminated this requirement. In contrast, TFIIA stimulated transcription when crude but not purified TFIID was incorporated into the transcription mixture. It was found that TFIIA removed a component that negatively atoregulated TFIID activity (Zawel and Reinberg, 1992). This component, termed Dr-2, bears similarity to other negative regulators of transcription, described in further detail in Section 1D.2.

The factor TFIIB has recently been implicated in the mediation of activation by factors bound at upstream activating sequences or enhancer elements. The experimental evidence and implications of this mechanism for transactivation are described in greater detail in Section 1D.2.

The cloning of the TATA binding protein (TBP) of the TFIID factor from mammalian cells (Kao *et al.*, 1990; Peterson *et al.*, 1990) showed it to be a 38kDa protein, with homology to proteins of similar function from yeast and *Drosophila.* Sequence comparison of these proteins illustrated that they shared a conserved 180 amino acid carboxyl-terminal region but contained variable amino-terminal regions. Pugh and Tjian (1990) found that the recombinant *Drosophila* TBP was able to support basal but not activated transcription, and presented data to suggest that upstream activation was mediated through coactivator proteins present in semi-purified mammalian and *Drosophila* TFIID fractions. In addition, an amino-terminal truncated form of the *Drosophila* TBP was unable to support activated transcription, suggesting that the variant amino-terminal region of TBP mediated this interaction with coactivators. Dynlacht *et al.* (1990) and Tanese *et al.* (1990) have purified these coactivator proteins or TBP associated factors (TAF's) from TFIID and illustrated that they mediate induction by certain activating factors. Whilst the amino-terminal region of TBP mediates the interaction with these TAF's, it is possible that the conserved carboxy-terminal domain

of TBP interacts with the CTD of the large subunit of RNA polymerase II (Usheva *et al.*, 1992; Pugh and Tjian, 1992).

The TFIID fraction of HeLa cells has been found to exist as two multicomponent complexes (Timmers and Sharp, 1991). One complex has a molecular weight of over 700kDa, whilst the other has a molecular weight of approximately 300kDa. The two complexes exhibit many different properties, including differences in the ability to support transcriptional activation. This evidence suggests that the TFIID complex is in a continual state of flux within mammalian cells, containing differing compositions of TBP-associated factors. Further evidence suggests that the previously identified TBP is also involved in RNA polymerase I and RNA polymerase III transcription (Comai *et al.*, 1992; Margottin *et al.*, 1991; White *et al.*, 1992; reviewed by Sharp, 1992).

The final basal transcription factor which has been extensively characterized over recent years is the factor TFIIH. This factor has been implicated as the specific kinase which phosphorylates the CTD of the large subunit of RNA polymerase II (Lu *et al.*, 1992; Feaver *et al.*, 1991; reviewed in Peterson and Tjian, 1992). Purification of the CTD kinase activity illustrated that the activity copurified with the TFIIH factor. In addition, an antibody directed against the 62kDa subunit of TFIIH inhibited both transcription and CTD kinase activity (Peterson and Tjian, 1992). Although this phosphorylation function of TFIIH is an important step in the initiation of transcription by RNA polymerase II, the TFIIH factor comprises five subunits, and may carry out additional functions, such as interaction with the additional factors TFIIE and TFIIJ.

1D.2. The Activation of RNA Polymerase II Transcription.

1D.2.1. The Mechanism of Action of Transactivating Proteins.

The activation of transcription by RNA polymerase II is mediated by upstream activating sequences within the promoter and the more distant enhancer elements, together with the protein *trans*-activating factors which bind to such sequences. The two types of regulatory elements are defined by their position within the promoter. Upstream regulatory sequences are situated relatively close to the TATA box of the promoter, whilst enhancer elements exert their effects regardless of position or orientation. Whilst upstream regulatory elements such as GC-rich boxes or CAATT boxes are present within a wide variety of genes, enhancer elements are often cell-type or promoter specific, for example the steroid hormone receptor enhancer element (reviewed in Maniatis *et al.*, 1987; Saltzman and Weinmann, 1989).

Initial results suggested that all *trans*-activating factors contained two domains, controlling the DNA-binding and transactivation functions of the protein (reviewed in Ptashne, 1988). Domain-swap experiments demonstrated that these domains were readily separable, and that a chimaeric protein containing DNA-binding and transactivation domains from two proteins could activate gene expression both *in vivo* and *in vitro* (for example, see Brent and Ptashne, 1985). Additional transcription activators have since been characterized which contain only one of these defined domains, for example, the herpes simplex virus transactivating protein, Vmw65, which contains just a transactivation domain (see Section 1B.2). Whilst the structure of the DNA-binding domains of many such proteins have been determined (see Section 1D.3), the transactivation domains are not so precisely defined. The transactivation domains investigated so far can be divided into three rough categories - the glutamine-rich domains, contained in the factor Sp1 (Courey and Tjian, 1988); the acidic domains, contained in the yeast factors GCN4 and GAL4 and the viral factor Vmw65 (Hope and Struhl, 1986; Ma and Ptashne, 1987; Triezenberg *et al.*, 1988); and the proline-rich domain, contained in the mammalian factor AP-2 (Williams *et al.*, 1988).

The mechanisms by which these proteins induce RNA polymerase II transcription are poorly understood. However, recent work has begun to provide clues as to some aspects of the activation mechanism. Most of this work has been carried out using acidic activation domains, normally represented by the transactivating domain of the viral protein Vmw65.

The earliest arguments suggested that activating proteins bound to a site within or proximal to the promoter, and contacted a component of the basic transcription machinery, with the intervening stretch of DNA looping out to allow such a connection to occur (Ptashne, 1988). This interaction may increase the rate at which the preinitiation complex assembles upon the activated promoter. This suggestion was supported by the observation of a phenomenom termed 'squelching', whereby high concentrations of an acidic transactivation domain sequestered some unknown component of the cellular transcription machinery, thus inhibiting transcription from a heterologous promoter (Gill and Ptashne, 1988).

Initial affinity chromatography experiments with the acidic transactivation domain of the Vmw65 protein illustrated that it specifically interacted with the TFIID fraction of the mammalian basal transcription machinery and also to the cloned yeast TATA-binding protein (Stringer *et al.*, 1990). Additional evidence illustrated that activation by the Sp1 *trans*-activating protein was mediated through adaptor or coactivator proteins present in semipurified TFIID fractions which mediated interactions between activating proteins and the TATA-binding protein (Pugh and Tjian, 1990; 1992). Further analysis demonstrated that these coactivator proteins were found associated with the TATA-binding protein in the TFIID fraction, and were termed TBP-associated factors (see Section 1D.1).

During an investigation of the 'squelching' phenomenom previously observed with acidic activator domains, Berger *et al.* (1990) observed two types of inhibition of heterologous promoters in yeast systems, which they termed *trans*-inhibition and *cis*-inhibition. The reasoning behind their proposed explanation for the two observed effects was intricate, and will not be outlined here. However, the conclusions of Berger *et al.* (1990) agreed with the observations of Pugh and Tjian (1990, 1992), in that the acidic activator domain was contacting an adaptor molecule which mediated the connection between upstream activating proteins and the basal transcription machinery. Berger *et al.* suggested that the member of the basal transcription machinery involved in this process interacted with DNA, and was in fact TFIID, although no evidence for the involvement of this particular factor was presented (reviewed in Lewin, 1990).

Kelleher *et al.* (1990) observed that the inhibition of transcription from heterologous promoters in yeast extracts by acidic activator domains could be relieved by the addition of a semipurified yeast fraction, which they suggested contained an adaptor molecule, analogous to the adaptor molecule proposed by Berger *et al.* (1990). However, as discussed by Ptashne and Gann (1990), this additional factor could also be a basal transcription factor required for an essential step following the interaction of the activator protein with TFIID. What this additional factor could be was not discussed.

To further complicate the issue, recent work by Lin and Green (1991) has demonstrated that an acidic activator domain acts through the recruitment of the basal transcription factor TFIIB into the preinitiation complex. TFIIB enters the preinitiation complex following the stable association of TFIID with the TATA box, and thus fits the requirements of the factor proposed by Ptashne and Gann (1990). Affinity chromatography experiments of Lin and Green (1991) demonstrated that the acidic activation domain interacted with both TFIID and TFIIB, but that the interaction with TFIIB was of a higher affinity. The factor TFIID was retained on the affinity column when chromatography was carried out under the conditions outlined by Stringer *et al.* (1990), suggesting that experimental conditions greatly influence the protein:protein interactions observed *in vitro*.

Such interactions of transactivating proteins with two target basal transcription factors could result in the functional co-operativity of multiple activators which is sometimes observed *in vitro* (reviewed in Roeder, 1991).

1D.2.2. The Role of Chromatin Structure in the Regulation of Transcription.

One criticism of the *in vitro* assays of transactivator proteins has been that the levels of transactivation observed *in vitro* (approximately 2-10 fold) do not match the transactivation levels of up to 100 fold or greater observed *in vivo*. One clue as to the reason for this discrepancy came from the observation that nucleosome formation *in vitro*

inhibited basal transcription. This inhibition could be overcome by the addition of the transcription factor TFIID or the formation of initiation complexes before assembly of nucleosomes on the DNA template (Workman and Roeder, 1987). The use of this *in vitro* nucleosome assembly system has been extremely useful in determining the probable action of transactivator proteins *in vivo*.

Studies using the pseudorabies virus IE protein, which is analogous to Vmw175, illustrated that the protein increased the interaction of the transcription factor TFIID with the TATA box (Abmayr *et al.*, 1988). Further studies illustrated that whilst in the absence of nucleosomes the IE protein exhibited a 2-fold stimulation of transcription, assembly of nucleosomes on the DNA template in the presence of PRV IE led to a 40 fold stimulation of transcription by the IE protein. This increase in induction of transcription by the transactivating protein was not due to an overall increase in transcription activity; rather, the assembly of nucleosomes had an inhibitory effect on basal transcription which was relieved by incorporation of the PRV IE protein in the assembly mix, a phenomenom termed antirepression. The transactivatory protein had little effect if it was added after the nucleosomes had been assembled, suggesting that the transactivator competed with nucleosomal proteins for binding sites on the DNA template. The exclusion of individual transcription factors illustrated that the factor TFIID was essential for this induction by the IE protein (Workman *et al.*, 1988).

Further analysis of the GAL4-Vmw65 chimaeric protein illustrated that it too resulted in a 100 fold induction of transcription if added at the same time as nucleosomal factors. This effect required the presence of the acidic activating domain, as the GAL4 DNA binding domain alone had no stimulatory activity. In addition, the GAL4-Vmw65 protein did not require the presence of basal transcription factors for the antirepression activity. Addition of GAL4-Vmw65 together with nucleosomal factors committed the template to a higher level of transcription activity when the basal transcription machinery was later added (Workman *et al.*, 1991).

Studies by Croston *et al.* (1990) illustrated that the histone H1 protein alone had an inhibitory effect on basal transcription, and that the factors Sp1 and GAL4-Vmw65 could relieve this repression, leading to similar levels of induction as observed *in vivo*. These factors were also able to relieve H1-mediated repression when added after the formation of H1-DNA complexes, a situation not observed with nucleosomal assembly assays (see above). However, these studies have been criticised, as the interaction of H1 with DNA does not represent the true nucleosomal structure *in vivo*. To answer these criticisms, Laybourn and Kadonoga (1991) carried out a detailed analysis of nucleosomal assembly *in vitro*. They isolated the core histone proteins and assembled H1-deficient nucleosomes on DNA templates *in vitro*. This nucleosomal structure repressed basal transcription only four fold, and it was found that Sp1 and GAL4-Vmw65 activated these templates to the same degree as obtained with naked DNA templates. The

subsequent addition of H1 in the presence or absence of Sp1 or GAL4-Vmw65 resulted in the observation of up to 200 fold induction of transcription by the transactivating proteins. It was also found that the transactivating proteins could not relieve the H1-mediated repression if added after the H1-nucleosomal structures had assembled on the DNA, an observation in agreement with the results of Workman *et al.* (1988, 1991).

It has been assumed from these studies that the activating proteins compete with nucleosomal factors for binding sites on the DNA template and thus displace nucleosomes from the promoter region. This action could take place during DNA replication *in vivo*, when newly synthesized DNA is being assembled into chromatin structures. The transactivating proteins therefore appear to play a dual role - the interaction with transcription factors to increase levels of basal transcription and the displacement of histones to relieve chromatin-mediated repression (reviewed in Kornberg and Lorch, 1991).

The *in vitro* models of such transcription systems may not represent the true situation *in vivo*, however, as the exact position of nucleosomes within the gene of interest both *in vitro* and *in vivo* would contribute to the transcriptional activity measured. Such maps of nucleosomal structure have not yet been obtained. However, the biochemical relevance of the regulation of gene transcription by chromatin has been demonstrated by several studies of histone mutants in yeast (reviewed in Kornberg and Lorch, 1991; Felsenfeld, 1992).

1D.2.3. Other Factors Involved in the Negative Regulation of Gene Transcription.

Several studies have also identified factors associated with the transcription factor TFIID which negatively regulate basal transcription. It appears that the interaction of such negative factors interferes with the association of TFIID with TFIIA and thus inhibits the formation of initiation complexes. The addition of such factors into *in vitro* transcription systems leads to an increase in the transactivation activity of several cellular factors, in manner similar to that observed with studies on the nucleosomal repression of transcription. The identification of these nuclear factors suggests that the negative regulation of gene transcription is an important process in the control of gene expression *in vivo* (Meisterernst *et al.*, 1991; Meisterernst and Roeder, 1991)

1D.3. DNA Binding Motifs of Transactivating Proteins.

In recent years, a number of studies have concentrated on the identification of and elucidation of the structure of protein DNA binding motifs. As a result, a vast amount of data has accumulated both on the sequence conservation within DNA binding motifs and the structure of such conserved motifs. Sequence analysis has led to the identification of a number of families of DNA-binding proteins. Predictive methods and structural analyses have helped to identify the prominent secondary structural features of such motifs. This section will concentrate on four predominantly helical motifs, as the sequence conservation and structure of these motifs are most relevant to the work presented in this thesis.

1D.3.1. The Prokaryotic Helix-Turn-Helix Motif.

A number of prokaryotic DNA-binding proteins interact with DNA through a highly conserved 20 amino acid motif termed the helix-turn-helix motif (reviewed in Pabo and Sauer, 1984). Sequence analysis has illustrated that this motif is widely conserved within prokaryotic DNA-binding proteins. Dodd and Egan (1987, 1990) developed a method for the detection and statistical evaluation of potential DNA binding motifs exhibiting homology with known prokaryotic helix-turn-helix motifs. This method is described in more detail in Section 3A.4.1. Use of this analysis led to the identification of eukaryotic DNA-binding proteins containing related motifs. A large majority of the proteins identified in this analysis contained the homeodomain DNA-binding motif, a domain with many similarities to the helix-turn-helix motif (see Section 1D.3.2).

X-ray crystallographic studies of the purified proteins identified the structural features of the helix-turn-helix motif in isolation, and model-building studies allowed the evaluation of the possible structures of the complex between protein and DNA (reviewed in Pabo and Sauer, 1984). Cocrytallization of these proteins with synthetic DNA binding sites led to the elucidation of the crystal structure of such complexes (for example, see Anderson *et al.*, 1985; Jordan and Pabo, 1988). In all of the analyses, it was found that the first helix of the motif lies above and across the major groove, whilst the second recognition helix fitted into the major groove, facilitating interactions with the DNA duplex. The conformation of this structure is stabilized by hydrophobic interactions between the two helices. Residues within the recognition helix make contacts with bases and phosphate groups of the DNA helix. This structure shows similarities to the complex formed between eukaryotic homeodomain proteins and DNA (see Section 1D.3.2).

The models generated by the studies described above allowed the prediction of amino acid residues which would play a role in determining the specificity of the interaction of helix-turn-helix proteins with DNA. These predictions were supported by mutagenesis studies of helix-turn-helix motifs, in which replacement of residues of one helix-turn-helix protein with those of another protein changed the DNA-binding specificity of the molecule (Wharton and Ptashne, 1985).

1D.3.2. The Homeodomain Motif.

The homeodomain is a 60-amino acid motif contained in a wide range of eukaryotic developmental genes, including the vertebrate Hox genes and *Drosophila* homeotic genes (Gehring, 1987; Tabin, 1991). In addition, the homeobox domain is contained within the POU-domain, first identified in three mammalian transcription factors and a developmental gene from *Caenorhabditis elegans* (Herr *et al.*, 1988). Upon the basis of amino acid homology, the homeodomain proteins fall into several subsections, with highest conservation exhibited between homeodomain proteins from the same species (Frain *et al.*, 1989).

Nuclear magnetic resonance studies demonstrated that the *Antennapedia* homeodomain contained three α-helices (Otting *et al.*, 1988). Sequence comparisons suggested that two of these α-helices formed a motif similar to the prokaryotic helix-turn-helix motif described in Section 1D.3.1 (Shepard *et al.*, 1984). Mutational analyses of *Drosophila* homeotic proteins demonstrated that residue 9 of helix 3 of the homeodomain motif determined the specificity of protein-DNA interactions (Hanes and Brent, 1989; Treisman *et al.*, 1989).

The crystal structure of a complex containing the *engrailed* homeodomain peptide bound to a DNA site has recently been determined (Kissinger *et al.*, 1990). This structure showed that the first two helices of the domain pack against each other and lie above the major groove, whilst the third recognition helix lies roughly perpendicular to the first two helices, within the major groove of the DNA helix. This structure exhibits some similarity with that of helix-turn-helix motif-DNA complexes, in which a protein helix lies above the DNA helix, and stabilizes a recognition helix, which lies within the major groove and facilitates contact with the DNA molecule. However, some differences exist between the two structures, in that the DNA exists in a slightly different conformation in the two complexes; that the homeodomain recognition helix is longer than that of the helix-turn-helix motif; and that a N-terminal arm of the homeodomain motif makes additional contacts with the DNA molecule. These additional contacts result in the differences observed in the position of the DNA helix in the homeodomain-DNA and repressor-DNA complexes (discussed in Kissinger *et al.*, 1990). In addition, whilst

the prokaryotic helix-turn-helix proteins bind to DNA sites as dimers, homeodomain proteins interact as monomeric units.

Analysis of the contacts made between protein residues and the DNA helix support the genetic data of Hanes and Brent (1989) and Treisman *et al.* (1989). In particular, a highly conserved sequence, WFQNKRR, within the recognition helix appears to be important for the interaction with DNA. Interestingly, this sequence also appears to be conserved within the regulatory proteins from four α-herpesviruses (this thesis).

1D.3.3. The Helix-Loop-Helix Motif.

The helix-loop-helix motif was first identified in two immunoglobulin enhancer binding proteins, E12 and E47; *Drosophila* developmental proteins; the *MyoD* muscle determination protein; and the *myc* protein family (Murre *et al.*, 1989a). These proteins were identified on the basis of sequence homology, and all exhibited the ability to bind to DNA and to dimerize. The helix-loop-helix domain consists of two amphipathic helices separated by a loop of variable length, which often contains one or more helix-breaking residues. Preceding this region is a domain containing a high proportion of basic residues. Hence, the helix-loop-helix motif can also be termed the basic helix-loop-helix motif (bHLH).

The helix-loop-helix motif was shown to be essential for the myogenic properties of *MyoD* (Tapscott *et al.*, 1988). In addition, Murre *et al.* (1989b) illustrated that dimerization of heterologous bHLH proteins required the intact helix-loop-helix domain, and that this heterodimerization resulted in the enhancement of DNA binding efficiency. Further mutational analysis demonstrated that the helix-loop-helix region mediated the dimerization of the proteins, whilst the basic domain mediated the interaction with DNA by protein dimers (Davis *et al.*, 1990). The heterodimerization of helix-loop-helix proteins could provide a mechanism for the fine control of their activity *in vivo*. This hypothesis is supported by the identification of two proteins, *extramacrochaetae* (Ellis *et al.*, 1990; Garrell and Modolell; 1990) and Id (Benezra *et al.*, 1990) which negatively regulated other helix-loop-helix proteins. Sequence analysis of these proteins demonstrated that they contained only the helix-loop-helix region, and lacked the basic domain. Such proteins may therefore interfere with the DNA binding activity of other proteins by the formation of heterodimers. This mechanism of control by heterodimerization may also apply to leucine zipper proteins, which again need to dimerize in order to interact with DNA (reviewed in Jones, 1990). Whilst it has been proposed that the bHLH motif is predominantly helical in nature, no direct structural information concerning the motif has been published.

1D.3.4. The Leucine Zipper Motif.

The leucine zipper motif was first identified in the CAATT binding factor, C/EBP, which exhibited sequence similarities to a region of the *myc, fos* and *jun* transforming proteins as well as to the yeast gene regulatory protein GCN4. The 28 amino acid motif of the C/EBP molecule when arranged on a schematic helical wheel was found to exhibit amphipathy, with a periodic spacing of leucine residues situated on one face of the helix. Sequence analysis illustrated that other proteins containing this motif also exhibited a periodic array of leucine residues. It was proposed that these conserved leucine residues interacted in a zipper-like manner to facilitate dimerization of proteins containing this motif (Landschulz *et al.*, 1988).

This domain also contains a region adjacent to the leucine rich region which has a strong concentration of basic amino acid residues. Mutational analysis has demonstrated that whilst the leucine-rich domain mediates dimerization, the basic motif is essential for the interaction with DNA, and that conservation of the spacing between the two domains is essential for the interaction with DNA (Neuberg *et al.*, 1989; Gentz *et al.*, 1989; Vinson *et al.*, 1989).

Computer model-building studies led to the prediction that leucine zipper motifs bound DNA through a "scissors grip" structure. The leucine zipper regions of the motif were proposed to dimerize and form a coiled coil structure, the helices of which would bifurcate at the basic motif. This would form a Y shaped molecule, the helical arms of which would track into the major groove of the DNA helix. This model was supported by chemical and DNase footprinting data, which illustrated that the leucine zipper motif protected a DNA sequence with exact dyad symmetry and of the correct spacing to fit the computer-generated model (Vinson *et al.*, 1989).

X-ray crystallographic analysis of a complex formed between the GCN4 basic-leucine zipper peptide and DNA has illustrated that the motif does in fact bind to DNA as an uninterrupted α-helix, the leucine-rich domain forming a coiled coil, whilst the basic domain helices diverge to track into the major groove of the DNA duplex. The coiled coil region of the dimer is stabilized by the formation of salt bridges, whilst residues within the basic domain form contacts with the DNA molecule. This model does not support the prediction of Vinson *et al.* (1989) that the basic α-helix contains a kink to allow further interactions with the major groove of the DNA duplex (Ellenberger *et al.*, 1992).

1D.4. A Model for Transactivation by Vmw175.

At present, information concerning the possible mechanisms of transactivation by Vmw175 is limited. Whilst mutational analyses of the protein suggest that the DNA binding function of the protein is essential for transactivation, the protein does not seem to require the presence of specific DNA binding sites in order to activate viral or cellular promoters (see Sections 1B and 1C). Preliminary evidence has also suggested that interaction between Vmw175 and cellular factors may modify the DNA binding properties of Vmw175, questioning the validity of some measurements of *in vitro* DNA binding activity (Smith and DeLuca, 1992). In fact, the only function of Vmw175 which appears to require the presence of a specific, high-affinity DNA binding site for Vmw175 is the autoregulation of the IE3 gene. This autoregulation appears to be mediated by a strong interaction between Vmw175 and the cap site of the IE3 promoter, possibly leading to the steric hindrance of the movement of the basal transcription machinery on the promoter.

The activation of viral genes appears to require the interaction of Vmw175 with DNA, in a manner which does not require that the binding sites be located at fixed positions within the promoter. The protein does not exhibit a strong sequence specificity and may interact at a number of sites within the promoter, thus acting as a viral enhancer-like protein. Transactivation may be brought about by the interaction of Vmw175 with basal transcription factors, preliminary evidence for which has been obtained (Smith and DeLuca, 1992; Pizer *et al.*, 1992). This would provide some similarity between the mechanism of action of Vmw175 and other cellular transactivating proteins (see Section 1D.2.1). In cases where a mutant Vmw175 molecule may interact only weakly with upstream promoter sequences, interactions with DNA-bound cellular factors may help to assemble the activating complex on the promoter template.

Studies on a protein analogous to Vmw175, the pseudorabies virus IE protein, have illustrated that this protein may act through an interaction with the cellular transcription factor TFIID (Abmayr *et al.*, 1988). This may also be the case with Vmw175, as mutational analysis has illustrated that transactivation by Vmw175 requires the presence of a functional TATA box (Everett, 1984a; Coen *et al.*, 1986; Johnson and Everett, 1986a, b; Homa *et al.*, 1988; 1991). Preliminary data that Vmw175 interacts directly with the TATA-binding protein (Pizer *et al.*, 1992) together with the data of Pugh and Tjian (1990; 1992), suggest that Vmw175 may act as a co-activator molecule (see Section 1D.1.3).

Additional studies on the mechanism of action of the pseudorabies IE protein also illustrated that the protein counteracted the repression of transcription mediated by the assembly of nucleosomes on the DNA template (Workman *et al.*, 1988). This

antirepression property may be an important feature of the *in vivo* action of transactivating proteins (see Section 1D.2.2). It would be interesting to investigate the ability of Vmw175 to exhibit such antirepression activity. The structure of lytic HSV-1 DNA has not been stringently investigated. If such DNA were packaged into nucleosomes, this may represent a means by which viral gene expression is limited in the early stages of infection, and one function of Vmw175 may be to release the transcription machinery from such constraints. Alternatively, Vmw175 may act to prevent the formation of nucleosomes on viral DNA.

The replication of viral DNA would displace nucleosomes from the template, thus allowing further binding of the cellular transcriptional machinery to the genome. Interestingly, viral genes which do not require prior DNA replication for expression contain binding sites for additional cellular factors (see Section 1B). The multiple effects of viral and cellular factors may allow the expression of genes normally silenced by chromatin-mediated repression. The investigation of the positions of nucleosomes within promoter regions of viral genes would yield additional information on the role of nucleosome assembly in the control of viral gene transcription.

1E. The Aims of the Research Presented in this Thesis.

The initial aim of the work presented in this thesis was to identify and to carry out a detailed mutational analysis of regions within the DNA binding domain of Vmw175 thought to be involved in contacting the DNA molecule. Target regions were identified by sequence analysis of the protein in comparison to other viral and cellular transactivators, together with the analysis of information obtained from previous mutational analyses of Vmw175.

Mutant forms of Vmw175 created by this study were expressed as DNA binding domain polypeptides in bacteria, and assayed for DNA binding and dimerization properties *in vitro*. It was hoped that these assays would lead to the identification of regions and/or residues of the DNA binding domain of Vmw175 essential for its DNA binding and dimerization properties. In addition, such analyses may lead to the separation of the DNA binding and dimerization properties of the domain.

In order to determine whether the ability of Vmw175 to bind to DNA correlates with its ability to transactivate viral promoters, mutant forms of the protein were introduced into the HSV-1 genome. The mutant Vmw175 molecules used in this part of the study included previously constructed insertional mutants (Paterson, 1989) as well as

point and deletion mutants which were constructed during the course of this work. Analysis of the ability of such recombinant viruses to replicate in tissue culture would give an indication of the transactivation activity of each mutant protein.

✳ derived from BHK-21 cells

Chapter 2. Materials and Methods.

2A. Materials.

2A.1. Cells

(i) BHK-21 clone 13 (Macpherson and Stoker, 1962), a fibroblastic cell line originally derived from baby hamster kidney cells, was generally used for growth of HSV-1 strain 17 syn$^+$, transfection assays and recombination experiments.

(ii) M64A cells (obtained from Dr. N. D. Stow of this Institute) express Vmw175 and were used to complement the growth of the virus D30EBA (Paterson and Everett, 1990) which contains a lethal mutation in IE gene 3. M64A cells are equivalent to the cell line M65 described by Davison and Stow (1985) and are a mixed cell line derived from BHKtk$^-$ cells which had been cotransformed with the HSV-1 large Hind III-Xho I subfragment of Xho Ic, from which ori$_S$ had been deleted, and a fragment containing the HSV thymidine kinase (tk) gene, and selcted for thymidine kinase gene expression.

(iii) RR1 cells (obtained from Dr. P. A. Johnson, UCSD Medical School) also express Vmw175 and were used to complement the growth of the virus D30EBA. These cells were also used to quantitate D30EBA viral preparations by means of a plaque assay. As the cell line has not been transformed with DNA fragments containing HSV-1 ori$_S$ or 'a' sequences, this results in a reduced risk of recombination between the virus D30EBA and Vmw175 coding sequences. A DNA fragment carrying the neomycin resistance gene (neo^R) had been introduced with the Vmw175 coding sequences, allowing selection of transformed cells in medium containing the antibiotic Geneticin (G418). (Hirth *et al.*, 1982).

2A.2. Media.

(i) BHK cells were grown in Glasgow Modified Eagle's Medium (GMEM) (Busby *et al.*, 1964) supplemented with 100units/ml penicillin, 100µg/ml streptomycin, 0.2µg/ml amphotericin, 0.002% phenol red and 10% newborn calf serum. (=EC10).

(ii) M64A cells were grown in EF10 supplemented with 1ng/ml aminopterin, 5ng/ml thymidine, 15ng/ml hypoxanthine and 15ng/ml glycine (=HAT medium). [EF10=GMEM supplemented with 100units/ml penicillin, 100µg/ml streptomycin, 0.2µg/ml amphotericin, 0.002% phenol red and 10% foetal calf serum].

(iii) RR1 cells were grown in Dulbecco's Modified Eagle's Medium supplemented with 100units/ml penicillin, 100μg/ml streptomycin, 0.002% phenol red and 10% foetal calf serum. G418, final concentration 400μg/ml, was included in alternate passages.

2A.3. Viruses.

All viruses used in this study were derivatives of the HSV-1 Glasgow strain 17 syn$^+$, which forms non-syncytial plaques on BHK cells (Brown *et al.*, 1973). Derivative D30EBA contains a large deletion from nucleotides 246-3806 within both copies of the IE-3 gene (Paterson and Everett, 1990).

2A.4. Bacteria.

All plasmids used in this study were propagated in *E. coli* strains HB101 (*F$^-$; RamC1; ProA2; galK2; strA; recA;* Boyer & Roulland-Dussoix, 1969) or DH5 (*SupE44; hsdR17; recA1; endA1; gyrA96; thi-1; relA1;* Low, 1968). The pET vectors used in this study were grown in *E. coli* strain BL21 DE3 pLysS (*hsdS; gal(λcIts857; ind1; S*am7*; nin5; lac*UV5-T7 gene*1;* Studier & Moffatt, 1986) for induction of polypeptide expression, M13 bacteriophage vectors were propagated in *E. coli* strains JM101 (*supEΔ(lac-ProAB); F'[traD36; proAB$^+$; lacI^q; lacZ*ΔM15*],* ;Messing, 1979) or TG1 (*supE; hsdΔ5; thi; Δ(lac-proAB); F'[traD36; proAB$^+$; lacI^q; lacZ*ΔM15*],;* Gibson, 1984).

2A.5. Bacterial Culture Media.

HB101 and DH5 strains of bacteria were grown on L-Broth (10g/l NaCl, 10g/l Difco Bactopeptone, 5g/l yeast extract, pH7.5 prior to sterilization). All other strains were grown on 2YT Broth (5g/l NaCl, 16g/l Difco Bactotryptone, 10g/l yeast extract). TG1 strain was grown on glucose/minimal media plates (containing 1xM9 salts [10xM9 salts=60g/l Na_2HPO_4, 30g/l KH_2PO_4, 10g/l NH_4Cl, 5g/l NaCl]; 1mM Mg SO_4; 1mM thiamine.HCl; 0.1mM $CaCl_2$; 0.2% (w/v) glucose; 1.5% (w/v) agar) for selection for the F plasmid. Agar plates contained 1.5% (w/v) agar in L-broth and where appropriate the medium was supplemented with 100μg/ml ampicillin or 25μg/ml chloramphenicol. Top agar contained 10g/l Difco Bactotryptone, 8g/l NaCl, 8g/l agar.

2A.6. Plasmids.

The following plasmids were kindly provided by the acknowledged authors:
p175: (Everett, 1987). The HSV-1 IE-3 gene is cloned on a BamHI-SstI fragment spanning IR_S and IR_L of HSV-1 under the control of the SV40 early promoter and enhancer.
pI9-pI20: (Paterson & Everett, 1988a). Derivatives of p175 created by insertion of a 12bp EcoRI linker into SmaI, NaeI, AluI or RsaI sites within the IE-3 coding region.
pA585.4: (C. M. Preston, unpublished). A derivative of the plasmid pET8c (Studier *et al.*, 1990) created by removal of the HindIII site of the parent pBR322 vector and insertion of oligonucleotide linkers containing a HindIII restriction site between the NcoI/BamHI sites of the T7 promoter/terminator sequences.
p585T7a, b, c: (J. Tyler, unpublished). Derivatives of the plasmid pA585.4, created by the removal of the EcoRI site; removal of the PvuII/SphI fragment; and addition of the pUC19 multiple restriction site polylinker within the T7 promoter/terminator region. All three plasmids also contain NcoI/EcoRI oligonucleotide linkers within the NcoI site at coordinate 607 to enable expression of cloned fragments in all three reading frames. (See Figure 3C.1).

2A.7. Reagents.

Oligonucleotides were synthesized in this Institute by Dr. John McLauchlan using a Biosearch 8600 DNA Synthesizer.

Most analytical grade chemicals were purchased from BDH Chemicals UK or Sigma Chemical Co.; TEMED, gelatin and ammonium persulphate were obtained from Bio-Rad Laboratories; caesium chloride from Melford Laboratories Ltd.; boric acid from Koch-Light Laboratories; nitrocellulose paper from Scleicher and Schuell Inc.; Hybond-N^+ paper from Amersham LKB Ltd,; 0.25mm silica gel TLC plates were obtained from Camlab; unlabelled dNTPs and double stranded poly dI:polydC were obtained from Pharmacia; formamide from Fluka; and G418 (Geneticin) was from Gibco.

2A.8. Radiochemicals.

Radiochemicals were obtained from Amersham at the following specific activities:

5'[α-^{32}P] dNTPs, 3000Ci/mmol
5'[γ-^{32}-P] ATP, 5000Ci/mmol
L-^{35}S-methionine, around1000Ci/mmol

2A.9. Enzymes.

DNase I, RNase A and lysozyme were supplied by Sigma; Proteinase K, DNA polymerase holoenzyme, T4 polynucleotide kinase, T4 DNA ligase, and most restriction enzymes were purchased from Boehringer Mannheim. Klenow fragment polymerase was expressed by Dr. E. Telford and purified in the Institute. Sequenase™ T7 polymerase and sequencing kits were purchased from United States Biochemical.

2A.10. Antibodies.

The anti-ICP4 polyclonal antibody was a kind gift from Dr. K. Wilcox (Wu and Wilcox, 1990). Horse-radish peroxidase conjugated protein A was purchased from Sigma.

2A.11. Solutions.

CLB (cell lysis buffer): 0.5% SDS, 20mM Tris-HCl (pH 7.5), 2mM EDTA.

Coomassie Brilliant Blue stain: 2mg/ml Coomassie Brilliant Blue in methanol:H_2O:acetic acid (200:200:28)

50x Denhardt's: 1% Polyvinylpyrrolidone, 1% BSA, 1% Ficol

20x dNTPs: 1mM dATP, dCTP, dGTP, and dTTP.

5x gel loading buffer: 1x TBE, 1% SDS, 50% glycerol, plus bromophenol blue.

Gel soak I: 0.5M NaOH, 1.5M NaCl.

Gel soak II: 1.5M NaCl; 0.5M Tris-HCl pH 7.2; 0.001M EDTA.

Giemsa stain: 1.5% suspension of Giemsa in glycerol, heated at 56°C for 2h and diluted with an equal volume of methanol.

200x HAT: 0.2mg/ml aminopterin, 1mg/ml thymidine, 3mg/ml hypoxanthine, 3mg/ml glycine, pH to 10 to dissolve.

2x HBS: 280mM NaCl, 50mM HEPES, 1.5mM Na_2HPO_4, pH to 7.12.

Hybridisation mix: 6x SSC, 5x Denhardt's, 0.1% SDS, 50µg/ml denatured salmon sperm DNA.

5x kinase buffer: 350mM Tris-HCl (pH 7.5), 50mM $MgCl_2$, 25mM DTT.

5x ligase buffer: 250mM Tris-HCl (pH 7.5), 50mM $MgCl_2$, 100mM DTT, 50µg/ml BSA.

MonoS buffer A: 0.1M NaCl, 50mM HEPES pH7.5, 0.01% (w/v) CHAPS, 1mM PMSF, 0.1mM DTT.

MonoS buffer B: 1M NaCl, 50mM HEPES pH7.5, 0.01% (w/v) CHAPS, 1mM PMSF, 0.1mM DTT.

PBS-A: 170mM NaCl, 3.4mM KCl, 10mM Na_2HPO_4, 2mM KH_2PO_4 (pH 7.2).

PBSc: PBS-A plus $CaCl_2.H_2O$ and $MgCl_2.6H_2O$ at 1g/l.

Resuspension buffer: 0.1M NaCl, 50mM HEPES, 0.4% (w/v) CHAPS, 1mM PMSF, 0.1mM DTT.

RGB (SDS-PAGE resolving gel buffer): 181.5g Tris, 4g SDS, 1l H_2O, pH to 8.9 at RT with HCl.

SDS boiling mix: 1ml SGB, 0.8ml 25% SDS, 0.5ml β-mercaptoethanol, bromophenol blue.

SDS tank buffer: 6.32g Tris, 4g glycine, 1g SDS, 1l H_2O.

SGB (SDS-PAGE stacking gel buffer): 59g Tris, 4g SDS, 1l H_2O, pH to 6.7 at RT with HCl.

20x SSC: 174g/l NaCl, 88.2g/l Trisodium Citrate.

20x SSPE: 3.6M NaCl, 0.2M NaH_2PO_4, 0.02M EDTA, pH7.7.

STET: 8% sucrose, 5% Triton X-100, 50mM EDTA, 60mM Tris-HCl (pH 8).

Superose buffer: 0.5M NaCl, 50mM HEPES (pH 7.5), 0.01% (w/v) CHAPS, 1mM PMSF, 0.1mM DTT.

20x TBE: 2.5M Tris, 800mM boric acid, 54mM EDTA, not pH'd.

10x TBE (for gel retardation assays): 890mM Tris-HCl, 890mM Boric Acid, 25mM EDTA, not pH'd.

TBS (for transfections): 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 (pH 7.5), 1mM EDTA (pH 8).

TE/5: 2mM Tris-HCl pH 8; 0.02mM EDTA.

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

TEP: 1mM Tris-HCl (pH 7.5), 0.05mM EDTA.

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

Versene: 0.6mM EDTA in PBS-A plus 0.002% w/v phenol red.

2B:Methods.

2B.1. Tissue Culture.

BHK cells were routinely passaged in 850cm^2 roller bottles seeded with approximately 5x10^7 cells at 37°C in 100ml EC10 and an atmosphere of 5% CO_2 in air. Confluent monolayers were harvested in 20ml EC10 after two versene washes and brief trypsinization with Trypsin:versene (1:4). Cells were resuspended by pipetting and remained viable for at least five days stored at 4°C.

M64A and RR1 cells were routinely passaged in 175cm^2 (Nunclon) flasks in appropriate medium and harvested in 10ml medium. These cells did not remain viable at 4°C and for long term storage harvested cells were pelleted and resuspended in medium plus 5% glycerol and 15% foetal calf serum, aliquoted and frozen slowly to -140°C. Recovery was by rapid thawing and resuspension in growth medium.

2B.2.Preparation of Stocks of Infectious Virus.

An 80% confluent monolayer of BHK or RR1 cells was seeded with 10^6 pfu of virus per roller bottle in 40ml appropriate medium (a moi of 0.003pfu/cell) and incubated for 2-4 days until the cells exhibited obvious cpe. The cells were shaken into the medium and pelleted at 3000rpm for 15 min at 4°C. Cell associated virus (CAV) was prepared by sonicating the pellet in 2ml medium. Cell released virus (CRV) was pelleted from the supernatant at 12000rpm for 2 hours at 4°C and gently resuspended in 2ml supernatant and sonicated in a sonibath. Sterility checks were performed by streaking on blood agar plates and incubating at 37°C for 5 days. Virus stocks were stored at -70°C.

2B.3. Titration of Virus Stocks.

Appropriate indicator cells were seeded at 0.7x10^6 cells per 35mm plate in 2.5ml medium. The following day cells were infected with serial 10-fold virus dilutions in 0.3ml medium. After 1hr adsorption time (gently mixing every 10 minutes) plates were overlaid with the appropriate medium containing 5% human serum, to prevent secondary plaque formation, and incubated for 2-3 days prior to fixing with Cidex and staining with Giemsa stain. The plates were washed after 15 minutes and plaques counted using a dissection microscope.

2B.4. Preparation of Cell Released Viral DNA.

Viral particles in the culture medium were pelleted as described in Section 2B.2 above and resuspended in 1ml TE plus 100mM NaCl. SDS was then added to 1% v/v and viruses disrupted by gentle inversion at 37°C. DNA was then extracted by serial 30 minute extractions, with gentle inversion (twice with phenol, then once with phenol/chloroform and once with chloroform). DNA was precipitated with 0.05 volumes of 5M NaCl and 2.5 volumes of ethanol, washed twice, dessicated and resuspended overnight in TE. DNA was quantitated on agarose gels compared to known standards and stored at -20°C.

2B.5. Preparation of Total Infected Cell DNA.

10mm linbro wells were seeded with $1x10^5$ permissive cells and infected the following day with a moi of 0.001 pfu/cell. After 2-3 days, when cpe became apparent, the supernatant was removed and stored as CRV stock. 0.2ml of cell lysis buffer (CLB) containing 0.1mg/ml proteinase K was added to the cell monolayer and incubated at 37°C for 3-5 hours. The glutinous extract was transferred to 1.5ml Eppendorf tubes and 15µl 5M NaCl added. Serial 30 minute phenol then chloroform extractions were performed, with gentle inversion, followed by precipitation with 2.5 volumes of ethanol, resuspension in 0.3M Na Acetate pH5, a further ethanol precipitation, one wash with 70% ethanol, dessication and finally resuspension overnight at 4°C in 50µl TE.

2B.6. Analysis of Polypeptides.

2B.6.1 Pulse labelling of polypeptides.

Sub-confluent permissive cells ($1x10^5$ cells on a 10mm linbro well) were infected at a moi of 5 pfu/cell in 0.3ml medium. After a 1h adsorption period (shaking every 10 min), cells were overlaid with growth medium and infection allowed to proceed at 37°C. The next day cells were washed with PBSc and labelled for 30mins in 200µl PBS containing 50-100µCi [^{35}S] methionine. Cells were then washed and harvested in 100µl SDS boiling mix. Extracts were stored at -20°C prior to analysis by SDS polyacrylamide gel electrophoresis.

2B.6.2 SDS polyacrylamide gel electrophoresis.

Electrophoresis was carried out as described by Marsden *et al.*, 1978. Polypeptides were resolved on 0.75mm thick 7.5% or 12.5% single concentration polyacrylamide gels using a Bio-Rad Mini-Protean^{TM} II electrophoresis kit. Gel plates were washed with dilute Decon, distilled water, then ethanol and left to dry. Gels were made up as follows:

Solution	7.5% resolving gel	12.5% resolving gel	stacking gel
30% acrylamide (2.5% cross-linker)	2.5ml	4.17ml	0.4ml
resolving gel buffer	2.5ml	2.5ml	
stacking gel buffer			2.5ml
distilled H_2O	5ml	3.33ml	1.4ml
10% APS	80µl	80µl	20µl
TEMED	8µl	8µl	3µl

The resolving gel was overlaid with butan-2-ol and allowed to set. After washing, the stacking gel was prepared and overlaid, wells were created with a teflon comb.

Protein samples were boiled for 2 min in 1 or 2 volumes of SDS boiling mix prior to loading. Gels were run at a constant voltage of 200V until the dye front reached the bottom of the gel. Gels of [^{35}S]methionine labelled polypeptides were fixed for 1h in methanol: acetic acid: water (50: 70: 50) before drying down under vacuum at 80°C onto a sheet of Whatman 3mm paper and exposing to Kodak S XS-1 film for 3-24h. Gels of non-radiolabelled polypeptides were stained for 2 min with Coomassie Brilliant Blue before destaining overnight in protein gel destain.

2B.6.3 Western blot analysis.

The Western blotting technique was that of Towbin *et al.*, 1979, with several modifications. Protein extracts were electrophoresed on 12.5% SDS-polyacrylamide gels as described in Section 2.6.2 and then transferred onto nitrocellulose using a Bio-Rad mini Trans-blot^{TM} apparatus. Two foam pads and two sheets of Whatman 3mm chromatography paper were presoaked in transfer buffer (192mM Glycine, 25mM Tris-HCl pH 8.3, and 20% Methanol) along with the nitrocellulose sheet to be used in the transfer. The gel was laid on one presoaked filter paper, and the nitrocellulose sheet and second filter paper laid on top. This assembly was held between two foam pads and

placed within the transfer tank with the nitrocellulose sheet towards the anode. Proteins were transferred electrophoretically in transfer buffer at 250mA for 3h at room temperature.

Following transfer, the nitrocellulose was blocked in blocking buffer (3% gelatin, 1x TBS, 0.05% Tween-20) at 37°C for 1 hour, the buffer being changed after 30 min, and then washed twice for 5 min in 1x TBS, 0.05% Tween-20 at room temperature. Antibody solutions were made up to the desired concentration in 1% gelatin, 1x TBS, 0.05% Tween-20. The dilutions were added to the nitrocellulose and incubated overnight at room temperature.

The following day, six five minute washes were carried out in 1x TBS, 0.05% Tween-20 at room temperature and the blot then incubated for 1h at room temperature in a 1/1000 dilution of protein A:HRP conjugate (diluted in 1% gelatin, 1x TBS, 0.05% Tween-20). The filter was then washed twice in 1x TBS, 0.05% Tween-20 for 5 min at room temperature and once in 1x TBS for 5 min at room temperature before addition of the colour reagent. 20mls of an ice cold solution of 3mg/ml HRP colour developer reagent in methanol and 100mls of 0.015% H_2O_2 in 1x TBS were mixed and immediately added to the nitrocellulose filter. The filter was incubated in the dark for 10-30 min, until colour developed, and the reaction stopped by washing the filter in running water.

Alternatively, proteins were visualized with the Amersham ECL™ detection system, according to the manufacturer's instructions.

2B.7. Transfection of Cells with Viral DNA for Marker Rescue.

Subconfluent BHK cells were transfected on 35mm plates using the calcium phosphate precipitation method of Corsalo & Pearson (1981). 0.5µg of intact viral DNA plus a 10-fold molar excess of linearized plasmid and 2-5µg calf thymus DNA were made to a final volume of 140µl with TEP. 20µl of 2M $CaCl_2$ was added dropwise to the solution with frequent mixing, and this mixture in turn added dropwise to 160µl 2xHBS. After standing for 30 min the medium was removed from the cells and the precipitate added. The cells were then incubated for 40 min at 37°C or 38.5°C, growth medium replaced, and incubation continued for a further 3-4 h. The cells were washed once with growth medium, overlaid with 25% DMSO in 1x HBS for 30sec, washed a final time, and overlaid with growth medium. After 2-4 days incubation at 31°C or 37°C the cells exhibited extensive cpe, or plaques; the cells were then scraped into the medium and sonicated to release infectious virus. Virus stocks were stored at -70°C.

2B.8. Gel Retardation Assay.

The binding of DNA binding domain polypeptides to the IE gene 3 capsite was assayed as described by Muller (1987). The 45bp AvaI/BamHI and 135bp EcoRI/BamHI fragments spanning the IE-3 capsite (-17 to +27 and -108 to +27 respectively) were end-labelled by filling in using Klenow fragment polymerase as described in Section 2B.24.2 and purified on an 8% polyacrylamide gel. Binding reactions were carried out at the appropriate temperature for 20 minutes in a 20µl mix containing 1µg poly(dI).poly(dC), about 1ng probe (2000cpm) and 1-10µl protein extract in 10mM Tris-HCl pH 7.5, 1mM EDTA, 0.1M NaCl, 0.1% NP40. Samples were resolved on 4% polyacrylamide gels (3.3% cross linker) run in 0.5x TBE (prepared from 10x TBE stock), and dried down before autoradiography. Where binding reactions were carried out at 0°C, electrophoresis was carried out at 4°C.

2B.9. Large Scale Plasmid DNA Preparations.

Isolated single colonies on plates selective for a plasmid-borne resistance marker were inoculated into 3ml L-broth and grown overnight. These were added to 350ml L-broth in 2l flasks and grown shaking at 37°C overnight.

Plasmid DNA was prepared by the "maxi-boiling" technique (Holmes and Quigley, 1981). Cells were pelleted at 10000 rpm for 6min (Sorvall GS3 rotor) and resuspended in 17ml STET per 350ml culture. 2ml 10mg/ml lysozyme in STET was added for a 1min incubation prior to bringing the suspension to the boil for 45sec. The lysate was cleared by centrifugation at 20000 rpm for 45 min at 4°C (Sorvall SS34 rotor) and DNA precipitated from the supernatant by addition of 0.9 vol isopropanol. After pelleting at 3000 rpm at room temperature for 10 min (Beckman GPR centrifuge), the drained pellets were resuspended in TE and CsCl added to give a final density of 1.6mg/ml and EtBr to 0.5mg/ml. The solution was cleared by spinning at 3000 rpm for 10 min at room temperature after a 15min incubation on ice. The DNA was banded by centrifugation at 40000 rpm at 15°C for 16-48 h in the vertical TV865 rotor. The DNA was visualized by daylight or long-wave uv illumination and the lower, supercoiled plasmid DNA recovered with a large bore needle and syringe, after removal of the upper, linearized chromosomal DNA band. The DNA was extracted twice with butan-1-ol (TE saturated), dialysed against TE at RT for 3h, treated with 50µg/ml RNase (1h/65°C), then 50µg/ml proteinase K plus 0.1% SDS (1h/37°C) before extraction with phenol/chloroform then chloroform and precipitated twice with ethanol. The DNA was lyophilized, resuspended and its concentration determined by spectrophometry (OD_{260} 1.0 = 50µg DNA/ ml).

2B.10. Miniprep Plasmid DNA Preparation.

Single plasmid-transformed bacterial colonies were grown overnight in 2ml L-broth. 1.5ml stationary phase culture was pelleted at 12000 rpm for 15sec (benchtop microcentrifuge) and cells resuspended in 200µl STET. The cell suspension was immediately placed in a 100°C water bath for 45sec after addition of 5µl lysozyme (10mg/ml), then centrifuged at 12000 rpm for 10min. The pellet was removed and DNA precipitated and pelleted from the supernatant by addition of 0.9 vol isopropanol and spinning at 12000 rpm for 5min. The pellets were washed with 70% ethanol, dried and resuspended in 20µl TE. Miniprep DNA was stored at -20°C and 5µl was used for restriction enzyme analysis. (0.2µl 1mg/ml RNAase was added to digestions of miniprep DNA).

2B.11. Restriction Enzyme Digests.

Restriction enzyme (RE) digests were generally carried out in 20µl volumes of the appropriate RE buffer, using 0.5µg DNA and 1u enzyme for 1h. Most digests were carried out using the Boehringer Mannheim buffer system, with commercially supplied 10x concentration buffers.

EcoRI and BamHI double digestions were always carried out in buffer H. All digests were carried out at 37°C. Preparative rather than diagnostic RE digestions used increased amounts of DNA and RE under similar conditions, in 20-50µl volumes.

2B.12. Separation and Purification of Purified DNA Restriction Fragments.

2B.12.1 Non-denaturing agarose gels.

Slab gels of 0.5-1.5% agarose in 1x TBE were run submerged at up to 12V/cm. Sampled were loaded in 0.2 vol loading buffer. Gels were run for 1-3h, until the dye almost ran off the end, and then stained for 5min in 1mg/ml EtBr and examined and photographed with short-wave uv illumination (long-wave uv was used to reduce DNA damage when preparative gels were being analysed).

To purify a resolved DNA fragment from a gel the DNA was run onto a strip of DE-81 paper inserted in front of the desired band, (DE-81 paper was presoaked in 2.5M NaCl overnight, then stored in 1mM EDTA at 4°C). After electrophoresing the fragment onto the paper it was eluted by pulverization and agitation in 1.5M NaCl, 1x TE for 2-16 h at 37°C, strained through siliconized glass wool, clarified and then ethanol precipitated.

2B.12.2 Non-denaturing polyacrylamide gels.

In order to resolve accurately DNA fragments 40-400bp in length, 4-8% polyacrylamide (acrylamide:bisacrylamide, 19:1) gels were run in 1x TBE. Gels were prepared in 1.5mm thick vertical glass plate sandwiches, and cross-linked with 0.001vol TEMED, catalysed with 0.01vol 10% APS. Samples were loaded as above and electrophoresed at a maximum voltage of 16V/cm, and DNA visualized by autoradiography or EtBr staining. ^{32}P-labelled DNA fragments were cut out of the gel after detection by autoradiography and eluted by pulverization and agitation in TE for 16 h at 37°C. Acrylamide was removed from the mixture by straining through siliconized glass wool, and DNA purified by ethanol precipitation.

2B.12.3 Denaturing polyacrylamide gels.

0.35mm thick gels prepared to final concentration of 6-8% acrylamide (acrylamide:bisacrylamide, 19:1), 1x TBE and 7M urea were used to resolve the products of DNA sequencing reactions. DNA sequencing samples were denatured at 95°C for 2min and run at 40W (20cm wide gels) or 70W (40cm wide gels; BRL S2 sequencing gel apparatus) for 2-5h on gels which had been pre-run for 30min. Gels were dried down and exposed to Kodak X-Omat S XS-1 film for 16-24h at room temperature without an intensifying screen.

2B.12.4 Sephadex columns (gel exclusion chromatography)

Sephadex G-50 "nick" columns obtained from Pharmacia were pre-equilibrated with 2 column volumes of 1x TE, 0.1% SDS. Columns were loaded with 100-120µl sample (containing bromophenol blue dye) and the fastest eluted fractions (detected by scintillation counting) taken as purified DNA.

2B.13. DNA Ligations.

DNA ligations were performed in 1x ligase buffer, with 50-500ng DNA and 1-2u DNA ligase. DNA was heated to 65-70°C for 5min prior to transformation of competent bacteria.

2B.14. Preparation of Competent *E. coli.* strains HB101, JM101 and DH5.

A colony purified overnight culture of bacteria was diluted 1/100 into fresh L-broth and grown shaking at 37°C until reaching an OD_{540} of 0.3-0.6. Using sterile,

ice cold procedures throughout, cells were pelleted at 3000 rpm for 10 min (Beckman GPR centrifuge), resuspended in 1/4 vol 0.1M $MgCl_2$ and immediately pelleted again at 3000 rpm for 10 min. After resuspension in 1/20 vol 0.1M $CaCl_2$ the cells were left for 45min on ice, resuspended in the same volume of 0.1M MOPS, 50mM $CaCl_2$, 20% glycerol and left 20min before aliquoting into 100-500µl and snap freezing and storing at -70°C.

2B.15. Preparation of Competent *E. coli* strain TG1.

TG1 cells were streaked onto glucose/minimal media plates and incubated overnight at 37°C, following which they could be stored at 4°C for up to 4 weeks. A single colony was inoculated into 10ml 2YT broth and grown shaking at 37°C. The following day, the culture was diluted 1/100 into 40ml fresh 2YT and grown shaking at 37°C for 2h or until OD_{550} = 0.3. Using sterile, ice cold techniques throughout, the cells were pelleted at 3000 rpm for 5 min (Beckman GPR centrifuge) and resuspended in 1/2 vol 0.1M $CaCl_2$, 1mM $MgCl_2$. After 20min incubation on ice, the cells were again pelleted at 3000 rpm for 5min and resuspended in 1/10 vol 0.1M $CaCl_2$, 1mM $MgCl_2$. Competent cells were stored for at least 1h on ice before transformation.

2B.16. Preparation of Competent *E. coli* strain BL21 (DE3) plysS.

A glycerol stock of bacteria was streaked onto an agar plate containing 25µg/ml chloramphenicol and incubated overnight at 37°C. The following day a single colony was inoculated into 10ml 2YT + 25µg/ml chloramphenicol and grown shaking at 37°C for 2h. Using sterile, ice cold techniques throughout, cells were pelleted at 3000 rpm for 10 min (Beckman GPR centrifuge) and resuspended in 1/4 vol 0.1M $CaCl_2$. The cells were immediately pelleted at 3000 rpm for 10 min and resuspended in 1/20 vol 0.1M $CaCl_2$. Competent cells were stored for at least 1h on ice before transformation.

2B.17. Transformation of Competent *E. coli* strains HB101, DH5 and BL21 (DE3) plysS.

100-200µl cells were thawed on ice, 100-300ng of DNA (either ligation mix, miniprep plasmid or purified plasmid) added, and the mixture incubated on ice for 30min. The cells were heat shocked at 37°C for 1min, 500µl prewarmed L-broth added and the mixture incubated at 37°C for 30min prior to plating the mixture out onto plates containing appropriate antibiotic. Plates were incubated at 37°C overnight, following which they could be stored at 4°C for up to two weeks.

2B.18. Transformation of Competent *E. coli* strains TG1 and JM101.

300µl competent TG1 cells or 100µl competent JM101 cells were used for each transformation. Approximately 1-200ng M13 DNA was added and the mixture incubated on ice for 40min. TG1 cells were then heat shocked at 42°C for 45sec and placed on ice for 5min before plating out. 200µl of stationary phase TG1 or JM101 culture and 3-4ml molten top agar (at 42°C) was added to each tube and the mixture poured onto dry, prewarmed agar plates. After the top agar had set, plates were incubated inverted overnight at 37°C, after which they could be stored at 4°C for up to two weeks.

2B.19. Small-Scale Preparation of Replicative Form and Template DNA's from Bacteria Transformed with M13 Vectors.

200µl of a fresh overnight culture of *E. coli* JM101 bacteria was added to 20ml 2YT broth. Single, isolated plaques, freshly picked with a sterile pasteur pipette, were added to 1.5ml aliquots of this suspension in sterile universals. The cultures were grown shaking at 37°C for 5-6h. The final infected culture was transferred to a 1.5ml eppendorf vial and centrifuged twice at 12000 rpm for 2min, turning through 180° between spins. Single-stranded template DNA and replicative form (RF) DNA were then prepared as follows.

2B.19.1 Template DNA

800µl of supernatant was transferred to a tube containing 200µl of 20% PEG, 2.5M NaCl, the rest of the supernatant being stored as bacteriophage stock. The supernatant and PEG mixture was mixed and left at room temperature for 30min or at 4°C overnight. Precipitated bacteriophage were recovered by centrifugation at 12000 rpm for 10min in a bench top microcentrifuge. All the supernatant was carefully removed using a sterile, disposable pipette tip and the pellet resuspended in 100µl TE. 100µl phenol (equilibrated with TE) was added and the contents of the tube mixed well and then centrifuged at 12000 rpm for 10min. The phenol phase was removed and DNA precipitated from the aqueous phase with 1/10 vol 4M Na acetate and 2.5 vols ethanol. DNA was then collected by centrifugation at 12000 rpm for 10min, washed with 70% ethanol, lyophilized and resuspended in 20µl TE/5.

2B.19.2 RF DNA.

The pellet of infected bacteria from above was resuspended in 200µl STET and RF DNA prepared by the method described in Section 2B.10.

2B.20. Preparation of Large Scale Single-Stranded Template DNA from Bacteria Transformed with M13 Vectors.

For mutagenesis reactions, large scale preparations of single-stranded template DNA were carried out as follows.

100ml 2YT broth was inoculated with 1ml fresh overnight JM101 culture and 2-5µl bacteriophage stock. This mixture was then grown shaking at 37°C for 5-6h. Cells were pelleted by centrifugation at 6500 rpm for 30 min at 4°C (SS34 rotor) and the supernatant transferred to a clean tube. 0.2 vol of 20% PEG, 2.5M NaCl was added and the mixture left to stand at 4°C for 1h. Precipitated bacteriophage were recovered by centrifugation at 6500 rpm for 20 min at 4°C, and the supernatant discarded. The pellet was again centrifuged at 6500 rpm for 5 min and remaining supernatant removed with a sterile pasteur pipette. Precipitated bacteriophage were resuspended in 500µl TE and transferred to a clean 1.5ml Eppendorf tube. Any remaining cells were removed by centrifugation at 12K rpm/ 10min in a benchtop microcentrifuge and transfer of the supernatant to a fresh tube. 200µl 20% PEG/ 2.5M NaCl was added to the supernatant and the mixture left to stand for 15min at room temperature or overnight at 4°C. Precipitated bacteriophage were again collected by centrifugation at 12K rpm/ 10min and all supernatant carefully removed. The pellet was resuspended in 500µl TE and extracted once with phenol (equilibrated with TE) and twice with chloroform. The aqueous phase was then divided between two tubes and DNA precipitated with 1/10 vol 3M Na acetate (pH 5) and 2.5 vols ethanol. DNA was collected by centrifugation at 12000 rpm for 15 min, washed with 70% ethanol, lyophilized and resuspended in 50µl TE. DNA concentration was determined by spectrophometry (OD_{260} 1.0 = 40µg/ml single stranded DNA) and adjusted to give a final concentration of 1mg/ml. Single-stranded DNA was stored at 4°C.

2B.21. Dideoxy Sequencing of DNA.

Products of site-directed mutagenesis reactions and plasmid constructs were sequenced by the dideoxynucleotide chain termination method of Sanger *et al*, (1977). DNA was sequenced with either Klenow large fragment polymerase or Sequenase™ T7 polymerase. Both sequencing reactions utilized 1/10 of a template M13 DNA small scale preparation or 2-6µg CsCl purified plasmid DNA. Plasmid DNA was denatured

with 0.4M NaOH at room temperature for 10min before neutralization and precipitation with 1/10 vol 3M Na acetate (pH 5) and 3 vols ethanol.

For the Klenow LFP reactions, template DNA was annealled to 5pmoles of appropriate primer in a final volume of 1x sequencing buffer (10mM Tris-HCl pH 8.5, 10mM $MgCl_2$) at 37°C for 20-30 min. 2 units of Klenow LFP were added to the annealled DNA and 2µl aliquots dispensed into 4 corresponding T, C, G and A wells of a round bottomed 96-well microwell plate (Nunclon). To each well was added 2µl of the appropriate sequencing mix containing dNTP's and specific ddNTP's (see below), dATP at a final concentration of 0.09µM, and 4µCi [α-^{32}P] dATP. The tray was spun briefly and the reaction allowed to proceed at 37°C for 15min. 2µl of chase mix (4x cold dNTP's each at 0.25mM) were added to each well and the reaction continued at 37°C for a further 30min. The reaction was stopped by addition of 2µl formyl dyes mixture (0.1% bromophenol blue (w/v), 0.1% xylene cyanol (w/v), in deionized formamide). The plate was heated at 90°C for 2min and 2-4µl of each reaction resolved on adjacent lanes of a denaturing polyacrylamide sequencing gel (Section 12.3).

Sequencing solutions:

	dA-0	dT-0	dC-0	dG-0
5mM dTTP	20µl	1µl	20µl	20µl
5mMdCTP	20µl	20µl	1µl	20µl
5mM 7-deaza dGTP	20µl	20µl	20µl	1µl
10x seq buffer	50µl	50µl	50µl	50µl
H_2O	540µl	370µl	370µl	370µl

seq buffer = 10mM Tris-HCl (pH 8.5), 10mM $MgCl_2$

Nucleotide concentrations in sequencing mixes:

	dN-0 mixes	ddNTP	H_2O
Tseq	500µl	500µl (600µM ddTTP)	0µl
Cseq	500µl	105µl (140µM ddCTP)	395µl
Gseq	500µl	155µl (200µM ddGTP)	345µl
Aseq	500µl	250µl (140µM ddATP)	370µl

Sequenase™ reactions were carried out using a Sequenase™ Version 2.0 sequencing kit, according to manufacturers instructions.

2B.22. Oligonucleotide-Mediated Site Directed Mutagenesis.

Mutagenesis of the Vmw 175 DNA binding domain was performed with the Amersham oligonucleotide-directed *in vitro* mutagenesis system Version 2, following

the manufacturers instructions. The Amersham system is based on the protocol of Eckstein *et al* ., (Sayers *et al.*, 1988). The region of Vmw175 to be mutated was cloned into M13mp18 and single-stranded template DNA was prepared. Before carrying out the mutagenesis reaction, the mutagenic oligonucleotide was used in sequencing reactions with the template DNA. Provided the mutagenic oligonucleotide generated the correct sequence, the following protocol was carried out.

2B.22.1 Hybridization of oligonucleotides to template DNA and primer extension.

The mutagenic oligonucleotide was phosphorylated as described in Section 2B.24.1. 5μg of template DNA and approximately 4pmoles oligonucleotide were annealled in a final volume of 17μl, by heating to 70°C for 3min and then incubating at 37°C for a further 30min. The mixture was then placed on ice, a nucleotide mix containing a phosphothiorate derivative of the deoxyribonucleotide dCTP (dCTP αS) was added and the primer extension and ligation carried out using 3 units Klenow LFP and 3 units T4 DNA ligase at 15°C for 16 h.

2B.22.2 Filtration of heteroduplex replicative form molecules.

Single-stranded and heteroduplex RF molecules were separated by means of a single filtration through nitrocellulose, which binds and retains ssDNA molecules.

2B.22.3 Nicking and exonuclease reactions.

The resultant heteroduplex RF DNA molecules consist of the original non-mutant DNA strand annealled to a mutant DNA strand which contains phosphothiorate derivatives of the deoxyribonucleotide dCTP. The non-mutant DNA strand was then selectively nicked with the enzyme NciI, which cannot cleave the phosphothiorate substituted mutant DNA strand, and digestion with Exonuclease III.

The purified heteroduplex DNA was incubated with 4 units NciI at 37°C for 90 min followed by incubation with 50 units Exonuclease III at 37°C for 15 min. Enzymes were then inactivated by heating at 70°C for 15min.

2B.22.4 Repolymerization and ligation of the digested DNA.

The digested strand was filled-in and religated with 1.25 units DNA polymerase I and 1.5 units T4 DNA ligase at 15°C for 3 h. The final DNA molecules were purified by precipitation with 1 vol 4M ammonium acetate (pH 5.4) and 4 vols ethanol, washed with 70% ethanol, lyophilized and resuspended in 100μl TE.

2B.22.5 Transformation of competent *E. coli* strains JM101 or TG1 and screening for mutants.

20µl of the above DNA mix was used to transform TG1 or JM101 as described in Section 2.18. Plaques obtained were picked and small scale single-stranded template DNA preparations made. Plaques containing mutant phage were identified by dideoxy DNA sequencing, and large scale replicative form DNA preparations made from mutant clones, for further cloning reactions.

2B.23. Large Scale Replicative Form M13 DNA Preparations.

5µl bacteriophage stock was inoculated into 350ml L-broth containing 1ml fresh overnight JM101 culture, and grown shaking at 37°C overnight. Bacteria were then pelleted and replicative form DNA prepared by the method described in Section 2B.9.

2B.24. Radioactive Labelling of DNA.

2B.24.1. 5'-end labelling.

Purified oligonucleotides which did not contain 5' phosphate groups were 5'-end labelled using T4 polynucleotide kinase and [γ-^{32}P] dATP. Phosphorylation was carried out in a final volume of 10µl 1x kinase buffer containing 10 pmols oligonucleotide, 10 units T4 polynucleotide kinase and excess 5'-[γ-^{32}P] dATP (usually 10-30µCi) at 37°C for 45 min. The labelled oligonucleotides were purified by gel electrophoresis.

Phosphorylation of oligonucleotides for mutagenesis reactions was carried out in a similar fashion but using 1mM cold dATP and 50pmols oligonucleotide in the kination reaction. The reaction was stopped at 70°C.

2B.24.2. 3'-end labelling.

3'-ends of DNA fragments were labelled by filling in 5'-overhangs with Klenow LFP. Digests of 0.5-1µg of DNA were filled in after stopping the restriction digest at 65°C. LFP reactions were carried out in 1x REact 2 buffers, and used 2u enzyme incubated at 37°C for 20min with appropriate radioactive (5'-[α-^{32}P]) and non radioactive dNTPs followed by a 10 min chase with all four non radioactive dNTPs. LFP was inactivated by incubation at 65-70°C prior to second enzyme digestions or DNA purifications, as described in Section 2.12.2.

2B.24.3. Internal labelling by random priming.

Random prime labelling of DNA fragments was carried out as described by Feinberg and Vogelstein (1983). 50ng DNA was boiled for 5min in a final volume of 20µl then incubated for 3-6h at RT in a 30µl mix containing 10µl random prime buffer (concentrated solution containing dATP, dGTP, dTTP, Tris-HCl pH7.5, $MgCl_2$, and β-mercaptoethanol), 0.4µl 50mg/ml BSA, 1µl 1mM dGTP, 50µCi [α-^{32}P] dCTP and 2u Klenow LFP. The reaction was stopped by the addition of 80µl NT stop and labelled DNA purified on a Sephadex G-50 column (Section 12.4).

2B.25. Southern Transfer of DNA to Nitrocellulose.

Restricted DNAs for Southern transfer (Southern, 1975) were run on agarose gels in 1x TBE. Following electrophoresis the gel was shaken first in Gel Soak I for 45min then Gel Soak II for 45min. The DNA was blotted overnight onto presoaked Hybond-N^+ paper using 10x SSPE, a wick of Whatman 3mm paper and a weighted capillary stack of absorbent paper towelling. DNA was fixed onto the Hybond-N^+ paper by soaking in 0.4M NaOH for 1-60min prior to hybridization.

2B.26. DNA-DNA Hybridization.

Nitrocellulose filters were prehybridized in fresh hybridization mix, containing the appropriate concentration of formamide, for 2-4h at 65°C. 10^7-10^8 cpm of probe DNA was denatured by the addition of 0.04 vol 5M NaOH for 10min then neutralization with 5M HCl before adding to the filter bag and hybridizing overnight under the same conditions as for prehybridization. The filter was then washed twice at RT in 2x SSPE, 0.1% SDS for 20min each before autoradiography at -70°C with an intensifying screen.

2B.27. Preparation of Bacterial Extracts for Gel Retardation Analysis.

BL21 (DE3) pLysS competent cells were transformed with appropriate small scale plasmid DNA preparations or purified plasmid DNA as described in Section 2.17. The cells were plated out on agar plates containing chloramphenicol and ampicillin and incubated overnight at 37°C. The next day 4-6 colonies were inoculated into 50ml of 2YT + appropriate antibiotics, and the cultures grown shaking at 37°C until OD_{540} = 0.5 (approx. 4h). The cultures were then cooled to 26°C and IPTG added to a final concentration of 200µM. Expression of the cloned gene was induced by shaking incubation of the cultures at 26°C for a further 2h. The bacteria were pelleted by

centrifugation at 3000 rpm for 10 min (Beckman GPR centrifuge) and resuspended in 1/100 vol of TE, 1mM PMSF, 0.1mM DTT. Resuspended pellets were frozen overnight at -20°C.

2B.27.1. Preparation of Extracts Treated with Polyethylenimine.

The following day, the bacteria were thawed on ice and lysed by the addition of lysozyme to a final concentration of 100µg/ml and sodium deoxycholate to a final concentration of 0.5%, followed by incubation on ice for 10min. Following lysis, DNAase was added to a final concentration of 20µg/ml, RNAase added to a final concentration of 10µg/ml and $MgCl_2$ added to a final concentration of 5mM. After a 30 min incubation on ice, NaCl was added to a final concentration of 1M and digested nucleotides pelleted by addition of a final concentration of 0.2% polyethylenimine and centrifugation at 15000 rpm for 20 min (Sorvall SS34 rotor). The supernatant was removed to a fresh tube and saturated $(NH_4)_2SO_4$ added to a final concentration of 30%. After a 1h incubation on ice, precipitated proteins were pelleted by centrifugation at 10000 rpm for 10 min (Sorvall SS34 rotor). The $(NH_4)_2SO_4$ supernatant was removed and pelleted proteins resuspended in 150µl resuspension buffer. Insoluble proteins were removed by centrifugation at 10000 rpm for 15 min, and polypeptide content of final extracts determined by SDS-polyacrylamide gel electrophoresis followed by Western blot analysis.

2B.27.2. Preparation of Untreated Extracts.

The following day, the bacteria were thawed on ice and lysed by the addition of lysozyme to a final concentration of 100µg/ml and sodium deoxycholate to a final concentration of 0.5%, followed by incubation on ice for 10min. Using ice cold procedures throughout, NaCl was added to a final concentration of 2M and cell debris removed by centrifugation at 40000 rpm for 2h (Beckman TLA 100.2 rotor). Supernatant was removed to a fresh tube and $(NH_4)_2SO_4$ added to a final concentration of 30% saturation. After a 1h incubation on ice precipitated proteins were pelleted by centrifugation at 10000 rpm for 10 min (Sorvall SS34 rotor). All $(NH_4)_2SO_4$ supernatant was carefully removed and the pellets resuspended in 150µl resuspension buffer. Any insoluble proteins were removed by centrifugation at 10000 rpm for 5 min (Sorvall SS34 rotor), and the polypeptide content of the final soluble extract determined by SDS polyacrylamide gel electrophoresis and Western blot analysis.

2B.28. Preparation of Bacterial Extracts for FPLC Purification and Glutaraldehyde Cross-Linking Assays.

BL21 (DE3) pLysS competent cells were transformed with appropriate small scale plasmid DNA preparations or purified plasmid DNA as described in Section 2.17. The cells were plated onto agar plates containing chloramphenicol and ampicillin and incubated overnight at 37°C. The following day 10 colonies were inoculated into 1l 2YT + ampicillin and cultures grown shaking at 37°C until OD_{540} = 0.5. Expression of cloned genes was induced by cooling cultures to 26°C, adding IPTG to a final concentration of 100μM and continuing the incubation at 26°C for a further 2h. Bacteria were then pelleted by centrifugation at 5000 rpm for 5min (Sorvall GS3 rotor) and resuspended in 1/100 vol resuspension buffer. Resuspended pellets were frozen overnight at -20°C.

The following day, bacteria were thawed on ice and extracts made as follows, using ice cold procedures throughout. Bacteria were lysed by 3x 10sec sonications at 80-90W with a 5mm soniprobe, DNAase added to a final concentration of 20μg/ml, RNAase added to a final concentration of 10μg/ml and $MgCl_2$ added to a final concentration of 5mM. After a 30 min incubation on ice, NaCl was added to a final concentration of 1M and digested nucleotides pelleted by addition of a final concentration of 0.2% polyethylenimine and centrifugation at 15000 rpm for 20 min (Sorvall SS34 rotor). It was found that the addition of polyethylenimine resulted in the loss of some more insoluble polypeptides. In these cases, polyethylenimine was ommitted from the extraction prodcedure. The supernatant was removed to a fresh tube and saturated $(NH_4)_2SO_4$ added to a final concentration of 30%. After a 1h incubation on ice, precipitated proteins were pelleted by centrifugation at 10000 rpm for 10 min (Sorvall SS34 rotor). The $(NH_4)_2SO_4$ supernatant was removed and pelleted proteins resuspended in 5ml resuspension buffer per 1l starting culture. Insoluble proteins were removed by centrifugation at 10000 rpm for 15 min, and polypeptide content of final extracts determined by SDS-polyacrylamide gel electrophoresis followed by staining with Coomassie Brilliant Blue. Final soluble extracts were dialysed at 4°C against MonoS buffer A to remove remaining $(NH_4)_2SO_4$ before storage at -70°C.

2B.29. FPLC Chromatography.

Bacterial extracts were dialysed against 50mM HEPES pH7.5, 0.1M NaCl, 100mM PMSF, 0.1mM DTT and 0.01% CHAPS to remove ammonium sulphate, prior to ion exchange chromatography. Insoluble proteins were removed by centrifugation at 10000 rpm for 10 minutes at 4°C. Extracts were then loaded onto a 1ml Pharmacia MonoS column equilibrated in the above buffer. Chromatography was performed on an

automated Pharmacia FPLC system and polypeptides eluted with a 20ml 0.1-1M NaCl gradient.

Polypeptides that were eluted in flow-through fractions were further purified by the treatment of fractions with 0.2% polyethylenimine and removal of insoluble material by centrifugation at 15000 rpm for 20 min at 4°C (Sorvall SS34 rotor). The supernatant was removed to a fresh tube and saturated $(NH_4)_2SO_4$ added to a final concentration of 30%. After a 1h incubation on ice, precipitated proteins were pelleted by centrifugation at 10000 rpm for 10 min (Sorvall SS34 rotor).

Polypeptides analysed further by gel filtration chromatography were concentrated by centrifugation in Centricon 10 concentrators at 5000rpm and 4°C, (Sorvall SS34 rotor).

Concentrated polypeptides were loaded onto Pharmacia Superose 12 or Superdex 75 columns previously equilibrated in 50mM HEPES pH 7.5, 0.5M NaCl, 0.01% CHAPS, 1mM PMSF and 0.1mM DTT. Chromatography was performed on an automated Pharmacia FPLC system and polypeptides eluted in the above buffer.

2B.30. Glutaraldehyde Cross-Linking Assay.

Bacterial extracts containing DNA binding domain polypeptides were incubated in a reaction mixture containing 50mM HEPES pH7.5, 0.1M NaCl, 1mM PMSF and 0.01% or 0.02% glutaraldehyde (freshly diluted from a 25% glutaraldehyde stock in 50mM HEPES pH7.5, 0.1M NaCl). Cross-linking assays were carried out at various temperatures for 20 minutes and the reaction stopped by the addition of SDS boiling mix.

2B.31. Computing.

Computer analyses of DNA and protein sequences were performed on the MicroVAX II computer in this department, using the University of Wisconsin Genetics Computer Group software package (Devereux *et al.*, 1984).

Chapter 3. Results.

3A: Structural and Sequence Analysis of the DNA Binding Domain of Vmw175.

Expression of polypeptides encompassing portions of the Vmw175 protein in bacteria has confirmed the location of the DNA binding domain of Vmw175 (see Section 1C.6). The domain lies within the limits of a region spanning codons 275-490 (Wu and Wilcox, 1989; Pizer *et al.*, 1991; Everett *et al.*, 1991a, b). Similar studies of the homologous DNA binding and transactivating immediate early protein of VZV, the 140K protein, has shown that its DNA binding domain lies within the region spanning codons 417-646 (Tyler and Everett, 1993; Wu and Wilcox, 1991). In order to carry out a detailed mutational analysis of essential regions within the DNA binding domain of Vmw175, it is useful to review the homology of this region to the DNA binding region of the VZV 140K protein, and also to the known sequences of other α-herpesvirus transactivating proteins.

This section presents sequence alignments of these sequences and secondary structure predictions for homologous regions of the proteins. Previous mutational analyses of the DNA binding domain of Vmw175 are reviewed, and a portion of the DNA binding domain highlighted by these analyses is compared to several conserved motifs of other DNA binding protein families.

Computer analyses of DNA and protein sequences were performed on the MicroVAX II computer in this department, using the University of Wisconsin Genetics Computer Group software package (Devereux *et al.*, 1984).

3A.1. Regions of Homology Shared Between Known and Predicted DNA Binding Domains of the α-Herpesvirus Regulatory Proteins.

In order to determine the sequence homology between the DNA binding domains of Vmw175 and the varicella zoster virus 140K protein, and the transactivating immediate-early proteins of equine herpesvirus type 1 and pseudorabies virus, the polypeptide sequences were aligned using the GCG Pileup programme. The

Figure 3A.1. Alignment of the DNA Binding Domains of the α-Herpesvirus Regulatory Proteins.

```
 VZV 140K              417 ..........  ..........  ......RSIS
 PRV IE180             414 RPGAPLARAG  PPPSPPAPAA  APRPSASSAS
 EHV IE1               356 RPRPQTQTQA  PAEEAPQTAV  WDLLDMNSSQ
 HSV-1 Vmw175          252 ..........  ..........  ..RAAARTPA
Conserved amino acids      ----------  ----------  -------S--
Identical amino acids      ----------  ----------  ----------

 VZV 140K              421  GPDPRIRKTK RLAGEPGRQR QKSFSLPRSR
 PRV IE180             444  ATSSSAAASP APAPEPARPP RRKRRSTNNH
 EHV IE1               386  ATGAAAAAAS APAAASCAPG VYQR......
 HSV-1 Vmw175          260  ASAGRIERRR ARAAVAGRDA ..TGRFTAGQ
Conserved amino acids      A---------  A-A----R--  ----------
Identical amino acids      ----------  --A-------  ----------
                                                  region 1---->|<--
 VZV 140K              451  TPII....PP VSGPLMMP..  ....DGSPWP
 PRV IE180             474  LSLMADGPPP TDGPLLTP..  ....LGEPWP
 EHV IE1               410  ..........  ..EPLLTP..  ....SGDPWP
 HSV-1 Vmw175          288  PRRVELDADA TSGAFYARYR  DGYVSGEPWP
Conserved amino acids      ----------  --GPLL-P--  -----GEPWP
Identical amino acids      ----------  ----------  -----G-PWP
                           ----region 2
 VZV 140K              471  GSAPLPSNRV RFGPSGETRE  GHWEDEAVRA
 PRV IE180             497  GSDPPADGRV RYGGAGDSRE  GLWDEDDVRQ
 EHV IE1               422  GSDPPPMGRV RYGGTGDSRD  GLWDDPEIVL
 HSV-1 Vmw175          318  GAGPPPPGRV LYGGLGDSRP  GLWGAPEAEE
Conserved amino acids      GS-PPP-GRV  RYGG-GDSRE  GLWDD-EV--
Identical amino acids      G--P----RV  --G--G--R-  G-W-------

 VZV 140K              501  ARARYEASTE PVPLYVPELG  DPARQYRALI
 PRV IE180             527  AAARYRAAAG PVPVFIPEMG  DSRKQHEALV
 EHV IE1               452  AASRYAEAQA PVPVFVPEMG  DSTKQYNALV
 HSV-1 Vmw175          348  ARRRFEASGA PAAVWAPELG  DAAQQYALIT
Conserved amino acids      A--RY-A---  PVPVFVPE-G  D---QY-ALV
Identical amino acids      A--R------  P-----PE-G  D---Q-----

 VZV 140K              531  NLIYC.PDRD PIAWLQNPKL  TGVNSALNQF
 PRV IE180             557  RLIYSGAAGE AMSWLQNPRM  QAPDQRFNQF
 EHV IE1               482  RMVF..ESRE AMSWLQNSKL  SGQDQNLAQF
 HSV-1 Vmw175          378  RLLYT.PDAE AMGWLQNPRV  VPGDVALDQA
Conserved amino acids      RLIY-----E  AM-WLQNP-L  ---D--L-QF
Identical amino acids      ----------  ---WLQN---  --------Q-

 VZV 140K              560  YQKLLPPGRA .GTAVTGSVA  SPVPHVGEAM
 PRV IE180             597  CQRRVHAPHG HGSFITGSVT  PPLPHIGDAM
 EHV IE1               510  CQKFIHAPRG HGSFITGSVA  NPLPHIGDAM
 HSV-1 Vmw175          407  CFRISGAARN SSSFITGSVA  RAVPHLGYAM
Conserved amino acids      CQ----A-R-  -GSFITGSVA  -P-PHIGDAM
Identical amino acids      ----------  -----TGSV-  ---PH-G-AM

 VZV 140K              589  ATGEALWALP HAAAAVAMSR  RYDRAQKHFI
 PRV IE180             627  AAQDPLWALP HAVSAVAMSR  RYDRTQKTFI
 EHV IE1               540  AAGNALWALP HAAASVAMSR  RYDRTQKSFI
 HSV-1 Vmw175          437  AAGRFGWGLA HAAAAVAMSR  RYDRAQKGFL
Conserved amino acids      AAG--LWALP  HAAAAVAMSR  RYDR-QK-FI
Identical amino acids      A-----W-L-  HA---VAMSR  RYDR-QK-F-
```

```
                                        region 2---->|<------
 VZV 140K          619  LQSLRRAFAS MAYP..... ...EATGSSP
 PRV IE180         657  LQSLRRAYAD MAYPGRAADP RAGEATVEAL
 EHV IE1           570  LQSLRRAYAD MAYPRDEAGR PDSLAAVAGC
 HSV-1 Vmw175      467  LTSLRRAYAP LLARENAALT GAAGSPGAGA
Conserved amino acids  LQSLRRAYA- MAYP---A-- ----A-----
Identical amino acids  L-SLRRA-A- ---------- ----------
                       ---region 3
 VZV 140K          640  AARISRG... ..........  ......  646
 PRV IE180         687  CARVRAAFAA AQPGRVP...  ......  703
 EHV IE1           600  PAQAAAAAAS QQQPEAPAPS  VRVREA  625
 HSV-1 Vmw175      497  DDEGVAAVAA AAPGERAVPA  GYGAAG  522
Conserved amino acids  -A---AA-A- ----------  ------
Identical amino acids  ---------- ----------  ------
```

Legend to Figure 3A.1. The DNA binding domain of the VZV 140K protein, the IE proteins of EHV-1 and PRV and the I9X DNA binding domain polypeptide of HSV-1 Vmw175 were aligned using the GCG Pileup program. The sequence of the IE protein of PRV was obtained from Cheung (1989). The positions of amino acids conserved in three out of four proteins and amino acids identical between all four proteins are indicated.

largest DNA binding domain polypeptides tested for DNA binding activity encompass codons 252-523 of Vmw175 (the I9X polypeptide, Pizer *et al.*, 1991) and codons 417-646 of 140K (Tyler and Everett, 1993). The sequences encompassing these amino acids were aligned with the amino acid sequence of the EHV IE1 protein (Grundy *et al.*, 1989), and the PRV IE180 protein (Cheung, 1989).

The GCG Pretty programme was used to determine the amino acid residues conserved between at least three proteins, and the amino acid residues identical between all four proteins. The alignments of the four proteins within the region of the Vmw175 and 140K DNA binding domains is shown in Figure 3A.1, together with the sequence of conserved and identical amino acids.

3A.2. Structural Features of the Vmw175 DNA Binding Domain and Conservation of These Features in the Homologous α-Herpesvirus Transactivating Proteins.

Four types of secondary structural motifs have been identified within globular protein molecules - the α-helix, β-sheet, reverse turn and random coil. It is possible to predict, to a limited degree of certainty, the probability of specific areas of primary protein sequence forming such secondary structural motifs. The GCG programme Peptidestructure uses the predictive methods of Chou and Fashman (1974) and Garnier, Osguthorpe and Robson (1978) to predict the secondary structure of input primary amino acid sequence.

Figure 3A.2 shows the secondary structure predictions made for amino acid sequences from the PRV IE180, EHV-1 IE1 and VZV 140K proteins which are homologous to the portion spanning codons 442-485 of Vmw175. This region was chosen for further analysis as it is highly conserved between regulatory proteins of the α-herpesviruses, and previous mutational analyses have illustrated the importance of this region to the DNA binding activity of Vmw175 (see Section 3A.3). It can be seen that the first 15 residues of the motif in all four homologous proteins are predicted to exist in a helical conformation, both by the Chou and Fashman and the G-O-R predictive methods. This prediction of α-helical structure is probably due to the highly hydrophobic nature of this region of the proteins.

The secondary structure predictions for the portion of the motifs carboxy-terminal to this predicted helix are less clear cut. For example, within Vmw175, the Chou-Fashman method results in the prediction of a random coil and reverse turn followed by a 10 amino acid stretch with a low probability of α-helical structure and a 10 amino acid stretch with a high probability of α-helical structure. However, the G-O-

VZV 140K	594	LWALPHAAAA	VAMSRRYDRA	QKHFILQSLR	RAFASMAYPE	ATGS
CF-Pred.		.. .. HHHHHHHH	HHHHHH..hhh	hhhhhhhhhh	hhhhhhhh....	.. tt..
GOR-Pred.		HHHHHHHHHH	HHHHHHHHHH	HBBBBB.....T	TTT..... ..T	TTT..
PRV IE180	612	LWALPHAVSA	VAMSRRYDRT	QKTFILQSLR	RAYADMAYPG	RAAD
CF-Pred.		.. .. hhhhhhhh	hhhhhh .. .. tt	..... BBBBBhhh	hhhhhhh..TT	.....
GOR-Pred.		HHHHHHHHHH	HHHHHTTTTT	TTBBBBB....T	TTBBBBBB.. ..	.....
EHV IE1	545	LWALPHAAAS	VAMSRRYDRT	QKSFILQSLR	RAYADMAYPR	DEAG
CF-Pred.		.. .. hhhhhhhh	hhhhhh.. .. tt	hhhhhhhhhh	hhhhhhh..TT	t.....
GOR-Pred.		HHHHHHHHHH	HHHHHTTTTT	TTBBBBB....T	TTTBBBB.. .. T	TTTT
HSV-1 Vmw175	442	GWGLAHAAAA	VAMSRRYDRA	QKGFLLTSLR	RAYAPLLARE	NAAL
CF-Pred.		.. .. .. HHHHHHH	HHHHHH.....	.. t thhhhhhh	hh..HHHHHHH	HHHH
GOR-Pred.		.. HHHHHHHHH	HHHHHHHHTT	TTBBBBB....T	TT...HHHHHH	HHHH

Figure 3A.2. Secondary Structure predictions for the DNA Binding Domains of the α–Herpesvirus Regulatory Proteins.

The regions of the four α-herpesvirus regulatory proteins homologous to amino acids 442-485 of HSV-1 Vmw175 are shown, together with secondary structure predictions for these regions of the protein.

CF-Pred. = predictions made using the Chou-Fashman rules.

GOR-Pred. = predictions made using the Garnier-Osguthorpe-Robson predictive method.

H = high probability of helical structure. h = low probability of helical structure.

B = high probability of β-sheet structure. b = low probability of β-sheet structure.

T = high probability of reverse turn structure. t = low probability of reverse turn structure.

.. = random coil structure.

Secondary structure predictions were made using the GCG Peptidestructure program

R method results in a different prediction of a 15 amino acid stretch containing a turn, β-sheet and random coil followed by the last 10 amino acid helical structure. Secondary structure predictions for the other proteins also demonstrate that the proposed helix-turn-helix secondary structural motif predicted for Vmw175 by the Chou-Fashman method is not conserved in the other α-herpesvirus regulatory proteins.

3A.3. Identification of a Possible Motif Involved in the DNA Binding Function of Vmw175.

Two linker-scanning mutational analyses of Vmw175 have previously been carried out, in order to define functionally significant regions of the molecule (Paterson and Everett, 1988a, b; Paterson, 1989; Shepard *et al.*, 1989). As described in Section 1C.3 and illustrated in Figure 1C.3, the results of these studies were in general agreement. Analysis of one series of insertional mutations created by Shepard *et al.* (1989) suggested that two portions of region 2 of the protein contributed to the DNA binding, transactivation and autoregulation properties of Vmw175. The authors proposed that the region of Vmw175 spanning codons 445-487 contained a motif with a predicted helix-turn-helix secondary structure which is essential for the interaction with DNA. A high degree of homology is shown within most of this motif in the regulatory proteins of four α-herpesviruses (see Section 3A.1). In addition, a widely studied temperature sensitive mutant of Vmw175, *ts*K (described in Section 1C.3), which fails to support viral replication under restrictive conditions, contains a point mutation within this region. These pieces of evidence suggest that the portion of Vmw175 encompassing amino acids 445-487 is of functional significance, and therefore it was decided that this region was worthy of a more detailed mutational analysis.

3A.4. Analysis of Sequence Similarities Between the Vmw175 "Helix-Turn-Helix" Motif and Other Known DNA Binding Protein Consensus Motifs.

3A.4.1. Homology to the Prokaryotic Helix-Turn-Helix Motif.

The prokaryotic helix-turn-helix DNA binding motif was first discovered in the bacteriophage λCro, λcI and the *Escherichia coli* CAP proteins (for review, see Pabo and Sauer, 1984; discussed in Section 1D.3.1).

As the region spanning the Vmw175 codons 445-487 was proposed to contain a secondary structural helix-turn-helix motif (see Sections 3A.2 and 3A.3), an analysis of the similarity of this region to the consensus sequence for the prokaryotic helix-turn-helix motifs was carried out. In order to avoid confusion between the two motifs, the Vmw175 motif is referred to as the "helix-turn-helix" region or motif.

Dodd and Egan (1987, 1990) have reported a method for the location and statistical evaluation of potential DNA binding motifs exhibiting sequence homology with the prokaryotic helix turn helix motif, as illustrated by the λCro protein. This approach utilizes a master set of proteins exhibiting homology to the λCro motif to create a weight matrix indicating the probability of a particular amino acid occurring at each position of the motif. Favoured amino acids have positive weights and disfavoured amino acids have negative weights. Addition of the weights of amino acids under analysis yields a score representative of the probability of that particular sequence being a member of the master set versus it being a random sequence. For analysis of a whole protein, the weight matrix is applied to all segments of the protein to find the highest scoring segment.

The statistical significance of the method is increased with the number of protein sequences used to create the weight matrix. In order to increase the number of protein sequences contributing to the weight matrix, a master set of known prokaryotic helix turn helix motifs whose tertiary structure has been solved by X-ray crystallographic analysis has been used to scan protein databases. Protein sequences showing homology to the master set and yielding a statistically significant score upon application of the weight matrix have been incorporated, and additional information has been used to update the weight matrix.

Dr A. Bailey in this institute has developed a computer programme for the statistical analysis of protein sequences by the Dodd and Egan method. The weight matrix used for this analysis was derived from 198 proteins, all of which have been shown to bind to DNA. The weight matrix used in this programme is larger than that developed by Dodd and Egan (1990), and so should yield results of higher statistical significance. Table 3A.1 shows the segments of the immediate early proteins of HSV-1 and EHV-1 which resulted in a positive score upon analysis with this programme. Interestingly, the segment yielding the highest score in each protein aligns with the "helix-turn-helix" motif of Vmw175 (see Figure 3A.1).

Protein	Region (a)	Score (b)
HSV-1 Vmw175	445	394
HSV-1 Vmw175	81	337
HSV-1 Vmw175	827	177
HSV-1 Vmw175	256	115
EHV-1 IE1	548	306
EHV-1 IE1	372	179
EHV-1 IE1	477	29

Table 3A.1. Analysis of the Regions of the Regulatory proteins of α-herpesviruses Showing Homology to the Prokaryotic Helix-Turn-Helix Motif.
The regions of homology were determined by computer analysis, using a programme based on the statistical analysis of Dodd and Egan (1987, 1990).
(a) The first codon of the region homologous to the 20 amino acid prokaryotic helix-turn-helix motif. Codons are numbered starting from 1 at the ATG start codon.
(b) Statistical scores were determined as described by Dodd and Egan (1987, 1990). The average score for a prokaryotic helix-turn-helix motif was found to be around 2000 (Dodd and Egan, 1987).

3A.4.2. Homology to the Homeobox Motif.

The homeobox was identified originally in many Drosophila homeotic proteins. It consists of a highly conserved 60 amino acid sequence (reviewed in Gehring, 1987) which contains a tertiary helix turn helix DNA recognition motif (Kissinger *et al.*, 1990; discussed in Section 1D.3.2).

In order to determine whether the Vmw175 DNA binding domain contains a homeodomain motif, a consensus sequence for the homeodomain motif was derived. As the members of the homeodomain family fall into subgroups in which the proteins show higher homology to members of their particular subgroup than to other members of the homeobox family, two different approaches were taken to create consensus amino acid sequences representative of the homeodomain. Analysis of 15 homeodomain sequences as detailed in Frain *et al.*, (1989) with the GCG Pileup and Profilemake programmes led to the creation of a matrix indicating the probability of a particular amino acid occurring at each position of the 30 amino acid segment encompassing the three proposed helices of the motif. In addition, the same programmes were used to create a profile from the homeobox motifs of five *Drosophila* homeodomain proteins (Eve, En, H2.0, Ubx and Antp, described in Frain *et al.*, 1989). The GCG Profilegap programme was used to align these profiles with the

```
HSV-1 Vmw175   330  GGLGDSRPGL WGAPEAEEAR RRFEASG APA AVWAPELGDA AQQYALITRL LYTPDAEAMG
Homeobox Consensus  ..... ..... .....EQMDALE KEFAENQYPT SHRREELADE LNLTEAQI KV

HSV-1 Vmw175   390  WLQNPR VVPG DVALDQACFR ISGAARNSSS FITGSVARAV PHLGYAMAAG RFGWGLAHAA
Homeobox Consensus  WFQNRRM..... ..... ..... ..... ..... .....

HSV-1 Vmw175   450  AAVAMSRRYD RAQKGFLLTS LRRAYAPLLA RENAALTGAA GSPGAGADDE   500
Homeobox Consensus  ..... ..... ..... ..... .....
```

Figure 3A.3. Alignment Between the DNA Binding Domain of Vmw175 and a Consensus Sequence Derived from a Sample of Eukaryotic Homeobox Proteins.

The sequence of Vmw175 between amino acids 330-500 is shown, together with the aligned homeobox consensus sequence. The homeobox consensus was derived from a weight matrix compiled from the sequences of 15 homeobox proteins as described in Frain et al. (1989). The weight matrix and consensus sequence were compiled using the GCG Profile and Profilemake programmes. Alignments were made between the consensus sequence and the DNA binding domain of Vmw175 with the GCG Profilegap programme.

HSV-1 Vmw175 330 GGLGDSRPGL WGAPEAEEAR RRFEASGAPA AVWAPELGDA AQQYALITRL LYTPDAEAMG

Homeobox Consensus .....RTA FTRFQMMELE KEFHENRYLT RRRRIELAAA ..... LNLTEAQIKI

HSV-1 Vmw175 390 WLQNPRVVPG DVALDQACFR ISGAARNSSS FITGSVARAV PHLGYAMAAG RFGWGLAHAA

Homeobox Consensus WFQNRRM..... ....

HSV-1 Vmw175 450 AAVAMSRRYD RAQKGFLLTS LRRAYAPLLA RENAALTGAA GSPGAGADDE 500

Homeobox Consensus ....

Figure 3A.4. Alignment between the DNA Binding Domain of Vmw175 and a Consensus Sequence Derived from a Sample of Drosophila Homeobox Proteins.

The sequence of Vmw175 between amino acids 330-500 is shown, together with the aligned homeobox consensus sequence. The homeobox consensus was derived from a weight matrix compiled from the sequences of Drosophila homeobox proteins as described in Frampton et al. (1989). The weight matrix and consensus sequence were compiled using the GCG Profile and Profilemake programmes. Alignments were made between the consensus sequence and the DNA binding domain of Vmw175 with the GCG Profilegap programme.

amino acid sequence of the DNA binding domain of Vmw175 (Figures 3A.3 and 3A.4).

Both homeodomain profiles exhibited highest homology with a sequence spanning Vmw175 amino acids 353-396, with the *Drosophila* profile exhibiting homology to an extended sequence encompassing amino acids 337-396. This region lies outwith the "helix-turn-helix" motif but still within the limits of the DNA binding domain of Vmw175. Interestingly, a region within the homeobox consensus sequence which is important for DNA interaction (Kissinger *et al.*, 1990) showed high homology to a region spanning residues 390-395 of Vmw175. This protein sequence is also highly conserved within other α-herpesvirus regulatory proteins, suggesting a possible functional role for this region.

3A.4.3. Homology to the Helix Loop Helix Motif.

The eukaryotic helix loop helix motif consists of two short amphipathic helices separated by a loop sequence of between 6 and 21 amino acids in length and varying amino acid content. The helix loop helix motif has been shown to mediate dimerization of the protein, whilst a cluster of basic amino acids immediately preceding the first helix has been shown to be involved in the interaction with the DNA helix (Davis *et al.*, 1990; discussed in Section 1D.3.3).

A consensus sequence for the two helices of this motif has been derived by analysis of ten human and Drosophila proteins (Murre *et al.*, 1989) with the GCG Pileup and Pretty programmes. These two consensus sequences were aligned with the amino acid sequences for the proposed DNA binding domains of the regulatory proteins of HSV-1, VZV, PRV and EHV-1. The alignments were carried out using the GCG Bestfit programme, and are shown in Figures 3A.5 and 3A.6. The consensus sequence for the first helix exhibited low level homology to a region spanning Vmw175 codons 350-380, outwith the "helix-turn-helix" region. The consensus sequence for the second helix exhibited very limited homology with a region spanning Vmw175 codons 444-468, which is situated within the "helix-turn-helix" region.

3A.5. Representation of the Proposed "Helix-Turn-Helix" Motif on Helical Wheel Models.

As discussed in Section 3A.2, the region spanning codons 445-487 of Vmw175 has been predicted to contain two different secondary structure motifs upon application of secondary structure prediction programmes using both G-O-R and Chou and

HSV-1 Vmw175	330	GGLGDSRPGL WGAPEAEEAR RRFEASGAPA AVWAPELGDA AQQYALITRL LYTPDAEAMG	
Helix Consensus		.... ERRXANNARE RXRXRDINEA .....FXELGRM C.....	
HSV-1 Vmw175	390	WLQNPRVVPG DVALDQACFR ISGAARNSSS FITGSVARAV PHLGYAMAAG RFGWGLAHAA	
Helix Consensus		....	
HSV-1 Vmw175	450	AAVAMSRRYD RAQKGFLLTS LRRAYAPLLA RENAALTGAA GSPGAGADDE	500
Helix Consensus		....	

Figure 3A.5. Alignment Between the DNA Binding Domain of Vmw175 and a Consensus Sequence for the First Helix of the Helix Loop Helix Motif.

The sequence of Vmw175 between amino acids 330-500 is shown, together with the aligned consensus sequence for the first helix of the helix loop helix motif. The helix consensus sequence was compiled from the sequences of helix loop helix proteins published in Murre et al. (1989), using the GCG Pileup and Pretty programmes. Alignments were made between the Vmw175 sequence and the helix consensus sequence with the GCG Bestfit programme.

HSV-1 Vmw175	330	GGLGDSRPGL WGAPEAEEAR RRFEASGAPA AVWAPELGDA AQQYALITRL LYTPDAEAMG
Helix Consensus		....
HSV-1 Vmw175	390	WLQNPRVVPG DVALDQACFR ISGAARNSSS FITGSVARAV PHLGYAMAAG RFGWGLAHAA
Helix Consensus		....KLXILX
HSV-1 Vmw175	450	AAVAMSRRYD RAQKGFLLTS LRRAYAPLLA RENAALTGAA GSPGAGADDE 500
Helix Consensus		XAVXVIXXLE QQVRERNLN.. ....

Figure 3A.6. Alignment Between the DNA Binding Domain of Vmw175 and a Consensus Sequence for the Second Helix of the Helix Loop Helix Motif.

The sequence of Vmw175 between amino acids 330-500 is shown, together with the aligned consensus sequence for the second helix of the helix loop helix motif. The helix consensus sequence was compiled from the sequences of helix loop helix proteins published in Murre et al. (1989), using the GCG Pileup and Pretty programmes. Alignments were made between the Vmw175 sequence and the helix consensus sequence with the GCG Bestfit programme.

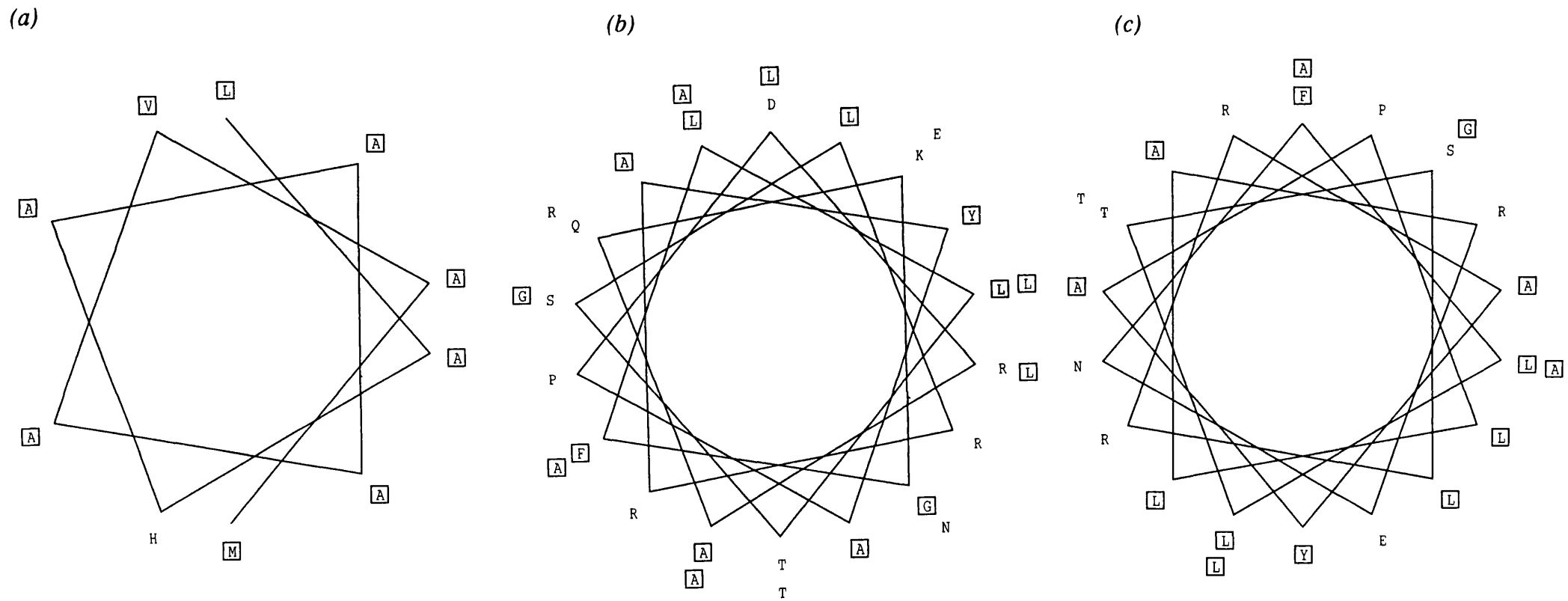


Figure 3A.7. Helical Wheel Models of Proposed Helical Regions of Vmw175.

Potential α-helices of Vmw175 were determined as described in Section 3A.2. The proposed α-helices were modelled on helical wheels using the GCG Helicalwheel programme. The helices shown incorporate the following codons of Vmw175:

(a) 445-454 *(b) 459-487* *(c) 465-487*

Boxes indicate hydrophobic residues.

Fashman rules. Both analyses predict the existence of an α-helix spanning amino acids 445-454, the analyses differ in the prediction of the limits of the second helix. Shepard *et al.*, (1989) proposed that an α-helix spans the region encompassing amino acids 459-487. However, the analysis presented in Section 3A.2 showed that this helix may be limited to a region spanning codons 465-487. The GCG Helicalwheel programme was used to represent the three proposed helices as helical wheels. The resultant models are illustrated in Figure 3A.7.

3A.6. Discussion.

Analysis of conserved residues within a family of proteins has long been accepted as a method useful in pinpointing regions of a protein which contribute to its *in vivo* activity, especially in the case of DNA binding proteins. In order to analyse Vmw175 in this way, the protein has to be aligned correctly with other members of the α-herpesvirus regulatory protein family. The alignment in Figure 3A.1 agrees well with previously published analyses (see, for example, Grundy *et al.*, 1989); and the region determined by this alignment which is proposed to encode the PRV IE180 DNA binding domain correlates with the observation that a bacterially expressed polypeptide spanning codons 448-696 of the PRV IE180 protein produces a DNAaseI footprint at the IE-3 capsite, and binds to several other sites in the VZV gene 62 promoter (Wu and Wilcox, 1991). The majority of conserved and identical amino acid residues fall within region 2 of the proteins. This is to be expected, as region 2 is one of the two most homologous regions of the α-herpesvirus transactivating proteins (McGeoch *et al.*, 1986). Of particular interest is one highly conserved stretch of amino acids spanning residues 442-475 of Vmw175. This region exhibits 92% conservation and 68% identity between the four proteins, suggesting that it may play a functional role in the activity of this region of the protein.

The importance of the less homologous region 1 to the activity of the DNA binding domain of the four proteins is unclear. Studies have shown that the I10X polypeptide encompassing codons 275-523 of Vmw175 exhibits DNA binding activity in gel retardation analysis, but fails to produce a DNAaseI footprint at the IE-3 capsite (R. Everett, personal communication). This polypeptide contains a smaller portion of region 1 as compared to the I9X DNA binding domain polypeptide, which is able to footprint the same DNA sequence. Only region 2 of the 140K protein is essential for the DNA binding activity of the molecule, as a polypeptide spanning codons 472-646 shows DNA binding activity in gel retardation analysis. But again, this protein does not footprint the same DNA sequences and incorporation of region 1 sequences into this

polypeptide are required to restore its footprinting activity (Tyler and Everett, 1993). The portions of region 1 incorporated into the larger DNA binding domain polypeptides of these proteins show limited homology to each other and to the two immediate early proteins of PRV and EHV-1 (see Figure 3A.1). The function of these amino terminal regions is unclear - it is possible that the specific amino acid content of the region is unimportant, but that addition of protein segments increases the conformational stability of the isolated DNA binding domain polypeptide in such a way that a DNaseI footprint becomes visible.

Predictive methods for determining the secondary structure of a protein are not entirely accurate. For example, Garnier *et al.* (1978) reported that their predictive method resulted in an average success rate of 62% when predictions were made for 26 proteins of known secondary conformation. Therefore, implications inferred about the secondary structure of primary amino acid sequence must be verified by nuclear magnetic resonance or X-ray crystallographic structural data before any conclusions can be made about the secondary and tertiary conformation of particular regions.

This observation is supported by the fact that the predictions obtained with the GCG Peptidestructure programme detailed in Section 3A.2 differ widely from a published observation that application of Chou and Fashman rules to the region of Vmw175 spanning codons 445-487 results in a clear secondary structure prediction of a helix-turn-helix motif, the turn being situated at amino acids 458-461 (Shepard *et al.*, 1989). It is possible that the region of the motif carboxy terminal to amino acid 441 may form a helical structural motif, but the probability for this occurring appears to be low. The differences in conclusions about the conformation of this motif may result from the varying methods of application of the Chou and Fashman rules, due to ambiguities in the descriptions of the predictive methods.

Comparison of conserved regions of interest of Vmw175 with other motifs of DNA binding proteins would show whether the protein falls into a larger family of conserved proteins. This would enable identification of amino acid residues essential for contact with the DNA helix. Unfortunately, the region encompassing amino acids 445-487, which is highly conserved and appears to be essential for the DNA binding activity of the protein, shows very little, if any, homology to recognised DNA binding motifs.

Although the segment spanning codons 445-487 yielded the highest score upon analysis with the Dodd and Egan programme to determine regions homologous to the prokaryotic helix turn helix motif (see Table 3A.1), the score determined for this region of the protein was lower than the limit considered to be statistically viable by Dodd and Egan. For example, Dodd and Egan (1987) found that if a protein scored between 1400-1499, then it had a 56% probability of being a Cro-like helix-turn-helix domain. This implies that the Vmw175 "helix-turn-helix" motif may be distantly related to the

prokaryotic helix turn helix motif, but is not a member of that protein motif family. In addition, the proposed "helix-turn-helix" motif within Vmw175 is larger than the prokaryotic 20 amino acid helix-turn-helix motif. As this region is highly conserved in all four α-herpesvirus regulatory proteins, it is not surprising that analysis of the EHV IE-1 protein with the Dodd and Egan programme results in a similar score for this region.

Alignment of other consensus motifs with the four α-herpesvirus regulatory proteins showed that the region of interest exhibits no obvious homology to the homeobox domain and limited homology to a portion of the helix loop helix domain.

However, the homeobox domain shows some homology to Vmw175 residues 353-396, and it is of interest to note that the highly conserved homeobox residues trp phe X asn X arg (shown in Gehring, 1987) which lie towards the carboxy terminal of the recognition helix, align with a highly conserved region of the four α-herpesvirus regulatory proteins. This region spans amino acids 390-395 of Vmw175, and exhibits some homology to both homeobox consensuses. Mutational analyses have illustrated that the glutamine residue within this sequence of *Drosophila* homeodomain proteins determines the DNA-binding specificity of the protein (Hanes and Brent, 1989; Treisman *et al.*, 1989). In addition, X-ray crystallographic analysis of the *Drosophila engrailed* homeotic protein has demonstrated that five out of the six amino acids in this short sequence either make direct contacts with the DNA duplex or stabilize the overall tertiary structure of the homeodomain. Whether this region of Vmw175 is of importance to the DNA binding activity of the protein has not been determined, although insertional mutations created in the vicinity of this region disrupt the DNA binding properties of Vmw175 (Paterson and Everett, 1988b; see Figure 1C.3).

Helix two of the helix loop helix motif exhibits some homology with a region spanning Vmw175 codons 444-468, however, the region proposed to be homologous to the first helix is 64 amino acids upstream. This results in two helices separated by over 60 amino acids of primary sequence, a region longer than the usual loop length of 6-21 amino acids, implying that Vmw175 is not a member of this class of DNA binding proteins.

Finally, if the region of Vmw175 encompassing amino acids 445-487 does contain two α-helices, it would be interesting to investigate the nature of the helices. It has been shown that the helix loop helix and leucine zipper motifs contain amphipathic helices , and that many DNA binding proteins contact DNA through basic amino acid residues (Jones, 1990; Vinson *et al.*, 1989). Therefore, the GCG Helicalwheel programme was used to determine the existence of such amphipathic helices within the Vmw175 "helix-turn-helix" region. Figure 3A.7 illustrates the observation that no helix exhibited amphipathic properties. The first helix contains predominately hydrophobic residues, and whilst the second helix contains a higher proportion of hydrophilic polar

or charged amino acid residues, there appears to be no concentration of hydrophilic residues on any one face of the helix. Therefore, although the region may be predominately α-helical in nature, the helices exhibit very few characteristics common to other DNA binding proteins.

Although the region exhibits a low degree of homology to other classes of DNA binding proteins, it must be remembered that it exhibits an unusual high degree of conservation within the family of α-herpesvirus regulatory proteins, and that disruption of this region has an extremely deleterious effect on the properties of the molecule. The possibility exists that this region of Vmw175 encodes a DNA interaction motif, and further mutational analysis may indicate the importance of residues within this motif to the properties of the protein molecule.

3B. Oligonucleotide-Directed Mutagenesis of the DNA Binding Domain of Vmw175.

The primary aim of the work presented in this thesis was to carry out a detailed mutational analysis of a proposed "helix-turn-helix" motif within the DNA binding domain of Vmw175 in order to determine the function of the motif and its importance relative to the *in vitro* and *in vivo* properties of Vmw175. This section describes the construction of point and in-frame deletion mutations within the motif; later sections deal with the analysis of the effects of these mutations on the properties of Vmw175.

3B.1 Constructs M13T7I9 and M13T7I10.

The constructs M13T7I9 and M13T7I10 (obtained from R. D. Everett) were cloned from plasmids pI9 and pI10 (Paterson, 1989). Plasmids pI9 and pI10, based on the plasmid p175 (Everett, 1987b), carry the IE-3 coding region cloned on a BamHI-SstI fragment containing most of IR_S, the 'a' sequence and part of IR_L of HSV-1, linked to the SV40 early promoter and enhancer (see Figure 3B.1). The IE-3 nucleotide sequence is numbered from +1 at the first base of the ATG start codon in the open reading frame, and the Vmw175 amino acid sequence is numbered from +1 at the ATG start codon. Plasmids pI9 and pI10 contain in-frame insertional mutations within the IE-3 coding region at nucleotides 752 and 823 (codons 252 and 276) respectively.

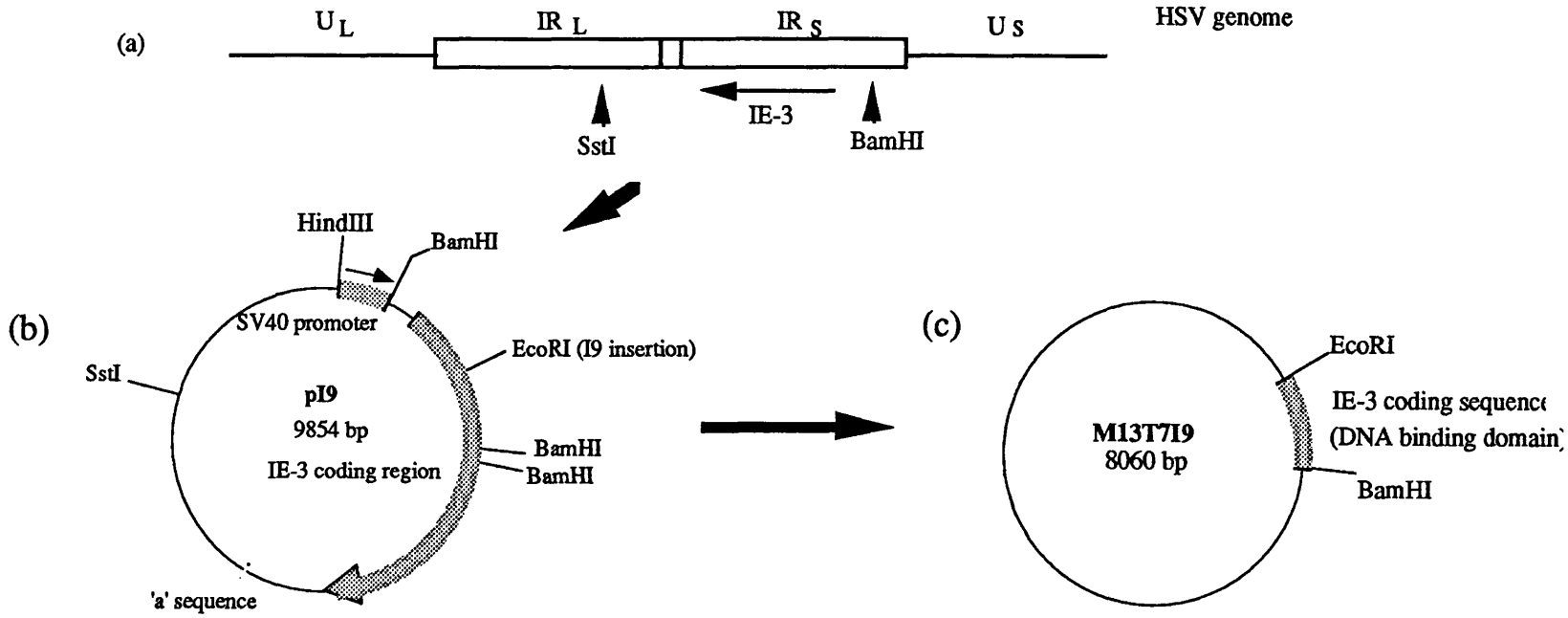


Figure 3B.1. Structure of constructs pI9 and M13T7I9.

(a) Diagrammatical representation of a portion of the HSV genome, showing the position of the IE-3 gene (not to scale).

(b) The plasmid pI9 (Paterson and Everett, 1988a) is based on the plasmid p175 (Everett, 1987b), which carries the IE-3 gene cloned on a BamHI-SstI fragment including parts of HSV-1 IR_L and IR_S as shown. The BamHI site is situated 26bp downstream of the IE-3 transcriptional start site, which in these plasmids is replaced by the SV40 early promoter and enhancer sequences cloned on a HindIII-BamHI restriction fragment. In addition, pI9 contain an EcoRI insertional mutation at codon 252. See text for full description of the insertional mutation.

(c) The bacteriophage M13T7I9 contains the EcoRI-BamHI restriction fragment of pI9, which encompasses Vmw175 codons 252-523. This fragment is inserted between the EcoRI and BamHI restriction sites of the vector M13mp18.

These mutations consist of the EcoRI dodecamer linker 5'CCCGAATTCGGG3' inserted at blunt ended SmaI and NaeI restriction sites within the IE-3 coding region. As p175 contains no other EcoRI restriction site, plasmids pI9 and pI10 contain unique EcoRI restriction sites at the point of insertion. The IE-3 coding region also contains a BamHI restriction site at nucleotide 1725; the EcoRI-BamHI fragments of pI9 and pI10, spanning nucleotides 725-1725 and 823-1725 respectively, encode protein sequences which include the DNA binding domain of Vmw175.

Constructs M13T7I9 and M13T7I10 contain the above EcoRI-BamHI restriction fragments from plasmids pI9 and pI10 cloned between the EcoRI and BamHI sites of the bacteriophage vector M13mp18. Single stranded DNA prepared from these vectors was used as a template for oligonucleotide-directed mutagenesis of selected portions of the DNA binding domain of Vmw175.

3B.2 Random Mutagenesis of the DNA Binding Domain of Vmw175.

The method employed, as described by Blacklow and Knowles (1991), utilizes "spiked" oligonucleotide preparations in directed mutagenesis to create a series of random point base pair mutations. Four overlapping oligonucleotides complementary to a region spanning the proposed helix-turn-helix motif (nucleotides 1332-1428 of the IE-3 coding region, see Figure 3B.2) were synthesized using a C precursor mix which had been contaminated with the A phosphoramidite at such a ratio as to produce one random C to A mutation per oligonucleotide molecule (McNeil and Smith, 1985). When used in oligonucleotide- directed mutagenesis, these oligonucleotides would result in the production of random G to T mutations in the upper strand of the DNA helix, throughout the helix-turn-helix region.

Site-directed mutagenesis was carried out using the Amersham *in vitro* mutagenesis system, based upon the protocol of Eckstein *et al.* (Sayers *et al.*, 1988), essentially as described in Section 2B.22. Mutagenesis was carried out with each of the four "spiked" oligonucleotides (LA15-18, see Figure 3B.2) and either M13T7I9 or M13T7I10 template DNA. The resultant M13 DNA was used to transform competent TG1 or JM101 cells. Individual clones were screened by dideoxy sequence analysis as described in Section 3B.4.1.

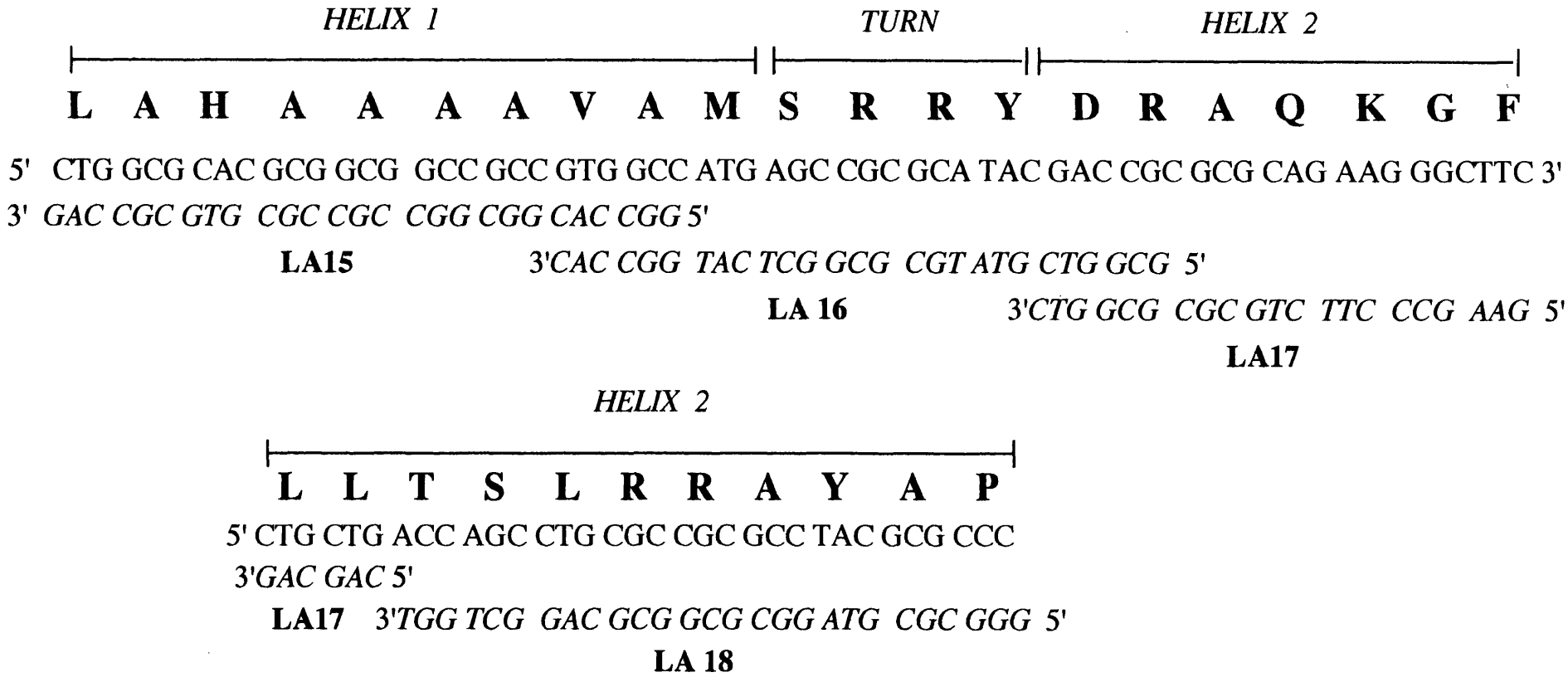


Figure 3B.2. Sequence of the "Helix-Tnrn-Helix" Motif Showing the Position of "Spiked" Oligonucleotides.

The amino acid sequence of Vmw175 spanning codons 445-476, containing the proposed helices and turn (Shepard et al., 1989) is shown in one letter code, below this is the nucleotide sequence spanning nucleotides 1333-1428. The sequences of "spiked" oligonucleotides used for random, oligonucleotide-directed mutagenesis are shown in italics.

3B.3. Creation of Internal Deletions Within the Helix-Turn-Helix Motif.

In order to determine the function, if any, of the individual helices within the proposed helix-turn-helix motif, in-frame deletion mutations were constructed. Two oligonucleotides (LA21 and LA22) were designed, which when used in an oligonucleotide-directed mutagenesis system would result in 7 amino acid and 10 amino acid deletions within helices 1 and 2, respectively (see Figure 3B.3). These oligonucleotides contain XhoI restriction sites, resulting in the incorporation of unique XhoI restriction sites within resultant mutant constructs. Mutagenesis was carried out with the Amersham *in vitro* mutagenesis system as described in Section 2B.22, using either LA21 or LA22 with M13T7I10 template DNA. The resultant M13 DNA was used to transform competent TG1 or JM101 cells. Individual clones were screened by restriction enzyme digestion and dideoxy sequence analysis, as described in Section 3B.4.2.

3B.4 Characterization of Mutations Created Within the DNA Binding Domain of Vmw175.

3B.4.1 Characterization of Point Mutations.

Resultant clones from random "spiked" oligonucleotide-directed mutagenesis were screened by dideoxy sequence analysis. Individual plaques were picked, and small-scale preparations of template DNA were made as described in Section 2B.19.2. Single stranded DNA was sequenced using either Klenow DNA polymerase or Sequenase^{TM} T7 DNA polymerase, with M13 universal primer or the primer LA19 (5'-CCCCTCGTCATCTGCGC-3') which anneals to M13T7I9/I10+ DNA, approximately 80-150bp downstream of the site of mutation. Resultant sequencing ladders were resolved on 6% or 8% denaturing polyacrylamide gels, and the sequence compared to that of wild type IE-3 coding sequence. Figure 3B.4 shows a representative analysis of nine clones from a round of mutagenesis with the LA18 "spiked" oligonucleotide. As can be seen, both wild type and mutant clones were obtained.

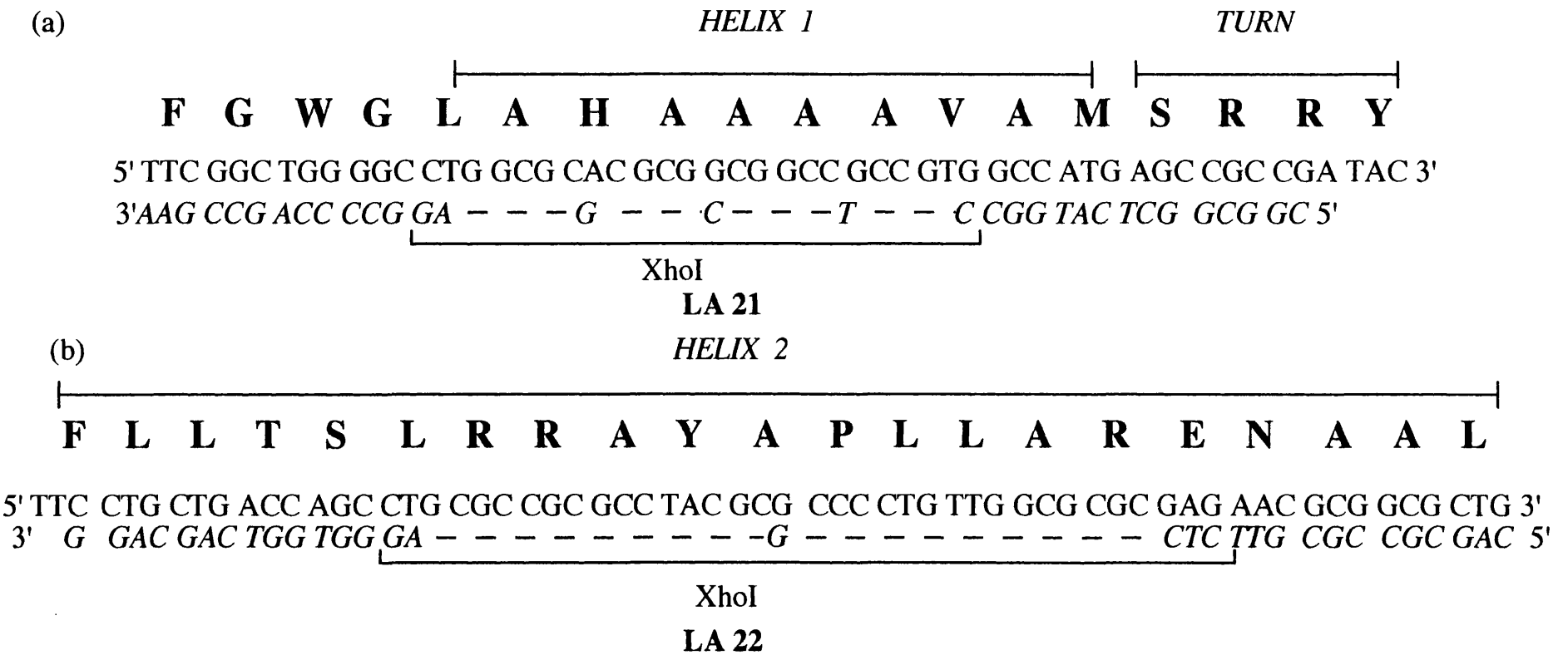
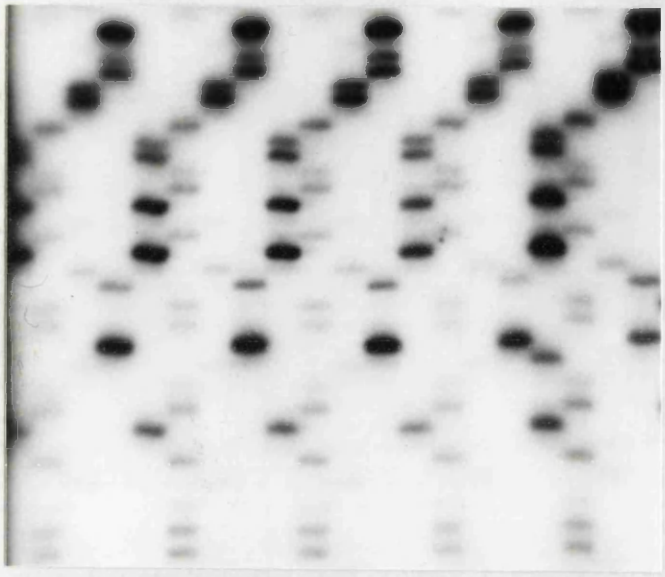
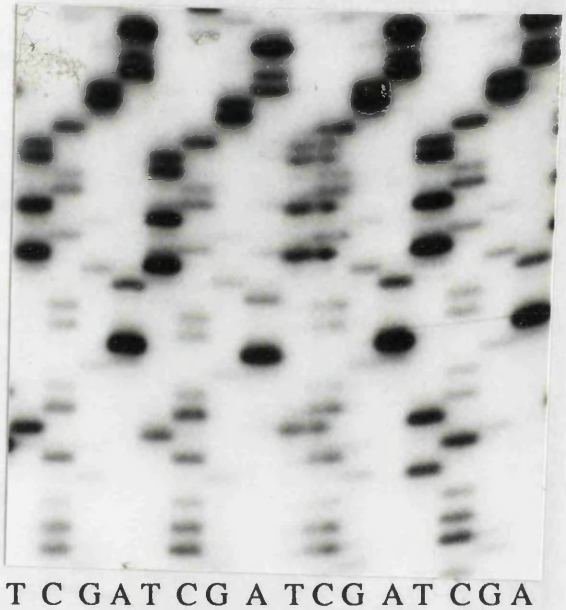


Figure 3B.3. Sequence of the "Helix-Turn-Helix" Motif Showing the Position of Oligonucleotides Used to Create Internal Deletions. *The amino acid sequence of Vmw175 from codons 441-458 (a) and 465-485 (b), is shown in single letter code, below this is the nucleotide sequence from nucleotides 1321-1374 (a) and 1393-1455 (b). The positions of proposed helices and turn (Shepard et al., 1989) are also indicated. The sequence of oligonucleotides used to create internal deletions is shown in italics, with XhoI sites underlined. Dashed lines denote the regions deleted from Vmw175 sequences.*



T C GA T C G A T C G A T C G A T C G A



T C G A T C G A T C G A T C G A

Figure 3B.4. Representative Sequence Analysis of Clones from In Vitro Site Directed Mutagenesis.

Clones generated by in vitro site-directed mutagenesis were sequenced by the dideoxy chain termination method, using Klenow DNA polymerase, as described in Section 2B.21. Radiolabelled fragments generated by the sequencing reaction were electrophoresed on a denaturing 8% acrylamide gel, and fragments visualized by autoradiography. The Figure illustrates sequences generated from 9 clones. Two of the clones illustrated contain point T to G mutations.
The sequence shown, starting from the top of the figure, reads:
5' CAGAAGGGCTTCCTGCTGACCAGCCTGCGCCC 3'

Table 3B.1 shows the number of mutations isolated and the mutation efficiency of each round of mutagenesis, and Figure 3B.5 shows sequence analysis of four single mutations situated in the proposed second helix, isolated from a round of

mutagenesis with LA18. The overall mutagenesis efficiency was 37%, lower than that normally achieved with the Amersham system. In total, 20 single base pair mutants, 5 double base pair mutants and 1 deletion mutant were isolated from four rounds of mutagenesis with the "spiked" oligonucleotides.

Spiked oligo-nucleotide (a)	Number clones sequenced	Number wild type clones	Number mutant clones	Mutation efficiency (b)	Number of mutants obtained (c)		
					Single	Double	Deln.
LA15	84	47	37	44%	6	1	1
LA16	52	43	9	17%	5	1	-
LA17	65	39	26	40%	4	2	-
LA18	28	15	13	46%	5	1	-
TOTAL	229	144	85	37%	20	5	1

Table 3B.1: Mutation Efficiency of Random Oligonucleotide-Directed Mutagenesis.
(a) The sequences of the spiked oligonucleotides used for in vitro mutagenesis reactions are shown in Figure 3B.2.
(b) The mutation efficiency for each oligonucleotide is represented as the percentage of clones screened which were found to contain mutations.
(c) Represents the number of individual clones obtained from one round of mutagenesis with the given oligonucleotide which result in either a single amino acid alteration, double amino acid alteration or base pair deletion. Several additional silent base pair mutations were also isolated.

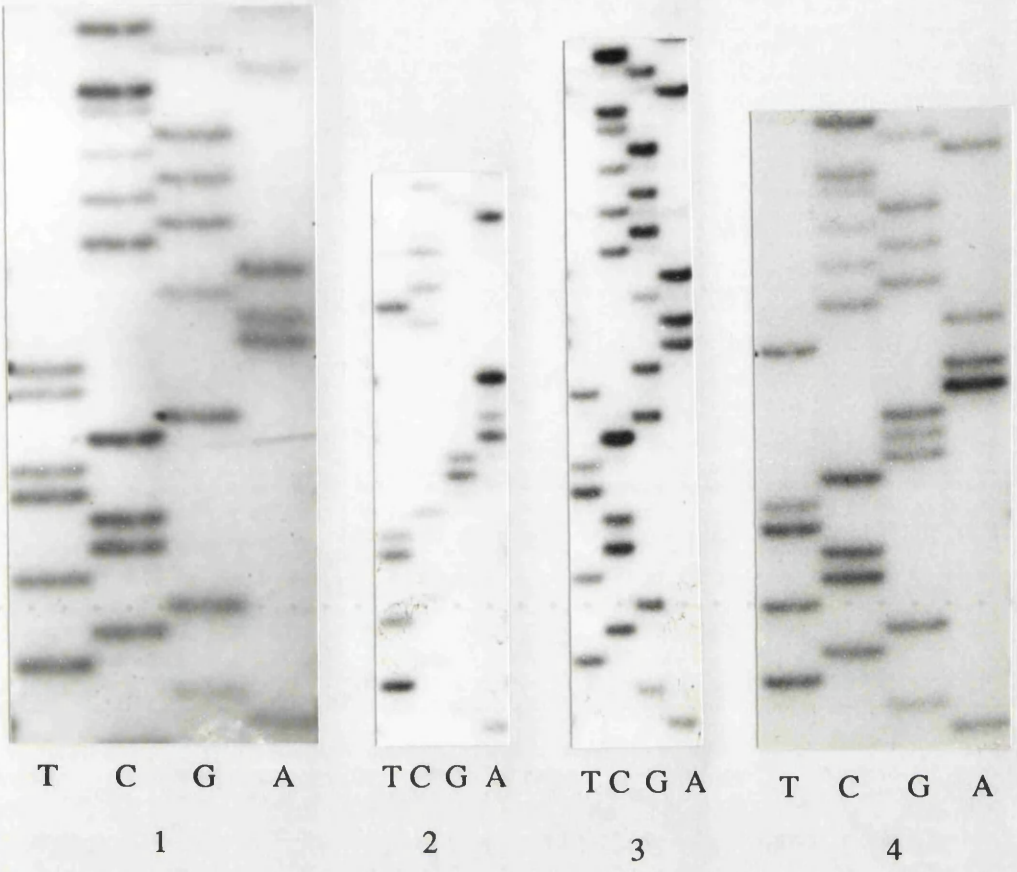


Figure 3B.5. Sequence Analysis of Four Mutant Clones Isolated from In Vitro Site-Directed Mutagenesis.

Clones generated by in vitro site-directed mutagenesis were sequenced by the dideoxy chain termination method, using a Sequenase Version 2.0 sequencing kit or Klenow DNA polymerase, as described in Section 2B.21. Radiolabelled fragments generated by the sequencing reaction were electrophoresed on a denaturing 8% acrylamide gel, and fragments visualized by autoradiography. The Figure illustrates sequences generated from four mutant clones.
The sequence of each mutant clone, starting from the top of the figure, reads:
1: 5' CGACCGCGCGCAGAATTGCTTCCTGCTGA 3'
2: 5' CGACCGCTCGCAGAAGGGCTTCCTGCTGA 3'
3: 5' CGACCGCGCGCAGAAGTGCTTCCTGCTGA 3'
4: 5' CGACCGCGCGCATAAGGGCTTCCTGCTGA 3'

3B.4.2. Characterization of Internal In-Frame Deletion Mutations.

Resultant clones from site-directed mutagenesis using the oligonucleotides LA21 and LA22 were initially screened by XhoI restriction enzyme digestion. From an initial screen of 12 clones, the mutation efficiency was found to be 75% with LA21 and 80% with LA22. Clones containing XhoI restriction sites were further analysed by dideoxy sequencing reactions, as described in Section 3B.4.1. A small number of clones were found to have large DNA rearrangements; the majority gave the correct sequence. Figure 3B.6 shows representative sequence analysis of clones obtained from mutagenesis with the LA22 oligonucleotide.

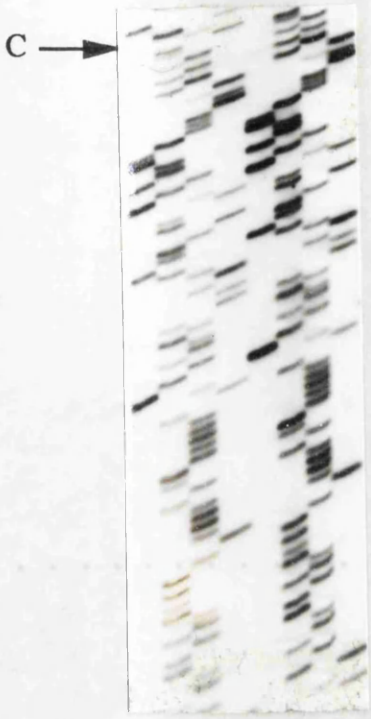


Figure 3B.6. Representative Sequence Analysis of Clones from In Vitro Site Directed Mutagenesis.

Clones generated by in vitro site-directed mutagenesis were sequenced by the dideoxy chain termination method, using a Sequenase Version 2.0 sequencing kit, as described in Section 2B.21. Radiolabelled fragments generated by the sequencing reaction were electrophoresed on a denaturing 8% acrylamide gel, and fragments visualized by autoradiography. The Figure illustrates sequences generated from 2 clones. The clones illustrated contain deletion mutations introduced by the mutagenic oligonucleotide LA22.
The sequence shown, starting from the arrowed C and proceeding downwards, reads:
5'
CCGCGCGCAGAAGGGCTTCCTGCTGACCAGCCTCGAGAACGCGGCGCTGA
CGGGGGCCGCGGGGAGCCCCGGCGCCGG 3'
The XhoI restriction site introduced by the mutagenesis is underlined.

3B.5. Discussion.

Random mutagenesis of the "helix-turn-helix" motif was achieved by the introduction of G to T mutations throughout the region encompassing nucleotides 1333-1428. This approach was used, rather than the introduction of specific amino acid alterations throughout the motif, as little was known about the function of portions of the region, and the Vmw175 motif fitted no known consensus sequences for other DNA binding proteins, which may have highlighted specific target residues for mutagenesis (see Section 3A.6). Of the twelve possible combinations of base pair alterations that could be introduced, the G to T alteration was used to carry out

mutagenesis of the amino acid sequence, as it was determined that alteration of individual G residues throughout the motif to T residues would result in the maximum number of mutations of conserved amino acids. In addition, alteration of G residues to T residues would not result in the introduction of stop codons within the target region.

Mutagenesis with the "spiked" oligonucleotides resulted in an overall mutation efficiency of 37%, which is lower than that normally achieved with the Amersham system. With other, non-"spiked" oligonucleotide mutagenesis reactions, the mutation efficiency was as expected (eg. 75% with LA21, see Section 3B.4.2). Therefore, it appears that this drop in efficiency is due solely to the use of the "spiked" oligonucleotides. Degradation of the "spiked" oligonucleotides LA15-18 was observed when the oligonucleotides were used as primers for dideoxy sequencing reactions. This degradation may have led to the reduced efficiency due to problems created during the annealling step of the mutagenesis protocol. In addition, the synthesis of "spiked" oligonuceotides results in a final preparation containing at least 50% of oligonucleotides with a wild type sequence. The use of such a mixture in oligonucleotide-directed mutagenesis would result in a drop in the mutation efficiency, as observed.

Although the mutation efficiency of the system was relatively low, most of the mutations which were expected to occur were isolated. Only 5 mutations were not isolated; the majority of these were situated towards the 5' or 3' termini of the mutagenic oligonucleotides. "Melting" of the termini of the oligonucleotides during mutagenesis may have led to inefficiencies in either primer utilization or ligation after synthesis of the completed second strand, resulting in the failure to obtain these mutations. Several double mutations were isolated, most of which duplicated single mutations previously isolated. In the majority of cases these were not analysed further. One base pair deletion mutant was isolated from a round of mutagenesis with the LA15 olignucleotide. This deletion resulted in a frameshift in the IE-3 coding region, and introduction of a stop codon at codon 454. Expression of this mutation would lead to truncation of the Vmw175 DNA binding domain before the proposed turn of the motif, and analysis of the *in vitro* properties of this mutant could possibly reveal the function of this portion of the Vmw175 DNA binding domain.

The majority of point mutations were analysed further. The mutations analysed for *in vitro* DNA binding activity are summarized in Figure 3B.7 and Table 3B.2.

In-frame deletion mutations were introduced within the motif in order to remove large regions of either proposed helix. Analysis of these mutations would provide information on the function of either helix, a technique which has proved useful in delineating the function of other helical protein motifs (Williams and Tjian, 1991). Two rounds of mutagensis were caried out to introduce these deletions into the

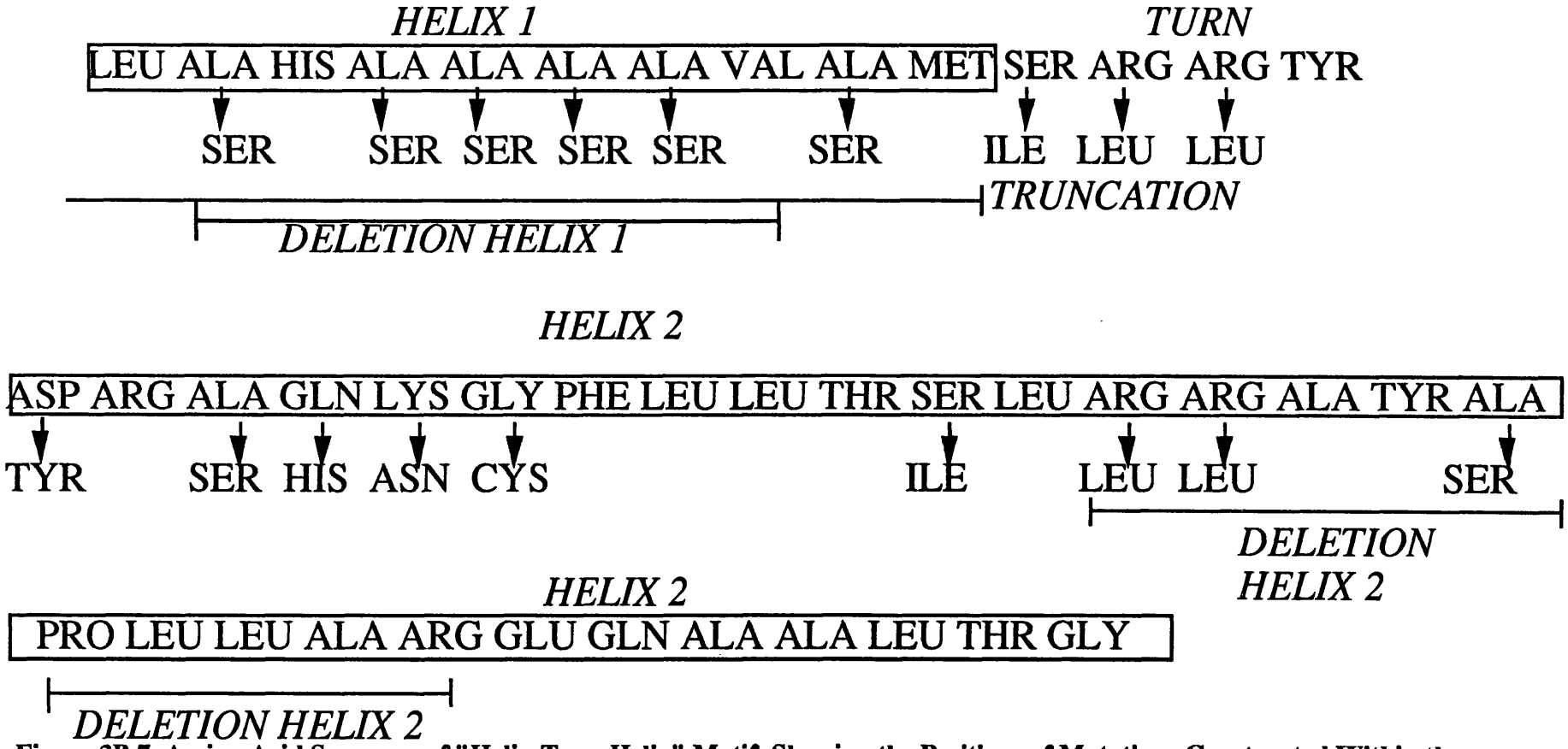


Figure 3B.7. Amino Acid Sequence of "Helix-Turn-Helix" Motif, Showing the Positions of Mutations Constructed Within the Motif.

The Vmw175 amino acid sequence spanning codons 445-487 containing the proposed helices and turn (Shepard et al., 1989) is shown in three letter code. The positions of point amino acid mutations are indicated by arrows; the limits of truncations and deletions are indicated by bars.

Table 3B.2: Mutants Isolated From Spiked Oligonucleotide-Directed Mutagenesis, Analysed for DNA Binding Activity.

Construct (a)	Site of G to T mutation (b)	Site of amino acid alteration (c)	Amino acid alteration (d)	Name (e)
LA15 no. 2	1351	451	Ala to Ser	451 A/S
LA15 no. 3	1351 1357	451 453	Ala to Ser Ala to Ser	451/3AA/SS
LA15 no. 5	1345	449	Ala to Ser	449A/S
LA15 no. 9	1336	446	Ala to Ser	446A/S
LA15 no. 36	1357	453	Ala to Ser	453A/S
LA15 no.74	1348	450	Ala to Ser	450A/S
LA16 no. 16	1357	453	Ala to Ser	453A/S
LA16 no. 39	1370	457	Arg to Leu	457R/L
LA16 no. 48	1375	459	Asp to Tyr	459D/Y
LA16 no. 49	1367	456	Arg to Leu	456R/L
LA17 no. 6	1381	461	Ala to Ser	461A/S
LA17 no. 29	1390	464	Gly to Cys	464G/C
LA17 no. 31	1386	462	Gln to His	462Q/H
LA17 no.35	1389 1390	463 464	Lys to Asn Gly to Cys	463/4KG/NC
LA18 no. 1	1415	472	Arg to Leu	472R/L
LA18 no.14	1412	471	Arg to Leu	471R/L
LA18 no. 19	1406	469	Ser to Ile	469S/I
LA18 no. 26	1423	475	Ala to Ser	475A/S
LA15 no. 57	1355(del)	452	Stop codon at 454	452del

Legend to Table 3B.2.

(a) The original name of the mutant clone, showing the oligonucleotide used for the mutagenesis reaction and the clone number isolated.

(b) The site of G to T mutation is shown as nucleotide number within the IE-3 open reading frame, numbered from +1 at the ATG start codon. In the case of the mutant 452del, the base pair mutation introduced was not a G to T mutation, but a base pair deletion.

(c) The site of amino acid alteration is shown as codon number within the IE-3 open reading frame, numbered from 1 at the ATG start codon

(d) The type of amino acid alteration introduced by the G to T mutation. Amino acids are represented as three letter codes.

(e) The final construct name, showing the position of mutation as codon number, followed by the amino acid alteration. Amino acids are represented as one letter codes.

motif, as described in Section 3B.3. The mutation efficiency achieved with the "deletion" oligonucleotides was close to the expected efficiency of the Amersham system, and deletion mutations of the correct sequence were obtained for further analysis.

3C. Expression and Purification of Wild Type and Mutant Polypeptides Encompassing the DNA Binding Domain of Vmw175

In order to investigate the *in vitro* properties of the mutations whose construction is described in Section 3B, the mutations were introduced into a bacterial expression system. Mutant polypeptides encompassing the DNA binding domain of Vmw175 expressed from this system provided a means of analysing the effects of the mutations on the *in vitro* properties of Vmw175. This section describes the expression and purification of wild type and mutant DNA binding domain polypeptides.

3C.1. The pET Expression System.

Previously, portions of the IE-3 coding region have been expressed in a bacterial T7 expression system in order to assess the *in vitro* properties of the isolated DNA binding domain of Vmw175 (Pizer *et al.*, 1991; Everett *et al.*, 1991a, b). The bacterial expression system used in these studies and for the work presented in this thesis is based on the pET expression system of Studier *et al.* (1990). The plasmid pA585.4 (obtained from C. M. Preston, Institute of Virology, Glasgow) is a derivative of the plasmid pET8c, constructed by the removal of the HindIII site of the parent pBR322 vector and the insertion of an oligonucleotide linker containing a HindIII restriction site between the NcoI and BamHI sites of the T7 promoter/terminator sequences (see Figure 3C.1). The plasmid series pA585T7 (obtained from J. Tyler, Institute of Virology, Glasgow) were further derivatives of pA585.4, created by removal of the SphI/PvuII restriction fragment; removal of the EcoRI restriction site; and the insertion of NcoI/EcoRI oligonucleotide linkers and the EcoRI/HindIII restriction fragment containing the pUC18 polycloning site between the NcoI and HindIII sites of pA585.4 (see Figure 3C.1). The plasmids pA585T7a and pA585T7b_2

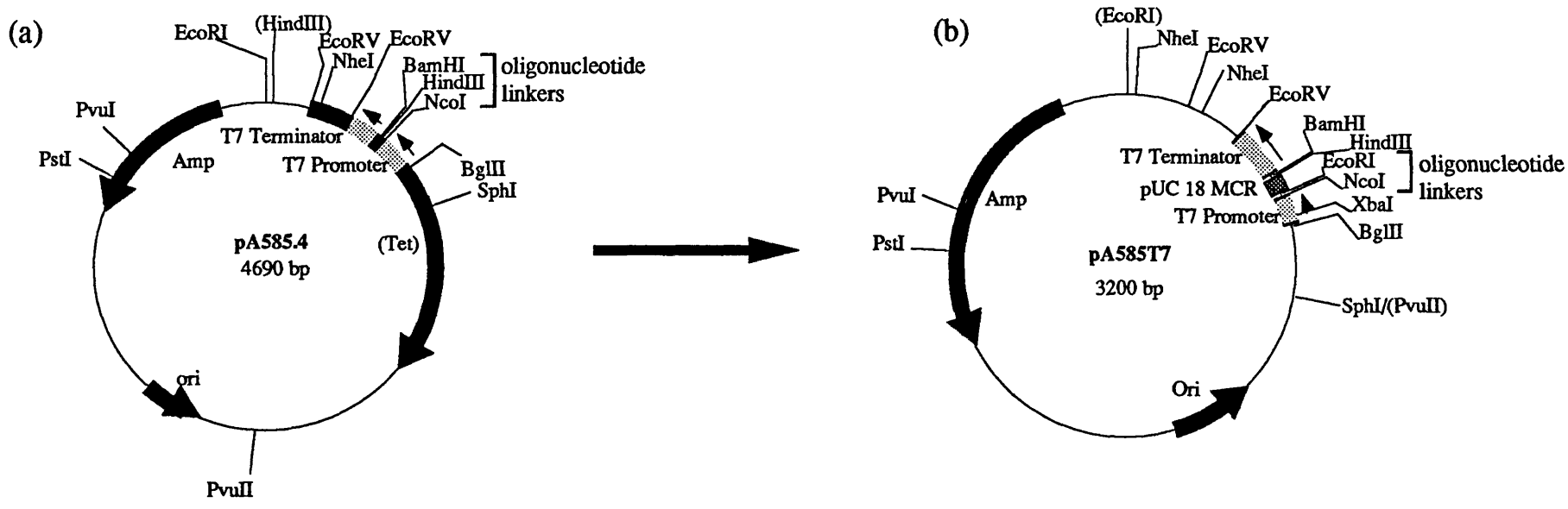


Figure 3C.1a. The pET Expression Vectors.

(a) The plasmid pA585.4, based on the vector pET8c (Studier et al., 1990), constructed by the removal of the pET8c HindIII restriction site and insertion of HindIII oligonucleotides between the NcoI and BamHI restriction sites between the T7 promoter and terminator regions.

(b) The plasmid pA585T7, based on pA585.4, constructed by the removal of the EcoRI site, removal of the SphI/PvuII restriction fragment and insertion of NcoI/EcoRI linkers and the pUC18 polycloning site between the NcoI and HindIII restriction sites between the T7 promoter and terminator regions.

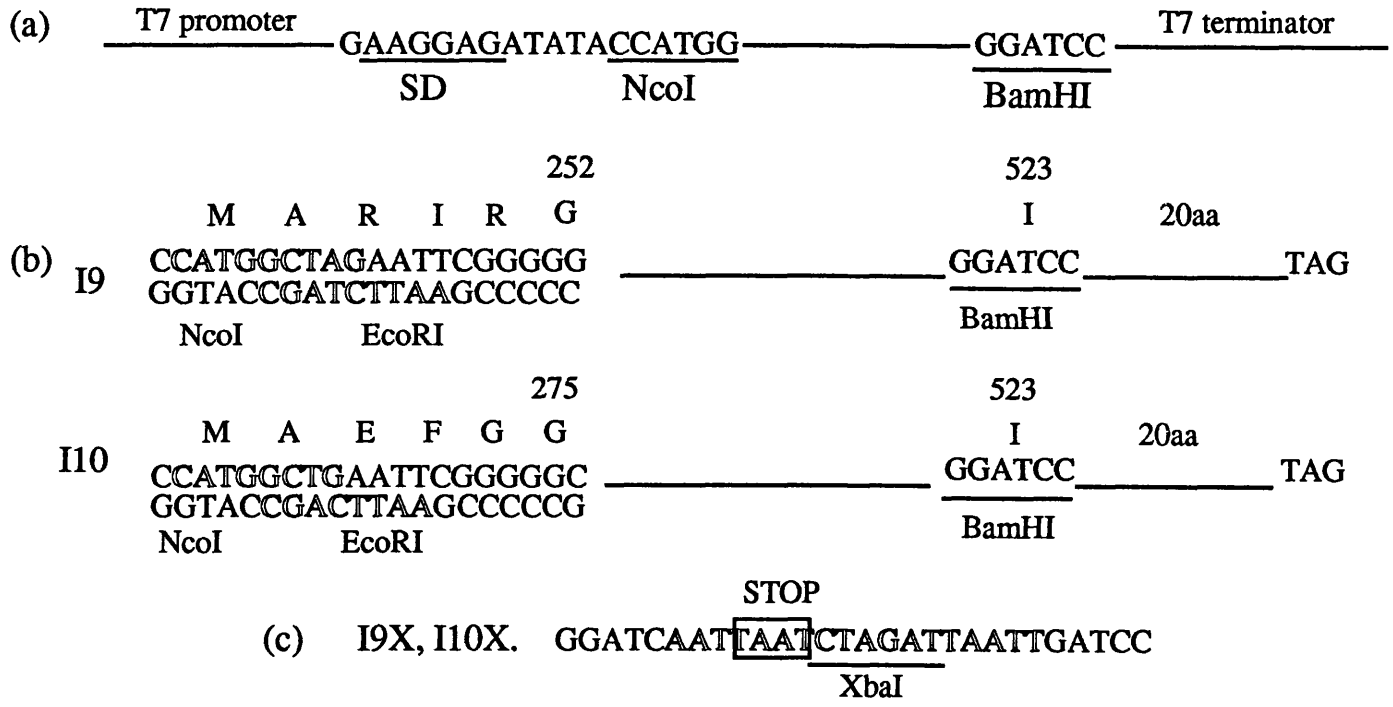


Figure 3C.1b. The pET Expression System.

(a) Diagrammatical representation of the pET T7 promoter sequences at the translation start site. The Shine-Dalgarno sequence, the NcoI site at the ATG start codon and the BamHI site of the T7 terminator are underlined.

(b) Sequences of the translation start sites of constructs JI9 and JI10. NcoI/ EcoRI linker sequences are shown in outline. The top line shows the amino acid sequence of the expressed polypeptides, in single letter code. Numbers represent the codon within the IE-3 open reading frame.

(c) The sequence of the XbaI stop codon oligonucleotide linkers inserted into the BamHI sites of vectors pJI9 and pJI10. The linker sequence is shown in outline and the XbaI restriction site is underlined.

contain unique NcoI/EcoRI oligonucleotide linkers enabling expression of cloned coding regions in two reading frames. The plasmids pJI9 and pJI10 contain the I9/I10 EcoRI/BamHI restriction fragments encompassing IE-3 nucleotides 752-1725 and 823-1725, as described in Section 3B.1, cloned into the EcoRI and BamHI restriction sites of pA585T7b_2 and pA585T7a respectively, thus enabling expression of portions of the IE-3 coding region in bacteria. pJI9 expresses codons 252 to 523 and pJI10 expresses codons 276 to 523 of Vmw175, both with an additional five residues encoded by vector sequences at the N-terminus, and another 20 spurious residues at the C-terminus (see Figure 3C.1b).

Mutations created as described in Section 3B were transferred as EcoRI/BamHI restriction fragments into plasmids pJI9 or pJI10. In order to remove the 20 C-terminal vector encoded residues, "X" versions of the plasmids were made by insertion of XbaI blunt-ended oligonucleotide linkers containing stop codons in all three reading frames into the blunt-ended BamHI restriction site immediately preceding the T7 terminator sequences (see Figure 3C.1). The sequence containing the proposed mutation was verified by dideoxy sequence analysis after each step of the cloning procedure.

The cloned Vmw175 coding sequences within these plasmids also contained a unique NruI restriction site, situated at residue 920 of the IE-3 open reading frame. This restriction site provided a means by which mutations could be transferred between different expression vectors. Mutations were transferred between I9 and I10 plasmids as NruI/BamHI fragments, whereas mutations were transferred between I9X and I10X plasmids as NruI/XbaI fragments.

3C.2. Production of Bacterial Extracts Containing Wild Type and Mutant Polypeptides for Gel Retardation Analysis.

Mutations were transferred to the relevant pA585T7 based vector and resultant vectors transfected into the *E. coli* strain BL21(DE3)pLysS. This bacterial strain carries a lysogen (DE3) which directs the transcription of T7 RNA polymerase upon induction with IPTG, and therefore the expression of the cloned polypeptide.

The expression of cloned polypeptides, and the preparation of bacterial extracts containing these polypeptides was carried out as described in Sections 2B.27 and 2B.28. Various methods were employed to obtain bacterial extracts containing mutant and wild type forms of the Vmw175 DNA binding domain, due to problems with the solubility and stability of expressed proteins, as detailed below.

3C.2.1 Induction and Extraction of I9 Polypeptides.

Initially, mutations were introduced into the pJI9 vector and expression from these vectors induced to produce mutant and wild type forms of the I9 polypeptide. SDS-polyacrylamide gel electrophoresis of samples taken during the extraction of wild type I9 polypeptide showed that the I9 polypeptide yield was relatively low, due mainly to insolubility of the protein and possible association of I9 with DNA in the cell debris. As the induction of I9 expression was poor, the concentration of I9 polypeptide in the final extract was low.

It was found that mutant polypeptides in general were expressed to lower quantities and were less soluble than their wild type counterparts. Therefore, it was concluded that, with the exception of the mutant 452 del, expression of mutations in the background of the I9 polypeptide would yield insufficient quantities of mutant polypeptide for *in vitro* analysis (see Tables 3C.1 and 3C.2).

3C.2.2. Induction and Extraction of I10 Polypeptides.

Mutant polypeptides were next expressed as I10 DNA binding domain polypeptides. SDS-PAGE analysis of an I10 extraction procedure is shown in Figure 3C.2. Comparison of tracks 1 and 2 shows that the induction of I10 polypeptide was still low; however, less I10 polypeptide was lost in the centrifugation of cell debris or in the final solubilization of ammonium sulphate precipitated proteins.

The induction and solubility of mutant I10 polypeptides were lower than that of the wild type polypeptide, as shown by SDS-polyacrylamide gel elctrophoresis of final bacterial extracts (see Figure 3C.3). The solubility of mutant polypeptides was extremely variable, as illustrated in Tables 3C.1 and 3C.2.

Table 3C.1. The Expression and Purification of Mutant Vmw175 DNA Binding Domain Polypeptides.

Mutant Name (a)	Expressed as (Polypeptide) (b)	Method of Extraction (c)	Further Purification (d)	Solubility (e)	Properties Determined (f)
451A/S	I9	2B.27.1	-	++	DNA binding
451/3AA/SS	I9	2B.27.1	-	+/-	DNA binding
446A/S	I9	2B.27.1	-	+++	DNA binding
449A/S	I9	2B.27.1	-	++	DNA binding
453A/S	I9	2B.27.1	-	+++	DNA binding
452del	I9	2B.27.1	-	+++	DNA binding
450A/S	I9	2B.27.1	-	+	DNA binding
472R/L	I10	2B.27.1	-	++	DNA binding
471R/L	I10	2B.27.1	-	+++	DNA binding
475A/S	I10	2B.27.1	-	+	DNA binding
469S/I	I10	2B.27.1	-	-	N/D
453A/S	I9X	2B.27.2	-	+++	DNA binding
463/4KG/NC	I10X	2B.27.2	-	+++	DNA binding
459D/Y	I9X	2B.27.2	-	++	DNA binding
456R/L	I9X	2B.27.2	-	++	DNA binding
464G/C	I10X	2B.27.2	-	+++	DNA binding
462Q/H	I10X	2B.27.2	-	+++	DNA binding
i12	I10X	2B.27.2	-	-	N/D
461A/S	I10X	2B.27.2	-	+++	DNA binding
DH1	I10X	2B.27.2	-	+	DNA binding
457R/L	I10X	2B.27.2	-	+	DNA binding'
ri12	I10X	2B.27.2	-	+	DNA binding
DH2	I10X	2B.27.2	-	++	DNA binding

Legend to Table 3C.1.

(a) The final construct name, showing the position of mutation as codon number, followed by the amino acid alteration. Amino acids are represented as one letter codes.
(b) The version of the DNA binding domain of Vmw175 used to express the individual mutations. The various constructs are described in the text.
(c) The method of extraction is indicated by the relevent Section number.
(d) Indicates whether the polypeptide was purified further by ion exchange chromatography.
(e) The qualitative solubility of the polypeptide, as measured by the amount of polypeptide in the final soluble extract, as determined by Western Blot analysis. The solubility of individual mutant polypeptides was measured by comparing to an internal wild type polypeptide control. - = barely detectable levels of polypeptide; ++++ = level of wild type polypeptide.
(f) The properties measured using the soluble extract shown. DNA binding activity was determined by gel retardation analysis as described in Section 3D. N/D = properties not determined due to insolubility of the polypeptide.

Table 3C.2. The Expression and Purification of Mutant Vmw175 DNA Binding Polypeptides.

Mutant Name (a)	Expressed as (polypeptide) (b)	Method of Extraction (c)	Further Purification (d)	Solubility (e)	Properties Determined (f)
469S/I	I10	2B.28	-	-	N/D
472R/L	I10	2B.28	-	+++	DNA binding
471R/L	I10	2B/28	-	+++	DNA binding
475A/S	I10	2B.28	-	+	DNA binding
451A/S	I10	2B.28	-	++	DNA binding Cross-linking
451/3AA/SS	I10	2B.28	-	+	DNA binding Cross-linking
449A/S	I10	2B.28	-	+++	DNA binding Cross-linking
446A/S	I10	2B.28	-	++	DNA binding Cross-linking
453A/S	I10	2B.28	-	+	DNA binding Cross-linking
450A/S	I10	2B.28	-	+	DNA binding Cross-linking
452del	I9	2B.28	-	-	N/D
452del	I10	2B.28	-	-	N/D
472R/L	I10	2B.28	MonoS	+++	Cross-linking
451A/S	I10	2B.28	MonoS	+	Cross-linking
451/3AA/SS	I10	2B.28	MonoS	+++	Cross-linking
472R/L	I10X	2B.28	-	++	DNA binding
451A/S	I10X	2B.28	-	+	DNA binding
451/3AA/SS	I10X	2B.28	-	++	DNA binding
471R/L	I10X	2B.28	-	+	DNA binding
469S/I	I10X	2B.28	-	-	N/D
449A/S	I10X	2B.28	-	++	DNA binding
446A/S	I10X	2B.28	-	+++	DNA binding
453A/S	I10X	2B.28	-	++	DNA binding
450A/S	I10X	2B.28	-	+++	DNA binding
tsK	I10X	2B.28	-	+	DNA binding

Mutant Name (a)	Expressed as (polypeptide) (b)	Method of Extraction (c)	Further Purification (d)	Solubility (e)	Properties Determined (f)
456R/L	I10X	2B.28	MonoS	+++	Gel Filtration
463/4KG/NC	I10X	2B.28	MonoS	++	Gel Filtration
DH1	I10X	2B.28	MonoS	+	N/D

Legend to Table 3C.2.

a) The final construct name, showing the position of mutation as codon number, followed by the amino acid alteration. Amino acids are represented as one letter codes.
(b) The version of the DNA binding domain of Vmw175 used to express the individual mutations. The various constructs are described in the text.
(c) The method of extraction indicated by the relevent Section number.
(d) Indicates whether the polypeptide was purified further by ion exchange chromatography.
(e) The qualitative solùbility of the polypeptide, as measured by the amount of polypeptide in the final soluble extract, as determined by Western Blot analysis. The solubility of individual mutant polypeptides was measured by comparing to an internal wild type polypeptide control. - = barely detectable levels of polypeptide; ++++ = level of wild type polypeptide.
(f) The properties measured using the soluble extract shown. DNA binding activity was determined by gel retardation analysis as described in Section 3D. Dimerization activity was determined by glutaraldehyde cross-linking assays or gel filtration analysis as described in Section 3E. N/D = properties not determined due to insolubility of the polypeptide.

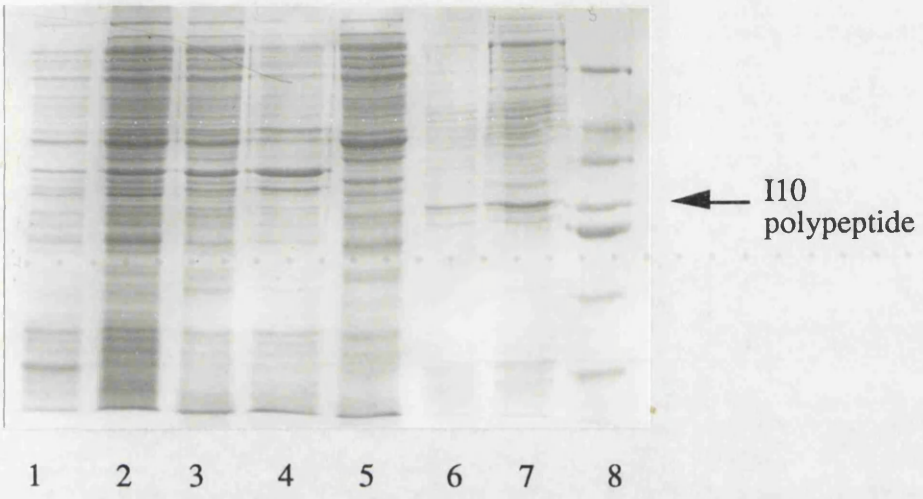


Figure 3C.2. SDS-Polyacrylamide Gel Electrophoresis of an Extraction of I10 Polypeptide.

Proteins were extracted from induced cultures of BL21(DE3)pLysS as described in Section 2B.27.2. Samples taken at particular stages during the extraction protocol were electrophoresed on 12.5% SDS-polyacrylamide gels and stained with Coomassie blue.

Track 1 : Polypeptide profile of uninduced BL21(DE3)pLysS bacteria.
Track 2 : Polypeptide profile of induced BL21(DE3)pLysS bacteria expressing I10 polypeptide.
Track 3 : 1/300 vol.supernatant after centrifugation of cell debris.
Track 4 : 1/100 vol. cell debris pellet.
Track 5 : 1/450 vol. supernatant after ammonium sulphate precipitation.
Track 6 : 1/2 vol. insoluble protein pellet.
Track 7 : 1/30 vol. final bacterial extract.
Track 8 : Molecular weight standards : 66kDa, 45kDa, 36kDa, 29kDa, 24kDa, 20.1kDa, 14.2kDa.

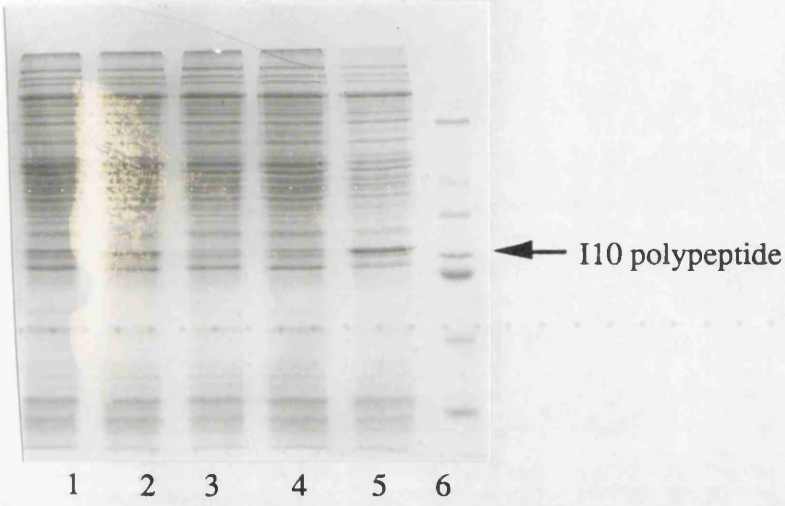


Figure 3C.3. SDS-Polyacrylamide Gel Electrophoresis of Mutant and Wild Type I10 Polypeptide Extracts.

Proteins were extracted from induced cultures of BL21(DE3)pLysS as described in Section 2B.27.2. 1/30 vol. of the final extracts were analysed on a 12.5% SDS-polyacrylamide gel and stained with Coomassie blue.
Track 1 : 472R/L I10
Track 2 : 471R/L I10
Track 3 : 469S/I I10
Track 4 : 475A/S I10
Track 5 : I10
Track 6 : Molecular weight standards : 66KDa, 45kDa, 36kDa, 29kDa, 24kDa, 20.1kDa, 14.2kDa.

3C.2.3. Induction and Expression of I9X and I10X Polypeptides.

Mutations were finally expressed within I9X and I10X polypeptides. Although mutant polypeptide induction and solubility were again lower than that of wild type, the final yield of protein was sufficient for gel retardation analysis. Figure 3C.4 shows a representative SDS-PAGE analysis of the extraction of the I10X polypeptide.

The solubility of mutant I9X and I10X polypeptides was extremely variable, as summarized in Tables 3C.1 and 3C.2.

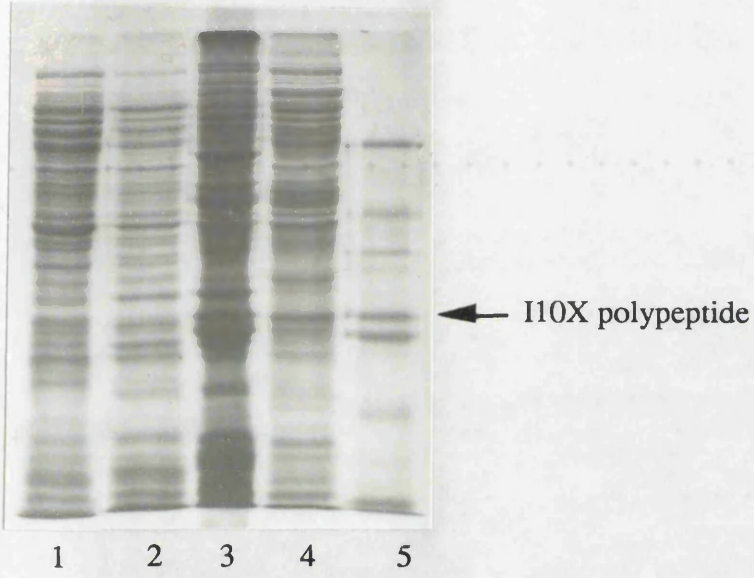


Figure 3C.4. SDS-Polyacrylamide Gel Electrophoresis of an Extraction of the I10X Polypeptide.

Proteins were extracted from induced bacterial cultures as described in Section 2B.27.1. Samples taken at various stages during the extraction protocol were electrophoresed on 12.5% SDS-polyacrylamide gels and stained with Coomassie blue.

Track 1: Polypeptide profile of induced BL21 (DE3) pLysS bacteria expressing the I10X polypeptide.
Track 2: 1/300 vol. supernatant following centrifugation of cellular debris.
Track 3: 1/2 vol. insoluble protein pellet.
Track 4: 1/30 vol. final extract.
Track 5: Molecular weight standards: 66kDa, 45kDa, 36kDa, 29kDa, 24kDa, 20.1kDa, 14.2kDa.

3C.2.4. Methods of Extraction of Mutant and Wild Type Polypeptides for Gel Retardation Analysis.

Various extraction methods were employed to obtain samples of mutant and wild type polypeptides for gel retardation analysis. Initially, extracts were obtained as described in Section 2B.27.1. This method was found to result in smeared complexes upon gel retardation analysis. This problem was resolved in two ways. Firstly, the

expression of DNA binding domain polypeptides as "X" versions (described in Section 3C.1) increased the stability of the polypeptides and resulted in clearly resolved complexes upon gel retardation analysis. Additionally, the method of extraction of the polypeptide was adapted for various purposes.

It was found that adaptation of the extraction protocol as detailed in Section 2B.27.2 resulted in extracts with relatively low yields of DNA binding domain polypeptides, but which gave clearly resolved complexes upon gel retardation analysis.

Finally, the extraction procedure was adapted to increase the purity of the expressed polypeptide. This adapted procedure is described in Section 2B.28, and was mainly used to prepare extracts for dimerization studies (see Section 3C.3). However, several mutant polypeptide extracts prepared in this way were analysed in gel retardation assays. The different methods used to prepare bacterial extracts containing wild type and mutant DNA binding domain polypeptides for gel retardation analysis are summarized in Tables 3C.1 and 3C.2. The results illustrated in Section 3D regarding gel retardation analysis of DNA binding domain polypeptides were obtained using bacterial extracts prepared as described in Sections 2B.27.2 and 2B.28.

3C.3. Production of Bacterial Extracts Containing Wild Type and Mutant Polypeptides for Dimerization Studies.

Mutations were tested for their ability to disrupt the dimerization properties of the I9X or I10X polypeptides, using the extraction protocol described in Section 2B.28.

Figure 3C.5 shows SDS-polyacrylamide gel elctrophoresis of samples taken during the extraction of selected mutant and wild type I9X and I10X polypeptides, and illustrates the improved yield and purity achieved with this extraction procedure. Certain mutant polypeptides, for example the mutant 459D/Y, were found to be insoluble under these conditions, and could not be analysed further. As with all other extraction procedures, the solubility of mutant polypeptides was found to be variable. Table 3C.2 summarizes the results obtained with this extraction procedure.

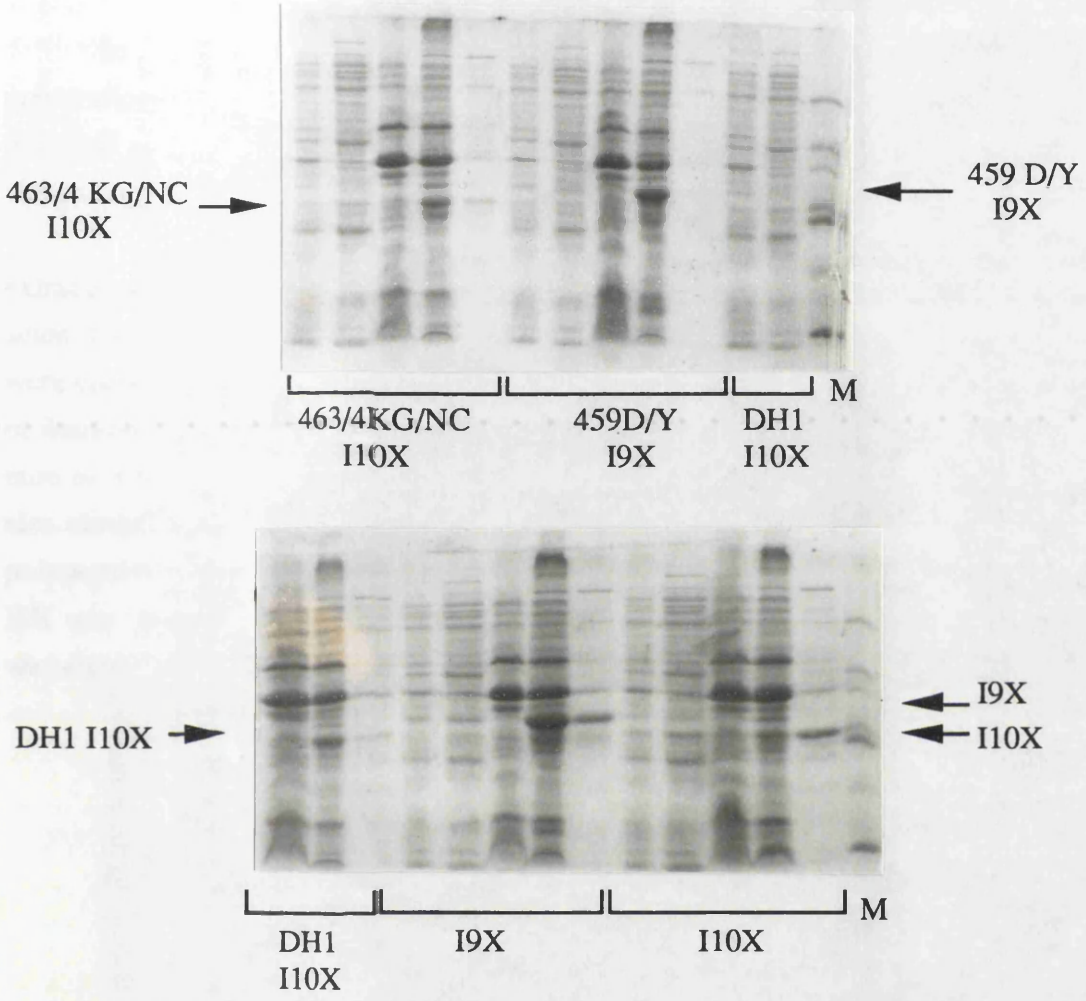


Figure 3C.5. SDS-Polyacrylamide Gel Electrophoresis of Mutant and Wild Type I9X and I10X Polypeptides.

Proteins were extracted from induced cultures as described in Section 2B.28. Samples taken at particular stages during the extraction protocol were electrophoresed on 12.5% SDS-polyacrylamide gels and stained with Coomassie blue.

For each polypeptide shown, the following samples were analysed:
Track 1: Polypeptide profile of induced BL21 (DE3) pLysS bacteria expressing the relevant polypeptide.
Track 2: 1/300 vol. supernatant after centrifugation of cell debris.
Track 3: 1/100 vol. cell debris pellet.
Track 4: 1/2 vol. insoluble protein pellet.
Track 5: 1/30 vol. final extract.

M: Molecular weight standards: 66kDa, 45kDa, 36kDa, 29kDa, 24kDa, 20.1kDa, 14.2kDa.

3C.4. Further Purification of Bacterial Extracts Containing Wild Type and Mutant Polypeptides for Dimerization Studies.

In order to carry out gel filtration chromatography, it was necessary to obtain preparations of various polypeptides at purities of 90% or greater. Bacterial extracts obtained as described in Section 2B.28 were further purified by FPLC ion exchange chromatography, as detailed in Section 2B.29.

Initially, ion exchange chromatography was carried out on bacterial extracts which had been treated with polyethylenimine to remove digested nucleic acids. Under these conditions, it was found that wild type I9X and I10X polypeptides were eluted in sharp peaks at 0.8M and 0.5M NaCl respectively. SDS-PAGE analysis of fractions from this purification showed that peak fractions were approximately 90% pure on Coomassie blue stained gels (see Figure 3C.6). Some I10X polypeptide was also eluted in flow-through fractions, as seen in Figure 3C.6, tracks 1 and 2. Mutant polypeptides were eluted from the MonoS column at various ionic strengths. 456R/L I9X was eluted at 0.8M NaCl, 472R/L I10 was eluted at 0.25M NaCl and 449A/S I10 was eluted at 0.28M NaCl.

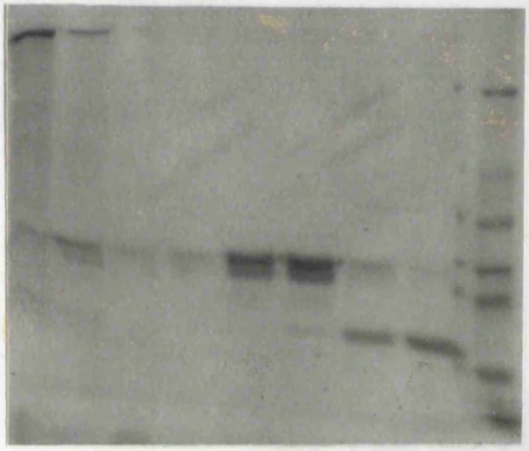


Figure 3C.6. SDS-Polyacrylamide Gel Electrophoresis of Fractions from Ion Exchange Chromatography of the I10X Polypeptide.

I10X protein was extracted from induced BL21 (DE3) pLysS cultures as described in Section 2B.28. Following dialysis, the extract was loaded on to a Pharmacia MonoS ion exchange column and the polypeptide eluted with a salt gradient, as described in Section 2B.29. 2ml fractions were collected and 20ul of each fraction analysed by electrophoresis on a 12.5% SDS- polyacrylamide gel, followed by staining with Coomassie blue.
Tracks 1-2: Flowthrough fractions.
Tracks 3-8: Peak fractions.
Track 9: Molecular weight standards: 66kDa, 45kDa, 36kDa, 29kDa, 24kDa, 20.1kDa, 14.2kDa.

Mutant polypeptide extracts which had not been treated with polyethylenimine during the extraction procedure were eluted from the MonoS column in the flow-through fractions (see Figure 3C.7). Mutant polypeptides eluted in this way were further purified as described in Section 2B.29. Final fractions from this procedure were again approximately 90% pure on Coomassie blue stained gels (see Figure 3C.8).

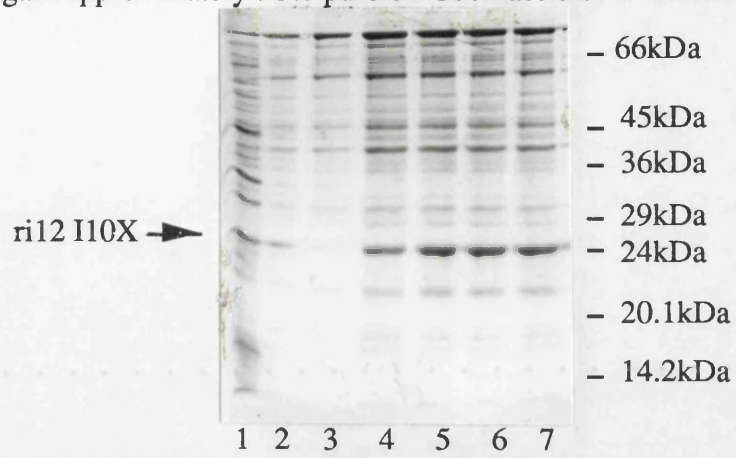


Figure 3C.7. SDS-Polyacrylamide Gel Electrophoresis of fractions from Ion exchange Chromatography of the ri12 I10X Polypeptide.

ri12 I10X protein was extracted from induced BL21 (DE3) pLysS cultures as described in Section 2B.28, omitting treatment with polyethylenimine. Following dialysis, the extract was loaded onto a Pharmacia MonoS ion exchange column, and the protein eluted with a salt gradient, as described in Section 2B.29. 1ml fractions were collected and 20ul of each fraction analysed by electrophoresis on a 12.5% SDS-polyacrylamide gel, followed by staining with Coomassie blue.
Track 1: Input extract.
Tracks 2-7: Flowthrough fractions.

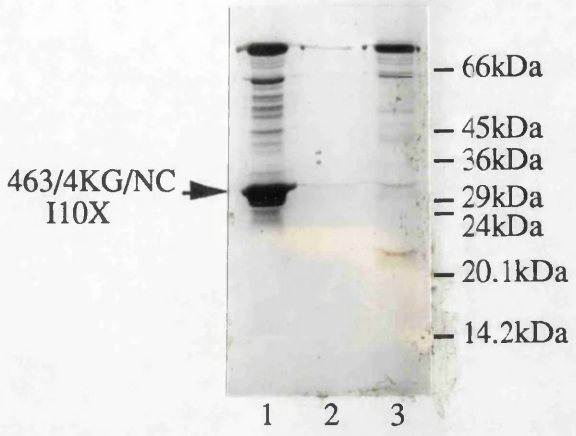


Figure 3C.8. SDS-Polyacrylamide Gel Electrophoresis of a Purification of a Mutant I10X Polypeptide Following Ion Exchange Chromatography

463/4KG/NC I10X polypeptide was extracted from induced BL21 (DE3) pLysS cultures as described in Section 2B.28, omitting treatment with polyethylenimine. Following dialysis, the extract was loaded onto a Pharmacia MonoS ion exchange column and the protein eluted with a salt gradient as described in Section 2B.29. The majority of the protein was eluted in the flow-through fractions. Protein was purified further from these flowthrough fractions as described in Section 2B.29. Samples taken at various stages during this purification protocol were analysed by electrophoresis on 12.5% SDS-polyacrylamide gels, followed by staining with Coomassie blue.
Track 1: Purified fraction following concentration.
Track 2: Purified fraction before concentration.
Track 3: Flowthrough fraction.

3C.5. Discussion.

Expression of wild type and mutant versions of the DNA binding domain of Vmw175 led to many problems, mainly due to the low induction and solubility of the expressed polypeptides. Initially, it was found that induction of bacterial cultures transformed with the pJI9 vector led to low levels of expression of the wild type I9 polypeptide. This could have been due to an inherent instability of the isolated polypeptide. Extraction of the polypeptide from bacterial cells led to a further reduction in yield, as the I9 polypeptide was precipitated both during the centrifugation of cell debris and the centrifugation of insoluble ammonium sulphate precipitated proteins. Loss of I9 polypeptide during centrifugation of cell debris may have been caused by the incomplete disruption of the interactions between the I9 polypeptide and cellular DNA. Loss of I9 polypeptide during both centrifugation steps would also occur due to the low solubility of the expressed protein. As shown in Section 3C.2.2, expression of I10 polypeptides led to an improvement in the final yield of polypeptide, as the I10 protein appeared to be slightly more soluble. It would have been preferable to test the

effects of the mutations on the DNA binding properties of the I9 polypeptide, as the I9 polypeptide had been shown to bind to DNA with a specificity similar to that of the whole Vmw175 protein, and to produce a DNAaseI footprint at the Vmw175 consensus binding site situated at the IE-3 capsite; whereas I10 polypeptide exhibited a slightly more relaxed specificity in binding ability as measured by gel retardation analysis, and did not produce a pronounced DNAaseI footprint at the IE-3 capsite (Pizer *et al.*, 1991; Everett *et al.*, 1991; Everett, unpublished observations). Also, previous studies had shown that the I9 insertion mutation produced no detectable alteration in the transactivation, repression and *in vitro* DNA binding properties of Vmw175 (as measured by transient transfection and gel retardation assays); whereas the I10 mutation led to slight reductions in these properties, suggesting that the site of the I10 mutation may be within a region important for the conformational stability of the whole protein and for the isolated DNA binding domain (Paterson, 1989). However, due to the improved yield of I10 polypeptide, it was thought necessary to express the mutations within I10 instead of I9 polypeptides, in order to obtain sufficient polypeptide concentrations for gel retardation analysis.

Ongoing studies showed that gel retardation analysis with I9 and I10 polypeptides produced smeared complexes which were not easily resolvable, presumably due to degradation of the polypeptides. Removal of the 20 C-terminal amino acids by insertion of a stop codon within the BamHI restriction site at codon 523 of the IE-3 coding sequence of the pJI9 and pJI10 vectors, to yield I9X and I10X polypeptides, resulted in a clearly resolved protein-DNA complex upon gel retardation electrophoresis (see section 3D for gel retardation analysis). Therefore, mutations were expressed within I9X or I10X polypeptides.

It was observed consistently throughout these extraction procedures that the expression and purification of all mutant polypeptides were lower than that of the wild type polypeptide. This observation correlates closely with results obtained in transfection experiments with plasmids containing insertional mutations within the Vmw175 DNA binding domain. In the case of the insertional mutants I14-I18, situated at IE-3 codons 329, 337, 373, 398 and 438 respectively, it was found that cells transfected with versions of p175 containing these mutations produced lower quantities of mutant Vmw175 protein compared to transfections with wild type p175. It was suggested that this low yield of mutant protein may again have been caused by lower protein solubility or reduced stability of expressed proteins (Paterson, 1989). These solubility and stability problems led to a wide variation in polypeptide content of bacterial extracts containing mutant versions of DNA binding domain polypeptides.

Tables 3C.1 and 3C.2 summarize the methods employed to produce bacterial extracts containing mutant and wild type DNA binding domain polypeptides for DNA binding and dimerization assays. These tables also illustrate the relative

solubility of mutant DNA binding domain polypeptides as compared to wild type polypeptides, and demonstrate the variable solubility of the mutant proteins. These differences in solubility cannot be attributed to specific sequence changes, as it was found that the solubility of a mutant polypeptide as measured by the final yield of the protein in the bacterial extract varied between subsequent inductions, as well as with different expression systems and extraction procedures. However, certain mutations drastically decreased the solubility of the polypeptides consistently, such that the properties of such mutant proteins could not be determined. This is true of the mutants 469S/I and i12. As each mutant polypeptide was compared to an internal control wild type polypeptide for the determination of qualitative solubility, these tables do not illustrate the differences in yield observed with different extraction procedures. The main advantages obtained from the adaptation of extraction protocols described in Sections 3C.2 and 3C.3 were an improvement in the purity and stability of the final polypeptide extracts rather than a quantitative increase in the yield of certain polypeptides.

The adaptation of the polypeptide extraction procedure described in Section 3C.3 greatly improved the purity of I9X and I10X polypeptides. The one-step purification of polypeptides upon ion exchange chromatography yielded peak fractions of sufficient purity for further analysis by glutaraldehyde cross-linking and gel filtration chromatography. The loss of some wild type and mutant polypeptides in the flow-through during such purification procedures was presumably due to over-saturation of the column's binding capacity. In contrast, when nucleic acids had not been removed from polypeptide extracts by precipitation with polyethylenimine, the majority of the polypeptide was eluted in the flow-through. This change in behaviour of the polypeptides during ion exchange chromatography was presumably due to the presence of nucleic acids in the bacterial extracts which interfere with the binding of the polypeptides to the ion exchange matrix. The observation that the wild type I9X polypeptide was eluted from a MonoS column at a much higher ionic strength than previously noted for the I10X polypeptide (around 0.8M NaCl for I9X, 0.29M NaCl for I10X) implies that the I9X polypeptide binds more stongly to the negative ionic matrix of the MonoS column. It is interesting to note that the region removed from the I10X polypeptide which is present in the I9X polypeptide (codons 253-275) contains a stretch of basic amino acids which may contribute to its strong interaction with the negative ionic matrix of the MonoS column.

3D. Determination of the *In Vitro* DNA Binding Activity of Wild Type and Mutant Polypeptides Encompassing the DNA Binding Domain of Vmw175.

One approach used to investigate the function of the "helix-turn-helix" motif of the DNA binding domain of Vmw175 involved the analysis of the DNA binding activities of the mutant polypeptides whose construction, expression and extraction has been described in Sections 3B and 3C. In order to determine the DNA binding activities of these polypeptides, gel retardation analysis was carried out using a DNA probe which encoded the Vmw175 consensus DNA binding site situated at the capsite of its own IE-3 gene.

In addition to this analysis, the DNA binding activity of another mutant polypeptide, designated ri12, was determined. The ri12 mutation, described in Section 1C.3, is situated towards the N-terminus of the DNA binding domain of Vmw175, outwith the limits of the "helix-turn-helix" motif.

These assays would be useful in determining the contribution, if any, of these regions of the protein towards the *in vitro* DNA binding properties of the isolated DNA binding domain of Vmw175.

3D.1. Western Blot Analysis of Bacterial Extracts.

In order to ensure that equivalent amounts of DNA binding domain polypeptides were assayed for DNA binding activity, the Vmw175 polypeptide content of bacterial extracts was quantified by Western blot analysis. The polypeptides were detected with a polyclonal anti-Vmw175 antibody, as described in Section 2B.6.3.

Initially, Western analysis was carried out on expressed I9 polypeptides containing various mutations. It was found that the mutation 452del led to the expression of a truncated polypeptide of 22kDa molecular weight, which is close to the predicted molecular weight of the I9 polypeptide truncated at codon 454 (Figure 3D.1, track 2) Additional polypeptides of higher molecular weight than the I9 polypeptide were also detected (see Figure 3D.1). Analysis of uninduced BL21(DE3)pLysS extracts showed that these bands were due to interaction of the anti-Vmw175 antibody with bacterial proteins. The I9 protein was found to be unstable; degradation products were observed in extracts even without previous incubations at room temperature or 37°C

(see Figure 3D.1, track 9). The degradation products were presumably due to proteolysis of the protein during the extraction procedure.

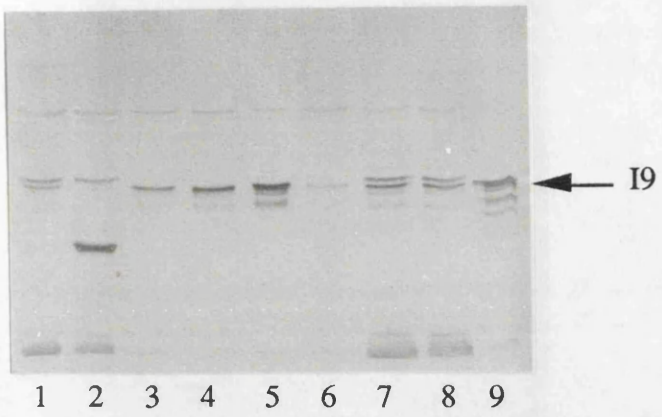


Figure 3D.1. Western Blot Analysis of Mutant and Wild Type I9 Polypeptides.

Polypeptides were extracted from induced BL21 (DE3) pLysS cultures as described in Section 2B.27.1. Extracts were electrophoresed on 12.5% SDS-polyacrylamide gels, proteins transferred onto nitrocellulose and immunoblotted as described in Section 2B.6.3. Proteins were immunoblotted with a polyclonal rabbit antibody raised to a bacterial fusion protein containing the Vmw175 DNA binding domain, used at a 1:1000 dilution.

Track 1: 450A/S I9 *Track 2: 452del I9*
Track 3: 449A/S I9 *Track 4: 453A/S I9*
Track 5: 446A/S I9 *Track 6: 451/3AA/SS I9*
Track 7: 451A/S I9 *Track 8: 451A/S I9*
Track 9: I9

Mutations were also assayed for their ability to interfere with the DNA binding ability of I9X and I10X polypeptides. Again, Western blot analysis was used to determine the volume of bacterial extracts containing equivalent quantites of Vmw175 polypeptide. The I10X extracts were frequently observed to contain a degradation product (see Figure 3D.2) and mutant and wild type I9X polypeptides still showed evidence of proteolysis (see Figure 3D.3). Several mutant I9X polypeptides yielded large quantities of a degradation product approximately 1000-2000Da smaller than the I9X polypeptide (Figure 3D.3, compare tracks 3, 9, 11 and 12). Interestingly, wild type I9X extracts did not appear to be contaminated with this particular degradation product.

Certain mutant polypeptides, such as the mutant 469S/I I10X, 453/5AS/SI I9X and i12 I10X were consistently expressed to very low levels (see Figure 3D.2, track 1 and Figure 3D.3, tracks 2 and 8). The low yield of polypeptides may have been due to their insolubility during the extraction procedure or instability during expression, as discussed in Section 3C.5. Due to these problems, several mutants could not be analysed further.

The amounts of polypeptide in each extract were roughly quantitated by densitometric analysis of Western blots, and equivalent amounts analysed in gel retardation experiments (see Section 3D.4).

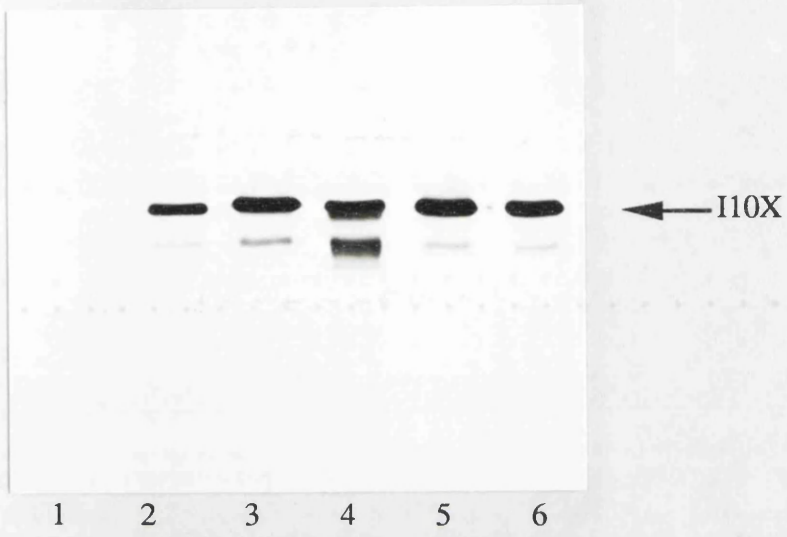


Figure 3D.2. Western Blot Analysis of Mutant and Wild Type I10X Polypeptides.

Polypeptides were extracted from induced BL21 (DE3) pLysS cultures as described in Section 2B.27.1. Extracts were electrophoresed on 12.5% SDS-polyacrylamide gels, proteins transferred onto nitrocellulose and immunoblottedas described in Section 2B.6.3. Proteins were immunoblotted with a polyclonal rabbit antibody raised to a bacterial fusion protein containing the Vmw175 DNA binding domain, used at a 1:100 dilution.

Track 1: 469S/I I10X *Track 2: 471R/L I10X*
Track 3: 451/3AA/SS I10X *Track 4: 451A/S I10X*
Track 5: 472R/L I10X *Track 6: I10X*

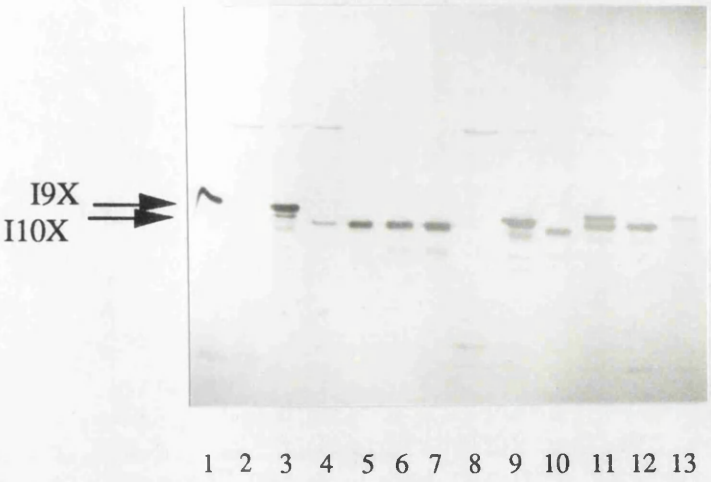


Figure 3D.3. Western Blot Analysis of Mutant and Wild Type I9X and I10X Polypeptides.

Polypeptides were extracted from induced BL21 (DE3) pLysS cultures as described in Section 2B.27.1. Extracts were electrophoresed on 12.5% SDS-polyacrylamide gels, proteins transferred onto nitrocellulose and immunoblotted as described in Section 2B.6.3. Proteins were immunoblotted with a polyclonal rabbit antibody raised to a bacterial fusion protein containing the Vmw175 DNA binding domain, used at a 1:1000 dilution.

Track 1: 457R/L I9X
Track 2: 469S/I I9X
Track 3: 451A/S I9X
Track 4: DH1 I10X
Track 5: 461A/S I10X
Track 6: 464G/C I10X
Track 7: 462Q/H I10X
Track 8: i12 I10X
Track 9: 453A/S I9X
Track 10: 463/4KG/NC I10X
Track 11: 459D/Y I9X
Track 12: 456R/L I9X
Track 13: I9X

3D.2. Detection of Further Degradation of I9X and I10X Polypeptides Upon Incubation at Room Temperature or 37°C.

Due to the degradation of polypeptides observed in previous Western blot analyses, experiments were carried out to determine the stability of expressed polypeptides at room temperature and 37°C, in order to ensure that any loss in binding activity observed upon gel retardation analysis was not due to protein degradation. Bacterial extracts were incubated at room temperature or 37°C for 20min in gel retardation buffer. SDS boiling mix was added and polypeptides detected as described in Section 2B.6.3.

Analysis of the mutant polypeptide ri12 I10X in this way showed that incubation of the polypeptide at room temperature or 37°C did not lead to further degradation than that seen in the original extract (see Figure 3D.4). Analysis of other mutant polypeptides resulted in similar results, and so it was concluded that any loss of

DNA binding activity observed was not due to proteolysis during the incubation of protein with DNA.

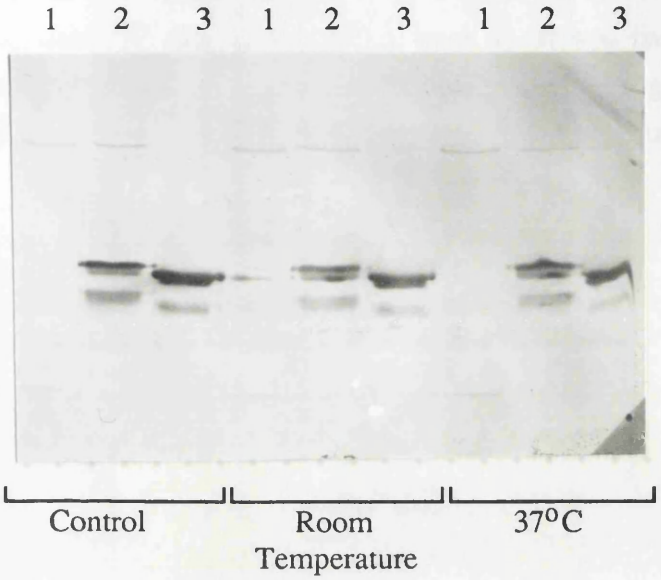


Figure 3D.4. Western Blot Analysis of i12 and ri12 I10X Polypeptides.

i12 and ri12 I10X polypeptides were extracted from induced BL21 (DE3) pLysS cultures as described in Section 2B.27.1. The extracts were incubated at room temperature or 37°C for 20 minutes in gel retardation buffer and then electrophoresed on 12.5% SDS-polyacrylamide gels. Proteins were transferred onto nitrocellulose and immunoblotted as described in Section 2B.6.3. Proteins were immunoblotted with a polyclonal rabbit antibody raised to a bacterial fusion protein containing the Vmw175 DNA binding domain, used at a 1:1000 dilution.
At each temperature, the following extracts were analysed:
Track 1: i12 I10X
Track 2: ri12 I10X
Track 3: I10X.

3D.3. Wild Type DNA Binding Domain Polypeptides Bind to a Consensus Vmw175 Binding Site.

Gel retardation analysis was carried out on extracts from non-induced BL21(DE3)pLysS cultures and induced cultures of BL21(DE3)pLysS which expressed the wild type I10X polypeptide. Proteins were tested for their ability to bind to 45bp AvaI-BamHI or 130bp EcoRI-BamHI DNA restriction fragments, both of which span the IE-3 capsite. Extracts were prepared as described in Section 3C, and equivalent amounts of polypeptide incubated for 20min at room temperature with the DNA probe as described in Section 2B.8.

Uninduced bacterial extracts produced a high mobility complex with the IE-3 capsite probe, designated "bac" (see Figure 3D.5, track 2). The presence of the I10X polypeptide led to the formation of a specific, lower mobility I10X-DNA complex designated "P" (see Figure 3D.5, track 1). It was frequently observed that bacterial extracts containing induced Vmw175 DNA binding domain polypeptides produced both "bac" and "P" complexes upon gel retardation analysis.

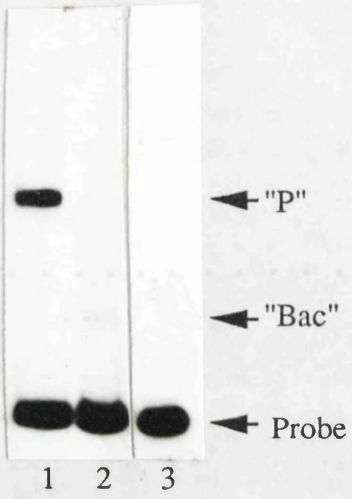


Figure 3D.5. Gel Retardation Analysis of the Wild Type I10X Polypeptide.

Bacterial extracts were incubated with a radiolabelled probe spanning the IE-3 capsite for 20 minutes at room temperature. The conditions of the binding reaction are described in Section 2B.8. Free and bound probe were separated by electrophoresis on a non-denaturing 4% polyacrylamide gel. Complexes were visualized by autoradiography.
Track 1: Bacterial extract containing I10X polypeptide.
Track 2: Bacterial extract containing no induced polypeptide.
Track 3: Free probe.
"P": Complex formed between I10X polypeptide and probe.
"Bac": Complex formed between bacterial protein(s) and probe.

Incubation of the DNA probe with high concentrations of the I10X polypeptide led to the formation of a low mobility complex which migrated only a short distance into the gel. Titration of the I10X extract led to the resolution of a ladder of specific I10X-DNA complexes, presumably formed due to multiple protein-DNA interactions (see Figure 3D.6). At higher dilutions, the formation of the specific primary complex "P" was again observed (see Figure 3D.6, tracks 6 and 7).

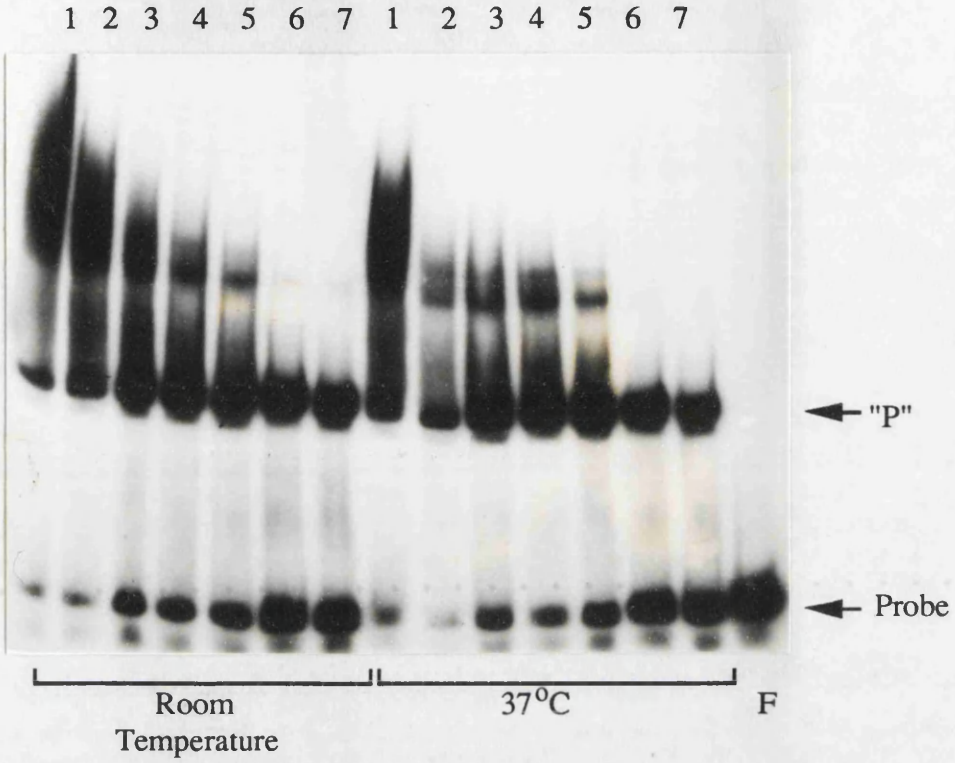


Figure 3D.6. Gel Retardation Analysis of the Wild Type I10X Polypeptide.

Bacterial extracts were incubated with a radiolabelled probe spanning the IE-3 capsite for 20 minutes at room temperature or 37°C. The conditions of the binding reaction are described in Section 2B.8. Free and bound probe were separated by electrophoresis on a non-denaturing 4% acrylamide gel. Complexes were visualized by autoradiography.
At each temperature, the same dilutions of the bacterial extract were analysed:
Track 1: undiluted extract. *Track 2: 1:2 dilution.*
Track 3: 1:5 dilution *Track 4: 1:10 dilution.*
Track 5: 1:15 dilution. *Track 6: 1:20 dilution*
Track 7: 1:30 dilution. *F: Free probe*
"P": Complex formed between I10X polypeptide and the probe.

Data presented in Figure 3D.12(b) repeats data presented in Figure 3D.6.

3D.4. Gel Retardation Analysis of Truncated and Wild Type DNA Binding Domain Polypeptides.

The mutation 452del was transferred to the pJI9 vector for analysis of its DNA binding properties. Extracts containing the mutant and wild type versions of the I9 polypeptide were subjected to Western blot analysis in order to roughly determine their polypeptide concentration (shown in Figure 3D.1). Equivalent amounts of mutant

and wild type polypeptide were tested for their ability to bind to a consensus Vmw175 DNA binding site. The incubation with DNA was carried out at 0°C for 20min. as described in Section 2B.8.

It was found that the I9 polypeptide containing the 452del mutation produced a truncated polypeptide as expected (see Figure 3D.1, track 2), and showed no DNA binding activity at 0°C (Figure 3D.7, track 1). The I9 polypeptide at similar concentrations produced a clearly resolvable DNA-protein complex (Figure 3D.7, track 2). The faint, higher mobility band appears to be a secondary complex formed between I9 and DNA; it was produced with high concentrations of I9 polypeptide, whereas high concentrations of bacterial extracts containing no induced protein produced only a higher mobility complex ("bac", see Figure 3D.1).

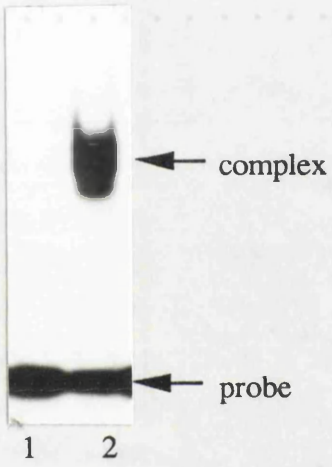


Figure 3D.7. Gel retardation Analysis of a Wild Type and a Truncated I9 Polypeptide.

Bacterial extracts were incubated with a radiolabelled probe spanning the IE-3 capsite for 20 minutes at 0 °C. The conditions of the binding reaction are described in Section 2B.8. Free and bound probe were separated by electrophoresis on a non-denaturing 4% polyacrylamide gel. Complexes were visualized by autoradiography.
Track 1: 452del I9.
Track 2: I9.

3D.5. Gel Retardation Analysis of DNA Binding Domain Polypeptides Containing Point Amino Acid Mutations.

3D.5.1. Certain Point Amino Acid Mutations Result In a Temperature Sensitive DNA Binding Phenotype.

Mutant I9X and I10X polypeptides were quantified by Western blot analysis, and equal amounts of the polypeptides assayed for DNA binding activity.

Incubation of the protein with DNA was carried out at room temperature or 37°C for 20min as described in Section 2B.8.

Incubation of wild type I10X polypeptide with DNA at 37°C had no effect on the formation of the "P" complex. Some extracts produced a doublet of complexes upon gel retardation analysis, designated "P" and "P' ". The "P' " complex appeared with both mutant and wild type polypeptides and was found to be due to the presence of a higher concentration of a degradation product of the I10X polypeptide, which is obviously able to interact with DNA.

Certain point mutant polypeptides showed a DNA binding activity indistinguishable from wild type polypeptides; that is, they produced similar amounts of complex both at room temperature and at 37°C (see Figure 3D.8, compare tracks 1, 3 and 5 at both temperatures). Although it was observed in several instances that point mutant polypeptides exhibited slightly lower DNA binding activity at room temperature than the wild type polypeptide, this reduction was variable between different extracts of the same polypeptide, and therefore difficult to quantitate.

Analysis of the binding activity of polypeptides at room temperature and 37°C showed that certain mutant polypeptides exhibited greatly reduced binding activity at 37°C; that is they exhibited a temperature sensitive DNA binding phenotype (Figure 3D.8, compare tracks 2 and 4 at both temperatures). A summary of the DNA binding activities of the mutant polypeptides is given in Table 3D.1. Interestingly, the single amino acid mutation *ts*K, which results in a temperature-sensitive phenotype when present within the viral IE-3 open reading frame (Preston, 1979a;Davison, 1984) also results in a temperature-sensitive DNA binding phenotype when expressed within a DNA binding domain polypeptide (see Table 3D.1).

Table 3D.1. DNA Binding Properties of Mutant Polypeptides.

Construct (a)	Site of G to T mutation (b)	Site of amino acid alteration (c)	Type of alteration (d)	Name (e)	DNA binding activity (f)
LA15 no. 2	1351	451	Ala to Ser	451A/S	WT
LA15 no. 3	1351 1357	451 453	Ala to Ser Ala to Ser	451/3AA/SS	TS
LA15 no. 5	1345	449	Ala to Ser	449A/S	TS
LA15 no. 9	1336	446	Ala to Ser	446A/S	WT
LA15 no. 36	1357	453	Ala to Ser	453A/S	TS
LA15 no. 74	1348	450	Ala to Ser	450A/S	WT
LA16 no. 16	1357	453	Ala to Ser	453A/S	TS
LA16 no. 39	1370	457	Arg to Leu	457R/L	TS
LA16 no. 48	1375	459	Asp to Tyr	459D/Y	no binding
LA16 no. 49	1367	456	Arg to Leu	456R/L	no binding
LA17 no. 6	1381	461	Ala to Ser	461A/S	WT
LA17 no. 29	1390	464	Gly to Cys	464G/C	WT
LA17 no. 31	1386	462	Gln to His	462Q/H	WT
LA17 no. 35	1389 1390	463 464	Lys to Asn Gly to Cys	463/4KG/NC	no binding
LA18 no. 1	1415	472	Arg to Leu	472R/L	TS
LA18 no. 14	1412	471	Arg to Leu	471R/L	WT
LA18 no. 19	1406	469	Ser to Ile	469S/I	N/A
LA18 no. 26	1423	475	Ala to Ser	475A/S	TS
LA15 no. 57	1355(del)	452	stop codon at 454	452del	no binding
LA21 no. 5	1334-1356	445-453	del	DH1	no binding
LA22 no. 9	1409-1441	470-481	del	DH2	no binding
*ts*K	1424 (C to T)	475	Ala to Val	*ts*K	TS

Legend to Table 3D.1.

(a) The original name of the mutant clone, showing the oligonucleotide used for the mutagenesis reaction and the clone number isolated.

(b) The site of G to T mutation is shown as nucleotide number within the IE-3 reading frame, numbered from +1 at the ATG start codon. In the case of the mutant 452del, the base pair mutation introduced was not a G to T mutation, but a base pair deletion.

(c) The site of amino acid alteration is shown as codon number within the IE-3 open reading frame, numbered from 1 at the ATG start codon.

(d) The type of amino acid alteration introduced by the G to T mutation. Amino acids are represented as three letter codes.

(e) The final construct name, showing the position of mutation as codon number, followed by the amino acid alteration. Amino acids are represented as one letter codes.

(f) The DNA binding phenotype of a DNA binding domain polypeptide containing the mutation indicated. DNA binding activity was measured by gel retardation analysis, using a DNA probe containing the consensus DNA binding site situated at the IE-3 capsite. WT = DNA binding activity indistinguishable from a wild type DNA binding domain polypeptide. TS = DNA binding activity reduced at temperatures of 37°C and above. No binding = no binding activity at any temperature.

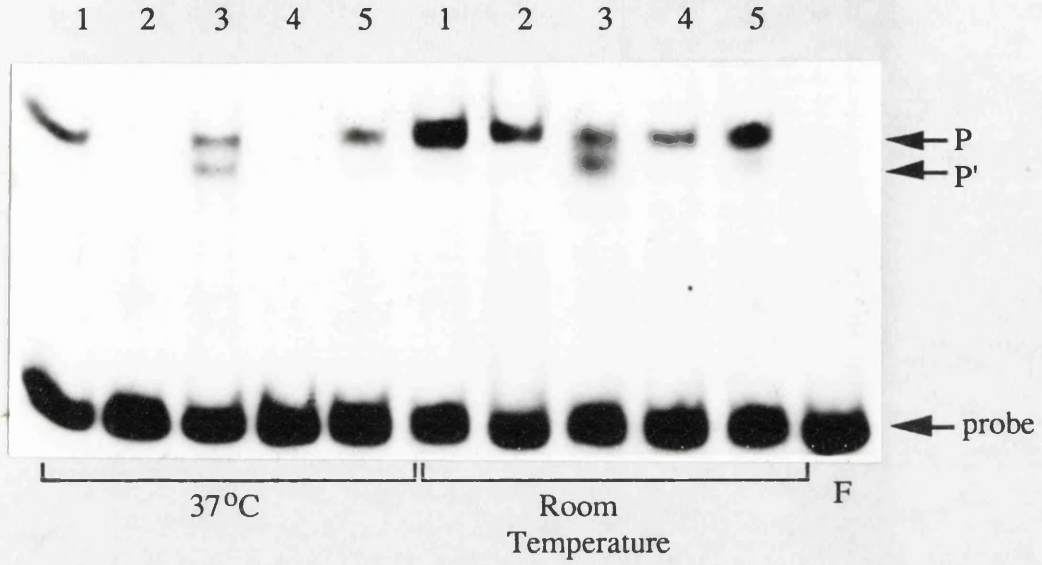


Figure 3D.8. Gel retardation Analysis of Wild Type and Mutant I10X Polypeptides.

Bacterial extracts were incubated with a radiolabelled probe spanning the IE-3 capsite for 20 minutes at room temperature or 37°C. The conditions of the binding reaction are described in Section 2B.8. Free and bound probe were separated by electrophoresis on a non-denaturing 4% acrylamide gel. Complexes were visualized by autoradiography.
At each temperature, the following polypeptides were analysed:
Track 1: I10X *Track 2: 472R/L I10X*
Track 3: 451A/S I10X *Track 4: 451/3AA/SS I10X`*
Track 5: 471R/L I10X *F: Free probe.*

3D.5.2. Certain Point Amino Acid Mutations Result in the Loss of DNA Binding Activity.

Gel retardation analysis of more point amino acid mutations within I9X or I10X polypeptides revealed that certain single and double amino acid mutations resulted in a loss of DNA binding activity.

Polypeptide extracts were analysed by Western blot assays, and equal amounts of polypeptide tested for their ability to bind to a consensus Vmw175 DNA binding site, as described previously. Western analysis showed that certain mutant I9X polypeptides produced large amounts of a degradation product 1000-2000Da smaller than the I9X polypeptide (Figure 3D.3, tracks 9, 11 and 12). In the case of the mutant 453A/S, it was found that the presence of this degradation product did not interfere with the DNA binding activity of the polypeptide (see Figure 3D.9, track 1). This

mutant exhibited a temperature sensitive DNA binding phenotype, as described in Section 3D.5.1. (Figure 3D.9, compare track 1 at both temperatures).

Other mutant I9X and I10X polypeptides showed no DNA binding activity at either temperature (see Figure 3D.9, tracks 2, 3 and 4 at both temperatures). This loss of DNA binding activity was not due to a reduction in the amount of polypeptide in the extract; the exact volumes of extracts used in this assay had been shown to contain roughly equivalent amounts of polypeptide (Figure 3D.3). Also, the loss of DNA binding activity was not due to degradation of the I9X polypeptide - one DNA binding negative mutant 463/4KG/NC was expressed within an I10X polypeptide which showed limited evidence of proteolysis (see Figure 3D.3, track 10). Western blot analysis of mutant polypeptides after incubation in gel retardation buffer had shown that no further proteolysis occurred during the incubation at room temperature or 37°C (described in Section 3D.3). Therefore, this loss of DNA binding activity is a true reflection of the effect of the mutation on the isolated DNA binding domain. A summary of the mutations resulting in an elimination of DNA binding activity is shown in Table 3D.1.

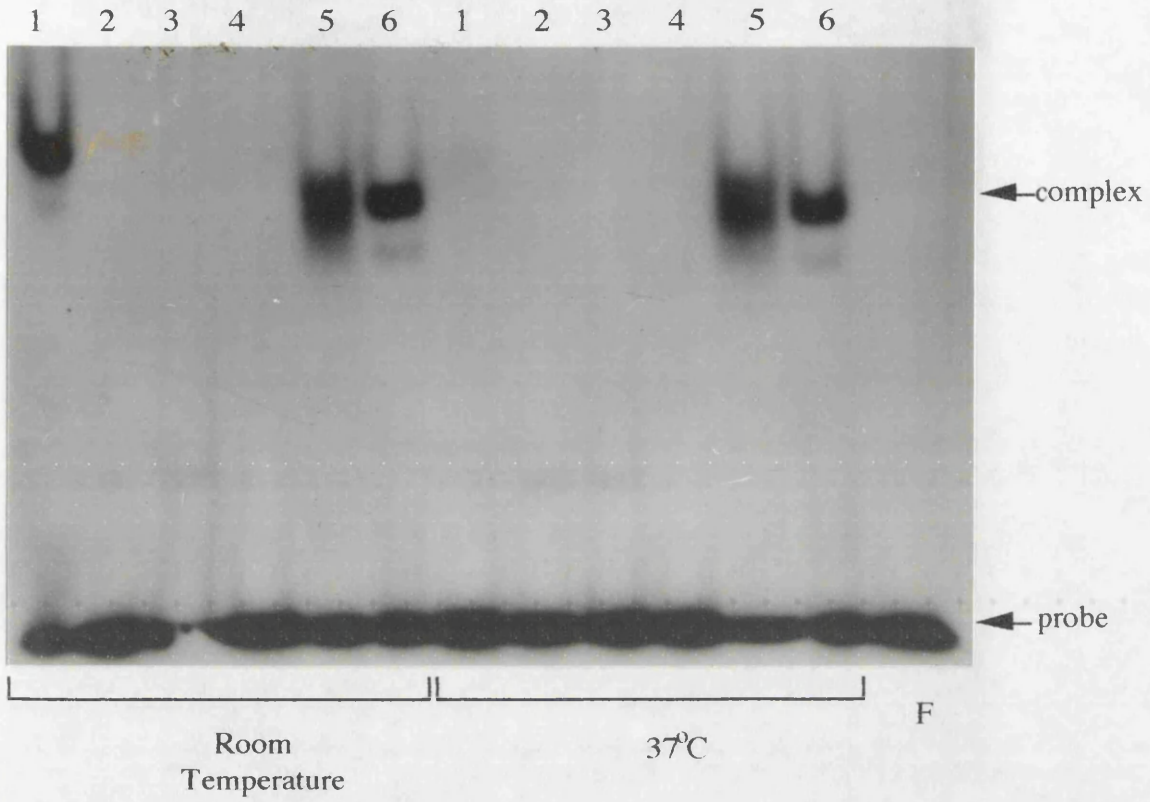


Figure 3D.9. Gel Retardation Analysis of Wild Type and Mutant I9X and I10X Polypeptides.

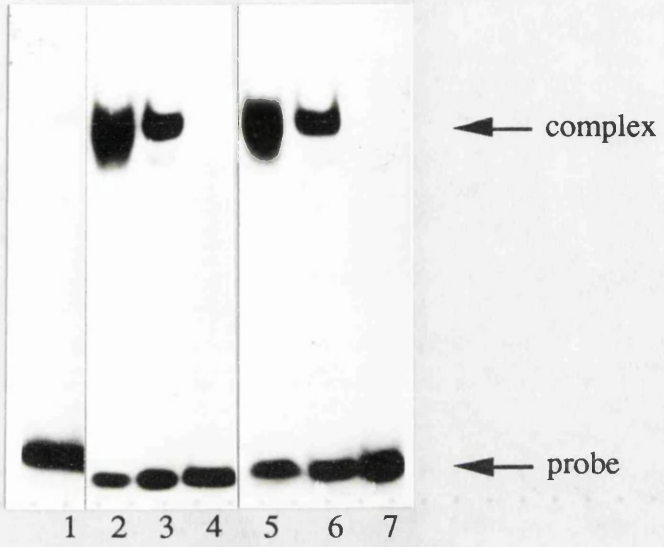
Bacterial extracts were incubated with a radiolabelled probe spanning the IE-3 capsite for 20 minutes at room temperature or 37°C. The conditions of the binding reaction are described in Section 2B.8. Free and bound probe were separated by electrophoresis on non-denaturing 4% acrylamide gels. Complexes were visualized by autoradiography.
At each temperature, the following polypeptides were assayed:
Track 1: 453A/S I9X *Track 2: 463/4KG/NC I10X*
Track 3: 459D/Y I9X *Track 4: 456R/L I9X*
Track 5: I9X *Track 6: I10X*
F: Free probe.

3D.6. Gel Retardation Analysis of DNA Binding Domain Polypeptides Containing Internal In-Frame Deletions.

Internal, in-frame deletions constructed as described in Section 3B.3 were expressed within I10X polypeptides, and bacterial extracts prepared as described in Section 2B.27.2. The concentration of mutant and wild type polypeptide in the bacterial extracts was quantified by Western blot analysis, and equivalent amounts of polypeptide assayed for their ability to bind to a consensus Vmw175 DNA binding site, as described above.

Removal of seven amino acids from the first proposed helix, or ten amino acids from the second proposed helix resulted in the loss of DNA binding activity of the I10X polypeptide at both room temperature and 37°C (Figure 3D.10). At both temperatures, the mutant polypeptides failed to form a specific "P" complex with DNA. The higher mobility complex observed upon incubation of the DH2 I10X extract with DNA was formed by the interaction of bacterial DNA binding proteins with DNA. (Figure 3D.10b, tracks 3, 4 and 5).

(a)



(b)

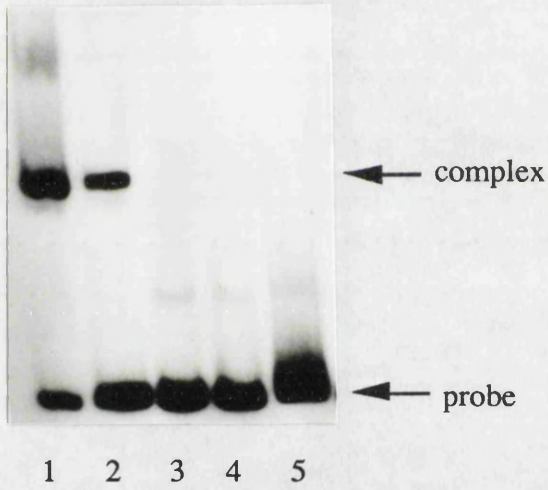


Figure 3D.10. Gel retardation Analysis of Wild Type and Deletion Mutant I9X and I10X Polypeptides.

Bacterial extracts were incubated with a radiolabelled probe spanning the IE-3 capsite for 20 minutes at room temperature or 37°C. The conditions of the binding reaction are described in Section 2B.8. Free and bound probe were separated by electrophoresis on non-denaturing 4% acrylamide gels. Complexes were visualized by autoradiography.

(a): Track 1: DH1 I10X room temperature. Track 2: I9X room temperature
Track 3: I10X room temperature. Track 4: DH1 I10X 37°C
Track 5: I9X 37°C Track 6: I10X 37°C
Track 7: Free probe.

(b): Track 1: I10X room temperature Track 2: I10X 37°C
Track 3: DH2 I10X room temperature Track 4: DH2 I10X 37°C
Track 5: Bacterial extract with no induced protein.

3D.7. Gel Retardation Analysis of DNA Binding Domain Polypeptides Containing the i12 and ri12 Mutations.

The i12 and ri12 mutations, described in Section 1C.3 are mutations within the N-terminal portion of the DNA binding domain of Vmw175.

BamHI "y" restriction fragment DNA from the virus vri12 was kindly provided by Dr. N. A. DeLuca (University of Pittsburgh School of Medicine). Dideoxy sequence analysis of this fragment revealed the presence of the i12 four amino acid insertion mutation at codon 320 of the IE-3 reading frame, and a second C to T mutation at nucleotide 1034 resulting in an alanine to valine substitution at codon 325. This observation differs from that made by Shepard and DeLuca (1991b), who stated that the ri12 molecule contained an alanine to valine substitution at codon 345, and no substitution at codon 325. There appears to be no explanation for this discrepancy. The i12 and ri12 mutations were cloned on NruI/BamHI restriction fragments inserted into the pJI10 expression vector, and X versions of these plasmids were constructed as described in Section 3C.1. The mutant polypeptides were expressed from these vectors, and bacterial extracts containing the polypeptides prepared as decribed in Section 2B.27

As illustrated in Figure 3D.3, the i12 I10X polypeptide was found to be expressed at very low levels, probably due to the low stability and solubility of the mutant polypeptide. However, the ri12 I10X polypeptide was expressed successfully to concentrations sufficient for gel retardation analysis, and Western blot analysis of the ri12 I10X polypeptide showed it to be stable upon incubation at room temperature and 37°C (see Figure 3D.4). Extracts containing ri12 I10X and wild type I10X polypeptides were quantified using Western blot analysis and equivalent amounts of polypeptide assayed for their ability to bind to a consensus Vmw175 DNA binding site, as described in Section 2B.8.

Initially the binding activity of both polypeptides was tested in the presence of varying amounts of non-specific DNA competitor, as it had been noted that Vmw175 molecules containing the ri12 mutations had been able to form complexes with the IE-3 capsite probe only in the presence of low concentrations of non-specific competitor (Shepard and DeLuca, 1991b). The ri12 I10X polypeptide formed specific "P" complexes with the IE-3 capsite probe in the presence of up to 5µg of polydI:polydC competitor (Figure 3D.11). At lower concentrations of polydI:polydC, it was found that both ri12 I10X and wild type I10X polypeptides formed low mobility complexes, presumably due to additional and perhaps non-specific protein-DNA interactions.

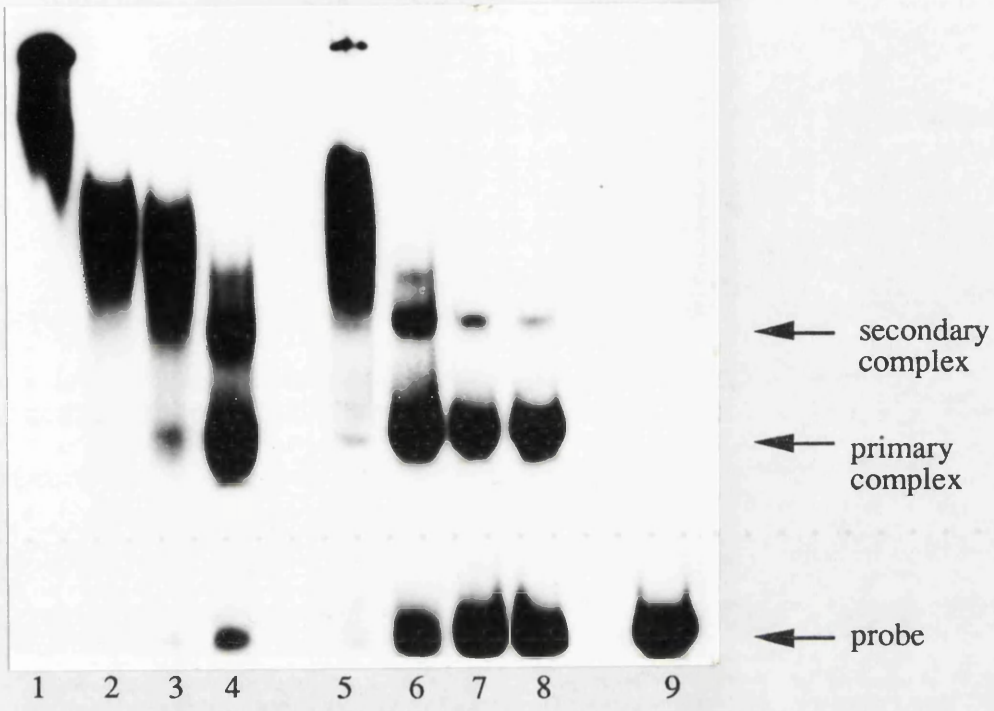


Figure 3D.11. Gel Retardation Analysis of Wild Type and ri12 I10X Polypeptides.

Bacterial extracts were incubated with a radiolabelled probe spanning the IE-3 caps for 20 minutes at room temperature with varying amounts of dIdC.The conditions of the binding reaction are described in Section 2B.8. Free and bound probe were separated by electrophoresis on a non-denaturing 4% polyacrylamide gel. Complexe were visualized by autoradiography.

Track 1: I10X with 0ug dIdC *Track 2: I10X with 1ug dIdC*
Track 3: I10X with 2ug dIdC *Track 4: I10X with 5ug dIdC*
Track 5: ri12 I10X with 0ug dIdC *Track 6: ri12 I10X with 1ug dIdC*
Track 7: ri12 I10X with 2ug dIdC *Track 8: ri12 I10X with 5ug dIdC*
Track 9: Free probe.

Analysis of the DNA binding properties of the ri12 I10X polypeptide at room temperature and 37°C in the presence of non-specific competitor DNA showed that the mutation resulted in a temperature sensitive DNA binding phenotype when expressed within the isolated DNA binding domain of Vmw175. When increasing concentrations of the polypeptide were incubated with an IE-3 capsite probe at room temperature, a ladder of primary and secondary complexes were seen upon gel retardation analysis. The number and amount of complexes formed were very similar to those formed with roughly equivalent concentrations of wild type I10X polypeptide. When the incubation temperature was raised to 37°C, the I10X polypeptide still produced the specific primary and secondary complexes with the IE-3 capsite probe, whereas the ri12 I10X polypeptide exhibited no specific DNA binding activity (Figure 3D.12).

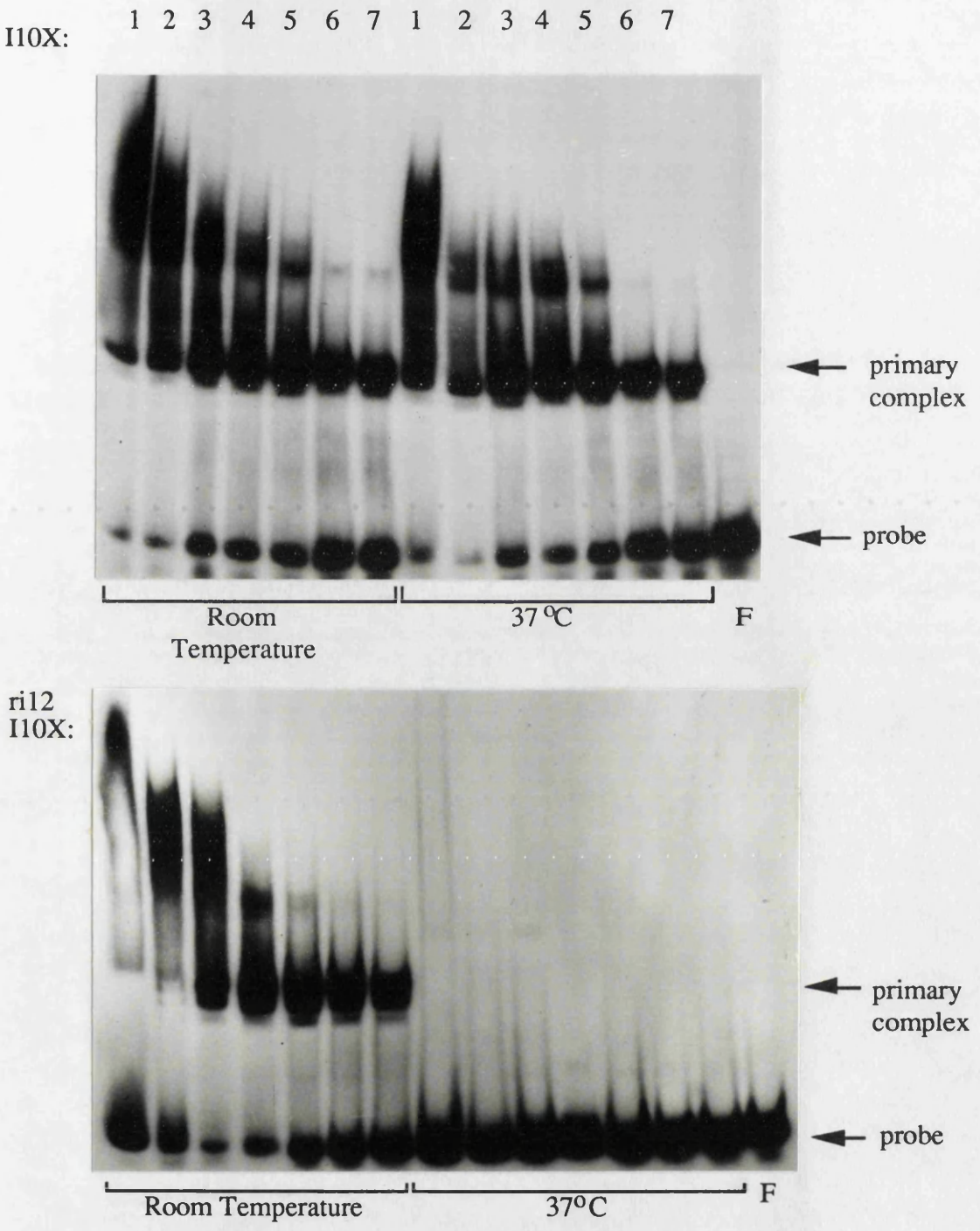


Figure 3D.12. Gel Retardation Analysis of Wild Type and ri12 I10X Polypeptides.

Bacterial extracts were incubated with a radiolabelled probe spanning the IE-3 capsite for 20 minutes at room temperature or 37°C. The conditions of the binding reaction are described in Section 2B.8. Free and bound probe were separated by electrophoresis on non-denaturing 4% acrylamide gels. Complexes were visualized by autoradiography.

At each temperature, the same dilutions of the bacterial extracts were assayed:

Track 1: undiluted extract — *Track 2: 1:2 dilution*
Track 3: 1:5 dilution — *Track 4: 1:10 dilution*
Track 5: 1:15 dilution — *Track 6: 1:20 dilution*
Track 7: 1:30 dilution — *F: Free probe.*

3D.8. Discussion.

It has previously been demonstrated by several studies that the isolated DNA binding domain of Vmw175 will bind to, and produce a DNAaseI footprint at, consensus Vmw175 DNA binding sites (Wu and Wilcox, 1989; Pizer *et al.*, 1991; Everett *et al.*, 1991; Everett, unpublished observations). These studies have also led to the determination of the limits of the Vmw175 DNA binding domain to within a region encompassing codons 262-490 of the IE-3 open reading frame.

Gel retardation analysis of the wild type I10X polypeptide (encoding Vmw175 codons 275-523) presented in this thesis has shown that bacterial extracts containing wild type DNA binding domain polypeptides will produce specific complexes with DNA; that uninduced bacterial extracts will produce higher mobility complexes; and that both types of complex were frequently observed to be formed with bacterial extracts containing Vmw175 DNA binding domain polypeptides. These observations suggest that bacterial DNA binding proteins are not recruited into the Vmw175 DNA binding domain-DNA complex, and show that formation of the "P" complex is specific to the presence of a Vmw175 polypeptide.

In vitro studies have shown that the isolated DNA binding domain of Vmw175 binds to DNA as a dimer (J. Tyler, personal communication). The formation of multiple secondary protein-DNA complexes with higher concentrations of Vmw175 DNA binding polypeptides as seen in Figure 3D.2 could be due to two types of interactions - the interaction of multiple protein dimers with DNA or the interaction of protein dimers with more than one DNA molecule leading to the formation of higher order complexes. Additional binding studies with highly purified DNA binding domain polypeptides and DNA probes of differing lengths showed that the secondary complexes were produced only with longer DNA probes, suggesting that multimer complexes occur through the binding of additional proteins on DNA flanking the authentic binding site (R. Everett, unpublished observations). The possibility of cooperativity in the formation of these higher order complexes has not yet been investigated.

Analysis of mutant DNA binding domain polypeptides showed that mutations within the "helix-turn-helix" motif lead to three different DNA binding phenotypes: (1) the ability to bind to DNA both at room temperature and 37°C; (2) the ability to bind to DNA only at temperatures of 33°C and lower (the temperature sensitive phenotype); and (3) no detectable DNA binding activity at any temperature. The DNA binding properties of all the mutations analysed are shown in Table 3D.1.

The temperature sensitive phenotype was probably caused by the destabilizing effect of the mutation on the secondary structure of the polypeptide

molecule at 37°C, resulting in a reduction in its ability to interact with DNA. Analysis of the occurence of temperature sensitive mutations within the "helix-turn-helix" motif revealed no definite pattern of occurence. No particular amino acid alteration resulted in a consistently greater disruption of the DNA binding activity. For example, 10 out of the 18 point mutations analysed contained an alanine to serine alteration. However, this alteration resulted in both a wild type and a temperature sensitive phenotype in DNA binding activity, depending on the position of the alteration. As no particular amino acid substitution of itself showed a detrimental effect, it is likely that the temperature sensitive phenotype was caused by the position at which the mutation was made. Analysis of the temperature sensitive mutations within the motif showed that they were not concentrated within any stretch of the motif.

Temperature sensitive mutations were not concentrated on certain faces of the proposed helices. Figure 3D.13 shows the position of the mutations and their DNA binding phenotypes, represented on the helical wheel models discussed in Section 3A.5. In helix 1, a temperature sensitive and a wild type Ala-Ser mutation at codons 453 and 446 respectively are juxtaposed on the same face of the helix, implying that there is no one face more sensitive to mutation. The single Ala-Ser mutation at codon 451 showed a wild type DNA binding phenotype whereas a double Ala-Ser mutation at codons 453 and 451 showed a temperature sensitive DNA binding phenotype. It was assumed that the temperature sensitive phenotype of the double mutant was due to the Ala-Ser mutation at codon 453, since the Ala-Ser mutation at codon 451 by itself had no effect on the DNA binding properties of the polypeptide.

The situation with the second helix is more complex, due to the different types of amino acids present in the helix and the differing amino acid alterations introduced (see Figure 3D.13b). Many DNA binding proteins seem to contact DNA through clusters of basic amino acids (Davis *et al.*, 1990; Busch and Sassone-Corsi, 1990). In this helix, although there is not an obvious concentration of basic amino acids on any one face, interesting mutations to study would be ones which involve changes of basic amino acids to neutral or even acidic residues. Several such changes have been introduced and their effects range from elimination of DNA binding activity (see further discussion), through loss of DNA binding activity at higher temperatures to negligible changes in DNA binding activity. As discussed later, mutations which result in the loss of DNA binding activity are situated towards the proposed turn of the motif, and their detrimental effect on DNA binding activity may be due more to their position within the motif rather than the type of amino acid mutation. In the case of the mutations 471R/L and 472R/L, they lead to wild type and temperature sensitive DNA binding profiles respectively, suggesting that neither amino acid is essential for the interaction of protein with DNA, and that the mutations solely affect the secondary structure of the polypetide molecule.

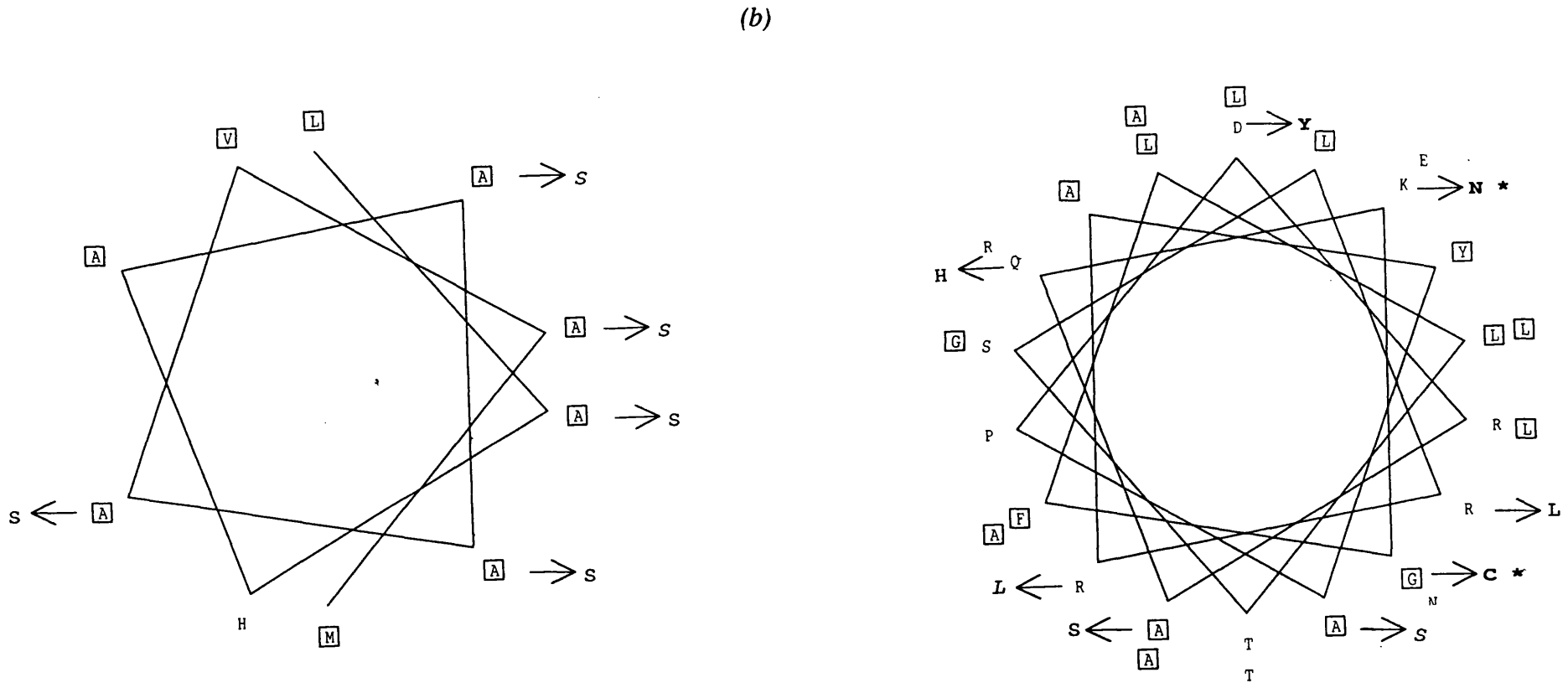


Figure 3D.13. Mutations Created Within α-Helical Regions of Vmw175.

*Proposed α-helical regions of Vmw175 were modelled on helical wheels using the GCG Helicalwheel programme. The two helices incorporate codons (a) 445-454 and (b) 459-487 of Vmw175. Boxes indicate hydrophobic residues. Mutations created within these regions are indicated with arrows. Italics indicate mutations which exhibited a temperature sensitive DNA binding phenotype, whilst bold indicates mutations which exhibited no DNA binding activity. * = double mutation.*

Another interesting observation is that an individual mutation at codon 464 (464G/C) results in a temperature sensitive phenotype, whereas the double mutant carrying both this mutation and a mutation at codon 463 (463K/N) results in elimination of DNA binding activity. A possible explanation of this observation is that the accumulation of these mutations results in the destabilization of the local secondary structure sufficient to disrupt the interactions with DNA, a situation not observed with the other double mutant that has been analysed. Alternatively, the elimination of DNA binding activity may have been caused by the single mutation at codon 463, which was not isolated and therefore could not be investigated. As discussed in Section 3A.6, analysis of the Vmw175 amino acid sequence by the GCG Peptidestructure programme produces different secondary structure predictions to those published previously (Shepard *et al.*, 1989). The double mutation discussed above could be situated in the proposed turn of the motif or within the proposed second helix, depending on the method used to predict the secondary structure of the proposed "helix-turn-helix" motif.

The overall conclusion is that no one face of either helix is more sensitive to mutation, a phenomenom observed with other helical DNA binding motifs (Busch and Sassone-Corsi, 1990). This implies that the designation of the proposed helices may be incorrect. An alternative explanation is that, whilst existing in the predicted secondary structure, neither helix presents a face for interaction with DNA.

Several point amino acid mutations, the deletion mutation leading to the truncation of the expressed polypetides at codon 454 and the two internal, in-frame deletion mutations result in the elimination of the DNA binding activity of the polypeptides. These mutations may be exerting their effects in one of two ways. The mutation could result in a disruption of the local secondary structure of the polypeptide, sufficient to interfere with its interaction with DNA. Alternatively, the point mutations and deletions may lead to alteration or removal of residue(s) within the polypeptide which is/are essential for contact with the DNA helix, as seen with mutational analyses of other DNA binding proteins (Gentz *et al.*, 1989; Davis *et al.*, 1990; Marmorstein *et al.*, 1992). It is difficult to distinguish between the two mechanisms without a means to probe the secondary structure of the mutant and wild type DNA binding domain polypetides. It is interesting to note that the point mutations which eliminate the DNA binding activity of the polypeptide cluster within or near to the turn of the proposed "helix-turn-helix" motif (see Figure 3D.13), as predicted by Shepard *et al.* (1989). This apparent sensitivity of the turn region to mutation may represent a conformational effect of the mutations on the motif, or may imply that residues within this region are important in the interaction of the domain with DNA.

The internal deletions of portions of the proposed helices will lead to a reorientation of residues on a different face of the helix, and probably lead to a gross

conformational change in this region of the "helix-turn-helix" motif. However, deletions removing codons 162-229 or 209-236 of the IE-3 reading frame have a much less significant effect on the transcription and repression activities of Vmw175, as measured by transient transfection assays (Paterson, 1989). The fact that small in-frame deletions, truncations and even point mutations can disrupt the binding of the isolated DNA binding domain illustrates the importance of the region encompassing codons 445-482 towards the conformational stability and DNA binding properties of Vmw175.

The result obtained with the i12 I10X and ri12 I10X polypeptides further complicate the situation. The fact that an insertion mutation and a point mutation approximately 120 amino acids N-terminal to the "helix-turn-helix" motif can disrupt the DNA binding properties of the I10X polypeptide suggests that this portion of the DNA binding domain is also involved in the interaction of the protein with DNA. Another possible explanation for these observations is that the mutations disrupt a portion of the DNA binding domain responsible for maintaining the secondary and tertiary conformational stability of the molecule. This hypothesis is supported by previous suggestions that the region encompassing codons 263-338 may be involved in stabilising low affinity interactions of the DNA binding domain with DNA (Shepard *et al.*, 1989). In addition, previous mutational analyses have demonstrated that mutations within this region of Vmw175 disrupt the *in vitro* DNA binding and transactivation properties of the protein (Paterson and Everett, 1988a, b; Paterson, 1989; Shepard *et al.*, 1989; see Figure 1C.3).

3E. Determination of the *In Vitro* Dimerization Activity of Wild Type and Mutant Polypeptides Encompassing the DNA Binding Domain of Vmw175.

Vmw175 has been purified from infected cells as a homodimeric complex (Metzler and Wilcox, 1985). In addition, bacterial expression and *in vitro* translation studies have illustrated that the isolated DNA binding domain of Vmw175 exists as a dimer in solution (Everett *et al.*, 1991b; J. Tyler, personal communication), implying that the DNA binding domain of Vmw175 also contains a potential dimerization motif.

In order to fully assess the effects of various mutations on the *in vitro* properties of the isolated DNA binding domain of Vmw175, the dimerization properties of certain mutant and wild type polypeptides were investigated.

3E.1. The Dimerization Properties of a Polypeptide Containing the Mutation 456R/L.

The point mutation 456R/L, situated close to or within the proposed turn of the "helix-turn-helix" motif has been shown to eliminate the DNA binding activity of the isolated DNA binding domain of Vmw175 (see Section 3D.5.2).

In order to determine the dimerization properties of this mutant polypeptide, gel filtration chromatography was carried out on purified preparations of mutant and wild type I9X polypeptides. The polypeptides were extracted from induced cultures (1 litre) as described in Section 2B.28, and mutant and wild type polypeptides purified from the final extracts by chromatography on a Pharmacia MonoS ion exchange column as described in section 2B.29. Both mutant and wild type polypeptides were eluted in a sharp peak at 0.8M NaCl. Peak fractions were pooled and concentrated by centrifugation in Centricon 10 concentrators. Concentrated samples appeared to be approximately 95% pure on Coomassie blue stained acrylamide gels (see Figure 3E.1).

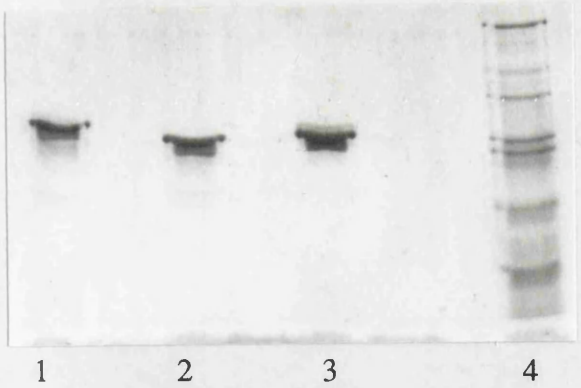


Figure 3E.1. SDS-Polyacrylamide Gel Electrophoresis of Purified Wild Type and Mutant I9X and I10X Polypeptides.

Polypeptides were extracted from induced BL21 (DE3) pLysS cultures as describea in Section 2B.28 and purified by chromatography on a Pharmacia MonoS ion exchange column, as described in Section 2B.29. Peak fractions containing DNA binding domain polypeptides were pooled and concentrated by centrifugation in Centricon 10 concentrators. Concentrated proteins were analysed by electrophores on 12.5% SDS-polyacrylamide gels followed by staining with Coomassie blue.
Track 1: I9X
Track 2: I10X
Track 3: 456R/L I9X
Track 4: Molecular weight standards: 66kDa, 45kDa, 36kDa, 29kDa, 24kDa, 20.1kDa, 14.2kDa.

The concentrated preparations of polypeptide were fractionated on a 20ml Superose 12 columnas described in Section 2B.29. 1ml fractions were collected and the peak fractions analysed by electrophoresis on 12.5% SDS polyacrylamide gels followed by detection by Western blot analysis. The wild type I9X polypeptide was eluted at a peak position of 12.98ml, as determined by the uv monitor and the 456R/L I9X polypeptide was eluted at 12.99ml (Figure 3E.2), showing that the behaviour of the mutant polypeptide was indistinguishable from that of the wild type polypeptide upon gel filtration chromatography analysis. A "shoulder" peak was observed prior to elution of the major peak for both wild type and mutant polypeptides. This shoulder peak could possibly contain tetrameric forms of the polypeptides. The molecular weight standards bovine serum albumin (BSA, native molecular weight 66kDa), carbonic anhydrase (CA, native molecular weight 29kDa) and cytochrome C (Cyt. C, native molecular weight 13kDa) were eluted at peak positions of 12.59, 14.37 and 15.31mls respectively. Analysis of these molecular weight standards enabled the construction of a calibration curve for the Superose 12 column. From this, the molecular weights of I9X and 456R/L I9X were calculated to be 57kDa, close to the predicted dimeric molecular weight of 56kDa (see Figure 3E.3).

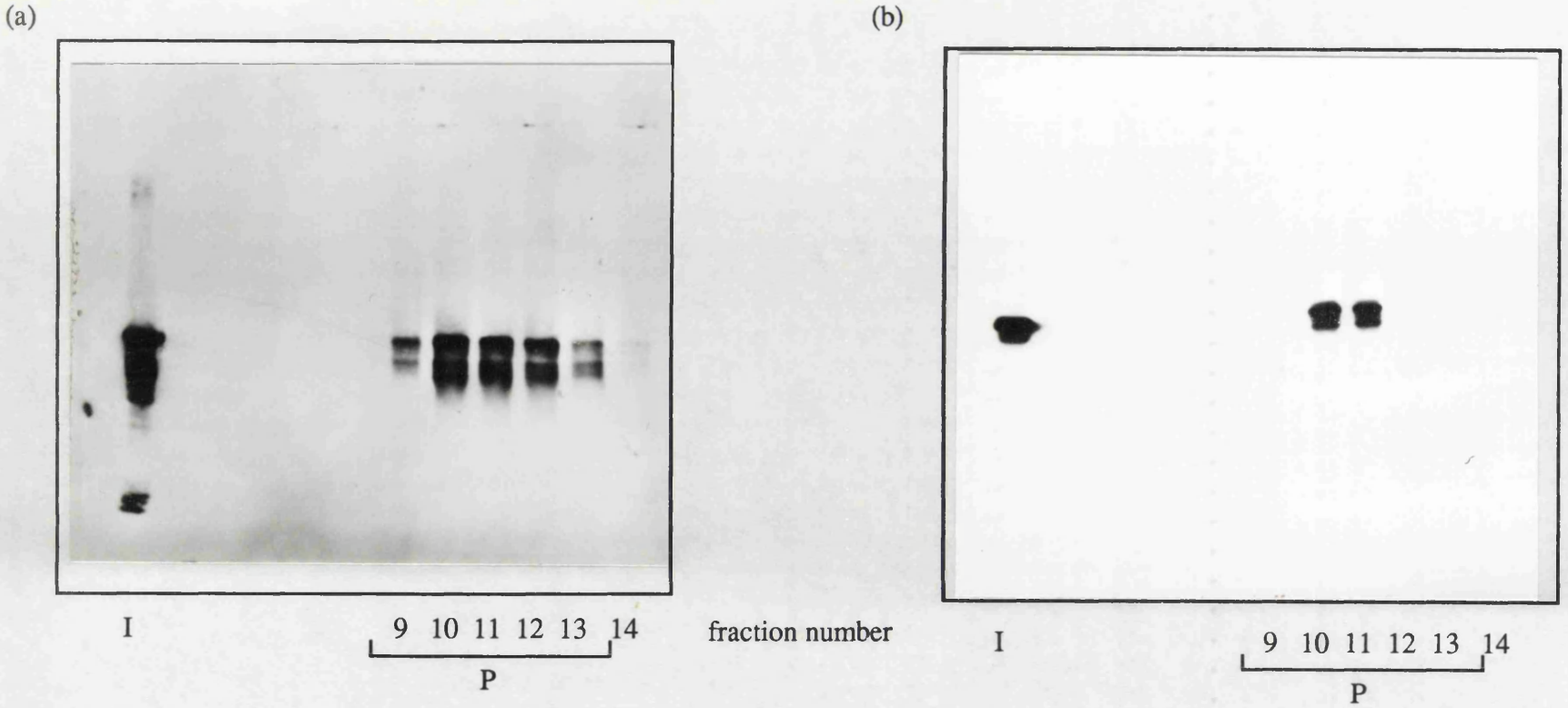


Figure 3E.2. Western Blot Analysis of Fractions from Gel Filtration Chromatography of the I9X and 456R/L I9X Polypeptides.

Polypeptides were purified as described in the text, and loaded onto a Pharmacia Superose 12 column. Peak fractions eluted from this column were analysed by Western blot analysis. Aliquots of the fractions were electrophoresed on 12.5% SDS-polyacrylamide gels, proteins transferred onto nitrocellulose and immunoblotted as described in Section 2B.6.3. Proteins were immunoblotted with a polyclonal rabbit antibody raised to a bacterial fusion protein containing the Vmw175 DNA binding domain, used at a 1:1000 dilution. Proteins were detected using the Amersham ECL detection system, as described in Section 2B.6.3.

(a) I9X polypeptide.
I: Input fraction.
P: Peak fractions.

(b) 456R/L I9X polypeptide.
I: Input fraction.
P: Peak fractions

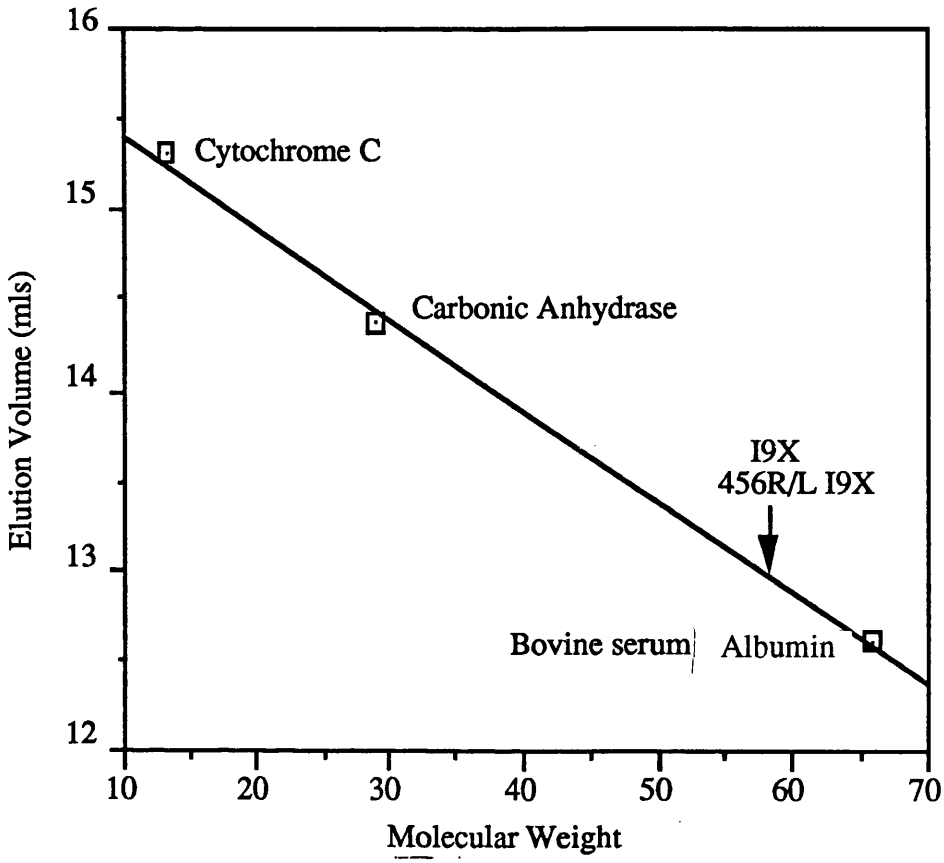


Figure 3E.3. Standard Curve for the Superose 12 Gel Filtration Column.

50ng of molecular weight standards as indicated were analysed on a 20ml Superose 12 ge filtration column. The column was equilibrated in 50mM HEPES pH7.5, 0.5M NaCl, 0.01% CHAPS, 0.1mM DTT and 1mM PMSF and proteins eluted in the same buffer. The graph shows the elution volume of the proteins, together with the elution volume of wild type and the mutant 456R/L form of the I9X polypeptide. The molecular weights of the I9X polypeptides were estimated from the standard curve, and were found to be 57KDa.

3E.2. The Dimerization Properties of a Polypeptide Containing the Mutation 463/4 KG/NC.

The double point mutation 463/4 KG/NC also contains mutations situated close to or within the proposed turn of the "helix-turn-helix" motif. This mutation resulted in the loss of the DNA binding activity of the isolated DNA binding domain of Vmw175 (see Section 3D.5.2).

In order to determine the effect of this mutation on the dimerization properties of the I10X polypeptide, gel filtration chromatography was carried out on purified preparations of the wild type and mutant polypeptides. The polypeptides were extracted from 1 litre induced bacterial cultures as described in section 2B.28, with the ommission of the treatment with polyethylenimine. Mutant and wild type polypeptides

were purified from the final extracts by chromatography on a Pharmacia MonoS ion exchange column, as descibed in Section 2B.29. In this experiment, the polypeptides were eluted in the flow through fractions from the MonoS column; this behaviour was discussed in Section 3C.5. The polypeptides were further purified from the flow through fractions as described in Section 2B.29. Final preparations were shown to be approximately 80-90% pure on Coomassie blue stained acrylamide gels (see Figure 3C.8).

The concentrated polypeptides were fractionated on a Superdex 75 column as described in Section 2B.29. 1ml fractions were collected and peak fractions analysed by electrophoresis on 12.5% SDS-polyacrylamide gels and detection by Western blot analysis. The wild type I10X polypeptide was eluted at 9.84mls and 463/4 KG/NC I10X eluted at 9.82mls (see Figure 3E.4), showing that the behaviour of the mutant polypeptide was indistinguishable from that of the wild type polypeptide upon gel filtration chromatography analysis. The molecular weight standards bovine serum albumin (BSA, native molecular weight 66kDa), carbonic anhydrase (CA, native molecular weight 29kDa) and cytochrome C (Cyt. C, native molecular weight 13kDa) were eluted at 9.53, 11.81 and 13.48mls respectively. Analysis of these molecular weight standards enabled the construction of a calibration curve for the Superdex 75 column. From this, the molecular weights of I10X and 463/4 KG/NC were calculated to be 60kDa, close to the predicted dimeric molecular weight of 52kDa (see Figure 3E.5).

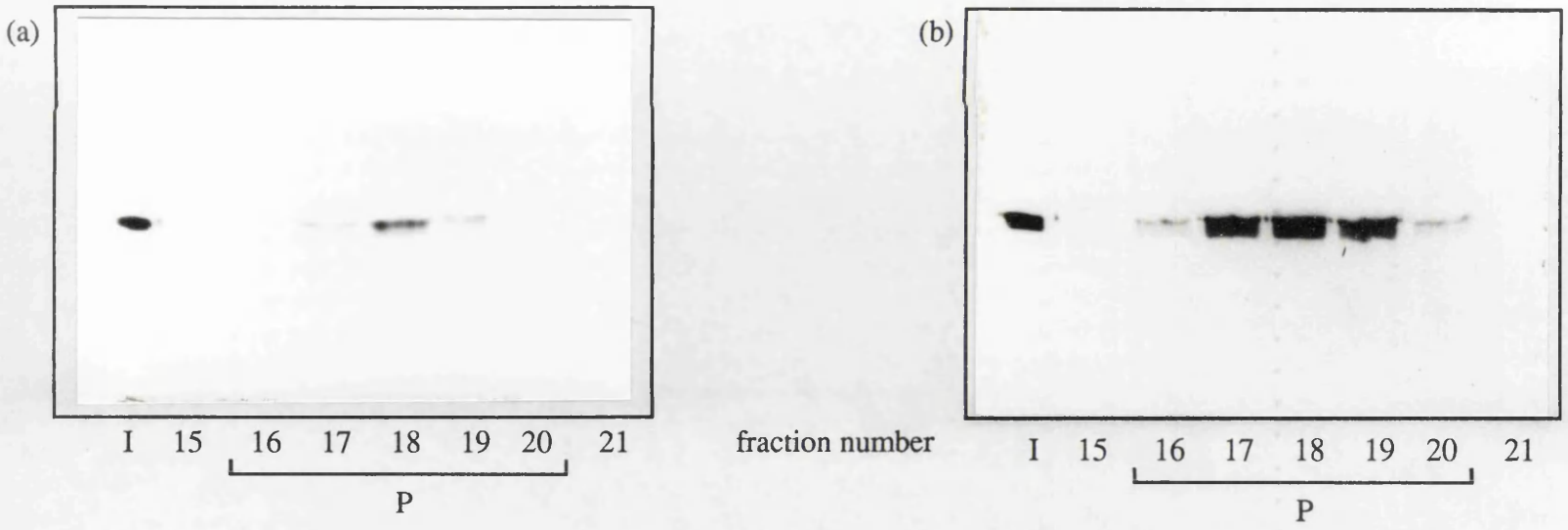


Figure 3E.4. Western Blot Analysis of Fractions from Gel Filtration Chromatography of the I10X and 463/4KG/NC I10X Polypeptides.

Polypeptides were purified as described in the text, and loaded onto a Pharmacia Superdex 75 column. Peak fractions eluted from this column were analysed by Western blot analysis. Aliquots of the fractions were electrophoresed on 12.5% SDS-polyacrylamide gels, proteins transferred onto nitrocellulose and immunoblotted as described in Section 2B.6.3. Proteins were immunoblotted with a polyclonal rabbit antibody raised to a bacterial fusion protein containing the Vmw175 DNA binding domain, used at a 1:1000 dilution. Proteins were detected using the Amersham ECL detection system, as described in Section 2B.6.3.

(a) I10X polypeptide.
I: Input fraction.
P: Peak fractions.

(b) 463/4KG/NC I10X polypeptide.
I: Input fraction.
P: Peak fractions

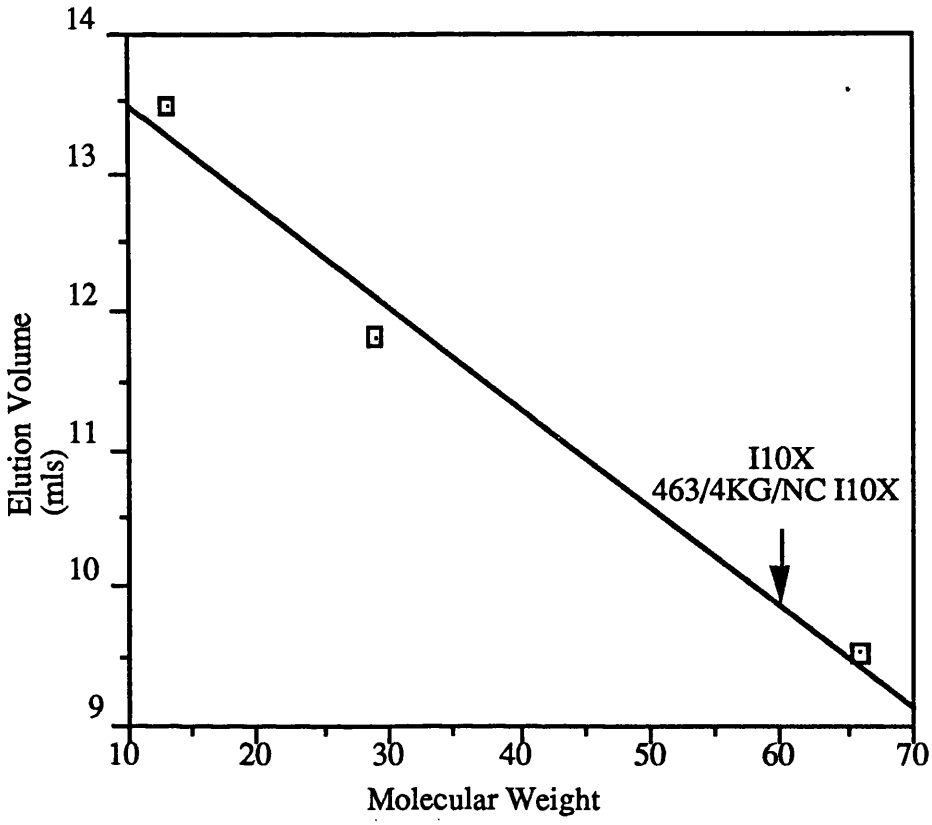


Figure 3E.5. Standard Curve for the Superdex 75 Gel Filtration Column.
50 ng of the molecular weight standards indicated were analysed on a 20 ml Superdex 75 gel filtration column. The column was equilibrated in 50mM HEPES pH7.5, 0.5M NaCl, 0.01% CHAPS, 0.1mM DTT and 1mM PMSF and proteins eluted in the same buffer. The graph shows the elution volumes of the proteins, together with the elution volume of wild type and the mutant 463/4KG/NC form of the I10X polypeptide. The molecular weights of the I10X polypeptides were estimated from the standard curve, and were found to be 60KDa.

3E.3. Investigation of the Dimerization Properties of a Polypeptide Containing the Mutation DH1.

The mutation DH1 is an internal, in-frame deletion of seven amino acids from the first proposed helix of the "helix-turn-helix" motif. Analysis of the DNA binding properties of an I10X polypeptide containing this mutation showed that the deletion eliminated the DNA binding properties of the isolated DNA binding domain of Vmw175 (see Section 3D.6).

In order to determine the dimerization properties of this mutation, it was necessary to obtain relatively pure preparations of the DH1 I10X polypeptide for glutaraldehyde cross-linking assays or gel filtration chromatography. As the DH1 I10X

polypeptide had exhibited low solubility upon previous extractions, the polypeptide was extracted from a 1 litre induced bacterial culture in the absence of polyethylenimine, as described in Section 2B.28. A high proportion of the polypeptide was lost at the centrifugation stages of the extraction procedure, however the final soluble extract contained a sufficient quantity of protein for glutaraldehyde cross-linking or gel filtration chromatography analysis. The soluble extract was dialysed and insoluble proteins removed by centrifugation as detailed in Section 2B.29. SDS-polyacrylamide gel electrophoresis of the pellet and supernatant fractions from this centrifugation step showed that the DH1 I10X polypeptide had precipitated and the soluble extract contained a very low concentration of the mutant polypeptide (see Figure 3E.6, tracks 1 and 2). The soluble protein extract was loaded onto a Pharmacia MonoS ion exchange column and polypeptides eluted with a 20ml 0.1-1M NaCl gradient. It was expected that the DH1 I10X polypeptide would be eluted in the flow-through fractions from this column, as had been previously observed with preparations of purified polypeptides extracted in the absence of polyethylenimine (see Figure 3C.6). However, due to the low concentration of DH1 I10X polypeptide in the input sample (Figure 3E.6, track 2), the concentration of DH1 I10X in the flow through fractions was too low for further purification (see Figure 3E.6, tracks 3-6). After several attempts to purify this mutant polypeptide, it was concluded that the low solubility of the polypeptide prevented the purification of sufficient quantities of the protein for glutaraldehyde cross-linking or gel filtration analysis.

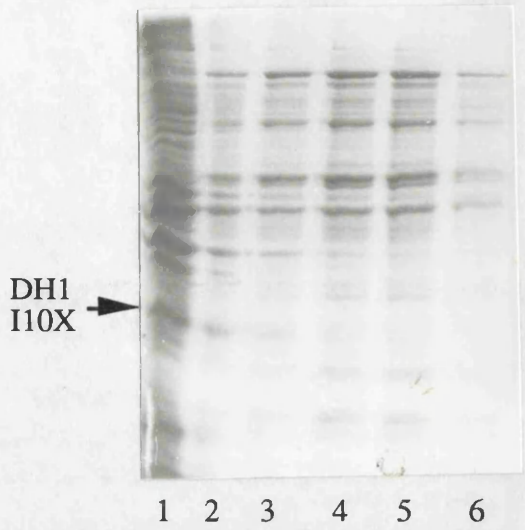


Figure 3E.6. SDS-Polyacrylamide Gel Electrophoresis of fractions from Ion exchange Chromatography of the DH1 I10X polypeptide.

DH1 I10X polypeptide was extracted from induced BL21 (DE3) pLysS cultures as described in Section 2B.28, ommitting treatment with polyethylenimine. Following dialysis, the extract was loaded onto a Pharmacia MonoS ion exchange column, and the protein eluted with a salt gradient, as described in section 2B.29 Fractions were analysed by electrophoresis on a 12.5% SDS-polyacrylamide gel followed by staining with Coomassie blue.
Track 1: Insoluble proteins following dialysis.
Track 2: Soluble input fraction.
Tracks 3-6: Flowthrough fractions.

3E.4. The Dimerization Properties of Polypeptides Containing Point Mutations Resulting in a Temperature Sensitive DNA Binding Defect.

Bacterial extracts of point mutant I10X polypeptides which exhibited temperature sensitive or wild type DNA binding phenotypes and wild type I10X polypeptide were prepared from 100ml induced bacterial cultures, as described in Section 2B.27.1. The dimerization properties of these polypeptides were investigated using a glutaraldehyde cross-linking assay. Glutaraldehyde is an aldehyde polymer which forms covalent Schiff base linkages with lysine residues of protein molecules, leading to the covalent cross-linking of oligomeric protein molecules in solution. Such oligomeric molecules may be identified by SDS-polyacrylamide gel electrophoresis analysis of the reaction mixture. This method was utilized to analyse the dimerization properties of point mutant polypeptides which exhibit temperature sensitive defects in DNA binding activity, as it allows for the assay of dimerization activity at various temperatures.

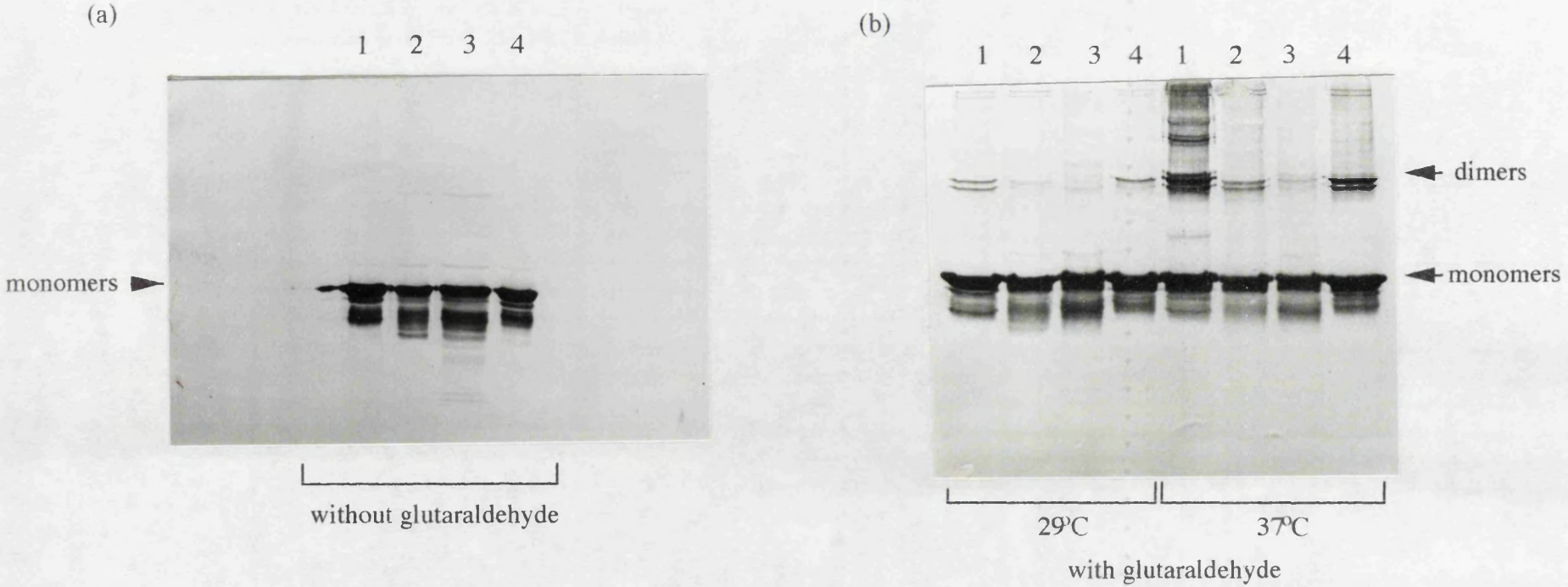


Figure 3E.7. Glutaraldehyde Cross-Linking Assay of Wild Type and Mutant Vmw175 DNA Binding Domain Polypeptides.
Bacterial extracts containing point mutant and wild type DNA binding domain polypeptides were prepared as described in Section 2B.27. Extracts were incubated with or without 0.01% glutaraldehyde in a buffer containing 50mM HEPES pH7.5; 0.1M NaCl; 1mM PMSF for 30 min. at 29 C or 37 C. The cross-linking reaction was stopped by the addition of SDS sample loading buffer and the samples were boiled for 5 min. Extracts were electrophoresed on 12.5% SDS-polyacrylamide gels, proteins transferred onto nitrocellulose and immunoblotted as described in Section 2B.6.3. Proteins were immunoblotted with a rabbit polyclonal antibody raised to a bacterial fusion protein containing the Vmw175 DNA binding domain, used at a 1:1000 dilution.
The following polypeptides were assayed:
Track 1: 472R/L I10X *Track 2: 471R/L I10X*
Track 3: tsK I10X *Track 4: I10X.*

Bacterial extracts containing containing mutant and wild type I10X polypeptides were incubated with glutaraldehyde at 29°C or 37°C as described in Section 2B.30. The samples were boiled for 5 min before electrophoresis on 12.5% SDS-polyacrylamide gels. The experiments were carried out at low protein concentrations using extracts containing insufficient amounts of I10X polypeptide for visualization on Coomassie blue stained acrylamide gels. Therefore, monomeric and dimeric forms of the polypeptides were visualised by Western blot analysis as decribed in section 2B.6.3.

Western blot analysis of proteins incubated in the absence of glutaraldehyde illustrated the extent of degradation of the mutant and wild type polypeptides (Figure 3E.7a).The extent of protein degradation was unchanged between reactions carried out at 29°C and reactions carried out at 37°C (Figure 3E.7b). Cross-linked dimers of the I10X polypeptides were formed at both 29°C and 37°C, regardless of whether the mutant polypetpide exhibited a temperature sensitive or wild type DNA binding phenotype. As shown in Figure 3E.7b, only a small percentage of the monomeric polypeptide molecules were incorporated into the cross-linked dimeric molecules, and the amount of cross-linked dimeric molecule formed with the various polypeptide extracts was not consistent. Possible explanations for these observations are discussed in section 3E.6.

3E.5. The Dimerization Properties of a Polypeptide Containing the Mutation ri12.

The double mutation ri12 consists of a four amino acid insertion mutation at codon 320 and a valine to alanine substitution at codon 345 of the IE-3 open reading frame. Analysis of DNA binding domain polypetides containing the ri12 mutation showed that the mutation resulted in a temperature sensitive defect in DNA binding activity (see Section 3D.7).

In order to determine the effects of this mutation on the dimerization activity of the I10X polypeptide at various temperatures, glutaraldehyde cross-linking analysis was carried out. The mutant and wild type polypeptides were extracted from 1 litre induced bacterial cultures in the absence of polyethylenimine, as described in Section 2B.28. Mutant and wild type I10X polypeptides were purified from the final bacterial extracts by chromatography on a Pharmacia MonoS ion exchange column, as described in Section 2B.29. The polypeptides were eluted from this column in the flow through fractions and polypeptides further purified as described in Section 2B.29.

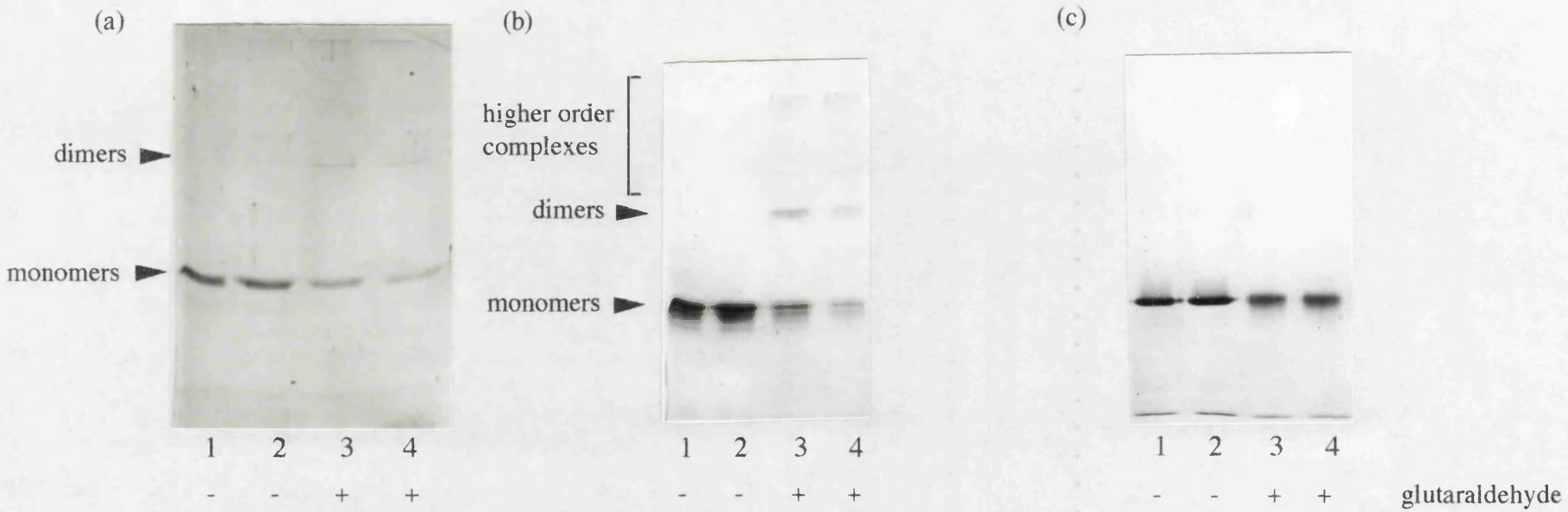


Figure 3E.8. Glutaraldehyde Cross-Linking Analysis of Wild Type and Mutant I10X Polypeptides.

Bacterial extracts containing ri12 mutant or wild type I10X polypeptides were prepared as described in Section 2B.28. Proteins were incubated with or without 0.01% glutaraldehyde in a buffer containing 50mM HEPES pH 7.5; 0.1M NaCl; 1mM PMSF for 30 minutes at 29°C or 37°C. The cross-linking reaction was stopped by the addition of SDS sample loading buffer and the samples were boiled for 5 min. Extracts were electrophoresed on 12.5% SDS-polyacrylamide gels, proteins transferred onto nitrocellulose and immunoblotted as described in Section 2B.6.3. Proteins were immunoblotted with a rabbit polyclonal antibody raised to a bacterial fusion protein containing the Vmw175 DNA binding domain, used at a 1:1000 dilution.

(a): ri12 *(b) I10X* *(c) Carbonic anhydrase*

Tracks 1 and 3: 29°C *Tracks 2 and 4: 37°C*

Polypeptides were incubated with glutaraldehyde at 29°C or 37°C as described in Section 2B.30. Samples were boiled for 5 min before electrophoresis on 12.5% SDS-polyacrylamide gels. Monomeric and oligomeric forms of the polypeptide molecules were visualised by staining with Coomassie blue.

The incubation of the wild type I10X polypetide and carbonic anhydrase in the presence of glutaraldehyde led to an increase in protein degradation (Figure 3E.8b and c, compare tracks 1 and 2 with tracks 3 and 4). Cross-linking analysis of the I10X polypeptide at 29°C and 37°C led to the visualization of a ladder of bands upon SDS-polyacrylamide gel electrophoresis (Figure 3E.8b, tracks 3 and 4). These ladders were not seen when the polypeptide was incubated in the absence of glutaraldehyde. Incubation of the monomeric protein carbonic anhydrase with glutaraldehyde did not result in the formation of cross-linked molecules of higher molecular weight (Figure 3E.8c, tracks 3 and 4).

Glutaraldehyde cross-linking analysis of the ri12 I10X polypeptide led to the visualization of higher order bands, corresponding to cross-linked dimers of the mutant polypeptide, which were not observed when the polypeptide was incubated in the absence of glutaraldehyde (Figure 3E.8a, compare tracks 1 and 2 with tracks 3 and 4). The dimeric molecules were formed at both 29°C and 37°C, indicating that the higher incubation temperature did not disrupt the dimerization activity of the ri12 I10X polypeptide.

3E.6. Discussion.

Analysis of bacterially expressed polypeptides has shown that the isolated DNA binding domain of Vmw175 exists as a dimer in solution (Everett *et al.*, 1991b). In addition, *in vitro* translated polypeptides encompassing the DNA binding domain of Vmw175 bind to a DNA probe as dimers upon gel retardation analysis (Roberts and Hayward, 1991). Although the *in vitro* translation studies suggested that within the DNA binding domain dimer molecule there may only be a requirement for one DNA binding functional monomer to interact with the DNA probe (Roberts and Hayward, 1991), the existence of an interaction between DNA binding domain monomeric molecules and DNA has not been shown to occur. In fact, it is possible that disruption of the dimerization activity of the DNA binding domain of Vmw175 may lead to the disruption of the DNA binding activity of the molecule. Therefore, investigation of the dimerization activity of the mutant polypeptides which exhibited a defect in DNA binding activity may yield information about the mechanism of the association of Vmw175 with DNA.

Sections 3E.1 and 3E.2 present results obtained from gel filtration analysis of two mutant DNA binding domain polypeptides. These polypeptides contain point mutations within the proposed "helix-turn-helix" motif, and exhibit no DNA binding activity at any temperature, enabling the analysis of dimerization activity to be conducted at room temperature by FPLC analysis. Although gel filtration chromatography is not a means by which to calculate the exact molecular weight of a protein molecule, comparison of the migration of the mutant polypeptides relative to the migration of known dimeric molecules of the same monomeric molecular weight can reveal the native form of the mutant molecules in solution. If the wild type and mutant polypeptides are eluted from the gel filtration column in the same elution volume, then it can be concluded that the mutant polypeptide also exists as a dimer in solution.

Analysis of the migration of of the mutant I9X polypeptide 456R/L as compared with the wild type I9X polypeptide on a Pharmacia Superose 12 column revealed that the two polypeptides were eluted in sharp peaks at nearly identical elution volumes (12.98 and 12.99mls respectively), indicating that both polypeptides exist in the same native state in solution. From a standard curve constructed from the elution volumes of known molecular weight standards, the molecular weight of the two polypeptides was estimated to be 57kDa, which is closer to the predicted dimeric molecular weight (56kDa) than to the predicted monomeric molecular weight (28kDa). Similar results were obtained from analysis of the migration of the mutant I10X polypeptide 463/4 KG/NC as compared with the migration of the wild type I10X polypeptide on a Pharmacia Superdex 75 column. The polypeptides were eluted as sharp peaks of nearly identical elution volumes (9.82 and 9.84mls respectively) and their molecular weight was estimated to be 60kDa, closer to the predicted dimeric molecular weight (52kDa) than to the predicted monomeric molecular weight (26kDa).

Therefore, the two mutant polypeptides which have lost the ability to bind to DNA still retain dimerization activity. This fact suggests that the DNA binding activity and the dimerization properties of the isolated DNA binding domain of Vmw175 are functions of separable regions of the polypeptide. As the mutants retain the ability to dimerize, the secondary structure of the polypeptide has not been grossly altered by the point mutations, implying that the effect of the mutations has been to alter only the local secondary conformation of the region spanning codons 445-487, or to alter the essential residues of the "helix-turn-helix" motif. The fact that mutations that obviously eliminate the function of this "helix-turn-helix" region do not affect the dimerization activity of the polypeptide suggests that portions of the polypeptide outwith the region spanning codons 445-487 are responsible for the dimerization function of the DNA binding domain of Vmw175.

In order to attempt to define the functions of the proposed helical regions of the "helix-turn-helix" motif further, it was necessary investigate the dimerization properties of polypeptides containing deletions within the two helices, which have been previously shown to eliminate the DNA binding activity of the DNA binding domain polypeptide (see Section 3D.6). Unfortunately, the deletion of these regions appeared to drastically reduce the solubility of the isolated DNA binding domain polypeptide. In order to investigate the dimerization activity of the I10X polypeptide mutant DH1, which contains a deletion removing seven amino acids from the first helix of the proposed motif, it was necessary to obtain relatively pure peparations of the polypeptide for cross-linking or gel filtration analysis. Section 3E.3 illustrated the problems encountered in the preparation of such samples. Due to the low solubility of the deletion mutant polypeptide, soluble extracts prepared from induced bacteria contained insufficient amounts of the polypeptide for purification on ion exchange chromatography, and were too impure for glutaraldehyde cross-linking analysis.

In order to investigate the effects of other point mutations within the "helix-turn-helix" motif and the mutation ri12 on the dimerization activity of the I10X polypeptide, it was necessary to conduct the dimerization assays at various temperatures as several of the point mutations and the ri12 mutation eliminate the DNA binding activity of the I10X polypeptide at temperatures of 37°C and above. This made it difficult to analyse the dimerization activity of the polypeptides by gel filtration chromatography, and it was decided to carry out the analyses using a glutaraldehyde cross-linking assay.

The results of the analysis of several point mutant polypeptides, which contain mutations within the "helix-turn-helix" region, are presented in Section 3E.4. It was found that for the mutant polypeptides investigated, molecules of a molecular weight corresponding to dimers of the DNA binding domain polypeptides were irreversibly cross-linked at both 29°C and 37°C, regardless of whether the point mutation resulted in a wild type or temperature sensitive DNA binding phenotype.

As noted in Section 3E.4, glutaraldehyde cross-linking of various polypeptides resulted in a wide variety in the amount of dimeric cross-linked molecules produced from different extracts, and as a whole, the percentage of monomeric molecules converted to cross-linked dimeric molecules was low. The variety in the rate of formation of cross-linked dimeric molecules between different extracts is probably due to variation in the concentration of other bacterial proteins present in the reaction mix. As the Vmw175 DNA binding domain monomeric molecules were visualized by immunodetection methods, the amounts of bacterial proteins present in the extracts could not be shown. However, Coomassie blue stained SDS polyacrylamide gels of the bacterial extracts showed that the concentration of Vmw175 polypeptides as compared to the concentration of bacterial proteins varied from extract to extract. The presence of

bacterial proteins would interfere with the specific cross-linking of Vmw175 molecules, as lysine residues on the bacterial proteins would react with the glutaraldehyde polymer, resulting in a lower concentration of unreacted glutaraldehyde in the reaction mix. The observation that a low percentage of Vmw175 monomeric molecules were converted to cross-linked dimeric molecules could be explained by the occurence of intramonomeric as well as intermonomeric linkages. That is, a high percentage of the glutaraldehyde cross-linking reagent reacted with and formed covalent bonds between lysine residues within one monomeric molecule, leading to a lower rate of formation of covalent bonds between lysine residues of the two monomeric subunits of a dimeric molecule.

The reaction of glutaraldehyde with carbonic anhydrase, which exists as a monomer in solution, did not lead to the formation of higher order cross-linked complexes. Therefore, it can be concluded that the higher order Vmw175 DNA binding domain polypeptide molecules were formed by specific cross-linking of dimers in solution. Therefore, at least for the mutant polypeptides analysed in this section, the defect in DNA binding activity at higher temperatures is not due to a dimerization defect. As discussed earlier, this again implies that the point mutations are disrupting the local secondary conformation of the "helix-turn-helix" region, and that disruption of the conformation of this region leads to the elimination of DNA binding activity but not to the elimination of dimerization activity. Cross-linking analysis was not carried out on other temperature sensitive point mutant polypeptides as the low solubility of the proteins made the preparation of relatively pure extracts difficult to achieve. However, the results obtained with the three point mutant polypeptides add support to the hypothesis that portions of the Vmw175 DNA binding domain outwith the region spanning codons 445-487 must contribute to the dimerization activity of the isolated polypeptide.

The mutation ri12 is situated within the N-terminal region of the Vmw175 DNA binding domain, and it is possible that this mutation leads to a local conformational change at higher temperatures, which might disrupt dimerization and explain the inability of the molecule to bind to DNA at higher temperatures. Again, glutaraldehyde cross-linking analysis was carried out to investigate the dimerization activity of the ri12 I10X polypeptide at 29°C and 37°C. It was found that samples of the mutant polypeptide of approximately 90% purity formed specific cross-linked molecules of a molecular weight corresponding to the dimeric molecular weight of the polypeptide upon glutaradehyde cross-linking analysis. Therefore, at temperatures at which the DNA binding activity of the polypeptide was eliminated, the ri12 mutation failed to disrupt the dimerization of the DNA binding domain of Vmw175. This result implies, as with previous results obtained with mutations within the "helix-turn-helix" motif, that the portion of the polypeptide containing the ri12 mutation (ie.

encompassing codons 320-345) does not contribute to the dimerization properties of the isolated DNA binding domain, and that portions of the polypeptide outwith the region spanning codons 320-245 are responsible for the dimerization function of the DNA binding domain of Vmw175.

Two other hypotheses may still account for the involvement of either the "helix-turn-helix" motif or the N-terminal portion of the DNA binding domain of Vmw175 in the dimerization interface of the protein. Firstly, the mutation introduced may be sufficiently disruptive to the dimerization interface to affect the DNA binding activity of the isolated polypeptide, but not sufficiently disruptive to eliminate the association of polypeptides into dimeric molecules. In the case of the mutation ri12, it should be remembered that the original mutation from which this pseudo-revertant developed, the i12 mutation, had a more disruptive influence on the DNA binding capacity of Vmw175 (Shepard *et al.*, 1989, Shepard and DeLuca, 1991b). Multiple attempts at expressing an I10X polypeptide containing the i12 mutation yielded very low quantities of the polypeptide - could it be that the polypeptide is unable to form stable dimers, and so is rapidly degraded? Other DNA binding domain polypeptides containing mutations within this region were also very difficult to express (R. D. Everett, personal comm.).

Secondly, the dimerization interface of the polypeptide molecule may be composed from different regions of primary sequence to form a "discontinuous" dimerization interface. Studies of protein antigenicity have shown that antigenic epitopes are composed of both continuous and discontinuous sequences. In order to test these hypotheses, it would be useful to isolate and identify mutations which disrupt the dimerization activity of the DNA binding domain polypeptide of Vmw175. At present, the existence of such mutations has not been reported, and as discussed above, the expression of dimerization deficient mutant polypepetides may be difficult to achieve.

3F. The Strategy for the Construction of Recombinant HSV-1 Viruses Containing In-Frame Point, Insertion and Deletion Mutations Within the IE-3 Open Reading Frame.

The aim of the work presented in this section was to introduce mutations, whose *in vitro* properties have previously been determined, into both copies of the IE-3 coding sequence of HSV-1 strain 17 syn^+. The construction of these viruses would

yield information on the ability of mutant Vmw175 molecules to support viral growth in tissue culture, and possibly lead to the further analysis of the transactivation and autoregulation properties of mutant Vmw175 molecules which are able to support the growth of HSV-1 in non-complementing cell lines.

3F.1. Complementing Cell Lines M64A and RR1, and Propagation of the Mutant Virus D30EBA.

In order to study the effects of mutations on the *in vivo* properties of the Vmw175 protein, it was necessary to introduce them into both copies of the IE-3 open reading frame of HSV-1. As a functional Vmw175 molecule is essential for the viral life cycle (for example, see Watson and Clements, 1978; Preston, 1979), any viruses containing lethal mutations within the IE-3 gene must be propagated on cell lines which express complementing levels of Vmw175 from the cellular genome. Previous workers (Paterson, 1989; Paterson and Everett, 1990) utilized the cell line M64A, which is similar to the cell line M65 (Davidson and Stow, 1985), to propagate and purify virus strains containing deleterious mutations within the IE-3 reading frame.

It was decided to utilize a similar strategy to construct a series of recombinant viruses containing insertional, point and deletion mutations within the IE-3 open reading frame. The virus strain D30EBA (Paterson, 1990) is based on the HSV-1 strain 17 syn$^+$ and contains large deletions encompassing nucleotides 246-3806 within both copies of the IE-3 open reading frame. Cotransfection of D30EBA viral DNA with linear plasmid DNA carrying mutant forms of the IE-3 coding sequence results in the production of recombinant viruses carrying mutant copies of the IE-3 gene. The effect of the mutations on the properties of the Vmw175 molecule determines the viability of the recombinant viruses in non-complementing cell lines.

Early preparations of D30EBA viral stocks propagated on the cell line M64A contained 0.001-0.01% of contaminating wild type virus, presumably formed through recombination events between D30EBA IE-3 coding sequences and wild type IE-3 coding sequences within the cellular genome. In order to reduce the background contamination with wild type HSV-1, D30EBA viral stocks used for the experiments presented in this thesis were propagated on the cell line RR1 (obtained from Dr. P. A. Johnston, University of California Medical School). As discussed in Section 3F.4, this cell line results in a reduced occurence of recombination betwen mutant and wild type IE-3 coding sequences, lowering the contamination of D30EBA stocks with wild type HSV-1. As discovered previously (Paterson, 1989), titration of D30EBA stocks on M64A cells was difficult due to the low growth of viral plaques and occasional loss of

complementing cellular activity, and similar problems were encountered when titrating stocks of D30EBA on the RR1 cell line. However, stocks of D30EBA of a titre estimated at 10^7 pfu/ml on RR1 cells contained wild type HSV-1 contamination at a titre of less than 10^2 pfu/ml on BHK cells, ie. contamination at less than 10^{-3}%. Pulse labelling of polypeptides after a 16 hour infection of RR1 and BHK cells with either HSV-1 strain 17 syn^+ or D30EBA showed that D30EBA in the absence of complementing Vmw175 had limited genetic expression of immediate early genes (excluding Vmw175) and (rather interestingly) large amounts of the large subunit of ribonucleotide reductase (see Figure 3F.1, track 6). However, the RR1 cell line allowed for the replication of D30EBA, leading to the expression of early and late genes, as seen in infections with the wild type virus (Figure 3F.1, compare track 3 with tracks 2 and 5).

As mentioned above, mutant IE-3 sequences were introduced into HSV-1 by cotransfection of D30EBA viral DNA with linear plasmid DNA carrying mutant IE-3 coding sequences. In order to avoid problems associated with recombination events, cotransfections were carried out in BHK cells. This would lead to the isolation of recombinant viruses containing mutations which did not disrupt the full *in vivo* activity of Vmw175. Any lethal mutations within the IE-3 open reading frame would lead to the formation of non-viable recombinant viruses; therefore, this strategy does not allow for the isolation of such mutant strains of HSV-1.

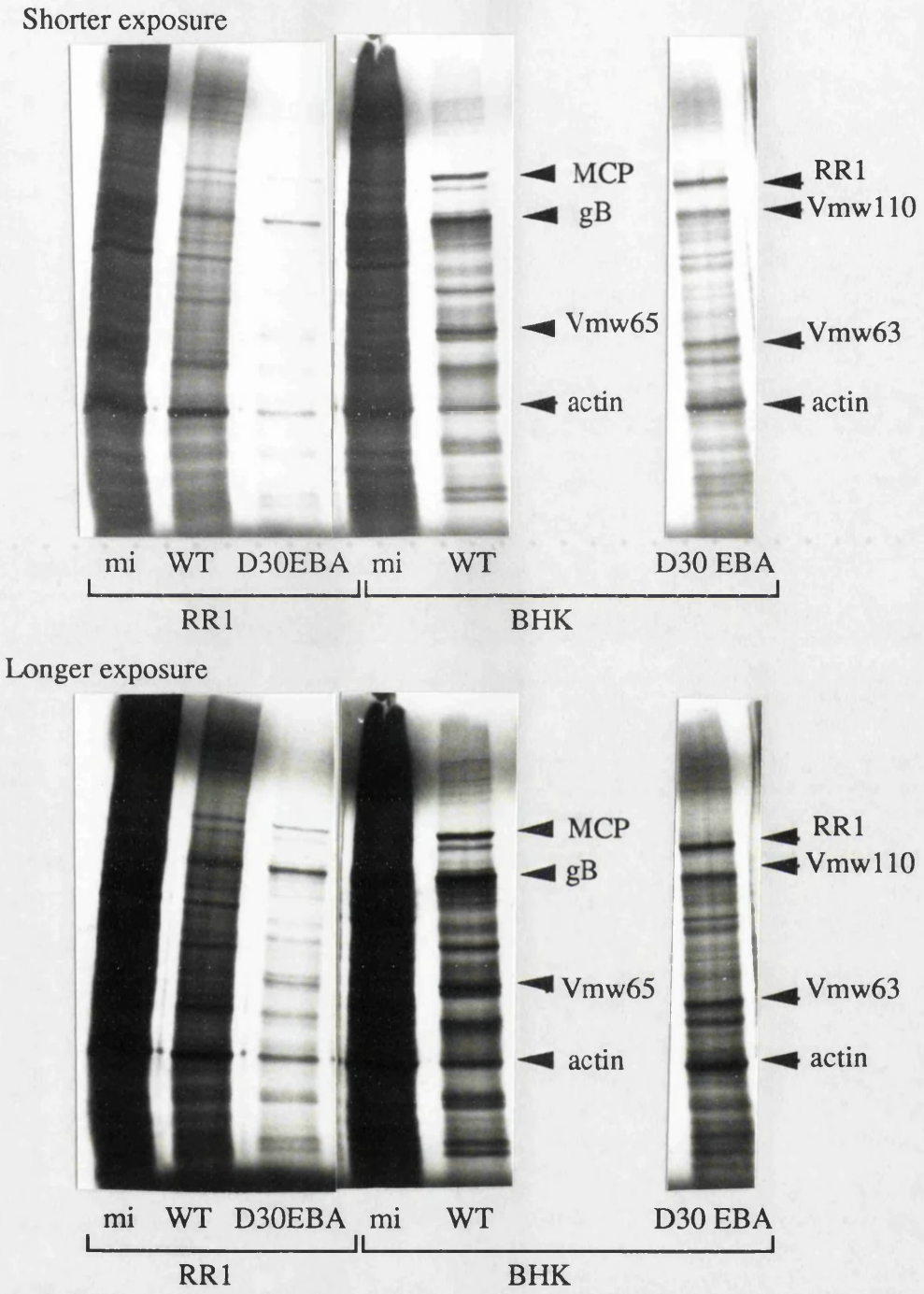


Figure 3F.1. Pulse Labelling of Polypeptides Synthesized During the Infection of Complementing and Non-Complementing Cell Lines with Wild Type and Mutant Strains of HSV-1.

10mm limbro wells were seeded with 1 x 10^5 RR1 or BHK cells per well and infected the following day at a m. o. i. of 5 pfu/cell. The infection was allowed to proceed for 16h. at 37°C and the pulse labelling of polypeptides was carried out as described in Section 2B.6.1. Extracts were electrophoresed on a 12.5% SDS--polyacrylamide gel, and polypeptides detected by autoradiography of the dried gel.
The positions of several polypeptide products of late and immediate-early HSV-1 genes are indicated adjacent to lanes 5 and 6, respectively.

3F.2. Construction of Plasmids Containing Mutations Within the IE-3 Coding Region.

The construction of a series of insertional mutations within the IE-3 coding sequence of the plasmid p175 has been described (Paterson, 1989). In order to study the effects of the point and in-frame deletion mutations on the *in vivo* properties of Vmw175, these mutations were also introduced into the plasmid p175.

The strategy for the construction of a series of plasmids, the pGL series, containing various point and deletion mutations within the IE-3 coding sequence, is represented diagrammatically in Figure 3F.2. Firstly, site directed mutagenesis techniques were used to remove the BamHI restriction site situated at nucleotide 1725 of the IE-3 coding region within the plasmid pI11. The plasmid pI11 is based on the plasmid p175 and contains a four amino acid insertional mutation and unique EcoRI restriction site at codon 292 of the IE-3 open reading frame. A single base pair change was introduced within pI11 which resulted in the removal of the BamHI restriction site at nucleotide 1725 without introducing an alteration in the IE-3 amino acid sequence. In order to carry out the mutagenesis, the SstI/EcoRI fragment encompassing nucleotides 876 to 2044 of the IE-3 reading frame was cloned into the vector M13 mp18. Single stranded DNA prepared from this vector was used as a template for site directed mutagenesis. The oligonucleotide LA20 (5' GCCTCCAGGATGCCGCGGCA 3')was used to introduce the base pair change, and mutagenesis was carried out using the Amersham *in vitro* mutagenesis system as described in Section 3B. Resultant M13 DNA obtained from the mutagenesis reaction was transfected into competent TG1 bacteria. The clones were screened by restriction enzyme digestion for the loss of the 155bp BamHI-BamHI restriction DNA fragment and putative mutant clones were further analysed by dideoxy sequencing reactions. The BamHI/SstI fragment from possible mutant constructs was cloned into the M13 mp18 vector and the subcloned fragment sequenced with universal M13 primer and Sequenase™ T7 DNA polymerase, as described in Section 2B.21. DNA ladders were electrophoresed on 8% denaturing acrylamide gels and the sequence compared to the wild type IE-3 coding sequence. Figure 3F.3 illustrates this mutagenesis and shows representative sequence analysis of a mutant clone containing the G to C mutation at nucleotide 1722.

Once the mutation had been introduced into the M13 vector, the mutant SstI/EcoRI DNA fragment was cloned into pI10 to yield the plasmid pLI10I11. The plasmid pI10 was digested with SstI in the presence of 40µg/ml ethidium bromide to yield a high percentage of singly cut, linear molecules. Linear plasmid DNA was isolated from non-denaturing agarose gels, as described in Section 2B.12.1 and digested with EcoRI. Approximately 400ng of this digested DNA was ligated to

approximately 400ng of the mutant EcoRI/SstI resriction fragment. The ligation mix was transfected into competent DH5 bacteria and resultant clones screened by restriction enzyme digestion for the presence of three SstI retriction sites and the loss of the 155bp BamHI-BamHI restriction fragment.

In order to construct the pGL plasmids, the pLI10I11 plasmid was digested with BamHI and PstI and the 6800bp DNA restriction fragment isolated from a non-denaturing agarose gel. The 2400bp NruI/PstI restriction fragment was isolated in the same way from plasmid p175 and the 600bp NruI/BamHI fragments were isolated from plasmids containing the mutations of interest. The three restriction fragments were ligated and transfected into competent DH5 bacteria. Resultant clones were screened by restriction enzyme analysis in order to isolate the clones containing all three restriction fragments. Positive clones were also checked for the absence of the BamHI site at nucleotide 1725 and the absence of an I10 or I11 EcoRI restriction site within the IE-3 coding sequence. Positive clones identified by restriction enzyme digestion were then subjected to dideoxy sequence analysis with the LA19 primer, as described in Section 3B.4.1, in order to verify the presence of the point or deletion mutation.

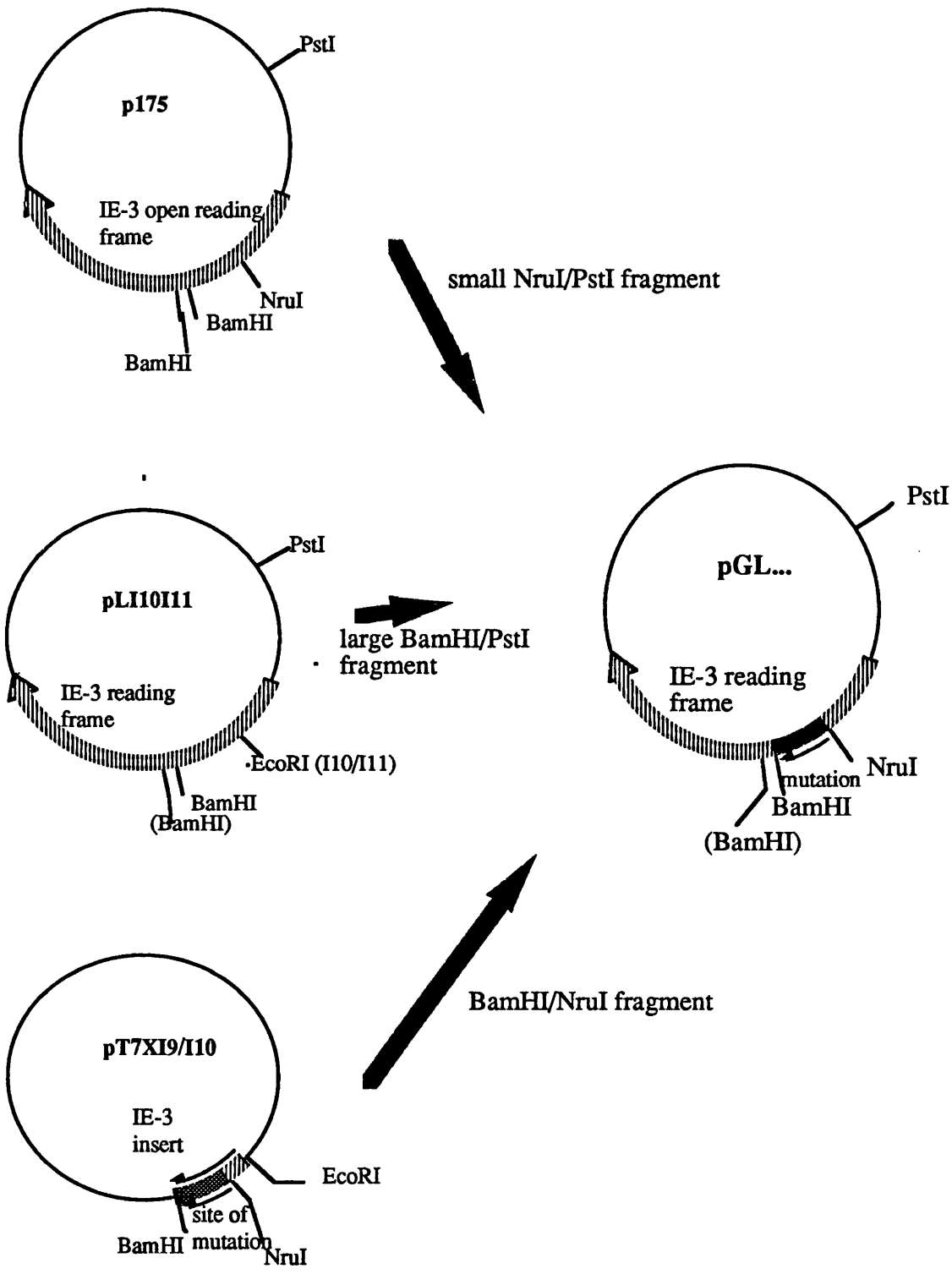


Figure 3F.2. Strategy for the Construction of pGL Plasmids.

The plasmids p175 and pLI10I11 are described in the text. Plasmid pT7XI9/I10 represents the T7 expression plasmids containing the point base pair or deletion mutation of interest. pGL plasmids are constructe by ligation of the DNA restriction fragments indicated.

(a) Sequence of original pI10 plasmid.

C R G I L E A

TGC CGC GGG ATC CTG GAG GCG

BamHI

(b) Sequence resulting from mutagenesis with LA20.

C R G I L E A

TGC CGC GGC ATC CTG GAG GCG

(c)

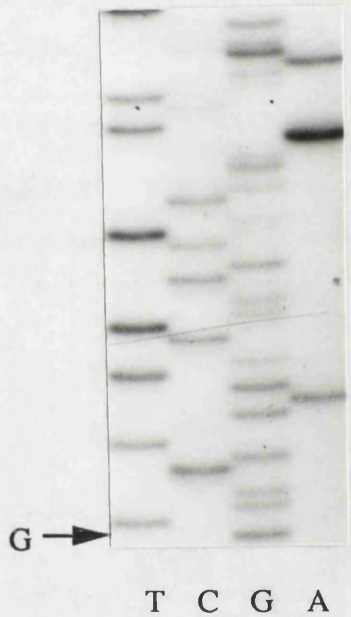


Figure 3F.3. Removal of a BamHI Site from the IE-3 Open Reading Frame by Site-Directed Mutagenesis.

(a) The sequence of the IE-3 open reading frame spanning codons 572-578 (nucleotides 1714-1734), showing the position of the BamHI site.
(b) The above sequence showing the base pair change introduced (in outline).
(c) Representative sequence analysis of a clone containing the base pair mutation. The clone was sequenced by the dideoxy chain termination method, using Klenow DNA polymerase, as described in Section 2B.21. Radiolabelled fragments generated by the sequencing reaction were electrophoresed on a denaturing 8% acrylamide gel, and fragments visualized by autoradiography. The sequencing reaction appeared to generate a lower proportion of fragments terminated at C residues. Therefore, bands appearing in the C track represent doublets created by two adjacent C residues.
The sequence shown, starting from the arrowed G and proceeding upwards, reads: 5'GTG GCC GTC GAG TGC CTG GCC GCC TGC CGC GGC ATC CTG GAG GCG

Mutated BamHI site

3F.3. Marker Rescue of the Vmw175 Deletion Mutant HSV-1 Virus D30EBA.

The strategy utilized to construct the recombinant viruses involved the marker rescue of the virus D30EBA by cotransfecting in non-complementing cell lines viral DNA with linear plasmid DNAs containing wild type or mutant IE-3 coding sequences. Plasmid DNA was linearised by digestion with PstI and cotransfected with HSV-1 D30EBA viral DNA into BHK cells by the calcium phosphate precipitation method as described in Section 2B.7. 3-4 days after the transfection, when viral cytopathic effect was visible, cells were scraped into the medium and sonicated to release viral paricles. The virus stock was then titrated on BHK cells. Initially, the cotransfection was carried out with wild type p175 DNA or control pUC19 DNA. Cotransfections with p175 DNA would indicate whether the IE-3 coding sequences were necessary and sufficient for the rescue of the deletion mutant D30EBA. Cotransfections with pUC19 DNA would indicate whether the D30EBA viral preparation was contaminated with wild type HSV-1 17^+ DNA, due to recombination between D30EBA sequences and wild type IE-3 coding sequences during the propagation of the virus on RR1 cells, as described in Section 3F.1.

Table 3F.1 shows the titres of HSV-1 virus obtained from transfections of D30EBA DNA with either p175 or pUC19 DNA on 50mm plates seeded at subconfluence with BHK cells. Marker rescue with p175 consistently yielded titres of around 10^5-10^7 pfu/ml of wild type HSV-1, indicating that the IE-3 coding sequences were necessary and sufficient for the replication of the virus D30EBA on BHK cells. In contrast, marker rescue with pUC19 DNA consistently resulted in no detectable virus growth on subsequent titrations; that is, any wild type virus produced from the cotransfection had a titre of 10^2 pfu/ml or less. Therefore, any contaminating wild type virus present within the D30EBA stock was present at such a low level that it could be assumed that viral replication detected during cotransfection experiments with mutant Vmw175 sequences was due to the replication of the recombinant virus containing mutant IE-3 coding sequences, rather than the proliferation of any wild type contaminating genome.

Co-transfection experiment number.	Titre of progeny virus obtained from co-transfection with p175	Titre of progeny virus obtained from co-transfection with pUC19.
1	$7x10^5$pfu/ml	$<1x10^3$pfu/ml
2	$5x10^6$pfu/ml	$<1x10^3$pfu/ml
3	$6x10^5$pfu/ml	$<1x10^2$pfu/ml
4	$7.5x10^7$pfu/ml	$<1x10^2$pfu/ml

Table 3F.1. Marker Rescue of the HSV-1 Deletion Mutant D30EBA.

Marker rescue was carried out by cotransfecting viral DNA from the HSV-1 Vmw175 deletion mutant D30EBA with linearized p175 or pUC19 DNA into BHK cells. 3-4 days post transfection, transfected cells were scraped into the medium and sonicated to release virus particles. Progeny virus was then titrated on BHK cells. The above results were obtained from four consecutive transfection experiments.

Representative titres from cotransfection experiments with mutant IE3 genes are given in Table 3F.2. These illustrate the wide variability of titres obtained with certain mutant p175-based plasmids that can support viral growth in tissue culture.

Cotransfection Experiment Number.	Cotransfected DNA	Titre of Progeny Virus.
1	pI9	$3x10^5$pfu/ml
1	pI10	$2x10^3$pfu/ml
1	pI11	$1x10^3$pfu/ml
2	pI9	$5x10^5$pfu/ml
2	pI10	$1.2x10^6$pfu/ml
2	pI11	$1.5x10^4$pfu/ml
3	pI9	$3.5x10^5$pfu/ml
3	pI10	$1x10^4$pfu/ml
3	pI11	$8x10^4$pfu/ml

Table 3F.2. Representive Titres from Consecutive Rounds of Cotransfection Experiments.

Cotransfection experiments were carried out with viral DNA from the Vmw175 deletion mutant D30EBA and linearized plasmids containing mutant forms of the IE-3 reading frame, as indicated. Transfections were carried out on BHK cells. 3-4 days post transfection, cells were scraped into the medium and viral particles released by sonication. Progeny virus obtained from the cotransfection experiment was titrated on BHK cells. The above results were obtained from three consecutive cotransfection experiments.

3F.4. Discussion.

Davidson and Stow (1985) initially reported that propagation of HSV-1 IE-3 mutant viruses on recombinant cells containing genomic HSV origin of replication sequences led to noticeable recombination events betwen HSV sequences in the cellular genome and sequences within the mutant HSV-1 viruses. For example, after a 16 hour infection at the non permissive temperature with the virus tsK, (which contains a temperature sensitive mutation within the IE-3 gene), on cell lines containing the HSV-1 origin of replication, "a" and Vmw175 coding sequences, progeny virus stocks were contaminated with approximately 1.3% wild type HSV-1. Davidson and Stow proposed that this recombination was due to the propagation of wild type Vmw175 sequences linked to the HSV-1 origin of replication within the cellular genome upon infection, followed by recombination between mutant and wild type IE-3 coding sequences. M65 cell lines infected in the same way did not yield recombinant wild type virus in the progeny stocks; Davidson and Stow suggested that this lack of recombination was due to the fact that M65 cells did not contain HSV-1 origin of replication sequences. However, Paterson (1989) reported that propagation of the virus D30EBA in M64A cell lines resulted in contamination with wild type virus at 0.001-0.01%, and that the passage of recombinant viruses containing mutant IE-3 coding sequences in the M64A cell line resulted in the emergence of wild type recombinant viruses. The recombination events described by Paterson (1989) were suggested to be due to recombination via highly repetitive "a" sequences contained within the genome of M64A cells. The RR1 cell line was created by transformation of BHK cells with DNA fragments containing only Vmw175 coding sequences, and did not contain either the HSV origin of replication or the "a" sequence mentioned above. This may explain the lower rate of emergence of wild type recombinant viruses observed during the propagation of the deletion mutant HSV-1 virus D30EBA on RR1 cell lines.

Table 3F.2 illustrates the variability of virus titre obtained from repeated cotransfection experiments with plasmids carrying mutant IE-3 coding sequences which are able to support virus growth in tissue culture. This variability illustrates that the yield of recombinant virus was not necessarily representative of the ability of the mutant Vmw175 molecule to support viral growth in tissue culture. The only conclusion to be drawn from such cotransfection experiments, therefore, is whether the mutation sufficiently disrupts the function of the Vmw175 molecule such that it is unable to support the replication of HSV-1 in tissue culture.

The variability in virus yield could be due to several factors, such as inconsistencies in the efficiency of the transfection procedure or changes in the rate of recombination between D30EBA and mutant IE-3 coding sequences. However, titres of

viable recombinant viruses obtained with this protocol were sufficient for the further analysis of the growth of mutant viruses in tissue culture, as described in Section 3G.3.

The strategy employed for the construction of recombinant viruses enabled only a rough determination of the effects of the mutations on the *in vivo* properties of Vmw175. Mutant forms of Vmw175 could only be assayed for their ability to support viral growth in BHK cells. Any mutations leading to a gross disruption of the activity of the protein and therefore a non-viable recombinant virus could not be analysed further. Therefore, the strategy does not allow analysis of the properties of the mutant polypeptides, such as whether the mutation affects transactivation or autoregulation by the protein. However, comparison of the viability of mutant viruses with the characteristics of the mutations which have already been determined *in vitro* would yield information on the *in vitro* properties of the protein necessary for its full function in the viral lytic cycle.

3G. Characterization of Recombinant Viruses.

3G.1. The Ability of Mutant Vmw175 Molecules to Support Viral Growth in Tissue Culture.

As discussed in Section 3F, cotransfection experiments were carried out with viral DNA prepared from the Vmw175 deletion mutant HSV-1 virus D30EBA and plasmids carrying wild type or mutant IE-3 coding sequences. Cotransfections were carried out in BHK cells as described in Section 2B.7. This allowed for the determination of the ability of the mutant Vmw175 molecule to support viral growth on BHK cells, as measured by the production of a viable recombinant virus. Cotransfections were carried out at least twice for each mutant plasmid tested and each transfection experiment contained control transfections with wild type p175 DNA and pUC19 DNA. Results obtained from the control experiments were similar to those described in Section 3F.3.

Initially, cotransfection experiments were carried out with the pI series of plasmids, which contain four amino acid insertional mutations throughout the IE-3 open reading frame (Paterson and Everett, 1988a; 1988b; Paterson, 1989). Only plasmids containing mutations within the Vmw175 DNA binding domain (pI9-pI20)

were tested for their ability to support HSV-1 replication in tissue culture. These mutations had previously been assayed for their DNA binding, transactivation and autoregulation properties in transient transfection and gel retardation analyses (Paterson and Everett, 1988a; 1988b; Paterson, 1989). The results of these assays are presented in Table 3G.1, together with the results of the cotransfection experiments. This summary illustrates that the mutations which disrupt the functions of Vmw175 as measured *in vitro* when incorporated into the genome of HSV-1 result in the non-viability of the virus in tissue culture. The implications of this observation are discussed further in Section 3G.4.

Mutant Number	Site of Insertion (codon)	Trans-activation (a)	Auto-regulation (b)	DNA Binding Activity (c)	Relative Amounts of Vmw175 (d)	Viability of recombinant Virus (e)
I9	252	127(43.2)	+++	+++	263	+
I10	275	25.8(4.6)	-	++	142	+
I11	292	30.8(6.3)	++	+++	276	+
I12	310	20.0(4.4)	+	+/-	97	-
I13	324	10.0(2.9)	-	+++	226	-
I14	329	4.9(0.9)	-	-	38	-
I15	337	4.5(0.9)	+	-	18	-
I16	373	16.5(2.8)	-	-	18	-
I17	398	11.8(1.9)	-	+	39	-
I18	438	15.5(4.4)	-	+	65	-
I19	494	49.2(10.4)	+++	+++	170	N/D
I20	518	47.2(10.2)	+++	+++	85	+

Table 3G.1 Ability of Mutant Vmw175 Molecules Containing Insertional Mutations to Support HSV-1 Replication in the Non-Complementing BHK Cell Line.

(a) The activation of the gD promoter (using pgDCAT in conjunction with Vmw110 provided by p111) is given as a percentage of that obtained in parallel experiments with p175. Figures shown represent mean values from several experiments, with standard error of the mean given in brackets. Reproduced from Paterson, 1989.
(b) The qualitative ability to repress the IE-3 promoter of pIE3CAT compared to p175 in cotransfection titration experiments. Repression activity is scored from - (essentially no activity) to +++ (wild type activity). Reproduced from Paterson, 1989.
(c) The ability of nuclear extracts made from HeLa cells transfected with each of the mutants to bind to the IE-3 capsite. Binding activity is scored from visual inspection of autoradiographs. - no binding, +/- barely detectable, + to +++ increasing binding activity. Reproduced from Paterson, 1989.
(d) Relative amounts in the nuclear extracts used in the in vitro assays (a) to (c), determined by ELISA. p175 = 100. Reproduced from Paterson, 1989.
(e) The viability of recombinant viruses containing the mutant Vmw175 molecule in non-complementing cell lines, as determined by co-transfection experiments with the Vmw175 deletion mutant virus, D30EBA. N/D = not determined.

A series of plasmids containing point mutations and one plasmid containing an in-frame internal deletion within the DNA binding domain of Vmw175 were also tested in cotransfection experiments, as described above. These mutations had been previously assayed for their ability to disrupt DNA binding (data presented in Section 3D) and certain mutations had also been tested for their ability to disrupt dimerization (data presented in Section 3E). The introduction of these mutations into the viral

genome would provide a means for testing whether they disrupted Vmw175 function during the viral replicative cycle.

The mutations tested in the cotransfection experiments had all been shown to disrupt the DNA binding activity of the isolated DNA binding domain of Vmw175. Some point mutations and the deletion mutation had been resulted in the elimination of the DNA binding activity; other point mutations resulted in a temperature sensitive DNA binding phenotype (see Setion 3D). In order to determine whether the temperature sensitive DNA binding phenotype of certain point mutations was reflected in a temperature sensitive phenotype of the Vmw175 molecule when the mutations were introduced into the HSV-1 genome, cotransfections were carried out at 31°C and the resultant cell associated virus stocks were titrated on BHK cells at 31°C and 38.5°C. The results of these transfection experiments and the effects of the mutations on the DNA binding activity and, where applicable, the dimerization activity of the isolated DNA binding domain are summarized in Table 3G.2.

Mutant Number	Site of alteration (codon)	Type of alteration	DNA binding phenotype (a)	Dimerization phenotype (b)	Viability of recombinant virus (c)	
					31°C	38.5°C
LI	574	removal of BamHI site	N/D	N/D	+	+
DH1	445-453	del	no binding	N/D	-	-
463/4 KG/NC	463 464	Lys to Arg Gly to Cys	no binding	WT	-	-
459D/Y	459	Asp to Tyr	no binding	N/D	-	-
453A/S	453	Ala to Ser	TS	N/D	+	+
457R/L	457	Arg to Leu	TS	N/D	-	-

Table 3G.2. The Ability of Vmw175 Molecules Containing Point and Deletion Mutations to Support HSV-1 Replication in Non-Complementing Cell Lines.

(a) The DNA binding phenotype of DNA binding domain polypeptide containing the indicated mutation. DNA binding activity was determined by gel retardation analysis with a DNA probe containing the consensus DNA binding site situated at the IE-3 cap site. WT = DNA binding activity indistinguishable from that of a wild type DNA binding domain polypeptide. TS = DNA binding activity reduced at temperatures of 37°C and above. No binding = no binding activity at any temperature. N/D = not determined
(b) The dimerization phenotype of a DNA binding domain polypeptide containing the mutation of interest, as determined by gel filtration chromatography or glutaraldehyde cross-linking assays. N/D = not determined WT = dimerization properties indistinguishable from that of a wild type DNA binding domain polypeptide.
(c) The viability of recombinant viruses containing the mutant Vmw175 molecule in non-complementing cell lines, as determined by co-transfection experiments with the Vmw175 deletion mutant virus D30EBA.

Again, the results illustrate that the majority of mutations which disrupt the *in vitro* properties of the isolated DNA binding domain of Vmw175 when incorporated into the HSV-1 genome result in the non-viability of the recombinant viruses in non-complementing cell lines. The two temperature sensitive mutations 453A/S and 457R/L produced interesting results when introduced into the viral genome. Whilst one mutation resulted in a recombinant virus viable in BHK cells at both 31°C and 38.5°C, the second mutation resulted in a virus which was non-viable in a non-complementing

cell line. Section 3G.4 presents a discussion of possible reasons for these observations, and the implications of the results presented in Table 3G.2.

3G.2. Southern Blot Analysis of Recombinant Virus Genomes.

As the summary in Table 3G.1 shows, of the twelve insertional mutant Vmw175 coding sequences which were tested for the ability to support viral replication in tissue culture, only four mutant IE-3 genes were incorporated into viable recombinant viruses. In order to verify that the virus stocks obtained from these cotransfection experiments contained the mutations of interest, and were not contaminating wild type HSV-1, the viral DNA was subjected to Southern blot analysis.

Total infected cellular DNA was prepared from viral stocks as described in Section 2B.5. The DNA was digested with EcoRI and BamHI at 37°C for 16-20 hours and DNA fragments electrophoresed on 1.5% non-denaturing agarose gels at 8V/cm. The DNA fragments were transferred onto a solid membrane as described in Section 2B.25. Initial Southern blot analyses were carried out using nitrocellulose membranes. These were found to lead to a low efficiency in the detection of fragments spanning the IE-3 coding region, as previously noted by others (Davidson and Stow, 1985; Clements *et al.*, 1979). This low efficiency of detection was thought to be due to the high G + C content of the IE-3 coding sequences, and was improved by the use of Hybond N^+ charged membrane. Restriction fragments were probed by incubation with a radiolabelled BamHI "y" HSV-1 DNA restriction fragment, as described in Section 2B.26. The probe DNA was radiolabelled to a high specific activity using the random prime method, as detailed in Section 2B.24.3.

Southern blot analysis of the wild type HSV-1 genome in this way led to the identification of the uncut BamHI "y" fragment, which spans the 5' portion of the IE-3 open reading frame (see Figure 3G.1a, tracks 14, 18 and 24). The three mutations incorporated into the IE-3 open reading frame in recombinant HSV-1 virus genomes result in the incorporation of unique EcoRI restriction sites within the BamHI "y" fragment. As each mutation specifies a unique site at various positions within the IE-3 reading frame, digestion of the viral genome with EcoRI and BamHI and identification of the DNA fragments which hybridize to the radiolabelled BamHI "y" restriction fragment led to the detection of a pair of restriction DNA fragments characteristic of the particular insertional mutation. Figure 3G.1 shows Southern blot analysis of virus stocks obtained from multiple cotransfection experiments with the plasmids pI9, pI10 and pI11, and illustrates the detection of the two unique EcoRI/BamHI restriction fragments. Longer exposure of these Southern blots in principle would enable the detection of any contaminating wild type genomes present in the mutant virus stocks.

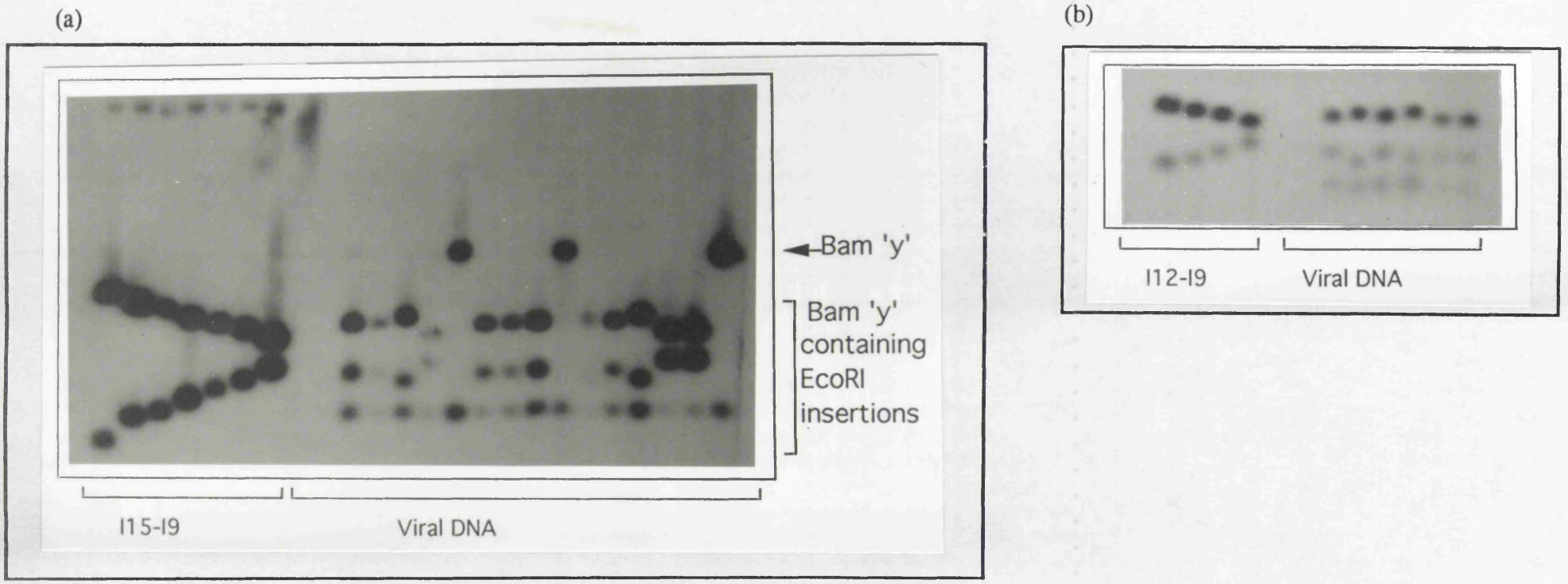


Figure 3G.1. Southern Blot Analysis of Viral DNA Obtained from Cotransfection Studies.

10mm limbro wells seeded with 5 x 10^5 BHK cells per well were infected with virus stocks at a m. o. i. of 5 pfu/cell. Infections were carried outt at 37°C for 2-3 days until viral cytopathic effect was observed. Total infected cellular DNA was prepared as described in Section 2B.5. The DNA was digested with EcoRI and BamHI at 37°C for 16-20 hours and DNA fragments electrophoresed on 1.5% non-denaturing agarose gels. The DNA was transferred to Hybond N^+ membrane, denatured and probed by incubation with a radiolabelled BamHI "y" restriction fragment probe as described in Section 2B.26.

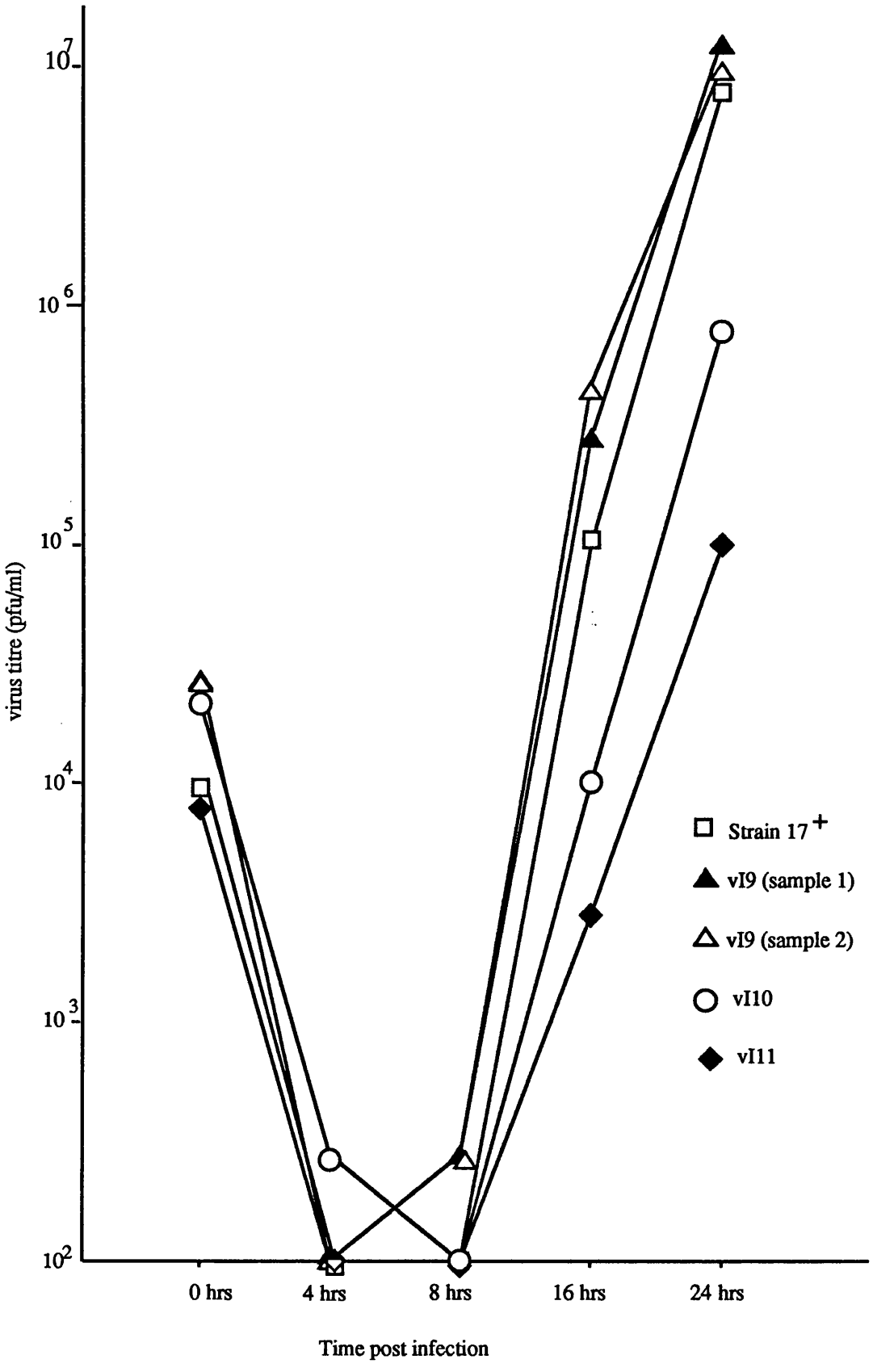


Figure 3G.2. Growth Curve Analysis of the HSV-1 Virus Strains 17syn$^+$, vI9, vI10 and vI11.

Legend to Figure 3G.2.

Linbro wells seeded at 10^5 BHK cells / well were infected with each viral strain at a multiplicity of infection of 0.01pfu/cell. Infections were carried out at 37°C for 24 hours. At the time points indicated, cells were scraped into the medium and sonicated to release viral particles. Progeny virus was titred on BHK cells. The graph represents the progeny virus titre for each virus strain at the time points indicated.

Stocks of viruses containing the I9-I11 mutations (vI9-vI11 respectively) had no detectable contaminating wild type HSV-1; they were analysed further as described in Section 3G.3.

3G.3. Growth Curve Analysis of the HSV-1 Recombinants vI9, vI10 and vI11.

The viruses vI9, vI10 and vI11, containing four amino acid insertional mutations at codons 252, 275 and 292 of the IE-3 open reading frame, were analysed for their rate of replication on BHK cells. Viruses were analysed by the determination of the titre of progeny virus obtained at specific time points during a 24 hour viral infection. 10mm linbro wells seeded with BHK cells were infected at a multiplicity of infection of 0.02 pfu/ cell. The infection was allowed to proceed at 37°C for 24 hours and at specific time points during the infection, cells were scraped into the medium and sonicated to release infectious virus particles. Resultant virus stocks were titrated on 10mm linbro wells seeded with BHK cells. The titres obtained over a 24 hour infection with HSV-1 strains 17 syn^+, vI9, vI10 and vI11 are shown in Figure 3G.2.

All four viral infections showed a drop in virus titre between 2-4 hours post infection. This is the eclipse phase of viral infection during which viral genomes are undergoing transcription and DNA replication prior to the production of infectious particles. In the period 6-24 hours post infection, during which time progeny viruses are accumulating in the cell or in the extracellular medium, distinctions can be made between the growth cycles of the four viruses. At 24 hours post infection, the virus vI9 yields progeny virus of a comparable titre to that obtained with wild type HSV-1 (see Figure 3G.2). However, titres of progeny virus obtained from a 24 hour infection with vI10 or vI11 are 10 and 100- fold lower, and accumulation of infectious vI10 and vI11 particles occurs at comparatively lower rates.

3G.4. Discussion.

A striking point of the results obtained with the cotransfection experiments described in Section 3G is that mutations that have previously been shown to disrupt the *in vitro* properties of Vmw175, when incorporated into the HSV-1 genome result in Vmw175 molecules incapable of complementing the deletion mutant D30EBA in tissue culture. This suggests that the conclusions drawn from previous *in vitro* assays

are valid, and that the measurement of the *in vitro* properties of Vmw175 gives a close approximation to the actual *in vivo* activity of the molecule.

In the studies of the insertional mutant Vmw175 molecules, one case does not entirely fit this hypothesis. The mutant Vmw175 molecule I13, which is apparently functional in DNA binding as measured by *in vitro* analysis (see Table 3G.1), is unable to support the growth of HSV-1 in tissue culture. However, two additional observations suggest that the I13 molecule is functionally defective. Firstly, *in vitro* analysis showed that the protein had reduced transactivation properties (see Table 3G.1). Secondly, the I13 mutation when introduced into a Vmw175 DNA binding polypeptide eliminated the DNA binding activity of the molecule (R. D. Everett, personal communication). Therefore, although the results obtained from the *in vitro* analysis of the Vmw175 I13 molecule suggest otherwise, it appears that the I13 molecule is functionally defective, and therefore cannot support the replication of HSV-1 in tissue culture.

Additionally, results from *in vitro* analysis suggested that the Vmw175 I10 molecule may be defective in autoregulation (see Table 3G.1). This mutation when introduced into the viral genome resulted in a viable recombinant virus. It would have been interesting to investigate the kinetics of Vmw175 expression during vI10 infection in BHK cells, in order to determine whether the protein does actually autoregulate expression of Vmw175 during the viral life cycle, and whether this function is essential for replication of HSV-1. However, time did not allow for this experiment to be done.

Previous experiments with the I9 and I10 DNA binding domain polypeptides have shown that the I9 polypeptide but not the I10 polypeptide can produce a DNAaseI footprint at the IE-3 capsite (R. D. Everett, unpublished observations). This situation is analogous to that observed with DNA binding domain polypeptides of the homologous VZV 140K protein (J. Tyler, personal communication). Therefore, it has been suggested that the region between the I9 and I10 insertions (spanning codons 252-275) is important for the stabilization of protein-DNA interactions. The observation that Vmw175 may bind to certain DNA probes in gel retardation analysis but not footprint the same sequences has been described by Papavassiliou and Silverstein (1990), and these authors also suggested that the loss of footprinting ability of Vmw175 was due to a reduction in the stability of the protein-DNA interactions. It is possible that the disruption of the autoregulation property of Vmw175 by the I10 insertion results from a reduction in the interaction of the whole Vmw175 molecule containing the I10 mutation with DNA, and that stable interactions are required for the autoregulation function, a theory also proposed by Roberts *et al.* (1988). The recombinant viruses vI9 and vI10 could be used to express Vmw175 molecules containing the I9 and I10 mutations for *in vitro* DNA footprinting experiments in order to investigate the footprinting activity of the Vmw175 I10 protein.

The results of the growth curve analysis of the vI9, vI10 and vI11 viruses suggests that the ability of the mutant molecules to support viral replication is closely related to the *in vitro* assay of their ability to transactivate, as summarized in Table 3G.1. The I10 Vmw175 molecule exhibits a slightly reduced transactivation activity and the I11 Vmw175 molecule exhibits a further reduction in transactivation as compared to wild type protein, and the ability of the recombinant viruses to replicate also follows this trend (see Figure 3G.1). It is clear from this data that as the mutations move further towards the centre of the DNA binding domain region, they exhibit a more disruptive effect on the functions of Vmw175.

With the point mutations and internal in-frame deletion mutations which have previously been shown to disrupt the function of Vmw175 *in vitro,* the correlation between their *in vitro* and *in vivo* activity is even more striking. The results of the cotransfection experiments show that the mutations also affect the properties of the whole Vmw175 protein in the context of a full viral infection. Because the mutations fail to complement the deletion mutant HSV-1 virus D30EBA, it could be predicted that they affect the transactivation and possibly the autoregulation properties of Vmw175. As the recombinant viruses containing these mutations are non-viable, they cannot be used as expression vectors for the mutant Vmw175 molecules. However, it would be possible to use the pGL series of plasmids in transient transfection and gel retardation experiments previously described by Paterson and Everett (1988a; 1988b) to study the effects of the mutations on the *in vitro* properties of Vmw175 and confirm that they correlate to their complementing activity, a situation observed with the insertional mutations.

With the two temperature sensitive point mutations tested, the results are not so easy to interpret. Whilst both mutations resulted in identical phenotypes upon *in vitro* analysis, incorporation of the mutations into the viral genome produced both a viable and a non-viable HSV-1 recombinant virus. It is clear that both temperature sensitive mutations somewhat disrupt the functions of Vmw175, but it is difficult to explain how one such disrupted molecule is rescued *in vivo* and can support the viral replicative cycle whilst a second mutant fails to do so. Interestingly, the temperature sensitive mutation which fails to support the viral life cycle is situated within the proposed turn of the "helix-turn-helix" motif, as predicted by Shepard *et al.* (1989). Whilst it is not clear that the secondary structure of the protein follows these predictions (see Section 3A), the mutations which most drastically affected the DNA binding activity of Vmw175 lay close to or within this region of the protein (see Section 3D.8). Therefore, it is possible that the mutation 457R/L has such a deleterious effect on the Vmw175 molecule that it is not able to support the transactivation of viral genes during HSV-1 infection. In contrast, the mutant Vmw175 molecule 453A/S, when in the cellular environment, may be stabilised by the interaction with other cellular or viral factors,

allowing it to function fully at higher temperatures. It would be interesting to determine the *in vitro* transactivation properties of these two mutant Vmw175 proteins, to test whether they correlate to the ability of the protein to support viral growth.

Other point mutations shown to result in a temperature sensitive DNA binding phenotype when expressed as isolated DNA binding domain polypeptides would possibly result in a temperature sensitive recombinant virus. This situation has already been shown to occur - the mutation tsK, which results in a temperature sensitive HSV-1 virus, consisits of an alanine to valine substitution at codon 475 (Davison *et al.*, 1984), and this mutation when cloned and expressed as an isolated DNA binding domain polypeptide exhibits a temperature sensitive DNA binding phenotype (see Section 3D).

These results point to the importance of region 2, and in particular the "helix-turn-helix" region within the Vmw175 molecule, and suggest that the intrinsic stability of this region is essential for the full function of the protein. The point mutations analysed in this section, with the exception of the mutant 453A/S, must result in a sufficient disruption of the "helix-turn-helix" motif to affect the function of the Vmw175 molecule.

The question of the contribution of the DNA binding activity of Vmw175 to its full range of functions *in vivo* is a controversial and complex topic. Amongst the wide range of Vmw175 mutant molecules studied, few mutations have been identified which result in the elimination of the DNA binding activity of the Vmw175 molecule as measured *in vitro* whilst showing little effect on the ability of Vmw175 to transactivate and autoregulate as measured by transient transfection experiments, or to support the viral replicative cycle (Paterson and Everett, 1988a; 1988b; Paterson *et al.*, 1990; DeLuca *et al.*, 1984; DeLuca and Schaffer, 1987; 1988; Shepard *et al.*, 1989; Shepard and DeLuca, 1991b). Only one mutation, the ri12 mutation (Shepard and DeLuca, 1991b), results in a Vmw175 molecule apparently exhibiting no DNA binding activity *in vitro* at any temperature, but fully able to support the viral replicative cycle in tissue culture. It should be remembered that the mutant virus from which this pseudo-revertant is derived, the virus vi12, contains a Vmw175 molecule exhibiting no *in vitro* DNA binding activity and is in turn unable to support viral growth (Shepard *et al.*, 1989).

Several temperature sensitive HSV-1 viruses which contain point mutations within the C-terminal portion of Vmw175 have been isolated (Paterson *et al.*, 1990; DeLuca *et al.*, 1984). Vmw175 molecules from these viruses exhibit a temperature sensitive defect in DNA binding activity *in vitro* but are partially permissive for early gene expression at the non-permissive temperature. However, they are unable to support the full viral replicative cycle. For a full interpretation of these results, it must be remembered that the DNA binding properties of Vmw175 are measured in artificial

situations which cannot reproduce the full cellular environment in which Vmw175 functions, and that the protein molecule may be stabilised by interactions with other cellular proteins, enabling the mutant molecule to overcome its defect in DNA binding activity. Some data exists to support this model of the mechanism of Vmw175 transactivation within the cell. This data also questions the validity of previous conclusions made about the ri12 mutation, and is discussed further in Chapter 4.

The cotransfection experiments presented in this thesis contain no examples of DNA binding deficient mutant Vmw175 molecules able to support viral growth. In fact, these results strongly suggest that the ability of Vmw175 to bind to DNA correlates closely with the ability of the protein molecule to function fully during the viral replicative cycle. Although the insertional mutation I13 appears to result in a binding proficient Vmw175 molecule that cannot support viral replication, analysis of the mutation within the I10X DNA binding domain polypeptide has suggested that the mutation may actually have a disruptive effect on this region of the protein (Everett, unpublished observations). Therefore, the majority of the data shows the importance of the integral stability of region 2 and the "helix-turn-helix" motif, and the DNA binding activity of Vmw175 for the full *in vivo* function of the protein.

Chapter 4. Discussion.

4A. Structural Motifs Within Vmw175.

As the work presented within this thesis primarily involves the study of the DNA binding activity of Vmw175, the structural analyses discussed here concentrate on the DNA binding domain of the protein.

As shown in Figure 3A.1, the DNA binding domain of Vmw175 is well conserved between the transcriptional regulatory proteins of the α-herpesviruses. This is not surprising, since the domain encompasses region 2, which was defined by McGeoch *et al.* (1986) as a conserved region between Vmw175 and the 140K protein of VZV. The highest conserved portion of the DNA binding domain encompasses amino acids 445-487, which has been suggested to encode a helix-turn-helix secondary structure motif (Shepard *et al.*, 1989), and is the region under study in this thesis. The high conservation of this portion of the protein suggests that this sequence plays an important role in the *in vivo* properties of Vmw175. As it is within the DNA binding domain of the protein, it is logical to suppose that the sequence plays a part in the DNA binding activity of the molecule, rather than mediating the transactivation function, or being a region modified by a post-translational reaction. *In vitro* studies have also illustrated that the DNA binding domain of Vmw175 exists as a dimer in solution (Everett *et al.*, 1991), therefore, the conserved region may also contain a dimerization motif.

As discussed in section 3A.6, the region encompassing amino acids 445-487 exhibits trivial homology to other well conserved protein DNA binding motifs. This could imply that the sequence, if it does encompass a DNA binding motif, represents a new, evolutionarily diverged motif. Two classes of protein DNA binding motifs, whilst consisting of diverse amino acid sequence, still retain a helix-turn-helix secondary structure motif, confirming that this structure is well suited to interactions with the DNA helix. In fact, virtually all DNA binding motifs, with exceptions such as the β barrel of TFIID (Nikolov *et al.*, 1992) contain helical regions, which track into the grooves of the DNA helix, and thus mediate specific protein-DNA contacts. In this light, it is interesting to note that the region between amino acids 445-457 of Vmw175 exhibits high predictions of α-helical structure, as do the relevant regions of other α-herpesvirus regulatory proteins.

It would also be of interest to investigate the functional relevance of sequences encompassing and adjoining amino acids 390-395 of Vmw175. This five amino acid sequence shows homology to evolutionarily conserved and functionally significant amino acids within the recognition helix of the homeobox motif. The rest of the homeobox consensus motif exhibits little homology to Vmw175; however, it is possible that amino acids 390-395 of Vmw175 contain an evolutionarily diverged version of the homeobox DNA recognition helix. This possibility could be investigated by further mutational analysis.

4B. The Effects of Point Mutations and Deletions on the *In vitro* Properties of Vmw175.

The method chosen to mutate the region of interest within Vmw175 limited the amino acid changes that could be introduced. Therefore, specific changes to amino acid side chains possibly involved in direct contacts between protein and DNA molecules could not be introduced. This leaves the possibility that amino acids that are involved in contacting the DNA molecule have not been detected, because some of the amino acid changes introduced are relatively conservative and might not disrupt protein-DNA contacts. Mutations created during the work presented in this thesis have been analysed both in the context of a bacterially-expressed DNA binding domain polypeptide *in vitro* and within the viral genome.

4B.1. Effects on DNA Binding Activity.

The results presented and discussed in section 3D illustrate the effects of the mutations upon the *in vitro* DNA binding activity of bacterially expressed DNA binding domain polypeptides of Vmw175. They are in agreement with previous insertional mutational analyses (Shepard *et al.*, 1989; Paterson and Everett, 1988a, b; Paterson, 1989). These results are significant in that they show that point mutations can disrupt the DNA binding properties of the Vmw175 DNA binding domain polypeptide. With all mutational analyses, there is the possibility that the effects seen are due to disruption of the secondary and tertiary structure of the protein, rather than the specific result of changing the amino acid sequence. Reducing the size of the mutation

introduced lowers the possibility of tertiary structure disruption occurring. Therefore, it is important that point amino acid mutations can eliminate the DNA binding properties of Vmw175 polypeptides *in vitro*. It would also be informative to investigate the effect of these mutations on the DNA binding properties of the whole protein.

The general conclusion of this study is that the mutations most deleterious to DNA binding activity are situated between amino acids 456-464, around the turn of a proposed helix-turn-helix motif encompassing amino acids 445-487 of Vmw175. This conclusion does not correlate with the fact that the majority of protein-DNA interfaces are α-helical. There is the possibility that the secondary structure proposals for Vmw175 are incorrect. If the suggested secondary structure does exist, there is still the possibility that the point mutations are destabilizing this region rather than inhibiting specific DNA contacts. In fact, whilst mutations within the region encompassing amino acids 456-464 abolish DNA binding activity, there is little evidence to suggest that the conserved tract between amino acids 445-487 is involved in contacting the DNA helix. These questions can only be answered by further structural studies. However, the results do point to a functional significance for this region of Vmw175.

The purpose of introducing in-frame internal deletions within this region was to investigate whether proposed α-helices are important for the DNA binding activity of the DNA binding domain. Introduction of such deletions abolished DNA binding activity, suggesting that the helices do in fact play some role in determining contact between Vmw175 and DNA. However, as mentioned above, it must be remembered that such mutations can disrupt the tertiary conformation and stability of the protein. Whilst such deletion mutant polypeptides were obtained in sufficient quantities to carry out DNA binding analyses, the polypeptides were relatively insoluble and quantities obtained were insufficient to carry out dimerization assays. These observations suggest that the deletions may have had some effect on the conformation and stability of the polypeptides. In order to investigate the function of proposed α-helical regions of the domain, insertion of proline residues within helices may lead to disruption of the helical structure whilst lessening the effect on overall conformation of the polypeptide.

Previous analyses of Vmw175 have suggested that multiple regions of the DNA binding domain function in the whole protein (Paterson and Everett, 1988a, b; Paterson, 1989; Shepard *et al.*, 1989; Shepard and DeLuca, 1991b). Whilst this study does not disagree with these conclusions, it highlights the importance of the carboxy-terminal region of the DNA binding domain of Vmw175.

An aspect of the mutant Vmw175 polypeptides not investigated within this study is the ability of the molecules to bind to DNA containing non-consensus Vmw175 binding sites, such as the sites within the thymidine kinase promoter / leader sequences (Papavassilliou and Silverstein, 1990). Shepard *et al.* (1989) used the study

of such non-consensus sites to suggest that certain mutations may affect the specificity of DNA binding by Vmw175. Whilst it can be assumed that mutations which abolish the ability of Vmw175 polypeptides to bind to consensus sites also abolish the ability of the protein to bind to non-consensus sites, it would be interesting to investigate the binding properties of mutant polypeptides which exhibit lowered DNA binding activity. Would such mutant polypeptides, such as the temperature sensitive mutant polypeptides, be able to bind to non-consensus sites under permissive or restrictive conditions?

4B.2. Effects on Dimerization.

The ability of Vmw175 to dimerize and the proposal that dimerization is a prerequisute for its DNA binding activity seems to be in contrast with studies defining a consensus DNA binidng site for the protein. All consensus sites proposed for Vmw175 exhibit asymmetry (DiDonato *et al.*, 1991; Everett *et al.*, 1991), whilst the majority of DNA binding proteins which bind to DNA as dimers bind to sites which exhibit palindromic symmetry (see, for example, Vinson *et al.*, 1989). Whilst preliminary results suggest that Vmw175 binds to DNA as a dimer, it is not clear exactly how each monomeric subunit contacts the DNA helix. These questions could be elucidated by further structural studies. Whatever the mechanism by which Vmw175 dimerizes and binds to DNA, several studies have illustrated that the dimerization properties of Vmw175 are significant *in vivo* (Shepard and DeLuca, 1989; 1991a; Shepard *et al.*, 1990).

The results presented and discussed in section 3E illustrate that point mutations in the region encompassing amino acids 445-487 of Vmw175 do not affect the dimerization properties of DNA binding domain polypeptides. It has therefore been possible to isolate mutants defective in DNA binding activity, but still able to dimerize. The design and expression of polypeptides defective in dimerization has so far proved unsuccessful (R. D. Everett, personal communication; this thesis). The isolation of such a polypeptide would be useful in delineating the dimerization motif within the DNA binding domain, and also in examining the functions of dimerization within the *in vitro* and *in vivo* properties of Vmw175. Until such a mutant is obtained, the region(s) of the DNA binding domain contributing to dimerization cannot be defined.

In conclusion, the results presented in this thesis have illustrated that the DNA binding activity of Vmw175 can be eliminated in the context of a dimerization-functional molecule.

4C. Effects of Mutations on the *In Vivo* Properties of Vmw175.

The insertional mutant proteins studies in this thesis have previously been well characterized *in vitro* (Paterson and Everett, 1988a, b; Paterson, 1989). Therefore, it is interesting to find that the results obtained *in vivo* and presented and discussed in section 3G mirror the *in vitro* properties of the mutant proteins. That is, mutants which are deficient in transactivation, autoregulation and DNA binding *in vitro* are unable to support the viral replicative cycle in cell culture. Therefore, the regions of Vmw175 involved in DNA binding, transactivation and autoregulation *in vitro* and support of the viral replicative cycle *in vivo* overlap. These results illustrate the validity of the previous *in vitro* studies in determining the activity of mutant polypeptides. In addition, they illustrate the importance of region 2 to the full range of properties of Vmw175. The fact that four insertional mutant forms of Vmw175, I9, I10, I11 and I20, can support viral replication suggests that the inability of other mutants to function fully is not due to a non-specific effect of the mutations on the conformation and stability, or cellular transport of Vmw175.

As the mutations appear to affect all activities of Vmw175 *in vitro*, it is difficult to attribute the inactivity of mutant proteins *in vivo* to the lack of any one property of the protein. However, as the mutations examined fall within a region of the protein which encompasses the DNA binding domain, rather than a domain solely involved in transactivation, the results suggest that the DNA binding properties of Vmw175 are essential for its function *in vivo*.

This conclusion is supported by results obtained with additional point and internal deletion mutant Vmw175 molecules. These mutations had previously been introduced into a bacterially expressed polypeptide encompassing the DNA binding domain of Vmw175 and assayed for DNA binding and dimerization activity (this thesis, see discussion above).

The analysis of point mutations and deletions which had previously been shown to disrupt DNA binding activity illustrated that these mutant proteins were unable to support the viral replicative cycle. Point mutations can therefore disrupt the function of the Vmw175 protein in the cellular environment, illustrating the importance of the region encompassing amino acids 456-464. As discussed in section 4B.1, these results may infer but do not prove the involvement of the complete conserved tract between amino acids 445-487 in the *in vivo* properties of Vmw175. The inability of a mutant polypeptide containing a deletion between amino acids 445-453 to support viral

replication suggests that this region too is essential for the function of Vmw175. However, the possibility that such a deletion results in the conformational instability of Vmw175 cannot be eliminated.

The results obtained with two point mutations which resulted in a temperature sensitive defect in DNA binding activity *in vitro* are more difficult to interpret. These mutations are present within the conserved tract, one falling inside and one outside the region encompassing amino acids 456-464. This latter region has been shown to be essential for the *in vitro* and *in vivo* functions of Vmw175 (see above). As discussed in Section 3G.4, it is interesting that whilst the two mutations result in a temperature sensitive DNA binding phenotype *in vitro*, the one within region 456-464 exhibits a more deleterious effect upon the *in vivo* function of Vmw175. The different phenotypes of the mutations may be explained by the fact that the mutations differ in their effect on the stability of the interaction between Vmw175 and DNA. Whilst this difference was not detected in *in vitro* assays, perhaps *in vivo* assays are able to detect these slight differences in activity. The two mutations involve different amino acid changes - the less deleterious mutation involves an alanine to serine change, whilst the second mutation involves mutation of a basic arginine residue to a non-polar leucine. This fact may also play a role in determining how the mutant proteins interact with DNA both *in vitro* and *in vivo*.

One surprising aspect of the results is that the mutation 457R/L results in a Vmw175 DNA binding domain polypeptide apparently able to bind to DNA at lower temperatures *in vitro*, but when introduced into the viral genome, results in a Vmw175 molecule unable to function at lower temperatures. It is feasible that within the cellular environment, additional protein-protein interactions may occur to stabilize the interaction of Vmw175 with DNA and perhaps cellular transcription factors, and that this mutation affects these additional interactions.

The suggestion that additional protein - protein interactions may affect the way that Vmw175 interacts with DNA *in vivo* is supported by preliminary results from Smith and DeLuca (1992). They found that the addition of general transcription factors TFIID and TFIIB to *in vitro* DNA binding assays of Vmw175 increased the affinity of the protein for DNA. Interestingly, they found that the mutant protein ri12, discussed in detail in section 1C.3, was able to bind to DNA in the presence of these transcription factors. During the viral replicative cycle, these transcription factors will be assembled on promoter regions and thus able to interact with Vmw175. It is therefore feasible that these protein - protein interactions will stabilize Vmw175 - DNA interactions within the cell as discussed above. These results appear to explain how a Vmw175 molecule apparently unable to bind DNA *in vitro* can support viral replication, being able to bind DNA within the cellular environment.

Whatever the mechanisms by which Vmw175 interacts with DNA within the cell, it appears apparent that the simple *in vitro* assay of Vmw175 DNA binding activity is insufficient to explain the full properties of certain mutant forms of Vmw175, both as presented in this thesis and as previously described by Paterson *et al.* (1990). However, from the results presented in this thesis, it can be concluded that the DNA binding activity of Vmw175 is extremely important, if not essential, for its full transactivation activity *in vivo*. In addition, the integrity of region 2 of the protein, and of a small conserved tract within the carboxy-terminal portion of this domain, is essential for its activity both *in vitro* and *in vivo*.

4D. Further Research Proposals.

There are still several unanswered questions concerning the mechanism by which Vmw175 controls HSV-1 gene expression. Firstly, whilst the results presented in this thesis have helped to delineate regions of the protein involved in the interactions with DNA, further structural studies could show exactly how the protein binds to DNA. These studies would also confirm whether amino acids conserved between the members of the family of transcriptional regulatory proteins of α-herpesviruses are involved in facilitating protein - DNA contacts. Such analysis of the complex formed between the Vmw175 DNA binding domain and DNA would also explain help explain how the protein dimerizes and binds to apparently asymmetric DNA sites.

Whilst the DNA binding domain of Vmw175 has been the subject of various studies, it is also of importance to determine which regions of the protein are involved in transactivation. It is generally accepted that transactivation domains act partly through the interaction with general transcription factors (see section 1D.2). With this in mind, it would be useful to determine how Vmw175 interacts with basal transcription factors, an interaction proposed by recent preliminary studies (Smith and DeLuca, 1992; Pizer *et al.*, 1992), and to delineate the regions of Vmw175 involved in contacting such factors. Whilst Pizer *et al.* (1992) proposed that the DNA binding domain mediates such interactions, this is in oppostion with preliminary results from DeLuca's group, and also fails to explain the function of other regions of Vmw175, including the conserved region 4.

It has been proposed that viral transactivators may act as "co-activators", mediating the interaction between the TATA binding protein and upstream factors

(reviewed by Lewin, 1990). As it has been suggested that Vmw175 interacts with the TATA binding protein, the possibility exists that Vmw175 acts as such a bridging molecule. Further analysis of interactions between Vmw175 and upstream factors, such as Sp-1 would be of use in such studies.

With these considerations in mind, the investigation of the function of the conserved region 4 of Vmw175 may yield some interesting results. This portion of the protein has been shown to be essential for the activation of viral late genes *in vivo* (DeLuca and Schaffer, 1987, 1988). The region could be involved in mediating protein - protein interactions, or in stabilizing the overall conformation of the protein molecule. In addition, mutational analysis of this region may define a separable transactivation domain, a domain not presently defined in Vmw175.

Recent studies on other DNA binding transcription factors has highlighted the importance of post-translational modifications, especially phosphorylation, in the control of such factors within the cell (Hunter and Karin, 1992). It has been shown that Vmw175 is phosphorylated *in vivo*, and that the phosphorylation state of the molecule varies during the viral replicative cycle (Wilcox *et al.*, 1980). Therefore, the further investigation of phosphorylation sites within Vmw175, and the effect of phosphorylation on the *in vivo* activity of Vmw175, would be extremely important in determining mechanisms of control of Vmw175 *in vivo*.

REFERENCES.

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., DALRYMPLE, M. A. RAMSEY, F. H., PRESTON, V. G. and PRESTON, C. M. (1988). Mutational analysis of the herpes simplex virus type 1 *trans*-inducing factor Vmw65. J. Gen. Virol. 69, 2595-2605.

ACE, C. I., DALRYMPLE, M. A., RAMSEY, F. H. and PRESTON, C. M. (1989). Construction and characterisation of a herpes simplex virus type 1 mutant unable to transinduce immediate early gene expression. J. Virol. 63, 2260-2269.

ALFORD, L. A. and BRITT, W. J. (1990). Cytomegalovirus, in: "Virology" vol. 2, pp. 1982-2010, Fields, B. N. and Knipe, D. M. (ed.s), Raven Press, New York.

ANDERSON, A. S., FRANCESCONI, A. and MORGAN, R. W. (1992). Complete nucleotide sequence of the Marek's disease virus ICP4 gene. Virology 189, 657-557.

ANDERSON, J. E., PTASHNE, M. and HARRISON, S. C. (1985). A phage repressor-operator complex at 7Å resolution. Nature 316, 596-601.

ANDIMAN, W., GRADOVILLE, L., HESTON, L., NEYDORFF, R., SAVAGE, M. E., KITCHINGMAN, G., SHEDD, D. and MILLER, G. (1983). Use of cloned probes to detect Epstein-Barr viral DNA in tissues of patients with noeplastic and lymphoproliferative diseases. Journal of Infectious Diseases. 48, 967-977.

ASHER, Y., KELLER, M. and BECKER, Y. (1969). Incorporation of lipids into herpes simplex virus particles. J. Gen. Virol. 4, 65-76.

ASO, T., VASAVADA, H. A., KAWAGUCHI, T., GERMINO, F. J., GANGULY, S., KATAJIMA, S., WEISSMAN, S. M. and YASUKOCHI, Y. (1992). Characterisation of the cDNA for the large subunit of the transcription factor TFIIF. Nature 355, 461-464.

BAER, R., BANKIER, A. T., BIGGIN, M. D., DEININGER, P. L., FARRELL, P. J., GIBSON, T. J., HATFIELD, G., HUDSON, G. S., SATCHWELL, S. C., TUFFNELL, P. S. and BARRELL, B. G. (1984). DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature 310, 207-211.

BARKER, D. E. and ROIZMAN, B. (1992). The unique sequence of the herpes simplex virus L component contains an additional open reading frame designated UL49.5. J. Virol. 66, 562-566.

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.

BATCHELOR, A. H. and O'HARE, P. (1990). Regulation and cell-type-specificity of a promoter located upstream of the latency-induced transcript of herpes simplex virus type 1. J. Virol. 64, 3269-3279.

BATCHELOR, A. H. and O'HARE, P. (1992). Localization of *cis*-acting sequence requirements in the promoter of the latency-associated transcript of herpes simplex virus type 1 required for cell-type-specific activity. J. Virol. 66, 3573-3582.

BATTERSON, W. and ROIZMAN, B. (1983). Characterisation of the herpes simplex virion-associated factor responsible for the induction of α genes. J. Virol. 46, 371-377.

BATTERSON, W., FURLONG, D. and ROIZMAN, B. (1983). Molecular genetics of herpes simplex virus. VIII. Further characterisation of a temperature-sensitive mutant, defective in release of viral DNA and in other stages of the viral reproductive cycle. J. Virol. 45, 397-407.

BEN-ZEEV, A. and BECKER, Y. (1977). Requirement of host cell RNA polymerase II in replication of herpes simplex virus in α-amanitin sensitive and resistant cell lines. Virology 76, 246-253.

BENEZRA, R., DAVIS, R. L., LOCKSHON, D., TURNER, D. L. and WEINTRAUB, H. (1990). The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell 61, 49-59.

BLAHO, J. A., MITCHELL, C. and ROIZMAN, B. (1993). Guanylylation and adenylylation of the α regulatory proteins of herpes simplex virus require a viral β or γ function. J. Virol. 67, 3891-3900.

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.

BLACKLOW, S. C. and KNOWLES, J. C. (1991). Spiked oligonucleotide mutagenesis. In: "Directed Mutagenesis: A Practical Approach" pp177-198 M.J. MacPherson (ed.) Oxford University Press, New York.

BLAHO, J. A., MICHAEL, N., KANG, V., ABOUL-ELA, N., SMULSON, M. E., JACOBSON, M. K., 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.

BOOY, F. P., NEWCOMB, W. W., TRUS, B. L., BROWN, J. C., BAKER, T. S. and STEVEN, A. C. (1991). Liquid-crystalline, phage-like packing of encapsidated DNA in herpes simplex virus. Cell 64, 1007.

BOYER, H. W. and ROULLAND-DUSOIX, D. (1969). A complementation analysis of the restriction modification of DNA in *Escherichia coli*. J. Mol. Biol. 41, 459-472.

BRENT, R. and PTASHNE, M. (1985). A eukaryotic transcription activator bearing the DNA specificity of a prokaryotic repressor. Cell 43, 248-254.

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., ROSS, N. L. J. and BINNS, M. M. (1988). Gene sequence and mapping data from Marek's disease virus and herpes virus of turkeys: implications for herpesvirus classification.

BURATOWSKI, S. and SHARP, P. A. (1990). Transcription initation complexes and upstream activation with RNA polymerase II lacking the C-terminal domain of the largest subunit. Mol. Cell. Biol. 10, 5562-5564.

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.

BUSBY, D. W. G., HOUSE, W. and MacDONALD, J. R. (1964). In: "Virological Techniques". Churchill, London.

BUSCH, S. J. and SASSONE-CORSI, P. (1990). Dimers, leucine zippers and DNA binding domains. Trends in Genetics 6(2), 36-40.

BUSH, M., YAGER, D. R., GAO, M., WEISSHART, K., MARCY, A. I., COEN, D. M. and CHALLBERG, M. D. (1991). Correct intranuclear localization of herpes simplex virus DNA polymerase requires the viral ICP8 DNA binding protein. J. Virol. 65, 1082-1089.

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.

CAI, W. and SCHAFFER, P. A. (1989). Herpes simplex virus type 1 ICP0 plays a critical role in the de novo synthesis of infectious virus following transfection of viral DNA. J. Virol. 63, 4579-4589.

CAI, W., GU, B. and PERSON, S. (1988). Role of glycoprotein B of herpes simplex virus type 1 in viral entry and cell fusion. J. Virol. 62, 2596-2604.

CALDER, J. M. and STOW, N. D. (1990). Herpes simplex virus helicase-primase: the UL8 subunit is not required for DNA-dependent ATPase and DNA helicase activities. Nucleic Acids Res. 18, 3573-3578.

CAMPADELLI-FIUME, G., QI, S., ARITABILE, E., FOA-TOMASI, L., BRANDIMARTI, R. and ROIZMAN, B. (1990). Glycoprotein D of herpes simplex virus encodes a domain which precludes penetration of cells expressing the glycoprotein by superinfecting herpes simplex virus. J. Virol. 64, 6070-6079.

CAMPADELLI-FIUME, G., FARABEGOLI, F., DI GAETA, S. and ROIZMAN, B. (1991). Origin of unenveloped capsids in the cytoplasm of cells infected with herpes simplex virus 1. J. Virol. 65, 1589-1595.

CAMPBELL, M. E. M., PALFREYMAN, J. W. and PRESTON, C. M. (1984). Identification of herpes simplex virus type 1 DNA sequences which encode a *trans*-acting polypeptide responsible for stimulation of immediate early transcription. J. Mol. Biol. 180, 1-19.

CHALLBERG, M. D. (1986). A method for identifying the viral genes required for herpesvirus DNA replication. Proc. Natl. Acad. Sci. (USA) 83, 9094-9098.

CHALLBERG, M. D. (1991). Herpes simplex virus DNA replication. Seminars in Virol. 2, 247-256.

CHAPMAN, C. J., HARRIS, J. D., HARDWICKE, M. A., SANDRI-GOLDIN, R. M., COLLINS, M. K. L. and LATCHMAN, D. S. (1992). Promoter independent activation of heterologous gene expression by the herpes simplex virus immediate-early protein ICP27. Virology 186, 573-578.

CHEE, M. S., BANKIER, A. T., BECH, S., BOHNI, R., BROWN, C. M., CENY, R., HORSNELL, T., HUTCHINSON III, C. A., KOUZARIDES, T., MARTIGNETTI, J. A., PREDDIE, E., SATCHWELL, S. C., TOMLINSON, P., WESTON, R. M. and BARRELL, B. G. (1990). Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Current Topics in Microbiology and Immunology 154, 125-170.

CHEN, J. and SILVERSTEIN, S. (1992). Herpes simplex viruses with mutations in the gene encoding ICP0 are defective in gene expression. J. Virol. 66, 2916-2927.

CHEUNG, A. K. (1989). DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies virus. Nucleic Acids Res. 17, 4637-4646.

CHOU, J. and ROIZMAN, B. (1990). The herpes simplex virus 1 gene for ICP34.5, which maps in inverted repeats, is conserved in several limited passage isolates but not in strain 17 syn^+. J. Virol. 64, 1014-1020.

CHOU, P. Y. and FASHMAN, U. D. (1974). Prediction of protein conformation. Biochemistry 13, 211-215.

CLEMENTS, J. B., WATSON, R. J. and WILKIE, N. M. (1977). Temporal regulation of herpes simplex virus type 1 transcription: location of transcripts on viral genome. Cell 12, 275-285.

CLEMENTS, J. B., McLAUGHLAN, J. and McGEOCH, D. J. (1979). Orientation of herpes simplex virus type 1 immediate early mRNA's. Nucleic Acids Res. 7, 77-91.

COEN, D., 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.

COMAI, L., TANESE, N. and TJIAN, R. (1991). The TATA binding protein and associated factors are integral components of the RNA polymerase I transcription factor, SLI. Cell 68, 965-976.

CORDEN, J. L. (1990). Tails of RNA polymerase II. Trends Biochem. Sci. 15, 383-387.

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 Genetics 7, 603-616.

CORTES, P., FLORES, O. nad REINBERG, D. (1991). Factors involved in specific transcription by RNA polymerase II: purification and analysis of transcription factor IIA and identification of transcription factor IIJ. Mol. Cell. Biol. 12, 413-421.

COSTANZO, F., CAMPADELLI-FIUME, G., FAO-TOMASI, L. and CASSAI, E. (1977). Evidence that herpes simplex virus DNA is transcribed by cellular RNA polymerase B. J. Virol. 21, 996-1001.

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

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.

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.

CRUTE, J. J. and LEHMAN, I. R. (1989). Herpes simplex-1 DNA polymerase. Identification of an intrinsic 5' to 3' exonuclease with ribonuclease H activity. J. Biol. Chem. 264, 19266-19270.

CRUTE, J. J. and LEHMAN, I. R. (1991). Herpes simplex virus-1 helicase-primase. Physical and catalytic properties. J. Biol. Chem. 266, 4484-4488.

CRUTE, J. J., TSURUMI, T., ZHU, L., WELLER, S. K., OLIVO, P. D., CHALLBERG, M. D., MOCARSKI, E. S. and LEHMAN, I. R. (1989). Herpes simplex virus 1 helicase-primase: a complex of three herpes-encoded gene products. Proc. Natl. Acad. Sci. (USA) 86, 2186-2189.

DARGAN, D. J. (1986). The structure and assembly of herpesviruses, in: "Electron microscopy of proteins; viral structure", vol. 5, pp. 359-437, Harris, J. R. and Horne, R. W. (ed.s), Academic Press, Orlando and London.

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.

DAVIS, M. G., KENNEY, S. C., KAMINE, J., PAGANO, J. S. and HUANG, E. S. (1987). Immediate-early region of human cytomegalovirus *trans*-activates the promoter of human immunodeficiency virus. Proc. Natl. Acad. Sci. USA 84, 8642-8646.

DAVIS, R. L., CHENG P-F., LASSAR, A. B. and WEINTRAUB, H. (1990). The MyoD DNA binding domain contains a recognition code for muscle specific gene activation. Cell 60, 733-746.

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. (1986). The complete DNA sequence of varicella-zoster virus. J. gen. Virol. 67, 1759-1816.

DAVISON, A. J. and TAYLOR, P. (1987). Genetic relations between varicella zoster virus and Epstein-Barr virus. J. Gen. Virol. 68, 1067-1079.

DAVISON, A. J. and WILKIE, N. M. (1981). Nucleotide sequences of the joint between the L and S segments of the herpes simplex virus types 1 and 2. J. Gen. Virol. 55, 315-331.

DAVISON, A. J. and WILKIE, N. M. (1983). Location and orientation of homologous sequences in the genomes of the herpesviruses. J. Gen. Virol. 64, 1927-1942.

DAVISON, M. D., RIXON, F. J. and DAVISON, A. J. (1992). Identification of genes encoding two capsid proteins (VP24 and VP26) of herpes simplex virus type 1. J. Gen. Virol. 73, 2709-2713.

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.

DE BRUYN KOPS, A. and KNIPE, D. M. (1988). Formation of DNA replication structures in herpes virus-infected cells requires a viral DNA binding protein. Cell 55, 857-868.

DE THE, G., GESSER, A., DAY, N. E., TUKEI, P. M., WILLIAMS, F. H., BEVI, O. P., SMITH, P. G., DEAN, A. G., BORNKAMM, G. W., FEORINO, P. and HENLE, W. (1978). Epidemiological evidence for a causal relationship between Epstein-Barr virus and Burkitt's lymphoma from Ugandan prospective study. Nature 274, 756-761.

DEAN, H. J., SHIEH, M.-T., TERHUNE, S. S. and SPEAR, P. G. (1992). Pleietropic effects, includibg altered cell tropisn, of gD mutations in HSV-1 mutants selected for resistance to gD-mediated interference. Abstract no. 33, 17th International Herpesvirus Workshop, Edinburgh, Scotland.

DEATLY, A. M., SPIVACK, J. G., LAVI, E. and FRASER, N. W. (1987). RNA from an immediate-early region of the type 1 herpes simplex virus genome is present in the trigeminal ganglia of latently infected mice. Proc. Natl. Acad. Sci. (USA) 84, 3204-3208.

DEB, S. and DEB, S. P. (1991). A 269-amino acid segment with a pseudo-leucine zipper and a helix-turn-helix motif codes for the sequence-specific DNA binding domain of herpes simplex virus type 1 origin-binding protein. J. Virol. 65, 2829-2938.

DEB, S. and DOELBERG, M. (1988). A 67-base-pair segment from the oriS region of herpes simplex virus type 1 encodes origin function. J. Virol. 62, 2516-2519.

DECKMAN, I. C., HAGEN, M. and McCANN III, P. J. (1992). Herpes simplex virus type 1 protease expressed in *Escherichia coli* exhibits autoprocessing and specific cleavage of the ICP35 assembly protein. J. Virol. 66, 7362-7367.

DEISS, L. P., CHOU, J. and FRENKEL, N. (1986). Functional domains of within the 'a' sequence involved in the cleavage-packaging of herpes simplex virus DNA.J. Virol. 59, 605-618.

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. Nucl. Acids Res. 15(11), 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(3), 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(3), 767-776.

DeLUCA, N. A., McCARTHY, A. M. and SCHAFFER, P. A. (1985). Isolation and characterisation 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.

DESAI, P. J., SCAFFER, P. A. and MINSON, A. C. (1988). Excretion of non-infectious virus particles lacking glycoprotein H by a temperature-sensitive mutant of herpes simplex virus type 1: evidence that gH is essential for virion entry. J. Gen. Virol. 69, 1147-1156.

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.

DEVEREUX, J., HAEBERLI, P. and SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12(1), 387-395.

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. and EVERETT, R. D. (1990a). 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.

DIXON, R. A. F. and SCHAFFER, P. A. (1980). Fine structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. J. Virol. 36, 189-203.

DOBSON, A. T., SEDARATI, F., DEVI-RAO, G., FLANAGAN, W. M., FARRELL, M. J., STEVENS, J. G., WAGNER, E. K. and FIELDMAN, L. T. (1989). Identification of the latency-associated transcript promoter by expression of a rabbit beta-globin mRNA in sensory nerve ganglia latently infected with a recombinant herpes simplex virus. J. Virol. 63, 3844-3851.

DODD, I. B. and EGAN, J. B. (1987). Systematic method for the detection of potential λCro-like DNA-binding regions in proteins. J. Mol. Biol. 194, 557-564.

DODD, I. B. and EGAN, J. B. (1990). Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. Nucleic Acids Res. 18, 5019-5026.

DODSON, M. S. and LEHMAN, I. R. (1991). Association of DNA helicase and primase activities with a subassembly of the herpes simplex virus 1 helicase-primase composed of the UL5 and UL52 gene products. Proc. Natl. Acad. Sci. (USA). 88, 1105-1109.

DODSON, M. S., CRUTE, J. J., BRUCKNER, R. C. and LEHMAN, I. R. (1989). Overexpression and assembly of the herpes simplex virus helicase-primase in insect cells. J. Biol. Chem. 264, 20835-20838.

DOERIG, C., PIZER, L. W. and WILCOX, C. L. (1991). An antigen encoded by the latency-associated transcript in neuronal cell cultures latently infected with herpes simplex virus type 1. J. Virol. 65, 2724-2727.

DOLAN, A., McKIE, E., McLEAN, A. R. and McGEOCH, D. J. (1992). Status of the ICP34.5 gene in herpes simplex virus type 1 strain 17 syn^+. J. Gen. Virol. 73, 971-973.

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.

EFSTATHIOU, S., MINSON, A. C., FIELD, H. J., ANDERSON, J. R. and WILDY, P. (1986). Detection of herpes simplex virus-specific DNA sequences in latently infected mice and in humans. J. Virol. 57, 446-455.

ELIAS, P., O'DONNELL, M. E., MOCARSKI, E. S. and LEHMAN, I. R. (1986). A DNA binding protein specific for an origin of replication of herpes simplex virus type 1. Proc. Natl. Acad. Sci. (USA) 83, 6322-6326.

ELIAS, P., GUSTAFSSON, C. M. and HAMMARSTEN, O. (1990). The origin binding protein of herpes simplex virus 1 binds co-operatively to the viral origin of replication. J. Biol. Chem. 265, 17167-17173.

ELKAREH, A., MURPHY, A. J. M., FICHTER, T., ESTRATIADIS, 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. E., BRANDI, C. J., STRUHL, K. and HARRISON, S. C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: crystla structure of the protein-DNA complex. Cell 71, 1223-1237.

ELLIS, H. M., SPANN, D. R. and POSAKONY, J. W. (1990). *extramacrochaetae*, a negative regulator of sensory organ development in Drosophila, defines a new class of helix-loop-helix proteins. Cell 61, 27-38.

EPSTEIN, M. A. (1962). Observations on the fine structure of mature herpes simplex virus and on the composition of its nucleoid.J. Exp. Med. 115, 1-9.

ESTRIDGE, J. K., KEMP, L. M., LATHANGUE, N. B., MANN, B. S., TYMS, A. S. and LATCHMAN, D. S. (1989). The herpes simplex virus type 1 immediate-early protein ICP27 is obligately required for the accumulation of a cellular protein during viral infection. Virology 168, 67-70.

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 transactivation by viral immediate-early gene products are not early-gene specific. Nucleic Acids Res. 12, 3037-3056.

EVERETT, R. D. (1984b). Transactivation of transcription by herpes simplex virus products; requirements 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 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 products Vmw175 and Vmw110. J. Mol. Biol. 202, 739-751.

EVERETT, R. D. (1989). Construction and characterisation 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 FENWICK, M. L. (1990). Comparative DNA sequence analysis of the host shutoff genes of different strains of herpes simplex virus: type 2 strain HG52 encodes a truncated UL41 product. J. Gen. Virol. 71, 1387-1390.

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., TEDDER, D. G., ELLIOTT, M. and LITMAN, B. (1991a). 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.

EVERETT, R. D., ELLIOTT, M., HOPE, G. and ORR, A. (1991b). Purification of the DNA binding domain of herpes simplex virus type 1 immediate-early polypeptide 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. (1991c). High level expresion 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. (1991d). Functional and genetic analysis of the role of Vmw110 in herpes simplex virus replication. In "Herpesvirus Transcription and its Regulation" pp. 49-76. E. K. Wagner (Ed.), LRC Press, Boca Raton, USA.

FABER, S. W. and WILCOX, K. W. (1986). 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.

FEAVER, W. J., GILEADI, O., LI, Y. and KORNBERG, R. D. (1991). CTD kinase associated with yeast RNA polymerase II initiation factor b. Cell 67, 1223-1230.

FEINBERG, A. P. and VOGELSTEIN, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132, 6-13.

FELSENFELD, G. (1992). Chromatin as an essential part of the transcriptional mechanism. Nature 355, 219-224.

FENWICK, M. L. (1984). The effects of herpes viruses on cellular macromolecular synthesis. In " Comprehensive Virology" vol. 19, pp. 359-390, Fraenkel-Conrat, H. and Wagner, R. R. (ed.s) Plenum Press, New York.

FENWICK, M. L. and CLARK, J. (1982). Early and delayed shut-off of host protein synthesis in cells infected with herpes simplex virus. J. Gen. Virol. 61, 121-125.

FENWICK, M. L. and EVERETT, R. D. (1990). Transfer of UL41, the gene controlling virion-associated host cell shutoff, between different strains of herpes simplex virus. J. Gen. Virol. 71, 411-418.

FENWICK, M. L. and WALKER, M. J. (1978). Suppression of the synthesis of cellular macromolecules by herpes simplex virus. J. Gen. Virol. 41, 37-51.

FIERER, D. S. and CHALLBERG, M. D. (1992a). Biochemical characterisation of the DNA binding domain of HSV-1 UL9. Abstract no. 1, 17th International herpesvirus Workshop, Edinburgh, Scotland.

FIERER, D. S. and CHALLBERG, M. D. (1992b). Purification and characterisation of UL9, the herpes simplex virus type 1 origin-binding protein. J. Virol. 66, 3986-3995.

FLANAGAN, W. M., PAPVASSILIOU, A. G., RICE, M., HECHT, L. B., SILVERSTEIN, S. and WAGNER, E. K. (1991). 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(2), 769-786.

FORRESTER, A., FARRELL, H., WILKINSON, G., KAYE, J., DAVIS-POYNTER, N. and MINSON, T. (1992). Construction and properties of a mutant herpes simplex virus type 1 with glycoprotein H coding sequences deleted. J. Virol. 66, 341-348.

FRAIN, M., SWART, G., MONACI, P., NICOSIA, A., STAMPFLI, S., FRANK, R. and CORTESE, R. (1989). The liver specific transcription factor LF-B1 contains a highly diverged homeobox DNA binding domain. Cell 59, 145-157.

FRAMPTON, J., LEUTZ, A., GIBSON, T. J. and GRAF, T. (1989). DNA binding domain ancestry. Nature 342, 134.

FREEMAN, J. L. and POWELL, K. L. (1982). DNA-binding properties of a herpes simplex virus immediate early protein. J. Virol. 44, 1084-1087.

FREEMONT, P. S., HANSON, I. M. and TROWSDALE, J. (1991). A novel cysteine-rich sequence motif. Cell 64, 483-484.

FRENKEL, N., JACOB, R. J., HONESS, R. W., HAYWARD, G. S., LOCKER, H. and ROIZMAN, B. (1975). Anatomy of herpes simplex virus DNA. III. Characterisation of defective DNA molecules and biological properties of virus populations containing them. J. Virol. 16, 153-167.

FRENKEL, N. LOCKER, H., BATTERSON, W., HAYWARD, G. S. and ROIZMAN, B. (1976). Anatomy of herpes simplex virus DNA. VI. Defective DNA originates from the S component. J. Virol. 20, 527-531.

FRENKEL, N., SCHIRMER, E. C., WYATT, L. S., KATSFANAS, G., ROFFMAN, E., DANOVICH, R. M. and JUNE, C. H. (1990). Isolation of a new herpesvirus from human CD4$^+$ T cells. Proc. Natl. Acad. Sci. USA 87, 748-752.

FRIEDMAN, A. D., TRIEZENBERG, S. J. and McKNIGHT, S. L. (1988). Expression of a truncated viral *trans*-activator selectively impedes lytic infection by its cognate virus. Nature 335, 452-454.

FULLER, A. O. and LEE, W.-C. (1992). Herpes simplex virus type 1 entry through a cascade of virus-cell interactions requires different roles of gD and gH in penetration. J. Virol. 66, 5002-5012.

FULLER, A. O. and SPEAR, P. G. (1987). Anti-glycoprotein D antibodies that permit adsorption but block infection by herpes simplex virus 1 prevent virion-cell fusion at the cell surface. Proc. Natl. Acad. Sci. USA 84, 5454-5458.

FULLER, A. O., SANTOS, R. E. and SPEAR, P. G. (1989). Neutralizing antibodies specific for glycoprotein H of herpes simplex virus permit viral attachment to cells but prevent penetration. J. Virol. 63, 3435-3443.

FURLONG, D., SWIFT, H. and ROIZMAN, B. (1972). Arrangement of herpesvirus deoxyribonucleic acid in the core. J. Virol. 10, 1071-1074.

GAFFNEY, D. F., McLAUGHLAN, J., LINDSAY, J. 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.

GARNIER, J., OSGUTHORPE, D. G. 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.

GARRELL, J. and MODOLELL, J. (1990). The Drosophila *extramacrochaetae* locus, an antogonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. Cell 61, 39-48.

GEHRING, W. J. (1987). Homeo boxes in the study of development. Science 236, 1245-1252.

GELB, L. D. (1990). Varicella zoster virus, in: "Virology" vol. 2, pp. 2011-2054, Fields, B. N. and Knipe, D. M. (ed.s), Raven Press, New York.

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. 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.

GELMAN, I. H. and SILVERSTEIN, S. (1987a). Herpes simplex virus immediate-early promoters are responsive to virus and cell trans-acting factors. J. Virol. 61, 2286-2296.

GELMAN, I. H. and SILVERSTEIN, S. (1987b). Dissection of immediate-early gene promoters from herpes simplex virus: sequences that respond to the virus transcriptional activators. J. Virol. 61, 3167-3172.

GENTZ, R., RAUSCHER III, F. J., ABATE, C. and CURRAN, T. (1989). Parallel association of Fos and Jun leucine zippers juxtaposes DNA binding domains. Science 243, 1695-1699.

GERSTER, T. and ROEDER, R. G. (1988). A herpesvirus trans-activating protein interacts with transcription factor OTF-1 and other cellular proteins. Proc. Natl. Acad. Sci. (USA) 85, 6347-6351.

GIBSON, T. J. (1984). Studies on the Epstein-Barr virus genome. PhD Thesis. University of Cambridge, England.

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

GODING, C. R. and O'HARE, P. (1989). Herpes simplex virus Vmw65-octamer binding protein interaction: a paradigm for combinatorial control of transcription. Virology 173, 363-367.

GOTTLIEB, J., MARCH, A. I., COEND. M. and CHALLBERG, M. D. (1990). The herpes simplex virus type 1 UL42 gene product: a subunit of DNA polymerase that functions to increase processivity. J. Virol. 64, 5976-5987.

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

GRUNDY, F. J., BALIMAN, 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.

HA, I., LANE, W. S. and REINBERG, D. (1991). Cloning of a human gene encoding the general transcription initiation factor IIB. Nature 352, 689-695.

HALLIBURTON, I. W., RANDALL, R. E., KILLINGTON, R. A. and WATSON, D. H. (1977). Some properties of recombinants between type 1 and 2 herpes simplex viruses. J. Gen. Virol. 36, 471-484.

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.

HARRIS, R. A. and PRESTON, C. M. (1991). Establishment of latency *in vitro* by the herpes simplex virus type 1 mutant *in*1814. J. Gen. Virol. 72, 907-913.

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.

HAY, J., GOTELES, G. J., KEIR, H. M. and SUBAK-SHARPE, J. H. (1966). Herpes virus specified ribonucleic acids. Nature 210, 3887-390.

HAY, R. T. and HAY, J. (1980). Properties of herpesvirus-induced "immediate-early" polypeptides. Virology 104, 230-234.

HAY, R. T. and HAY, J. (1980). Properties of herpesvirus-induced "immediate-early" polypeptides. Virology 104, 230-234.

HEINE, J. W., SPEAR, P. G. and ROIZMAN, B. (1972). Proteins specified by herpes simplex virus. VI. Viral proteins in the plasma membrane. J. Virol. 9, 431-439.

HERNANDEZ, T. R. and LEHMAN, I. R. (1990). Functional interaction between the herpes simplex-1 DNA polymerase and UL42 protein. J. Biol. Chem. 265, 11227-11232.

HERNANDEZ, T. R., DUTCH, R. E., LEHMAN, I. R., GUSTAFSSON, C. and ELIAS, P. (1991). Mutations in a herpes simplex virus type 1 origin that inhibit replication. J. Virol. 65, 1649-1652.

HEROLD, B. C., WuDUNN, D., SOLTYS, N. and SPEAR, P. G. (1991). Glycoprotein C of herpes simplex virus type 1 plays a principal role in the adsorption of virus to cells and in infectivity. J. Virol. 65, 1090-1098.

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.

HIGHLANDER, S. L., CAI, W. H., ARSON,S., LEVINE, M. and GLORIOSO, J. C. (1988). Monoclonal antibodies define a domain on herpes simplex virus glycoprotein B involved in virus penetration. J. Virol. 62, 1881-1888.

HILL, J. M., SEDARATI, F., JAVIER, R. T., WAGNER, E. K. and STEVENS, J. G. (1990). Herpes simplex virus latent phase transcription facilitates in vivo reactivation. Virology 174, 117-125.

HILL, T. M., SADLER, J. R. and BETZ, J. L. (1985). Virion component of HSV-1 (KOS) interferes with early shutoff of host protein synthesis induced by HSV-2 (186). J. Virol. 56, 312-316.

HIRTH, K. H., EDWARDS, C. A. and FIRTEL, R. A. (1982). A DNA-mediated transformation system for *Dictyostelium discoideum*.. Proc. Natl. Acad. Sci. USA 79, 7356-7360.

HO, D. Y. (1992). Herpes simplex virus latency: molecular aspects. In "Progress in Medical Virology" 39, pp. 76-115. J. L. Melnick (Ed.), Karger.

HOLMES, S. M. and QUIGLEY, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114, 193-197.

HOMA, F. L., OTAL, T. M., GLORIOSO, J. C. and LEVINE, M. (1986). Transcriptional control signals of a herpes simplex 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.

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. (1986). Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell 46, 885-894.

HOPE-SIMPSON, R. E. (1965). The nature of herpes zoster: a long term study and a new hypothesis. Proceedings of the Royal Society of Medicine. 58, 9-20.

HUNTER, T. and KARIN, M. (1992). The regulation of transcription by phosphorylation. Cell 70, 375-387.

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(6), 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(2), 565-574.

JACOB, R. J., MORSE, L. S. and ROIZMAN, B. (1979). Anatomy of herpes simplex virus DNA. XII. Accumulation of head to tail concatamers in nuclei of infected cells and their role in the generation of the four isomeric arrangements of viral DNA. J. Virol. 29, 448-457.

JOHNSON, D. C. and LIGAS, M. (1988). Herpes simplex viruses lacking glycoprotein D are unable to inhibit virus penetration: quantitative evidence for virus-specific cell surface receptors. J. Virol. 62, 4605-4612.

JOHNSON, D. C., BURKE, R. L. and GREGORY, T. (1990). Soluble forms of herpes simplex virus glycoprotein D bind to a limited number of cell surface receptors and inhibit virus entry into cells. J. Virol. 64, 2569-2576.

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/cap site region is sufficient for fully efficient regulated activity. Nucleic Acids Res. 14, 8247-8264.

JOHNSON, R. M. and SPEAR, P. G. (1989). Herpes simplex virus glycoprotein D mediates interference with herpes simplex virus infection. J. Virol. 63, 819-827.

JOHNSON, W. A., McCORMICK, C.. A., BRAY, S. J. and HIRSH, J. (1989). A neuron-specific enhancer of the Drosophila dopa decarboxylase gene. Genes Dev. 3, 676-686.

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.

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

JONES, P. C. and ROIZMAN, B. (1979). Regulation of herpes simplex virus 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.

JORDAN, S. R. and PABO, C. O. (1988). Structure of the lambda complex at 2.5Å resolution: details of the repressor-operator interactions. Science 242, 893-899.

KANER, R. J., BAIRD, A., MANSUKHANI, A., BASILICO, C., SUMMERS, B. D., FLORKIEWICZ, R. D. and HAJJAR, D. P. (1990). Fibroblast growth factor receptor is a portal of entry for herpes simplex virus type 1. Science 248, 1410-1413.

KAO, C. C., LIEBERMAN, P. M. SCHMIDT, M. C., ZHOU, Q., PEI, R. and BERK, A. J. (1990). Cloning of transcriptionally active human TATA binding factor. Science 248, 1646-1650.

KATAN, M., HAIGH, A., VERRIZER, C. P., VAN DER VLIET, P. C. and O'HARE, P. (1990). Characterisation of a cellular factor which interacts funtionally with Oct-1 in the assembly of a multi-component transcription complex. Nucleic Acids Res. 18, 6871-6880.

KATTAR-COOLEY, P. and WILCOX, K. W. (1989). Characterization of the DNA-binding properties of the herpes simplex virus regulatory protein ICP4. J. Virol. 63, 696-704.

KELLEHER III, R. J., 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.

KEMP, L. M. and LATCHMAN, D. S. (1988). Transduction 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.

KEMP, L. M., PRESTON, C. M., PRESTON, V. G. and LATCHMAN, D. S. (1986). Cellular gene induction during herpes simplex virus infection can occur without viral protein synthesis. Nucleic Acids Res. 14, 9261-9270.

KEMP, L. M., DENT, C. L. and LATCHMAN, D. S. (1990). Octamer motif mediates transcriptional repression of HSV immediate-early genes and octamer-containing cellular promoters in neuronal cells. Neuron 4, 215-222.

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 unrelicated template. J. Virol. 65, 6749-6760.

KIEFF, E. D., BACHENHEIMER, S. L. and ROIZMAN, B. (1971). Size, composition and structure of the deoxyribonucleic acid of herpes simplex virus subtypes 1 and 2. J. Virol. 8, 125-132.

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.

KNIPE, D. M., and SMITH, J. L. (1986). A mutant herpesvirus protein leads to a block in nuclear localisation of other viral proteins. Mol. Cell. Biol. 6, 2371-2381.

KNIPE, D. M., BATTERSON, W., NOSAL, C., ROIZMAN, B. and BUCHAN, A. (1981). Molecular genetics of herpes simplex virus. VI. Characterisation of a temperature-sensitive mutant defective in the expression of all early viral products. J. Virol. 38, 539-547.

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.

KOFF, A., SCHWEDES, J. F. and TEGTMEYER, P. (1991). Herpes simplex virus origin-binding protein (UL9) loops and distorts the viral replication origin. J. Virol. 65, 3284-3292.

KORNBERG, R. D. and LORCH, Y. (1991). Irresistible force meets immovable object: transcription and the nucleosome. Cell 67, 833-836.

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 α genes. Proc. Natl. Acad. Sci. (USA) 81, 4065-4069.

KRISTIE, T. M. and ROIZMAN, B. (1986a). α 4, the major regulatory protein of herpes simplex virus type 1, 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 cells proteins bind to the cis-acting site required for virion-mediated induction of herpes simplex virus type 1 α genes. Proc. Natl. Acad. Sci. (USA). 84, 71-74.

KRISTIE, T. M., LEBOWITZ, J. H. and SHARPE, P. A. (1989). The octamer-binding proteins form multi-protein-DNA complexes with the HSV α-TIF regulatory protein. EMBO J. 8, 4229-4238.

KUHN, J. E., KRAMER, M. D., WILLENBACHER, W., WIELAND, U., LORENTZEN, E. U. and BRAUN, R. W. (1990). Identification of herpes simplex virus type 1 glycoproteins interacting with the cell surface. J. Virol. 64, 2491-2497.

KWONG, A. D. and FRENKEL, N. (1987). Herpes simplex virus-infected cells contain a function(s) that destabilizes both host and viral mRNA's. Proc. Natl. Acad. Sci. (USA) 84, 1926-1930.

LADIN, B. F., BLANKENSHIP, M. I. and BEN-PORAT, T. (1980). Replication of herpesvirus DNA. V. Maturation of concatameric DNA of pseudorabies virus to genomic length is related to capsid formation. J. Virol. 33, 1151-1164.

LAMARCO, K. L. and McKNIGHT, S. L. (1989). Purification of a set of cellular polypeptides that bind to the purine-rich *cis*-regulatory element of herpes simplex virus immediate-early genes. Genes Dev. 3, 1372-1383.

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.

LATCHMAN, D. S., ESTRIDGE, J. K. and KEMP, L. M. (1987). Transcriptional induction of the ubiquitin gene during herpes simplex virus infection is dependent upon the viral immediate-early protein ICP4. Nucleic Acids Res. 15, 7283-7293.

LATHANGUE, N. B., SHRIVER, K., DAWSON, C. and CHAN, W. L. (1984). Herpes simplex virus infection causes the accumulation of a heat-shock protein. EMBO J. 3, 267-277.

LAWRENCE, G. L., CHEE, M., CRAXTON, M. A., GOMPELS, U. A., HONESS, R. W. and BARRELL, B. G. (1990). Human herpes virus 6 is closely related to human cytomegalovirus. J. Virol. 64, 287-299.

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

LEIB, D. A., NADEAN, K. C., RUNDLE, S. A. and SCHAFFER, P. A. (1991). The promoter of the latency-associated transcripts of herpes simplex virus type 1 contains a functional cAMP-response element: role of the latency-associated transcripts and cAMP in the reactivation of viral latency. Proc. Natl. Acad. Sci. (USA) 88, 48-52.

LEWIN, B. (1990). Commitment and activation at pol II promoters: a tail of protein:protein interactions. Cell 61, 1161-1164.

LIGAS, M. W. and JOHNSON, D. C. (1988). A herpes simplex virus mutant in which glycoprotein D sequences are replaced by β-galactosidase sequences binds to but is unable to penetrate into cells. J. Virol. 62, 1486-1494.

LIN, Y.-S. and GREEN, M. R. (1991). Mechanism of action of an acidic transcriptional activator in vitro. Cell 64, 971-981.

LIU, F. and ROIZMAN, B. (1991a). The promoter, transcriptional unit and coding sequence of herpes simplex virus 1 family 35 proteins are contained within and in frame with the UL26 open reading frame. J. Virol. 65, 206-212.

LIU, F. and ROIZMAN, B. (1991b). The herpes simplex virus 1 gene encoding a protease also contains within its coding domain the gene encoding the more abundant substrate. J. Virol. 65, 5149-5156.

LIU, F. and ROIZMAN, B. (1992). Differentiation of multiple domains in the herpes simplex virus 1 protease encoded by the UL26 gene. Proc. Natl. Acad. Sci. (USA) 89, 2076-2080.

LOCKSHON, D. and GALLOWAY, D. A. (1988). Sequence and structural requirements of a herpes simplex viral DNA replication origin. Mol. Cell. Biol. 8, 4018-4027.

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 alpha 47 gene. J. Virol. 58, 583-591.

LOPEZ, C. and HONESS, R. W. (1990). Human herpesvirus 6, in: "Virology" vol. 2, pp. 2055-2062, Fields, B. N. and Knipe, D. M. (ed.s), Raven Press, New York.

LOW, B. (1968). Formation of merodiploids in matings with a class of Rec⁻ recipient strains of *Escherichia coli* K12. Proc. Natl. Acad. Sci. USA. 60, 160.

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.

LYCKE, E., KRISTENSSON, K., SVENNERHOLM, B., VAHLNE, A. and SIEGLER, P. (1984). Uptake and transport of herpes simplex virus in neurites of rat dorsal root ganglia cells in culture. J. Gen. Virol. 55-64.

McCARTHY, A. M., McMAHAN, L. and SCHAFFER, P. A. (1989). Herpes simplex virus type 1 ICP27 deletion mutants exhibit altered patterns of transcription and are DNA deficient. J. Virol. 63, 18-27.

McGEOCH, D. J. (1989). The genomes of the human herpesviruses: contents, relationships and evolution. Annu. Rev. Microbiol. 43, 235-265.

McGEOCH, D. J., DOLAN, A., DONALD, S. and RIXON, F. J. (1985). Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. J. Mol. Biol. 181, 1-13.

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, 1727-1745.

McGEOCH, D. J., MOSS, H. W. M., McNAB, D. and FRAME, M. C. (1987). DNA sequence and genetic content of the HindIII *l* region in the short unique component of the herpes simplex virus type 2 genome: identification of the gene encoding glycoprotein G, and evolutionary comparisons. J. Gen. Virol. 68, 19-38.

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. (1988a). The complete sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. 69, 1531-1574.

McGEOCH, D. J., DALRYMPLE, M. A., DOLAN, A., McNAB, D., PERRY, L. J., TAYLOR, P. and CHALLBERG, M. D. (1988b). Structures of herpes simplex virus type 1 genes required for replication of virus DNA. J. Virol. 62, 444-453.

McKNIGHT, S. L. and KINGSBURY, R. C. (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 herpes simplex virus thymidine kinase gene: identification of an upstream control region. Cell 25, 385-398.

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.

McNEILL, J. B. and SMITH, M. (1985). *Saccharomyces cerevisiae CYC1* mRNA 5' end positioning: analysis by *in vitro* mutagenesis, using synthetic duplexes with random mismatch base pairs. Mol. Cell. Biol. 5, 3545-3551.

MacPHERSON, I. and STOKER, M. (1962). Polyoma transformation of hamster cell clones-an investigation of genetic factors affecting cell competence. Virology 16, 147-151.

MA, J. and PTASHNE, M. (1987). Deletion analysis of GAL4 defines two transcriptional activating segments. Cell 48, 847-853.

MACKEM, S. and ROIZMAN, B. (1982a). Differentiation between α promoter and regulator regions of herpes simplex virus type 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 regions 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.

MACNAB, J. C. M., ORR, A. and LATHANGUE, N. B. (1985). Cellular proteins expressed in herpes simplex virus transformed cells also accumulate on herpes simplex virus infection. EMBO J. 4, 3223-3228.

MANIATIS, T., GOODBURN, S. and FISCHER, J. A. (1987). Regulation of inducible and tissue-specific gene expression. Science 236, 1237-1245.

MARGOTTIN, F., DUJARDIN, G., GERARD, M., EGLY, J.-M., HUET, J. and SENTENAC, A. (1991). Participation of the TATA factor in the transcription of the yeast U6 gene by RNA polymerase C. Science 251, 424-426.

MARMORSTEIN, R., CAREY, M., PTASHNE, M. and HARRISON, S. C. (1992). DNA recognition by GAL 4. Structure of a protein-DNA complex. Nature (London), 356 408-414.

MARSDEN, H. S., CROMBIE, I. K. and SUBAK-SHARPE, J. H. (1976). Control of protein synthesis in herpes-virus 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., STOW, N. D., PRESTON, V. G., TIMBURY, M. C. and WILKIE, N. M. (1978). Physical mapping of herpes simplex virus induced polypeptides. J. Virol. 28, 624-642.

MARSDEN, H. S., CAMPBELL, M. E. M., HAARR, L., FRAME, M. C., PARRISS, D. S., MURPHY, M., HOPE, R. G., MULLER, M. T. and PRESTON, C. M. (1987). The 65,000-M_r DNA binding and trans-inducing proteins of herpes simplex virus type 1. J. Virol. 61, 2428-2437.

MARTINEZ, R., SHAO, L. amd WELLER, S. K. (1992). The conserved helicase motifs of the herpes simplex virus type 1 origin-binding protein UL9 are important for function. J. Virol. 66, 6735-6746.

MASUR, H. (1992). Clinical implications of herpesvirus infections in patients with AIDS. American Journal of Medicine 92(2A), 1S-36S.

MATTHEWS, R. E. F. (1982). Classification and nomenclature of viruses. Fourth report of the International Committee on Taxonomy of Viruses. Intervirology 17, 1-199.

MAVRAMARO-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., SILVER, S., HUBENTHAL-MOSS, J., McKNIGHT, J. L. C. and ROIZMAN, B. (1987). Regulation of herpes simplex virus genes: α gene sequence requirements for transient induction of indicator genes regulated by β or late (γ_2) promoters. Virology 149, 152-164.

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.

MELLERICK, D. M. and FRASER, N. W. (1987). Physical state of the latent herpes simplex virus genome in a mouse model system: evidence suggesting an episomal state. Virology 158, 265-275.

MESSING, J. (1979). A multipurpose cloning system based on a single stranded DNA bacteriophage M13. Recombinant DNA Technical Bulletin 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., 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 virus. Science 239, 1531-1534.

MILLER, G. (1990). Epstein-Barr virus, in: "Virology" vol. 2, pp. 1921-1958, Fields, B. N. and Knipe, D. M. (ed.s), Raven Press, New York.

MIRDA, D. P., NAVARRO, D., PAZ, P., LEE, P. L., PEREIRA, L. and WILLIAMS, L. T. (1992). The fibroblast growth factor receptor is not required for herpes simplex virus type 1 infection. J. Virol. 66, 448-457.

MOCARSKI, E. S. and ROIZMAN, B. (1982). Structure and role of the herpes simplex virus DNA termini in inversion, circularization and generation of virion DNA. Cell 31, 89-97.

MORSE, B. L., BUCHMAN, T. G., ROIZMAN, B. and SCAFFER, P. A. (1977). Analysis of herpes simplex virus DNA. IX. Apparent exclusion of some parental DNA arrangementsin the generation of intertypic (HSV-1 x HSV-2) recombinants. J. Virol. 24, 231-248.

MUGGERIDGE, M. I., COHEN, G. H. and EISENBERG, R. J. (1992). Herpes simplex virus infection can occur without involvement of the fibroblast growth factor receptor. J. Virol. 66, 824-830.

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.

MURCHIE, M.-J. and McGEOCH, D. J. (1982). DNA sequence analysis of an immediate-early gene region of the herpes simplex virus type 1 genome (map coordinates 0.950 to 0.978). J. Gen. Virol. 62, 1-15.

MURRE, C., McCAW, P. S. and BALTIMORE, D. (1989a). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless, MyoD,* and *myc* proteins. Cell 56, 777-783.

MURRE, C., McCAW, P. S., VAESSIN, H., CAUDY, M., JAN, L. Y., JAN, Y. N., CABRERA, C. V., BUSHKIN, J. N., HAUSCHKA, S. D., LASSAR, A. B., WEINTRAUB, H. and BALTIMORE, D. (1989b). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell 58, 537-544.

NAZERIAN, K. (1974). DNA configuration in the core of Marek's disease virus. J. Virol. 13, 1148-1150.

NEUBERG, M., SCHUERMANN, M., HUNTER, J. B. and MULLER, R. (1989). Two functionally different regions in Fos are required for the sequence-specific DNA interaction of the Fos/Jun protein complex. Nature (London) 338, 589-590.

NIKOLOV, D. B., HU, S. H., LIN, J., GASCH, A., HOFFMAN, A., HORIKOSHI, M., CHUA, N. H., ROEDER, R. G. and BURLEY, S. K. (1992). Crystal structure of TFIID TATA-box binding protein. Nature 360, 40-46.

NISHIOKA, Y. and SILVERSTEIN, S. (1978). Requirement of protein synthesis for the degradation of host mRNA in Friend erythroleukaemia cells infected with herpes simplex virus type 1. J. Virol. 27, 619-627.

NOTARIANI, E. L. and PRESTON, C. M. (1982). Activation of cellular stress protein genes by herpes simplex virus temperature-sensitive mutants which overproduce immediate-early polypeptides. Virology 123, 113-122.

NURSE, P. (1990). Universal control mechanism regulating the onset of M-phase. Nature 344, 503-508.

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. (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. and HAYWARD, G. S. (1987). Comparison of upstream sequence requirements for positive and negative regulation of a herpes simplex virus immediate-early gene by three virus-encoded trans-acting factors. J. Virol. 61, 190-199.

O'HARE, P., GODING, C. R. and HAIGH, A. (1988). Direct combinatorial interaction between a herpes simplex virus regulatory protein and a celllular octamer-binding factor mediates a specific induction of virus-mediated early gene expression. EMBO J. 7, 4231-4238.

OLIVO, P. D., NELSON, N. J. and CHALLBERG, M. D. (1988). Herpes simplex virus DNA replication: the UL9 gene encodes an origin binding protein. Proc. Natl. Acad. Sci. (USA) 85, 5414-5418.

OLSON, L. C., BUESCHER, E. L. and ARTENSTEIN, M. S. (1967). Herpesvirus infections of the human central nervous system. New England Journal of Medicine. 277, 1271-1277.

ORASKAR, A. A. and READ, G. S. (1987). A mutant of herpes simplex virus type 1 exhibits increased stability of immediate-early (*alpha*) mRNA's. J. Virol. 61, 604-606.

OTTING, G., QIAN, Y.-Q., MULLER, M., AFFOLTER, M., GEHRING, W. J. and WUTHRICH, K. (1988). Secondary structure determination for the *Antennapedia* homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix motif. EMBO J. 7, 4305-4309.

PABO, C. O. and SAUER, R. T. (1984). Protein-DNA recognition. Ann. Rev. Biochem. 53, 293-321.

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.

PAPAVASSILLIOU, A. G. and SILVERSTEIN, S. J. (1990). 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.

PATEL, R., CHAN, W. L., KEMP, L. M., LaTHANGUE, N. B. and LATCHMAN, D. S. (1986). Isolation of cDNA clones derived from a cellular gene transcriptonally induced by herpes simplex virus. Nucleic Acids Res. 14, 5629-5640.

PATERSON, T. (1989). A mutational analysis of the structure and function of the herpes simplex virus immediate early protein Vmw175. PhD thesis. University of Glasgow, Scotland.

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 binds to the cap site of its own promoter *in vitro* but fails to autoregulate *in vivo*. J. Gen. Virol. 71, 851-861.

PERDUE, M. L., COHEN, J. C., RANDALL, C. C. and O'CALLAGHAN, D. J. (1976). Biochemical studies of the maturation of herpesvirus nucleocapsid species. Virology 74, 194-208.

PEREIRA, L., WOLFF, M. H., FENWICK, M. J. 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.

PERRY, L. J., RIXON, F. J., EVERETT, R. D., FRAME, M. C. and McGEOCH, D. J. (1986). The IE110 gene of herpes simplex virus type 1: characterisation by mRNA mapping, DNA sequence, oligopeptide antiserum and mutational analysis. J. Gen. Virol. 67, 2365-2380.

PERSSON, R. H., BACHETTI, S. and SMILEY, J. R. (1985). Cells that constitutively express the herpes simplex virus immediate-early protein ICP4 allow efficient activation of viral delayed-early genes in trans. J. Virol. 54, 414-421.

PETERSON, M. G. and TJIAN, R. (1992). The tell-tail trigger. Nature 358, 620-621.

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, M. G., INOSTROZA, J., MAXON, M. E., FLORES, O., ADMON, A., REINBERG, D. and TJIAN, R. (1991). Structure and functional properties of human general transcription factor IIE. Nature 354, 369-373.

PIZER, L. I. and BEARD, P. (1976). The effect of herpes virus infection on mRNA in polyoma-virus transformed cells. Virology 75, 477-480.

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(3), 477-483.

PIZER, L. I., DOERIG, C., HOEFFLER, J. P. and WILCOX, K. W. (1992). *In vitro* "complex" formation between the viral transactivator ICP4 and the TATA binding protein (TBP) component of the cellular transcription factor TFIID. Abstract, 17th. International Herpesvirus Workshop, Edinburgh, Scotland.

POLVINO-BODNAR, M., ORBERG, P. D. and SCHAFFER, P. A. (1987). Herpes simplex virus type 1 oriL is not required for virus replication or for the establishment and reactivation of latent infection in mice. J. Virol. 61, 3528-3535.

POST, L. E. and ROIZMAN, B. (1981). A generalised technique for deletion of specific genes in large genomes: α gene 22 of herpes simplex 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.

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, 275-284.

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, 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., 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.

PRESTON, V. G. (1981). Fine-structure mapping of herpes simplex virus type 1 temperature-sensitive mutations within the short repeat region of the genome. J. Virol. 39, 150-161.

PRESTON, V. G. (1990). Herpes simplex virus activates expression of a cellular gene by specific binding to the cell surface. Virology 176, 474-482.

PRESTON, V. G., COATES, J. A. V. and RIXON, F. J. (1983). Identification and characterisation of a herpes simplex virus gene product required for encapsidation of virus DNA. J. Virol. 45, 1056-1064.

PRESTON, V. G., RIXON, F. J., McDOUGALL, I. M., McGREGOR, M. and AL KOBAISI, M. F. (1992). Processing of the herpes simplex virus assembly protein ICP35 near its carboxy terminal end requires the product of the whole UL26 reading frame. Virology 186, 87-98.

PRUIJN, G. J. M., VAN DRIEL, W. and VAN DER VLIET (1986). Nuclear factor III, a novel sequence-specific DNA binding protein from HeLa cells stimulating adenovirus DNA replication. Nature 322, 656-659.

PTASHNE, M. (1988). How eukaryotic transcriptional activators work. Nature 335, 683-689.

PTASHNE, M. and GANN, A. A. F. (1990). Activators and Targets. Nature 346, 329-331.

PUGH, B. F. and TJIAN, R. (1990). Mechanism of trancriptional activation by Sp1: evidence for coactivators. Cell 61, 1187-1197.

PUGH, B. F. and TJIAN, R. (1992). Diverse transcriptional functions of the multisubunit eukaryotic TFIID complex. J. Biol. Chem. 267, 679-682.

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.

RAAB-TRAUB, N., FLYNN, K., PEARSON, G., HUANG, A., LEVINE, P., LANIER, A. and PAGANO, J. (1987). The differentiated form of nasopharyngeal carcinoma contains Epstein-Barr virus DNA. International Journal of Cancer. 39, 25-29.

READ, G. S. and FRENKEL, N. (1983). Herpes simplex virus mutants defective in the virion-associated shutoff of host polypeptide synthesis and exhibiting abnormal synthesis of α (immediate-early) polypeptides. J. Virol. 46, 498-512.

REINBERG, D., HORIKOSHI, M. and ROEDER, R. G. (1987). Factors involved in specific transcription in mammalian RNA polymerase II: functional analyses 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.

RESNICK, J., BOYD, B. A. and HAFFEY, M. L. (1989). DNA binding by the herpes simplex virus 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 α 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.

RIXON, F. J. (1992). Structure and assembly of herpesviruses. In press.

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.

RIXON, F. J., CROSS, A. M., ADDISON, C. and PRESTON, V. G. (1988). The products of herpes simplex virus type 1 gene UL26 which are involved in DNA packaging are strongly associated with empty but not with full capsids. J. Gen. Virol. 69, 2879-2891.

RIXON, F. J., ADDISON, C. and McLAUCHLAN, J. (1992). Assembly of enveloped tegument structures (L particles) can occur independently of virion maturation in herpes simplex virus type1-infected cells. J. Gen. Virol. 73, 277-284.

ROBERTS, M. S. and HAYWARD, G. S. (1991). Studies into the mechanism of IE175 (ICP4) autoregulation. Abstract, 17th. International Hepesvirus Workshop, Asilomar, California.

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 capsite 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.

ROCK, D. L. and FRASER, N. W. (1983). Detection of HSV-1 genome in central nervous system of latently infected mice. Nature 302, 523-525.

ROCK, D. L. and FRASER, N. W. (1985). Latent herpes simplex virus type 1 DNA contains two copies of the virion joint region. J. Virol. 55, 8449-8452.

ROEDER, R. G. (1991). The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. TIBS 16, 402-408.

ROFFMAN, E., ALBERT, J. P., GOFF, J. P. and FRENKEL, N. (1990). Putative site for the acquisition of human herpesvirus 6 virion tegument. J. Virol. 64, 6308-6313.

ROIZMAN, B. (1979). The structure and isomerization of herpes simplex virus genomes. Cell 16, 481-494.

ROIZMAN, B. (1982). The family herpesviridae; general description, taxonomy and classification, in: "The Herpesviruses" vol. 1, pp. 1-23, Roizman, B. (ed.), Plenum Press, New York.

ROIZMAN, B. (1990). Herpesviridae - a brief introduction, in: "Virology" vol. 2, pp. 1787-1792, Fields, B. N. and Knipe, D. M. (ed.s), Raven Press, New York.

ROIZMAN, B. and FURLONG, D. (1974). The replication of herpesviruses, in: "Comprehensive Virology" vol. 3, pp. 229-403, Fraenkel, H. and Wagner, R. R. (ed.s), Plenum Press, New York.

ROIZMAN, B. and ROANE, P. R. (1964). The multiplication of herpes simplex virus. II. The relation between protein synthesis and the duplication of viral DNA in HEp-2 cells. Virology 22, 262-269.

ROIZMAN, B. and SEARS, A. E. (1990). Herpes simplex viruses and their replication, in: "Virology" vol. 2, pp. 1796-1843, Fields, B. N. and Knipe, D. M. (ed.s), Raven Press, New York.

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.

RUSSELL, J. and PRESTON, C. M. (1986). An in vitro latency system for herpes simplex virus 2. J. Gen. Virol. 67, 397-403.

RUSSELL, J., STOW, E. C., STOW, N. D. and PRESTON, C. M. (1987). Abnormal forms of the herpes simplex virus immediate-early polypeptide Vmw175 induce the cellular stress response. J. Gen. Virol. 68, 2397-2406.

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.

RUYECHAN, W. T. and WEIR, A. C. (1984). Interaction with nucleic acids and stimulation of the viral DNA polymerase by the herpes simplex virus type 1 major DNA-binding protein. J. Virol. 52, 727-733.

SACKS, W. R. and SCHAFFER, P. A. (1987). Deletion mutants in the gene encoding the herpes simplex virus type 1 immediate early protein ICP0 exhibit impaired growth in tissue culture. J. Virol. 61, 829-839.

SACKS, W. R., GREENE, C. C., ASCHMAN, D. P. and SCAFFER, 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.

SALAHUDDIN, S. Z., ABLASHI, D. V., MARKHAM, P. D., JOSEPHS, S. F., STURZENEGGER, S., KAPLAN, M., HALLIGAN, G., BIBERFIELD, P., WONG-STAAL, F., KRAMARSKY, B. and GALLO, R. C. (1986). Isolation of a new virus, HBLV, in patients with lymphoproliferative disorders. Science 234, 596-601.

SALTZMAN, 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 herpes virus regulatory protein appears to act post-transcriptionally by affecting mRNA processing. Genes Dev. 6, 848-863.

SANDRI-GOLDIN, R. M., GOLDIN, A. L., HOLLAND, L. E., GLORIOSO, J. C. and LEVINE, M. (1983). Expression of herpes simplex virus β and γ genes integrated in mammalian and their induction by an α gene product. Mol. Cell. Biol. 3, 2028-2044.

SANGER, F., NICKLEN, S. and COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 82, 4394-4398.

SAYERS, J. R., SCHMIDT, W. and ECKSTEIN, F. (1988). 5'-3' exonucleases in phosphothiorate-based oligonucleotide-directed mutagenesis. Nucleic Acids Res. 16(3), 791-802.

SCHEK, N. and BACHENHEIMER, S. L. (1985). Degradation of cellular mRNA's induced by a virion-associated factor during herpes simplex virus infection of Vero cells. J. Virol. 55, 601-610.

SCHROEDER, C. H., STEGMAN, B., LAUPPE, H. F.. and KAERNER, H. C. (1975). An unusual defective genotype derived from herpes simplex virus strain ANG. Intervirology 46, 270-284.

SEARS, A. E., HALLIBURTON, I. W., MEIGNER, B., SILVER, S. and ROIZMAN, B. (1985). Herpes simplex virus 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.

SEDARATI, F., IZUMI, K. M., WAGNER, E. K. and STEVENS, J. G. (1989). Herpes simplex virus type 1 latency-associated transcription plays no role in establishment or maintenance of a latent infection in murine sensory neurones. J. Virol. 63, 4455-4458.

SEKULOVICH, R. E., LEARY, K. and SANDRI-GOLDIN, R. M. (1988). The herpes simplex virus alpha protein ICP27 can act as a *trans*-repressor or a *trans*-activator in combination with ICP4 and ICP0. J. Virol. 62, 4510-4522.

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 +39 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.

SHELDRICK, P. and BERTHELOT, N. (1974). Inverted repetitions in the chromosome of herpes simplex virus. Cold Spring Harbor Symposia on Quantitative Biology 39, 667-678.

SHEPARD, A. A. and DELUCA, N. A. (1989). Intragenic Complementation among Partial Peptides of Herpes Simplex Virus Regulatory Protein ICP4. J. Virol. 63(3), 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(1), 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(2), 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(9), 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(8), 3916-3926.

SHEPARD, J. C., McGINNIS, W., CARRASCO, A. E., DeROBERTIS, E. M. and GEHRING, W. J. (1984). Fly and frog homeo domains show homologies with yeast mating type regulatory proteins. Nature 310, 70-71.

SHERMAN, G., GOTTLIEB, J. and CHALLBERG, M. D. (1992). The UL8 subunit of the herpes simplex virus helicase-primase complex is required for efficient primer utilization. J. Virol. 66, 4884-4892.

SILVER, S. and ROIZMAN, B. R. (1985). γ-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.

SMIBERT, C. A., JOHNSON, D. C. and SMILEY, J. R. (1992). Identification and characterisation of the virion-induced host shutoff product of herpes simplex virus gene UL41. J. Gen. Virol. 73, 467-470.

SMILEY, J. R., SWAN, H., PATER, M. M., PATER, A. and HALPERN, M. E. (1983). Positive control of the herpes simplex virus thymidine kinase gene requires upstream DNA sequences. J. Virol. 47, 301-310.

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 a viral early gene in infected fibroblasts. J. Virol. 61, 2368-2377.

SMILEY, J. R., PANNING, B. and SMIBERT, C. A. (1991). Regulation of cellular genes by HSV products. In "Herpesvirus Transcription and its Regulation", pp. 151-180, Wagner, E. K. (ed.), CRC Press, Boston.

SMILEY, J. R., JOHNSON, D. G., 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.

SMILEY, J. R., LAVERY, C. and HOWES, M. (1992). The herpes simplex virus type 1 *a* sequence serves as a cleavage / packaging signal but does not drive recombinational genomic isomerization when it is inserted into the HSV-2 genome. J. Virol. 66, 7505-7510.

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, C. L. and DELUCA, N. (1992). Protein-protein and protein-DNA complexes formed between DNA, ICP4 of HSV-1, TFIID and TFIIB. Abstract, 17th. International Herpesvirus Workshop, Edinburgh, Scotland.

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.

SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517.

SPEAR, P. G. (1985). Glycoproteins specified by herpes simplex viruses, in "The Herpesviruses" vol. 3, pp. 315-356, Roizman, B. (ed.), Plenum Press, New York.

SPEAR, P. G. and ROIZMAN, B. (1972). Proteins specified by herpes simplex virus. V. Purification and structural proteins of the virion. J. Virol. 9, 143-160.

STACKPOLE, C. W. (1969). Herpes-type virus of the frog renal adrenocarcinoma. I. Virus development in tumour transplants maintained at low temperature. J. Virol. 4, 75-93.

STEINER, I., SPIVACK, J. G., LIRETTE, R. P., BROWN, S. M., McLEAN, A. R., SUBAK-SHARPE, J. H. and FRASER, N. W. (1989). Herpes simplex virus type 1 latency-associated transcripts are evidently not essential for latent infection. EMBO J. 8, 505-511.

STEINER, I., SPIVACK, J. G., DESHMANE, S. L., ACE, C. I., PRESTON, C. M. and FRASER, N. W. (1990). A herpes simplex virus type 1 mutant containing a non-transinducing Vmw65 protein establishes latent infection *in vivo* in the absence of viral replication and reactivates efficiently from explanted trigeminal ganglia. J. Virol. 64, 1630-1638.

STEVENS, J. G., WAGNER, E. K., DEVI-RAO, G. B., COOK, M. L. and FRIEDMAN, L. T. (1987). RNA complementary to a herpes α gene mRNA is prominent in latently infected neurones. Science 235, 1056-1059.

STOW, N. D. (1992). Herpes simplex virus type 1 origin-dependent DNA replication in insect cells using recombinant baculoviruses. J. Gen. Virol. 73, 313-321.

STOW, N. D. and McMONAGLE, E. C. (1983). Characterisation of the TRs/IRs origin of DNA replication of herpes simplex virus type 1. Virology 130, 427-438.

STOW, N. D. and STOW, E. C. (1986). Isolation and characterisation 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., McMONAGLE, E. C. and DAVISON, A. J. (1983). Fragments from both termini of the herpes simplex virus type 1 genome contain sequences required for the encapsidation of viral DNA. Nucleic Acids Res. 11, 8205-8220.

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

STUDIER, F. W. and MOFFATT, B. A. (1986). Use of bacteriophage T7 polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. 189, 113.

STUDIER, F. W., ROSENBERG, A. H., DUNN, J. J. and DUBENDORFF, J. W. (1990).The use of T7 polymerase to induce expression of cloned genes. Methods Enzymology 185, 60-89.

SYDISKIS, R. J. and ROIZMAN, B. (1966). Polysomes and protein synthesis in cells infected with a DNA virus. Science 153, 76-78.

SYDISKIS, R. J. and ROIZMAN, B. (1967). The disaggregation of host polyribosomes in productive and abortive infection with herpes simplex virus. Virology 32, 678-686.

SZILAGYI, J. F. and CUNNINGHAM, C. (1991). Identification and characterisation of a novel non-infectious herpes simplex virus-related particle. J. Gen. Virol. 73, 661-668.

TABIN, C. J. (1991). Retinoids, homeoboxes and growth factors: toward molecular models for limb development. Cell 66, 199-217.

TACKNEY, C., CACHIANES, G. and SILVERSTEIN, S. (1984). Transduction of the chinese hamster ovary *aprt* gene by herpes simplex virus. J. Virol. 52, 606-614.

TANESE, N., PUGH, B. F. and TJIAN, R. (1991). Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. Genes Dev. 5, 2212-2224.

TAPSCOTT, S. J., DAVIS, R. L., THAYER, M. J., CHENG, P.-F., WEINTRAUB, H. and LASSAR, A. B. (1988). *MyoD*: a nuclear phosphoprotein required a *myc* homology region to convert fibroblasts to myoblasts. Science 242, 405-411,

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, 4662-4672.

TEDDER, D. G., EVERETT, R. D., WILCOX, K. W., BEARD, P. and PIZER, L. (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.

TIMBURY, M. C. and SUBAK-SHARPE, J. H. (1973). Genetic interactions between temperature-sensitive mutants of types 1 and 2 of herpes simplex virus. J. Gen. Virol. 18, 347-357.

TIMMERS, M. and SHARPE, P. A. (1991). The mammalian TFIID protein is present in two functionally distinct complexes. Genes Dev. 5, 1946-1956.

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 specificity of homeodomain proteins. Cell 59, 553-562.

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

TRIEZENBERG, S. J., LaMARCO, K. L., and McKNIGHT, S. L. (1988b). Evidence of DNA:protein interactions that mediate HSV-1 immediate-early gene activation by VP16. Genes Dev. 2, 730-742.

TUFARO, F., GRUENHEID, S., GATZKE, L. and MEADOWS, H. (1992). Herpes simplex virus infection and propagation in in murine L cells lacking heparan sulphate proteoglycans. Abstract no. 24, 17th. International Herpesvirus Workshop, Edinburgh, Scotland.

TYLER, J. K. and EVERETT, R. D. (1993). The DNA binding domain of the varicella-zoster virus gene 62 protein interacts with multiple sequences which are similar to the binding site of the related protein of herpes simplex virus type 1. Nucleic Acids Res. 19, 513-522.

USHEVA, A., MALDONADO, E., GOLDRING, A., LU, H., HOUBAVI, C., REINBERG, D. and ALONI, Y. (1992). Specific interaction between the non-phosphorylated form of RNA polymerase II and the TATA-binding protein. cell 69, 871-881.

VARMUZA, S. L. and SMILEY, J. R. (1985). Signals for site-specific cleavage of HSV DNA: maturation involves two separate cleavage events at sites distal to the recognition sequences. Cell 41, 793-802.

VINSON, C. R., SIGLER, P. B. and McKNIGHT, S. L. (1989). Scissors-grip model for DNA recognition by a family of leucine-zipper proteins. Science 246, 911-916.

VLACK, J. and PITHA, P. M. (1992). Herpes simplex virus type 1-mediated induction of human immunodeficiency virus type 1 provirus correlated with the binding of nuclear proteins to the NF-κB enhancer and leader sequences. J. Virol. 66, 3616-3623.

VLAZNY, D. A., KWONG, A. and FRENKEL, N. (1982). Site specific cleavage and packaging of herpes simplex virus DNA and the selective maturation of nucleocapsids containing full length viral DNA. Proc. Natl. Acad. Sci. (USA) 79, 1423-1427.

WADSWORTH, S., JACOB, R. J. and ROIZMAN, B. (1975). Anatomy of herpes simplex virus DNA. II. Size, composition and arrangement of inverted terminla repetitions. J. Virol. 15, 1487-1497.

WAGNER, E. K. (1985). Individual HSV transcripts: characterisation of specific genes. In: "The Herpesviruses" vol. 3, pp. 45-104, Roizman, B. (ed.), Plenum Press, New York.

WAGNER, E. K., DEVI-RAO, G., FELDMAN, L. T., DOBSON, A. T., ZHANG, Y. F., FLANAGAN, W. M. and STEVENS, J. G. (1988a). Physical characterisation of the herpes simplex virus latency-associated transcript in neurons. J. Virol. 62, 1194- 1202.

WAGNER, E. K., FLANAGAN, W. M., DEVI-RAO, G., ZHANG, Y. F., HILL, J. M., ANDERSON, K. P. and STEVENS, J. G. (1988b). The herpes simplex virus latency-associated transcript is spliced during the latent phase of infection. J. Virol. 62, 4577-4585.

WAGNER, M. J. and SUMMERS, W. C. (1978). Structure of the joint regions and the termini of the DNA of herpes simplex virus type 1. J. Virol. 27, 453-467.

WATSON, R. J. and CLEMENTS, J. B. (1978). Characterisation 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.

WEBER, P. C. and WIGDAHL, B. (1992). Identification of dominant-negative mutants of the herpes simplex virus type 1 immediate-early protein ICP0. J. Virol. 66, 2261- 2267.

WECHSLER, S. L., NESBURN, A. B., WATSON, R., SLANINA, S. M. and GHIASA, H. (1988). Fine mapping of the latency-related gene of herpes simplex virus type 1: alternative splicing produces distinct latency-related RNAs containing open reading frames. J. Virol. 61, 4051-4058.

WEINHEIMER, S. D. 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.

WEIR, H. M. and STOW, N. D. (1990). Two binding sites for the herpes simplex virus UL9 protein are required for efficient activity of the oriS replication origin. J. Gen. Virol. 71, 1379-1385.

WEIR, H. M., CALDER, J. M. and STOW, N. D. (1989). Binding of the herpes simplex virus type 1 UL9 gene product to an origin of viral DNA replication. Nucleic Acids Res. 17, 1409-1425.

WELLER, S. K., SPARADO, A., SCHAFFER, J. E., MURRAY, A. W., MAXAM, A. M. and SCAFFER, P. A. (1985). Cloning, sequencing and functional analysis of oriL, a herpes simplex virus type 1 origin of DNA synthesis. Mol. Cell. Biol. 5, 930-942.

WHARTON, R. P. and PTASHNE, M. (1985). Changing the binding specificity of a repressor by redesigning an α-helix. Nature 316, 601-605.

WHITE, P. J., JACKSON, S. P. and RIGBY, P. W. J. (1992). A role for the TATA-box-binding protein component of the transcription factor IID complex as a general RNA polymerase III transcription factor. Proc. Natl. Acad. Sci. (USA) 89, 1949-1953.

WIGHDAL, B., SMITH, C. A., TRAGLIA, H. M. and RAPP, F. (1984). Herpes simplex virus latency in isolated human neurones. Proc. Natl. Acad. Sci. (USA) 81, 6217-6221.

WILCOX, K. W., KOHN, A., SKLYANSKY, 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. (1963). Electron microscope studies on the architecture of animal viruses. Cold Spring Harbor Symposium of Quantitative Biology 27, 25-47.

WILDY, P., RUSSELL, W. C. and HORNE, R. W. (1960). The morphology of herpes virus. Virology 12, 204-222.

WILDY, R. J. (1990). Herpes simplex viruses, in: "Virology" vol. 2, pp. 1843-1888, Fields, B. N. and Knipe, D. M. (ed.s), Raven Press, New York.

WILLIAMS, T. and TJIAN, R. (1991). Characterisation of a dimerization motif in AP-2 and its function in heterologous DNA-binding proteins. Science 251, 1067-1071.

WILLIAMS, T., ADMAN, A., LUSCHER, B. and TJIAN, R. (1988). Cloning and expression of AP-2, a cell-type specific transcription factor that activates inducible enhancer elements. Genes Dev. 2, 1557-1569.

WONG, S. W. and SCHAFFER, P. A. (1991). Elements in the transcriptional regulatory region flanking herpes simplex virus type 1 oriS stimulate origin function. J. Virol. 65, 2601-2611.

WORKMAN, J. L. and ROEDER, R. G. (1987). Binding of transcription factor TFIID to the major late promoter during in vitro nucleosome assembly potentiates subsequent initiation by RNA polymerase II. Cell 51, 613-622.

WORKMAN, J. L., ABMAYR, S. M., CROMLISH, W. A. and ROEDER, R. G. (1988). Transcriptional regulation by the immediate early protein of pseudorabies virus during in vitro nucleosome assembly. Cell 55, 211-219.

WORKMAN, J. L., TAYLOR, I. C. A. and KINGSTON, R. E. (1991). Activation domains of stably bound GAL4 derivatives alleviate repression of promoters by nucleosomes. Cell 64, 533-544.

WOYCHIK, N. A. and YOUNG, R. A. (1990). RNA polymerase II subunit structure and function. TIBS 15, 347-351.

WU, C-L. and WILCOX, K. W. (1990). Codons 262 to 490 from the herpes simplex virus ICP4 gene are sufficient to encode a sequence-specific DNA binding protein. Nucleic Acids Res. 18(3), 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 recognise 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.

WUDUNN, D. and SPEAR, P. G. (1989). Initial interactions of herpes simplex virus with cells is binding to heparan sulphate. J. Virol. 63, 52-58.

YAMANISHI, K., OKUNO, T., SKHIROKI, K., TAKAHASHI, M., KONDO, T., ASANO, Y. and KURATA, T. (1988). Identification of human herpesvirus 6 as a causal agent for exanthem subitum. Lancet 1, 1065-1067.

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

ZAWEL, L. and REINBERG, D. (1992). Advances in RNA polymerase II transcription. Current Opinion in Cell Biol. 4, 488-495.

ZEHRING, W. A., LEE, J. M., WEEKS, J. R., JOKERST, R. S. and GREENLEAF, A. L. (1988). The C-terminal repeat domain of RNA polymerase II largest subunit is essential *in vivo* but is not required for accurate transcription initiation *in vitro*. Proc. Natl. Acad. Sci. (USA) 85, 3698-3702.

ZHANG, Y.-F. and WAGNER, E. K. (1987). The kinetics of individual herpes simplex virus type 1 transcripts. Virus Genes 1, 49-60.

ZHU, L. and WELLER, S. K. (1992). The six conserved helicase motifs of the UL5 gene product, a component of the herpes simplex virus type 1 helicase-primase, are essential for its function. J. Virol. 66, 469-479.

ZWAAGSTRA, J. C., GHIASA, H., SLANINA, S. M., NESBURN, A. B., WHEATLEY, S. C., LILLYCROP, K., WOOD, J., LATCHMAN, D. S., PATEL, K. and WECHSLER, S. L. (1990). Activity of herpes simplex virus type 1 latency-associated transcript (LAT) promoter in neuron-derived cells: evidence for neuron specificity and a large LAT transcript. J. Virol. 64, 5019-5028.

ZWAAGSTRA, J. C., GHIASA, H., NESBURN, A. B. and WECHSLER, S. L. (1991). Identification of a major regulatory sequence in the latency-associated transcript (LAT) promoter of herpes simplex virus type 1 (HSV-1). Virology 182, 287-297.

GLASGOW UNIVERSITY LIBRARY